Animal 16 (2022) 100476

Contents lists available at ScienceDirect

Animal
The international journal of animal biosciences

A genome-wide association study of mare fertility in the Pura Raza Español horse

N. Laseca [a,*], S. Demyda-Peyrás [b,c], M. Valera [d], M. Ramón [e], B. Escribano [f], D.I. Perdomo-González [d], A. Molina [a]

[a] Departamento de Genética. Facultad de Veterinaria, Universidad de Córdoba, Campus de Rabanales, CN-IV km 396, 14071 Córdoba, España
[b] Departamento de Producción Animal, Facultad de Ciencias Veterinarias, Universidad Nacional de La Plata, Calle 60 y 118 s/n, La Plata 1900, Argentina
[c] Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET) La Plata, La Plata 1900, Argentina
[d] Departamento de Agronomía, Escuela Técnica Superior de Ingeniería Agronómica. Universidad de Sevilla, Ctra. Utrera, Km 1, Sevilla, Spain
[e] Centro Regional de Selección y Reproducción Animal (CERSYRA), Av. del Vino, 10, 13300 Valdepeñas, Ciudad Real, Spain
[f] Departamento de Fisiología, Universidad de Córdoba, Campus de Rabanales, CN-IV km 396, 14071 Córdoba, Spain

ARTICLE INFO

Article history:
Received 10 December 2021
Revised 25 January 2022
Accepted 27 January 2022
Available online 2 March 2022

Keywords:
Association analysis
Candidate genes
Equine
Reproductive traits
Single-nucleotide polymorphisms

ABSTRACT

Despite the economic importance of fertility for the horse industry, few efforts have been made to achieve a better understanding of the genetic mechanisms underlying its control. This is probably due to the difficulty of obtaining reliable phenotypes and the complexity of modelling the environmental and management factors. This work is novel in that we propose to use reproductive efficiency (**RE**) as an indicator of mare fertility. To achieve this, we performed a genome-wide association study in the Pura Raza Español horse aimed at identifying genomic variants, regions, and candidate genes associated with fertility in mares. The dataset included 819 animals genotyped with the Affymetrix Axiom™ Equine 670 K single-nucleotide polymorphisms (**SNPs**) Genotyping Array and the deregressed breeding values for RE trait, obtained using a ssBLUP model, employed as pseudo-phenotypic data. Our results showed 28 SNPs potentially associated with RE, which explained 87.19% of the genetic variance and 6.61% of the phenotypic variance. Those results were further validated in BayesB, showing a correlation between observed and predicted RE of 0.57. In addition, 15 candidate genes (*HTRA3, SPIRE1, APOE, ERCC1, FOXA3, NECTIN-2, KLC3, RSPH6A, PDPK1, MEIOB, PAQR4, NM3, PKD1, PRSS21, IFT140*) previously related to fertility in mammals were associated with the markers and genomic regions significantly associated with RE. To our knowledge, this is the first genome-wide association study performed on mare fertility.

Implications

Reproductive traits are a critical factor for the profitability of equine farms. However, the information available is limited due to the difficulty in obtaining reliable phenotypes and the complexity of modelling the environmental factors. This is the first attempt to perform a genome-wide analysis study focused on mare fertility using a large cohort of horse genotypes and a pseudo-phenotypic value obtained by analysing more than 344 000 reproductive records of Pura Raza Español horses. We determined (and validated) the existence of several candidate regions and genes that might provide insights for genomic or marker-assisted selection in the mares reproduction.

Introduction

Fertility is a key factor in the economic success of livestock production systems. Nevertheless, the horse is probably the domestic species in which natural and artificial selection has had the least influence in fertility. Since its domestication about 8 000 years ago (Moazemi et al., 2020; Orlando, 2020), the horse has been used for work, warfare, leisure or sport, activities in which fertility is often considered less important than other traits (Palmer and Chavatte-Palmer, 2020). For that reason, it is not often included in breeding programmes as a selection objective. In addition, the difficulty of accounting for a criterion with sufficient heritability as a measure of reproductive aptitude in mares also hinders its inclusion in breeding programmes as well as scientific efforts. However, it has been demonstrated that fertility is still a critical factor for the profitability of horse farms (Gómez et al., 2020). In addition, the reproductive management and environmental condi-

* Corresponding author.
E-mail address: ge2lagan@uco.es (N. Laseca).

https://doi.org/10.1016/j.animal.2022.100476
1751-7311/

tions have a crucial effect on the reproductive efficiency of mares. Even though fertility is considerably lower in horses than in other domestic species (Perdomo-González et al., 2021), the existence of a genetic component affecting this trait is well-demonstrated (Gómez et al., 2020; Mantovani et al., 2020; Todd et al., 2020).

To date, few studies have been carried out which allow us to better understand the genomic mechanisms underlying fertility in mares (Laseca et al., 2021a; 2022). This is probably due to the lack of large, reliable phenotypic datasets, which impairs the modelling of environmental and management factors (e.g. age, nutrition, training, temperature at mating and breeding season, etc.) affecting this character. Furthermore, fertility is a complex polygenic trait with low heritability (Mahon and Cunningham, 1982) and is influenced by a large number of genes, each with a small absolute effect, and due to the high level of linkage disequilibrium between genomic variants, shown particularly in the Pura Raza Español (**PRE**) breed (Poyato-Bonilla et al., 2022), is it difficult to pinpoint causal variants for complex traits. For this reason, most of the reproductive traits usually employed to evaluate the fertility of mares (total foaling number, age at first foaling, average interval between first and second foaling, average inter-foaling interval, age at last foaling, and productive life (Gómez et al., 2020; Perdomo-González et al., 2021) tend to have low heritabilities. Therefore, its use as a selection criterion will only produce moderate results in terms of phenotypic improvement. However, Perdomo-Gonzalez et al., 2020 have recently developed a new phenotypic trait (reproductive efficiency, **RE**), which is able to estimate the fertility of a mare with great accuracy based on the analysis of pedigree records. This trait, which was validated using a large sample population of 300 000 foaling records, showed a moderate to high heritability (h_2 = 0.23) but of a greater magnitude than the classics trait previously used as reproductive criteria in mares (Sairanen et al., 2009; Wolc et al., 2009; Gómez et al., 2020). This parameter, which could be easily estimated in all the mares existing in a studbook, is an interesting candidate to be included in a breeding programme with the aim of improving the maternal fertility of the breed. However, it also allows its use in genomic studies aimed at determining genomic regions associated with fertility in mares.

