

## Genomics improves the reliability of Breeding Value Prediction of morphological and reproductive traits in the Pura Raza Español Horse: Preliminary results

C. Ziadi<sup>1</sup>, D. Perdomo-González<sup>2</sup>, M. Valera<sup>2</sup>, N. Laseca<sup>1</sup>, A. Encina<sup>3</sup>, P. Azor<sup>3</sup>, A. Rodríguez<sup>3</sup>, S. Demyda-Peyrás<sup>1</sup> and A. Molina<sup>1</sup>

<sup>1</sup>Department of Genetics, University of Córdoba. Córdoba, Spain

<sup>2</sup>Department of Agronomy, ETSIA, University of Sevilla, Sevilla, Spain

<sup>3</sup>Real Asociación Nacional de Criadores de Caballos de Pura Raza Española (ANCCE), Sevilla, Spain

Corresponding author: [anaencmar@gmail.com](mailto:anaencmar@gmail.com)

This study aimed to compare reliabilities for some conformational and reproductive traits between classical approach with pedigree-based REML and single-step genomic REML in PRE horse. Measurements from 5 zoometric traits were analysed: scapular-ischial length (SiL), length of back (LB), dorso-sternal diameter (DsD), thoracic perimeter (TP), and perimeter of anterior cannon bone (PACB). The following seven traits were considered in this study as measures of mare's fertility: age at first foaling (AFF), age at last foaling (ALF), average interval between foaling (AIF), total number of foalings (NF), interval between first and second foaling (IF12), productive life (PL) and reproductive efficiency (RE). The datasets consisted of 7152 conformation and 11,798 reproductive records and the pedigree included 41,888 animals. A total of 2916 animals were genotyped and 61,271 SNPs were included in the analysis. All analyses were performed using a multivariate model, separately for conformational and reproductive traits. The estimates of heritabilities were similar in both methodologies (0.34 to 0.64 for morphological traits, and 0.02 to 0.23 for the reproductive ones). Genetic correlations between EBVs and GEBVs varied from 0.59 (DsD) to 0.98 (SiL) for conformational traits and from 0.96 (IMPm) to 0.99 (PPm and EUPm) for reproductive traits.

A significant gain in reliabilities for ssGREML over REML evaluations has been observed in all conformational traits with an overall increase oscillating between 7.74% (LB) and 27.83% (DsD), being greater in genotyped animals (14.97% to 41.0%) compared to non-genotyped (6.84% to 26.13%). In general, when the reliability of the animal was previously very low, a greater gain in reliability has been observed, especially in the ungenotyped subpopulation. Similarly, considering only stallions with less than 40 controlled foals, this increase was much greater (7.92% to 26.43%). In the case of reproductive traits, a noticeable smaller increase was obtained for all traits, ranging from 1.99% for PL and RE to 4.29% for IF12. This work demonstrate the effectiveness of the genomic approach for the routine genetic evaluation of conformation and reproduction traits in the PRE breed..

*Keywords: single-step, estimated breeding value, reliability, PRE.*

### Abstract

## Introduction

The Pura Raza Española (PRE) horse is a native Spanish equine breed that has been officially recognized since the 15<sup>th</sup> century. Actually, the PRE is the most popular equine breed in Spain, representing the 70% of all registered equids. The total PRE population, 275,018 horses, are mainly located in Spain but also in other 67 countries (MAPA, 2023). In 2003, the genetic breeding program was approved to enhance the breed's conformation, functionality and reproductive traits, reduce inbreeding, and preserve its genetic heritage. More recently, the PRE breeding program has been updated in 2020 (MAPA, 2023). In order to ensure the PRE genetic integrity, various molecular tools such as blood groups, biochemical polymorphisms, and microsatellite markers have been used to establish paternity controls since the early 1980s. This has led to over 40 years of verified parental information, providing an accurate and reliable basis for genetic improvement efforts.