Nowadays, advances in high-throughput genotyping technologies (medium and high-density chips) and sequencing technologies have enabled us to identify many single-nucleotide polymorphisms (**SNPs**) associated with phenotypic traits (Laseca et al., 2021a). This genomic revolution has produced significant advances in our understanding of complex traits. Most of these associations were determined by performing genome-wide association studies (**GWAS**), which are now considered the most powerful tool to screen and determine (at least partially) the genomic architecture of qualitative and quantitative traits, thus improving the accuracy and persistency of genomic selection (VanRaden et al., 2017). In horses, several GWAS studies have recently been reported for several traits, such as conformation (Al Abri et al., 2018), racing performance (Pereira et al., 2018), gait (Fonseca et al., 2017), or diseases (Shrestha et al., 2020). However, despite the existence of GWAS in female reproductive traits in several livestock species, such as cattle (Keogh et al., 2021), pigs (Wang et al., 2018), sheep (Smołucha et al., 2021), and goat (Islam et al., 2020), these kinds of reports in horses have only been performed in stallions (Gottschalk et al., 2016; Gmel et al., 2021), and so far not in mares.

The Pura Raza Español is one of the oldest European horse breeds. Its breeding programme, which nowadays includes over 260 000 individuals from 63 countries, is managed by the Asociación Nacional de Criadores de Caballos de Pura Raza Española (**ANCCE**). Historically, PRE horses have been selected based mostly on aesthetic criteria. Nevertheless, since three decades ago, several additional criteria related to aptitude for sportive disciplines (mainly in classical dressage), as well as morphology (in relation to functionality through linear morphological qualification) have been included as selection objectives in the breeding programme. However, our research group has recently analysed several new traits focused on the fertility of individuals (Gómez et al., 2020; Perdomo-Gonzalez et al., 2020; 2021) which are nowadays being taken into consideration in the breeding selection criteria for fertility. This fact, together with the long, reliable pedigree (kept by the ANCCE since 1912), and the existence of a reliable, well-developed breeding scheme, makes this breed an interesting model to analyse the fertility of the horses from a genetic point of view.

The aim of this study was therefore to perform the first GWAS for reproductive efficiency traits in mares to identify genetic variants, genomic regions, and candidate genes associated with fertility, using deregressed breeding values for the reproductive efficiency trait as pseudo-phenotypic data. We hope this study will contribute to a better understanding of mare fertility in horses.

Material and methods

Phenotypic recording of the mare fertility

The fertility of 78 986 Pura Raza Español mares was determined using the reproductive efficiency trait defined as the percentual deviation between the optimal and actual parity numbers of the mare at each age. RE includes all the mare's recorded foaling throughout her lifetime, i.e. up to her last known age. This trait, which was developed and validated as an indicator of fertility in PRE mares, has produced the highest heritability (0.23) among several traits recently analysed in this breed (Perdomo-Gonzalez et al., 2020).

Reproductive efficiency was estimated individually by analysing all the information available in the PRE studbook (344 707 foaling records from 78 986 PRE breeding mares bred in 63 countries collected between 1970 and 2020). As a first step, we selected all the records of the mares belonging to studs whose main activity was the production of foals (more than 10 foals produced per stud per year). Secondly, we pruned out all the mares employed mainly for leisure or sports activities and kept only the individuals who had had their first foal between 4 and 7 years old, and who had intervals between first and second foaling and last and penultimate foaling of less than 5 years. The final dataset employed in RE estimation included 23 899 mares belonging to 8 133 studs.

Estimation of genetic parameters for reproductive efficiency

Reproductive efficiency breeding values ($\mathbf{RE_{EBV}}$) were estimated using an Single Step REML Animal model including the stud, year of birth and size of the herd of birth of the mare as fixed effects, age and inbreeding of the mare as lineal covariates, and additive effect of the mare and residual effect as random effects.

The extended pedigree of the mares (with all known generations) included 87 227 animals and was obtained from ANCCE (Asociación Nacional de Criadores de Caballos de Pura Raza Español) studbook. The number of maximum known generations of these mares was 16, with an average of 5.26 complete generations and 8.81 equivalent generations. In addition, a pedigree-genomic-based hybrid relationship matrix (matrix H) was constructed.

In the final step, RE_{EBV} showing an accuracy higher than 0.5 were deregressed to $\mathbf{RE_{dEBV}}$ (pseudo-phenotypes) following the procedure described by Garrick et al., 2009, in order to avoid the double imputation of the genetic effect on the GWAS. These values were used as pseudo-phenotypes in GWAS for fertility. All the genetic estimations, including RE_{EBV} and $RE_{dEBV,}$ were obtained

using the Renumf90, PreGSF90, Airemlf90, and Deproofsf90 modules from the BLUPF90 software family (Misztal et al., 2016).

Genotyping and quality control

In total, 819 horses belonging to the PRE studbook were selected for genotyping. The individuals were selected from 373 herds, based on a low average relatedness with the rest of the animals selected, a minimum of 60% of accuracy in the estimation of pseudo-phenotypes, and a balanced number of individuals with divergent fertility values (low and high).