Conformation traits are very important to this breed because of its relationship with the functionality since PRE horses are mainly used for dressage competitions (Sánchez-Guerrero *et al.*, 2016). So, a good conformation will determine the horse's final price. In the same sense, reproductive traits are a critical factor for the profitability of equine stud. To maintain a healthy and profitable equine operation, successful breeding and reproduction are crucial. The birth of a foal not only represents the continuation of the bloodline but also signifies a considerable investment of time and resources. For all these reasons, obtaining reliable genetic parameters and high accuracy of estimated breeding values is essential in genetic improvement programs, especially for traits with low heritabilities, such as the reproductive ones.

Historically, in animal populations, estimated breeding values have been calculated from phenotype and pedigree information and extended literature exists, however, genomic estimates are expected to be more accurate because they do not depend on the quality and completeness of the pedigree. The single-step genomic best linear unbiased prediction (ssGREML) method allows the inclusion of information from genotyped and non-genotyped relatives in the analysis, along with phenotypes, improving the accuracy of breeding values estimation (Lourenco *et al.*, 2020).

Then, this study aimed to compare reliabilities for some conformational and reproductive traits between classical approach with pedigree-based REML and single-step genomic REML in PRE horse breed.

## Material and methods

### Pedigree, phenotypic and genotypic data

The datasets used in this study were provided by the Royal National Association of Spanish Horse Breeders (ANCCE). Data consisted on 7152 conformation and 11,798 reproductive records. The total number of animals in the pedigree used for the genetic and genomic evaluations was 41,888. A total of 2,916 individuals were genotyped with medium density GGP Equine Array (NEOGEN), including over 70,000 evenly distributed SNPs. The raw genotype data was filtered using PLINK software v1.9: SNPs with call-rate > 0.95 were retained. The final genomic data included 61,271 SNPs located on autosomal and X chromosomes. Measurements from five conformation zoometric traits were analysed: scapular-ischial length (SiL), length of back (LB), dorso-sternal diameter (DsD), thoracic perimeter (TP), and perimeter of anterior cannon bone (PACB). Also, the following seven traits were considered as measures of mare's fertility: age at first foaling (AFF), age at last foaling (ALF), average interval between foaling (AIF), total number of foalings (NF), interval between first and second foaling (IF12), productive life (PL) and reproductive efficiency (RE).

The following animal model was fitted to analyze conformation and reproductive traits separately:

$$y = Xb + Za + e$$

where **y** is the vector of observations of conformation and reproductive traits; **b** is the vector of fixed effects (for conformation traits: sex, age, coat colour and geographical area; for reproductive traits: classical inbreeding coefficient and age at last foaling as linear covariables, coat colour, geographical area and birth stud size); **a** is the random additive genetic effect, and **e** is the random residual effect. **X** and **Z** are incidence matrices relating observations to fixed and random additive genetic effects, respectively.

The additive genetic effect was modeled using two kinds of genetic covariance structures: for the classical evaluation, a matrix **A** that denotes the pedigree-based additive genetic relationship was used, and, for ssGREML, the **A** matrix was replaced by **H**, the pedigree-genomic relationship matrix. **H** was derived by blending **A** with the genomic relationship matrix **G** calculated using (VanRaden, 2008) as follows::

$$G = 0.95 \frac{SS'}{2 \sum_{i=1}^n p_i(1-p_i)} + 0.05A$$

where, *n* is the number of SNP markers and *p<sub>i</sub>* is the allele frequency of marker *i*, **A** is the pedigree relationship matrix, and **S** is a centred incidence matrix of SNP markers.

Variance components, estimated breeding values (EBVs) and genomic estimated breeding values (GEBVs) were estimated applying a restricted maximum likelihood (REML) and ssGREML approaches for conventional and genomic evaluations, respectively. Analyses were performed with the HiBlup v1.3.1 (Yin *et al.*, 2023) and the BLUPF90+ (Lourenco *et al.*, 2022) programs for conformation and reproduction traits, respectively..

This study has established, for the first time, a genomic evaluation using a single-step approach with a combined relationship matrix for conformation and female fertility in PRE horse.

Heritabilities were similar in both methodologies being high for conformation and low for reproductive traits. Their estimates ranged between 0.34 to 0.64 (DsD and SiL, respectively) for conformation traits and 0.02 to 0.23 (IF12 and AFF-RE, respectively) for the reproductive ones. In the literature, high heritability estimates were reported for conformation traits in horse breeds (Solé *et al.*, 2014 in Menorca horse population; Gómez *et al.*, 2021 in PRE horse breed; Vosgerau *et al.*, 2022 in German Warmblood horses). However, for fertility, lower heritability values were observed (Mantovani *et al.*, 2020; Perdomo-González *et al.*, 2021).

The gain in reliability (R<sup>2</sup>) between REML and ssGREML estimates of conformation and reproductive traits is shown in Table 1. Our results indicated a gain for ssGREML over REML evaluations and that gain was even greater in the case of conformation traits. The overall increase (considering all the animals in the pedigree) oscillated between 7.8% and 27.83% for conformation, and between 1.99% and 4.29% for reproductive traits.

Moreover, for conformation traits, more detailed information about the gain for ssGREML over REML using some criteria (sex, number of foals per sire, etc) is provided in Table 2. This gain was greater in mares (from 7.8% to 29.2%) than in stallions, and in genotyped animals (from 15% to 41%) compared to non-genotyped (from 6.8% to

### Estimation of genetic parameters, EBVs and GEBVs

### Results and discussion

Table 1. Gain in reliability ( $R^2$ ) between REML and SSGREML estimates of conformation and reproductive traits in the PRE horse.

	Trait	$R^2_{REML}$	$R^2_{SSGREML}$	Gain(%)
Conformation traits	SiL	0.282	0.308	9.12
	LB	0.258	0.278	7.80
	DsD	0.221	0.283	27.83
	TP	0.229	0.283	23.35
	PACB	0.229	0.259	12.83
Reproductive traits	PPm	0.381	0.389	2.04
	EUPm	0.263	0.270	2.90
	IMPm	0.101	0.105	3.81
	NP	0.264	0.272	2.86
	I12m	0.076	0.079	4.29
	VPm	0.092	0.094	1.99
	ER	0.388	0.395	1.99

26.1%). In addition, stallions with less than 40 controlled foals or having a previously low REML reliability a greater gain was observed.

The impact of genomic breeding values (GEBVs) on the accuracy of EBVs has been analysed in previous studies (Haberland *et al.*, 2012; Vosgerau *et al.*, 2022). Haberland *et al.* (2012) observed that, for animals with a large number of progeny records available, additional gain in accuracy from GEBV is small. In the same way, in the study of Vosgerau *et al.* (2022) about withers height, the increase in reliability was greater for animals with a small number of offspring.

Figure 1 represented the plot of the estimates of reliability for single-step GREML ( $R^2_{SSGREML}$ ) versus reliability for REML ( $R^2_{REML}$ ) in the trait with the higher overall gain for conformation and fertility, DsD and IF12, respectively. It was clearly noticeable that when the reliability of the animal was previously very low, a greater gain was reached for the DsD trait (Figure 1.A). However, for IF12, the gain in reliability was even unfavorable for animals that already have low reliabilities estimates for REML (Figure 1.B).