Genomic DNA was isolated from blood or hair samples using a DNeasy Blood & Tissue extraction kit (Qiagen, Germantown, MD, USA). Next, we assessed the quality and quantity of nucleic acid by electrophoresis and spectrophotometry. Horses were genotyped with the Affymetrix Axiom™ Equine 670 K SNP Genotyping Array (Thermofisher, Spain), including 670 804 markers uniformly distributed across the entire genome (Schaefer et al., 2017). We then processed the raw genotype data following the "Best Practices Workflow" procedure in the Axiom Analysis Suite package v5.0 with default parameter (DishQC $\geq$ 0.82). Next, the quality control of the genotypes was performed using PLINK software v1.9 (Purcell et al., 2007). First, we removed SNPs located on sex chromosomes and those that showed a minor allele frequency < 0.01 and a call rate < 0.95. Finally, the SNP dataset was pruned by linkage disequilibrium using a window size of 50 SNPs, 5 SNPs shifted per step, and an r^2 threshold of 0.5. The final genomic dataset for GWAS analysis included 158 974 SNPs.

Genome-wide association analysis

A genome-wide association study was performed to test the relationship between individual SNPs and RE, using the deregressed RE breeding values as pseudo-phenotypes in the 496 individuals showing a RE_{dEBV} reliability > 0.75. Among these, 81.45% showed positive RE_{dEBV} values, while the remaining 18.55% showed negative values. GWAS analysis was implemented in GEMMA software (Zhou and Stephens, 2012), employing the following univariate linear mixed model:

$$\mathbf{y} = \mathbf{W}\boldsymbol{\alpha} + \mathbf{X}\beta + \boldsymbol{u} + \varepsilon$$

where **y** is an n-vector with pseudo-phenotypes; **W** is an incidence matrix of covariates (fixed effects) including a column of 1 s; $\boldsymbol{\alpha}$ is a vector of the corresponding coefficients including the intercept; **X** is an n-vector of marker genotypes; $\boldsymbol{\beta}$ is the effect size of the marker; **u** is an n-vector of random effects $\mathbf{u} \sim N(0, \lambda\tau - 1\ K)$, where $\tau - 1$ is the variance of the residual errors; λ is the ratio between two variance components; **K** is the genomic relationship matrix (estimated from the markers), and ε is the n-vector of errors. In addition, the model included a correction for population stratification based on the first ten components of a principal component analysis performed in PLINK v.1.9 as covariates.

To determine the existence of genomic inflation in the dataset, we estimated the λ inflation factor and performed a quantile–quantile (**QQ**) plot assessing the distribution of expected and observed SNP –log10(*P*-values). To do this, λ was calculated as the median or mean of the χ^2 test statistics divided by its theoretical median or mean under the null distribution (Devlin and Roeder, 1999), while the QQ plot was generated using *qqman* and *GWASTools* packages from R statistical environment V4.1.1 (R-Core-Team, 2021).

In the final step, the statistical significance of the SNP effect was calculated using a Wald test statistic and a *P*-value for each SNP implemented, using GEMMA software.

Statistical analyses of post-genome-wide association analysis

First, we performed a false discovery rate (**FDR**) correction for multiple testing at the chromosome level, using the methodology described by Benjamini and Hochberg, 1995. In addition, the proportion of phenotypic variance explained (**PVE**) by each significant SNP was estimated according to Velie et al. (2018) as follows:

$$PVE\ (\%) = \frac{2q(1-q)\beta^2}{S^2} \times 100$$

where $\boldsymbol{q}$ is the minor allele frequency of the SNP, $\boldsymbol{\beta}$ is the estimated effect of the SNP, and $\mathbf{S^2}$ is the sample phenotypic variance. Finally, we estimated the proportion of genetic variance explained by each SNP by substituting the phenotypic variance ($\mathbf{S^2}$) for the sample dEBV variance. All the statistical analyses were carried out in the R environment using the *tidyverse* and *data.table* packages.

Bayesian genomic prediction

The results were validated in the reference population (819 animals) by estimating the effects of significant SNPs and RE_{dEBV} using a BayesB whole-genome regression model (Xu et al., 2021), assuming the existence of a large number of loci with zero genetic variance and only a small proportion of loci with variance not equal to zero. The BayesB model was implemented in the *BGLR* package of R using a chain length of 50 000 iterations, and a burn-in of 10 000. Finally, the prediction accuracy was estimated using the average Pearson's correlation (**r**) coefficient between the predicted and observed values.

Functional analysis of genomic regions with significant association

To identify potential candidate genes associated with RE, we first measured all the genomic intervals located ±1 Mb upstream and downstream of each significant SNP. Next, all the genes located within those regions were retrieved using the *Ensembl BioMart* in the last available horse reference genome assembly (EquCab3.0. http://www.ensembl.org/Equus_caballus/Info/Index). Finally, the function of these genes and their putative relationship with fertility processes was established by performing an extensive review of the available literature in public databases, as well as in the DAVID V6.8 and Uniprot online resources.

Results

The deregressed breeding values for RE of the 496 PRE horses included in the GWAS ranged from −15.17 to 25.53, with an average of 5.35. Genome-wide association analyses detected 28 SNPs potentially associated to RE (*P*-value < 10^{-4}, Table 1). The markers were distributed across 17 different chromosomes (**ECA**), supporting the polygenic basis described in reproductive traits (Fig. 1). No evidence of data inflation was observed either in the QQ plot (Fig. 2) or the λ_{median} value (0.999), indicating a good concordance between observed and assumed distributions of the test statistics. It is worth mentioning that the clearest signals of association were observed on ECA17 and ECA20, but no significant association with RE was detected in those regions after FDR correction.

The average effect of each SNP estimated in absolute value was 3.01 (ranging from 1.28 to 8.44) and the average proportion of PV explained by the associated markers was 0.24 (ranging from 0.08 to 0.34) (Table 1). The sum of PV explained by the SNPs was 6.61%. However, the proportion of the estimated genetic variance of the 28 associated SNPs ranged from 0.99 to 4.48, with an average of 3.11 (Table 1).

Table 1
List of SNPs associated with reproductive efficiency in the Pura Raza Español horse.

SNP	ECA	Position (bp)	SNP effect	SE	*P*-value (*)[1]	Minor allele frequency	PVE (%)	GVE (%)
rs1142746103	1	30 242 107	2.02	0.51	$8.80\ 10^{-5}$	0.30	0.26	3.44
rs1148455393	1	81 957 660	−5.94	1.51	$9.41\ 10^{-5}$	0.02	0.22	2.94
rs1146383937	2	93 027 247	2.12	0.53	$6.99\ 10^{-5}$	0.31	0.3	3.88
rs68634906	3	81 730 767	1.83	0.47	$9.98\ 10^{-5}$	0.37	0.24	3.13
rs1140471861	3	117 731 559	−1.28	0.28	$6.07\ 10^{-6*}$	0.19	0.08	0.99
rs1149365591	4	24 893 574	2.95	0.72	$4.56\ 10^{-5}$	0.13	0.31	4.02
rs1139587939	5	58 267 844	−8.44	1.88	$9.10\ 10^{-6*}$	0.02	0.34	4.48
rs396994746	6	15 616 050	1.99	0.46	$1.79\ 10^{-5}$	0.49	0.30	3.95
no-rs	6	28 033 876	−2.1	0.53	$8.19\ 10^{-5}$	0.12	0.14	1.79
rs395182710	7	31 456 709	2.03	0.51	$7.07\ 10^{-5}$	0.29	0.26	3.42
rs68693538	7	90 469 286	−3.95	1.01	$9.96\ 10^{-5}$	0.04	0.19	2.44
rs1140299868	8	40 478 963	−2.54	0.57	$7.84\ 10^{-6*}$	0.09	0.16	2.14
rs1147356233	9	69 241 563	2.3	0.54	$3.88\ 10^{-5}$	0.27	0.3	3.91
rs1139996102	9	70 541 826	1.80	0.46	$8.37\ 10^{-5}$	0.45	0.25	3.23
rs397493565	10	16 020 355	5.88	1.30	$8.02\ 10^{-6*}$	0.02	0.24	3.15
rs1149964207	10	51 203 426	3.92	0.96	$5.42\ 10^{-5}$	0.07	0.30	3.97
rs1141581080	13	41 513 531	−2.55	0.56	$6.99\ 10^{-6*}$	0.09	0.16	2.17
rs1142582346	14	51 715 703	−3.55	0.90	$8.33\ 10^{-5}$	0.06	0.21	2.79
rs1139589610	17	389 302	−2.3	0.53	$1.88\ 10^{-5}$	0.20	0.26	3.40
rs396779351	17	1 070 230	−1.91	0.45	$3.08\ 10^{-5}$	0.32	0.24	3.18
rs1150725391	17	2 490 190	−5.66	1.43	$8.49\ 10^{-5}$	0.03	0.28	3.77
rs69122238	17	3 914 866	−1.76	0.44	$6.56\ 10^{-5}$	0.41	0.23	3.01
rs1141870973	17	9 360 896	−2.44	0.59	$4.13\ 10^{-5}$	0.17	0.25	3.34
rs1147007488	18	13 798 418	−2.49	0.56	$1.27\ 10^{-5}$	0.1	0.16	2.14
rs1148440103	20	32 416 135	−2.47	0.57	$1.59\ 10^{-5}$	0.09	0.15	2.04
rs1141483422	20	45 869 029	−3.14	0.76	$4.51\ 10^{-5}$	0.11	0.3	3.93
rs395430961	23	12 503 982	−3.06	0.76	$5.94\ 10^{-5}$	0.11	0.27	3.62
rs396082257	24	37 109 149	2.04	0.52	$9.29\ 10^{-5}$	0.23	0.23	3

Abbreviations: SNPs = single-nucleotide polymorphisms; ECA = equine chromosome; bp = base pairs; PVE = proportion of the phenotypic variance explained by each SNP; GVE = proportion of the genetic variance explained by each SNP
[1] (*) significant false discovery rate (FDR).

However, only five significant SNPs located in four different chromosomes were found after performing FDR correction (Fig. 1). For these, the estimated average effect of the significant SNPs in absolute value was 4.13. Interestingly, the most significant SNP *rs1140471861* (3:117 731 559, *P*-value = $6.07\ 10^{-6}$) explained the lowest proportion of the phenotypic variance. In contrast, *rs113958793* (located on ECA5) showed the largest estimated effect in absolute value and the strongest phenotypic proportion of estimated variance.

Candidate genes and validation

In total, we found 15 candidate genes previously linked with known biological processes, molecular functions and pathways related to fertility within the genomic intervals of 4 SNPs (Table 2), among which there were processes related to ovarian and oocyte development, but also processes linked to sperm physiology.

Finally, to validate our findings, we performed linear regression with a BayesB model on the reference population (819 animals), using the dataset of SNPs significantly associated with RE. The correlation obtained between the pseudo-phenotype and the phenotype predicted by the BayesB model was 0.57.