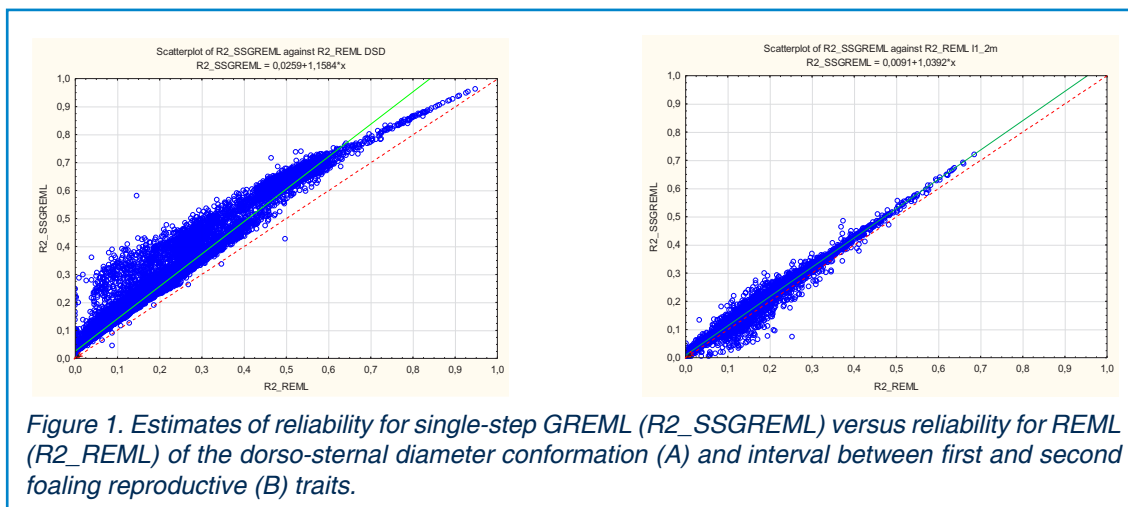
In general, this study demonstrated the effectiveness of the genomic approach for the routine genetic evaluations of conformation and reproduction traits in the PRE horse breed. The single-step method is an appealing approach for practical genomic prediction in PRE horses, because not many genotypes are available yet in this breed and animals without genotypes can by this way directly contribute to the estimation system..

Table 2. Gain in reliability ( $R^2$ ) between REML and ssGREML estimates of conformation traits based on different criteria in the PRE horse.

	Criteria	SiL	LB	DsD	TP	PACB
Sex	Stallions	9.08	7.75	26.1	22.1	12.2
	Mares	9.15	7.8	29.2	24.3	13.3
Stallion N°of foals	40 or more	3.5	2.2	15.2	13.3	4.9
	Less than 40	9.2	7.9	26.4	22.4	12.5
Genotyped	No	8.1	6.8	26.1	21.8	11.7
	Yes	16.2	15.0	41.0	35.1	21.6
Reliability	$\geq 0.6$	3.2	2.0	11.4	11.1	3.0
	$< 0.6$	35.6	25.9	44.7	38.8	24.5

Table 2. Gain in reliability ( $R^2$ ) between REML and ssGREML estimates of conformation traits based on different criteria in the PRE horse.

	Criteria	SiL	LB	DsD	TP	PACB
Sex	Stallions	9.08	7.75	26.1	22.1	12.2
	Mares	9.15	7.8	29.2	24.3	13.3
Stallion N°of foals	40 or more	3.5	2.2	15.2	13.3	4.9
	Less than 40	9.2	7.9	26.4	22.4	12.5
Genotyped	No	8.1	6.8	26.1	21.8	11.7
	Yes	16.2	15.0	41.0	35.1	21.6
Reliability	$\geq 0.6$	3.2	2.0	11.4	11.1	3.0
	$< 0.6$	35.6	25.9	44.7	38.8	24.5



The authors would like to thank the Royal National Association of Spanish Horse Breeders (ANCCE) for providing the data used in this study. This work has been financed with FEDER funds by the EQUIGENOM Operational Group (Ministry of Agriculture, Fisheries and Food, through the Spanish Agrarian Guarantee Fund, FEAGA).

This study has established, for the first time, a genomic evaluation using a single-step approach with a combined relationship matrix for conformation and female fertility in PRE horse. Our results demonstrated the effectiveness of the genomic approach for the routine genetic evaluation of conformation and reproduction traits in the PRE breed.

## Conclusion

The authors would like to thank the Royal National Association of Spanish Horse Breeders (ANCCE) for providing the data used in this study. This work has been financed with FEDER funds by the EQUIGENOM Operational Group (Ministry of Agriculture, Fisheries and Food, through the Spanish Agrarian Guarantee Fund, FEAGA).