Discussion

This study aimed to find an association between genomics and fertility in Pura Raza Español mares, in an attempt to identify genomic variants, regions, and candidate genes involved in the control

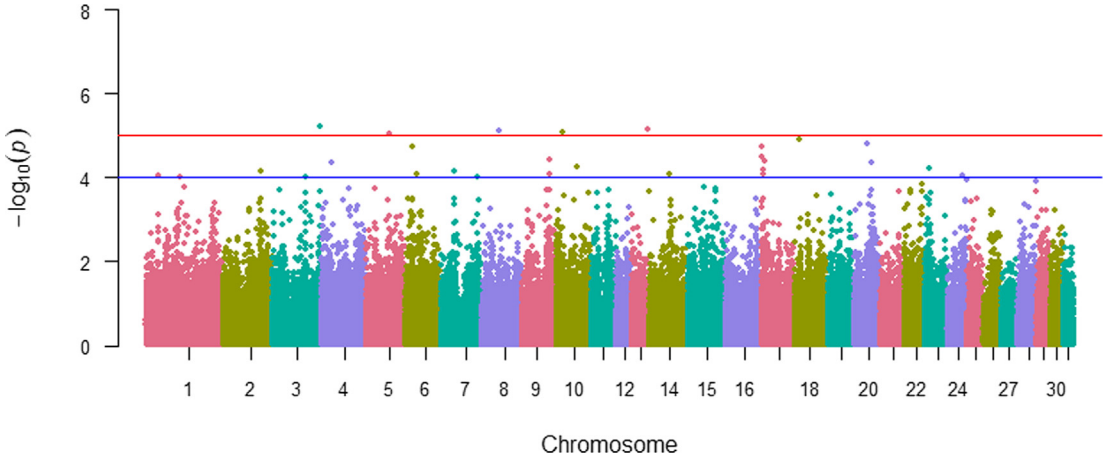

Fig. 1. Manhattan plot of the genome-wide association analyses for reproductive efficiency in Pura Raza Español horse. The red line shows a genome-wide significance threshold (−log10(*p*) = 5) and the blue line a genome-wide suggestive threshold (−log10(*p*) = 4). Abbreviations: p = *P*-value.

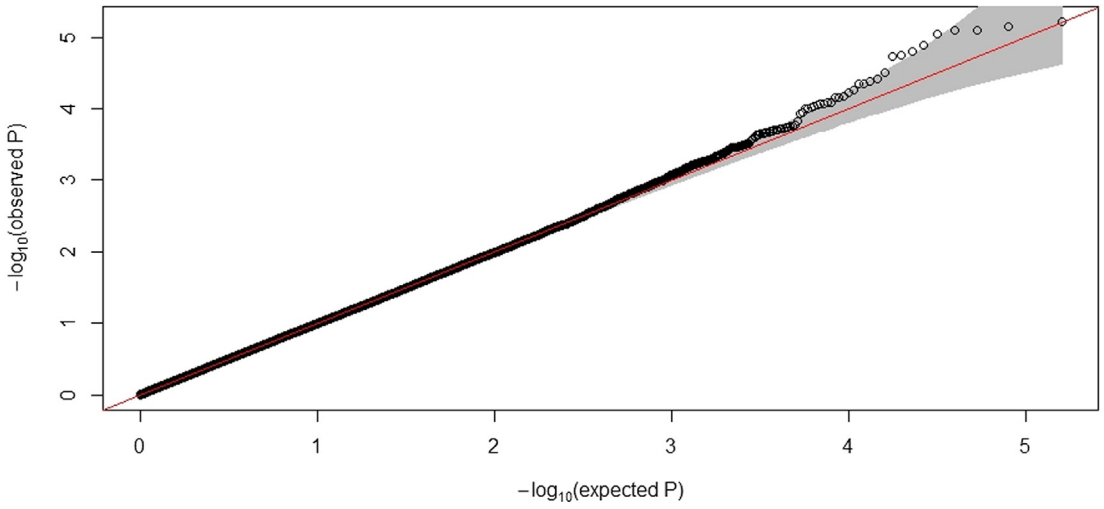

Fig. 2. Quantile-quantile plot corresponding to the *P*-values of the genome-wide association study for reproductive efficiency in Pura Raza Español horse.

of female reproduction in mares. Our aim was assessed using a genetic approach including deregressed breeding values of reproductive efficiency, a new fertility criterium, such as pseudo-phenotypes and high-density SNP genotyping. To our knowledge, this is the first time that a GWAS has been performed aimed at identifying SNPs markers associated with fertility in mares.

One of the major difficulties in the evaluation of the genetic effect on fertility in mares is how to quantify the phenotype accurately (Laseca et al., 2021a). It has been suggested that the small number of studies is due to a lack of large, reliable phenotypic datasets of traits associated with fertility in mares (such as ovulation or conception rate, among others). In our case, we were able to gather almost 80 000 phenotypic records of reproductive efficiency, a phenotypic trait associated with fertility which showed a considerable heritability (0.23) and high precision (0.8) in PRE (Perdomo-Gonzalez et al., 2020). This is in agreement with results reported in other species such as goats, in which heritability varied between 0.26 and 0.28 (Ziadi et al., 2021). This heritability obtained in PRE is considerably higher than those previously reported for similar traits, such as the pregnancy rate per cycle in Hanoverians (h_2 ranged from 0.07 to 0.13, (Distl, 2017)), which suggests that it contains a large proportion of the genetic variability in the fertility of this breed, and is therefore an interesting trait with which to evaluate the fertility of the mare at the population level.

The GWAS revealed 28 significant SNPs on 17 different chromosomes that explained 6.61% of the phenotypic variance. This result demonstrates that only a small proportion of the estimated phenotypic variance could be explained by specific SNPs, which fits in with the idea that fertility is a polygenic trait that does not depend on the influence of a major gene, but rather on many genes which each contribute a small effect. However, this low percentage of the phenotypic variance value explained by specific SNP markers was expected, since fertility traits are known for moderate heritability, while greater importance is attached to environmental effects, such as own reproductive management, age, nutrition, breeding season, climatic conditions, health and others (Onteru et al., 2012). In addition, the accurate modelling of the population structure in the dataset is also a major point to take into account, since the existence of strong family relationships is extremely common in livestock populations (van den Berg et al., 2019), but particularly in the PRE (Perdomo-González et al., 2020). However, the QQ plots and lambda values were accurate in the model employed in this analysis, giving high reliability to our results in terms of methodological procedures.

The link between female fertility and the X chromosome is controversial. For example, several reports have been able to determine the influence of the X chromosome in cattle fertility (Fortes et al., 2020), but all of these focused on specific traits in males. For this reason, sex chromosomes are commonly excluded from

Table 2
Candidate genes related to biological processes, molecular functions, and pathways of fertility found in the analysed horses.

SNP	ECA	Gene	Start gene (bp)	End gene (bp)	Related to
rs1140471861	3	*HTRA3*	117 767 023	117 803 782	Ovarian development, granulose cell differentiation and follicular luteinisation
rs1140299868	8	*SPIRE1*	40 539 360	40 796 466	Oocyte division
rs397493565	10	*NECTIN-2*	15 518 024	15 671 728	Organisation and reorganisation of cytoskeleton during spermiogenesis
		APOE	15 713 215	15 715 198	Production of androgens by theca cells, follicular maturation and steroidogenesis
		KLC3	16 083 814	16 100 636	Development of midpiece during spermiogenesis
		ERCC1	16 144 062	16 154 444	Female and male germ cells maturation and gametogenesis
		RSPH6A	16 453 159	16 471 388	Sperm capacitation
		FOXA3	16 507 566	16 513 710	The testicular germ cell and steroidogenesis
rs1141581080	13	*PAQR4*	41 020 286	41 022 630	Oestrus synchronisation
		PRSS21	41 174 424	41 178 932	Sperm capacitation, epididymal maturation and spermatozoa-oocyte interaction
		PDPK1	41 420 452	41 494 281	Survival of primordial follicles and activation of growing follicles.
		PKD1	41 880 905	41 926 116	Development and maintenance of male reproductive tract
		MEIOB	42 141 232	42 165 354	Meiotic homologous recombination and azoospermia
		NME3	42 228 271	42 229 120	Oogenesis and early embryonic development
		IFT140	42 375 191	42 459 360	Spermiogenesis and sperm flagella assembly

Abbreviations: SNPs = single-nucleotide polymorphisms; ECA = equine chromosome; bp = base pairs.

GWAS analyses which focus on fertility (Wang et al., 2018; Islam et al., 2020; Gmel et al., 2021; Keogh et al., 2021). In horses, we previously reported, in a preliminary study performed using sequencing technologies in a very reduced population of PRE mares, a weak (but significant) association between SNP markers located in ECAX and fertility (Laseca et al., 2021b). However, not all these markers were available in the SNP-array employed in this study, and therefore, we were able to perform a validation in a broader population. Nevertheless, we were unable to find any association between ECAX and RE in the present dataset using Axiom™ Equine 670 K SNP (data not shown). Our results were contradictory, reporting a null association between ECAX and RE, although they could be biased by differences in the methodology employed, which is why we decided that not to report it until it could be properly validated by analysing a large population of individuals with sequencing technology.

Finally, it is worth mentioning that the results obtained in the GWAS have been validated in a large PRE population genotyped by correlating the results predicted on the Bayes B model with the dEBV values. This methodology has been employed for important economic traits in different animal species, such as in US Limousin and Simmental beef cattle, with a correlation between 0.39 and 0.76 and between 0.29 and 0.65, respectively (Saatchi et al., 2012). Our results showed a high correlation (0.565) in comparison with previous reports, but the results must be taken with caution since no studies have yet been performed on fertility traits (or in horses) which might allow a more accurate comparison.

Candidate genes

Our analysis identified 5 SNPs associated with RE_{EBV} located on four different chromosomes (ECA3, ECA8, ECA10, and ECA13). However, the analysis of the genomic regions located within those SNP positions (±1Mb) revealed the existence within them of 15 candidate genes, previously related to fertility in mammals.

One of the genes located within the region of the most significant SNP (*rs1140471861*, ECA3) was *HtrA serine peptidase 3* (*HTRA3*). This gene is mainly involved in the development of the placenta and ovary, the differentiation of granulosa cells and the luteinisation of the follicle after ovulation (Nie et al., 2006). In addition, *HTRA3* has been proposed as a useful candidate for increasing litter size in pigs (Xundong et al., 2017), which suggests it may play an important role in female fertility.

On ECA8, we found the *spire type actin nucleation factor 1* (*SPIRE1)* gene located very close to the significant *rs1140299868* marker. This gene has been related to the oocyte division, as described by Pfender et al., 2011, who identified *SPIRE1* and *SPIRE2* as new essential factors in asymmetric oocyte division. They also suggested that *SPIRE1* and *SPIRE2* may cooperate with *FMN2* to nucleate actin filaments in mouse oocytes.

On ECA10, RE was associated with *rs397493565*, which is positioned near the *APOE* and *ERCC1* genes. *APOE* encodes a 34-kDa glycoprotein (Apolipoprotein E), which is involved in the physiological functions of the female gonads (Von Wald et al., 2010). However, more recently, Oriá et al., 2020 reported that an increase of *APOE* concentration in follicular fluid is negatively correlated with fertility due to a decrease in the production of mature oocytes. In addition, the role of *APOE* in steroidogenesis is also worth noting (Kacperczyk et al., 2021). *ERCC1* is primarily involved in recombination repair pathways in mammalian cells. However, Hsia et al. (2003) demonstrated that Ercc1-deficient female and male mice were infertile, which implies that the repair functions of *ERCC1* are necessary for both female and male germ cells at all stages of their maturation and that its role is therefore essential for normal oogenesis and spermatogenesis, since the premeiotic lesions and DNA damage observed are consistent with a general role for *ERCC1* repair functions throughout gametogenesis rather than with a specific requirement at the meiotic crossing-over.