## Acknowledgments

## References

- Gómez, M. D., A. Molina, M. J. Sánchez-Guerrero, and M. Valera.** 2021. Prediction of adult conformation traits from shape characteristics of Pura Raza Español foals. *Livestock Science* 253:104701. doi: <https://doi.org/10.1016/j.livsci.2021.104701>
- Haberland, A. M., U. König von Borstel, H. Simianer, and S. König.** 2012. Integration of genomic information into sport horse breeding programs for optimization of accuracy of selection. *Animal* 6(9):1369-1376. doi: <https://doi.org/10.1017/S1751731112000626>
- Lourenco, D., A. Legarra, S. Tsuruta, Y. Masuda, I. Aguilar, and I. Misztal.** 2020. Single-Step Genomic Evaluations from Theory to Practice: Using SNP Chips and Sequence Data in BLUPF90. *Genes* 11(7)doi:10.3390/genes11070790
- Lourenco, D., S. Tsuruta, Y. Masuda, M. Bermann, A. Legarra, and I. Misztal.** 2022. Recent updates in the BLUPF90 software suite. In: Proceedings of the 12th world congress on genetics applied to livestock production; July 3–8; Rotterdam, Netherlands
- Mantovani, R., F. Folla, G. Pigozzi, S. Tsuruta, and C. Sartori.** 2020. Genetics of Lifetime Reproductive Performance in Italian Heavy Draught Horse Mares. *Animals* 10(6)doi:10.3390/ani10061085
- MAPA.** 2023. Resolución de la Dirección General de Producciones y Mercados Agrarios, por la que se aprueba el Programa de Cría del Caballo de Pura Raza Española (PRE) y el Programa de Difusión de la Mejora. Available at: [https://www.mapa.gob.es/es/ganaderia/temas/zootecnia/report\\_1592238387802-resoluciondeladgporlaqueseapruebaelprogramadecriadelcaballopre\\_tcm30-539933.pdf](https://www.mapa.gob.es/es/ganaderia/temas/zootecnia/report_1592238387802-resoluciondeladgporlaqueseapruebaelprogramadecriadelcaballopre_tcm30-539933.pdf)
- Perdomo-González, D. I., A. Molina, M. J. Sánchez-Guerrero, E. Bartolomé, L. Varona, and M. Valera.** 2021. Genetic inbreeding depression load for fertility traits in Pura Raza Española mares. *J Anim Sci* 99(12)doi: 10.1093/jas/skab316
- Sánchez-Guerrero, M. J., A. Molina, M. D. Gómez, F. Peña, and M. Valera.** 2016. Relationship between morphology and performance: Signature of mass-selection in Pura Raza Español horse. *Livestock Science* 185:148-155.
- Solé, M., I. Cervantes, J. Gutiérrez, M. D. Gomez, and M. Valera.** 2014. Estimation of genetic parameters for morphological and functional traits in a Menorca horse population. *SPANISH JOURNAL OF AGRICULTURAL RESEARCH* 12:125-132. doi: 10.5424/sjar/2014121-4686
- VanRaden, P. M.** 2008. Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science* 91(11):4414-4423. doi: <https://doi.org/10.3168/jds.2007-0980>
- Vosgerau, S., N. Krattenmacher, C. Falker-Gieske, A. Seidel, J. Tetens, K. F. Stock, W. Nolte, M. Wobbe, I. Blaj, R. Reents, C. Kühn, M. von Depka Prondzinski, E. Kalm, and G. Thaller.** 2022. Genetic and genomic characterization followed by single-step genomic evaluation of withers height in German Warmblood horses. *Journal of applied genetics* 63(2):369-378. doi: 10.1007/s13353-021-00681-w
- Yin, L., H. Zhang, Z. Tang, D. Yin, Y. Fu, X. Yuan, X. Li, X. Liu, and S. Zhao.** 2023. HIBLUP: an integration of statistical models on the BLUP framework for efficient genetic evaluation using big genomic data. *Nucleic Acids Research* 51(8):3501-3512. doi: 10.1093/nar/gkad074