Interestingly, we also found genes related to male fertility on ECA10, although their role in female fertility is still unknown. It should be remembered that our knowledge of the genes involved in mare fertility is practically null, while there are numerous association studies in male fertility. It is therefore not known to what extent these genes described in males may be involved in metabolic pathways related to female fertility. This fact has already been reported in the human species, where the functionality of each gene is better known (even more so in female fertility). For example, the *APOE* gene, which was detected in the present study as associated to mare fertility, has been described as affecting male and female fertility pathways in human beings (Kacperczyk et al., 2021).

For instance, gene *FOXA3,* located near to the *rs397493565* SNP, is crucial for male fertility. Recently, Kim et al. (2021) reported that the overexpression of *FOXA3* in mouse primary Leydig cells resulted in decreased production of testosterone, and suggested the role of *FOXA3* in the regulation of steroidogenic genes in Leydig cells and fine-tuning steroidogenesis in the testis. In addition, we found genes related to spermatogenesis, such as the *NECTIN-2* gene, which is related to the organisation and reorganisation of the cytoskeleton during spermiogenesis (Bronson et al., 2017); *the kinesin light chain3* (*KLC3*) gene was also reported by Zhang et al. (2012) for its role in the development of the midpiece during spermiogenesis and for being involved in the normal function of spermatozoa. Finally, the *radial spoke head 6 homolog A* (*RSPH6A*) gene encodes a candidate protein mediating signalling processes in the sperm flagellum, which means that *RSPH6A* is involved in sperm capacitation (Paudel et al., 2019).

On ECA 13, we found *rs1141581080* to be significantly associated with RE. This marker is located very close to the *3 phosphoinositide dependent protein kinase 1* (*PDPK1*) gene, which has been associated with premature ovarian failure due to a massive primordial follicle activation in the knockout mouse (Reddy et al., 2009). However, the position of *rs1141581080* was close to the *meiosis-specific with OB domain* (*MEIOB*), which encodes a protein involved in meiotic homologous recombination, which is essential for sexual reproduction (Guo et al., 2020) and plays a key role in the repairing system required after the formation of double-strand breaks during the early stages of meiosis and crossing-over formation in late meiotic recombination (Guo et al., 2020).

In addition, we found an association between *rs1141581080* and RE. Two genes previously associated with fertility were located close to the marker: *the progestin and adipoQ receptor family member 4* (*PAQR4)* and *nucleoside diphosphate kinase 3* (*NME3)*. The former has been identified in a study of differential gene expression in goat ovaries of goats which were treated for oestrus synchronisation (Sun et al., 2018), while the latter has been associated with the process of oogenesis and early embryo development in zebrafish (Desvignes et al., 2011).

As on chromosome 10, we found candidate genes related to male fertility associated with significant SNP on chromosome 13, although we are currently unaware of their involvement in female fertility. One of these genes was *polycystin 1* (*PKD1*), which has been reported in one of the few fertility association studies carried out in stallions as a high impact variant in this gene, and considered a potentially deleterious factor for stallion fertility (Schrimpf et al., 2016).

We also found an interesting gene related to sperm capacitation, *the serine protease testisin* (*PRSS21*) gene. Curiously, a study in stallion spermatozoa reported that testisin appears to form part of the zona pellucida-binding complex in stallion spermatozoa and

may be involved in the proteolytic cascade that prepares the sperm surface for interaction with the oocyte (Swegen et al., 2019). They therefore suggested that testisin is an important candidate protein with potential roles in epididymal maturation, capacitation events, and spermatozoa–oocyte interaction. Furthermore, Stafuzza et al. (2020) performed a genome-wide association study for age at puberty in young Nelore bulls and detected, among others, *PRSS21* as a putative candidate gene.

Finally, *the intraflagellar transport protein 140 homologs* (*IFT140*) gene was located near to the *rs1141581080* on ECA13. In 2018, a study in mice demonstrated that *IFT140* is a key regulator for male fertility and normal spermiogenesis in mice (Zhang et al., 2018). Moreover, they reported that it not only plays a role in sperm flagellar assembly but is also involved in the critical assembly of proteins that interface between the germ cell plasma and the Sertoli cell.

Conclusion

This is the first GWAS study to assess the causes involved in the genetic control of fertility in mares. The analysis of a large dataset revealed the presence of 5 SNPs significantly associated with reproductive efficiency, a reproductive trait associated with fertility. We also found 15 candidate genes previously associated with female fertility in other species which are potentially related to the biological control of fertility in mares. These candidate genes might provide knowledge for genomic or marker-assisted selection in this trait by assigning greater weight to the genetic markers located in these regions. However, further studies are required to provide important knowledge on the understanding of metabolic routes for horse reproductive traits and to confirm these SNP associations and candidate genes in other horse breeds.

Ethics approval

Not applicable.

Data and model availability statement

None of the data were deposited in an official repository. The data supporting the findings of this study are available from Asociación Nacional de Criadores de Caballos de Pura Raza Español (ANCCE). Restrictions apply to the availability of these data, which were used under licence for this study. The data are available from the authors with the permission of ANCCE.

Author ORCIDs

Nora Laseca: https://orcid.org/0000-0003-3753-6725
Sebastián Demyda-Peyrás: https://orcid.org/0000-0003-3286-2441
Mercedes Valera: https://orcid.org/0000-0003-1742-550X
Manuel Ramón: https://orcid.org/0000-0003-4179-9894
Begoña Escribano: https://orcid.org/0000-0002-0966-0641
Davinia Isabel Perdomo González: https://orcid.org/0000-0003-2618-105X
Antonio Molina: https://orcid.org/0000-0002-9566-6600

Author contributions

The authors' contributions are as follows: **A. Molina, S. Demyda-Peyrás, N. Laseca** and **M. Valera** conceived and designed the study; **N. Laseca, S. Demyda-Peyrás, D. Perdomo-Gonzalez, M. Valera** and **B. Escribano** obtained the data and conducted the research; **N. Laseca, S. Demyda-Peyrás, M. Ramon** and **A. Molina** analysed and interpreted the data; **N. Laseca, S. Demyda-Peyrás** and **A. Molina** wrote the manuscript. All authors revised the manuscript and approved the final version of the manuscript.

Declaration of interest

The authors declare no conflict of interest.

Acknowledgements

The authors wish to thank the Asociación Nacional de Criadores de Caballos de Pura Raza Español (ANCCE) and Yeguada de La Cartuja Hierro del Bocado for their collaboration in this study and for providing the biological samples and genealogical data used in this study. Funding for open access charge: Universidad de Córdoba / CBUA.

Financial support statement

This study was financed by the AGL-2017-84217-P Research project from the Ministerio de Economia, Industria y Competitividad of the Spanish Government and by the PICT2018-0227 grant (Foncyt, Argentina, Sebastian Demyda-Peyrás PI). Nora Laseca is funded by an FPI from the Ministry of Science, Innovation, and Universities (PRE 2018-083492).

References

Al Abri, M.A., Posbergh, C., Palermo, K., Sutter, N.B., Eberth, J., Hoffman, G.E., Brooks, S.A., 2018. Genome-Wide Scans Reveal a Quantitative Trait Locus for Withers Height in Horses Near the ANKRD1 Gene. Journal of Equine Veterinary Science 60, 67–73. https://doi.org/10.1016/j.jevs.2017.05.008.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society 57, 289–300.

Bronson, R., Mikhailik, A., Schwedes, J., Gnatenko, D., Hatchwell, E., 2017. Detection of candidate nectin gene mutations in infertile men with severe teratospermia. Journal of Assisted Reproduction and Genetics 34, 1295–1302. https://doi.org/10.1007/s10815-017-0985-4.

Desvignes, T., Fauvel, C., Bobe, J., 2011. The nme gene family in zebrafish oogenesis and early development. Naunyn-Schmiedeberg's Archives of Pharmacology 384, 439–449. https://doi.org/10.1007/s00210-011-0619-9.

Devlin, B., Roeder, K., 1999. Genomic Control for Association Studies. Biometrics 55, 997–1004. https://doi.org/10.1111/j.0006-341X.1999.00997.x.

Distl, O., 2017. Sperm quality in the stallion and genomic markers for infertility and conception success. Zuchtungskunde 89, 29–38.

Fonseca, M.G., Ferraz, G.D.C., Lage, J., Pereira, G.L., Curi, R.A., 2017. A Genome-Wide Association Study Reveals Differences in the Genetic Mechanism of Control of the Two Gait Patterns of the Brazilian Mangalarga Marchador Breed. Journal of Equine Veterinary Science 53, 64–67. https://doi.org/10.1016/j.jevs.2016.01.015.

Fortes, M.R.S., Porto-Neto, L.R., Satake, N., Nguyen, L.T., Freitas, A.C., Melo, T.P., Scalez, D.C.B., Hayes, B., Raidan, F.S.S., Reverter, A., Boe-Hansen, G.B., 2020. X chromosome variants are associated with male fertility traits in two bovine populations. Genetics Selection Evolution 52, 46. https://doi.org/10.1186/s12711-020-00563-5.

Garrick, D.J., Taylor, J.F., Fernando, R.L., 2009. Deregressing estimated breeding values and weighting information for genomic regression analyses. Genetics Selection Evolution 41, 55. https://doi.org/10.1186/1297-9686-41-55.

Gmel, A.I., Burger, D., Neuditschko, M., 2021. A Novel QTL and a Candidate Gene Are Associated with the Progressive Motility of Franches-Montagnes Stallion Spermatozoa after Thaw. Genes 12, 1501. https://doi.org/10.3390/genes12101501.

Gómez, M.D., Sánchez, M.J., Bartolomé, E., Cervantes, I., Poyato-Bonilla, J., Demyda-Peyrás, S., Valera, M., 2020. Phenotypic and genetic analysis of reproductive traits in horse populations with different breeding purposes. Animal 14, 1351–1361. https://doi.org/10.1017/S1751731120000087.

Gottschalk, M., Metzger, J., Martinsson, G., Sieme, H., Distl, O., 2016. Genome-wide association study for semen quality traits in German Warmblood stallions. Animal Reproduction Science 171, 81–86. https://doi.org/10.1016/j.anireprosci.2016.06.002.

Guo, R., Xu, Y., Leu, N.A., Zhang, L., Fuchs, S.Y., Ye, L., Wang, P.J., 2020. The ssDNA-binding protein MEIOB acts as a dosage-sensitive regulator of meiotic recombination. Nucleic Acids Research 48, 12219–12233. https://doi.org/10.1093/nar/gkaa1016.

Hsia, K.-T., Millar, M.R., King, S., Selfridge, J., Redhead, N.J., Melton, D.W., Saunders, P.T.K., 2003. DNA repair gene Ercc1 is essential for normal spermatogenesis and oogenesis and for functional integrity of germ cell DNA in the mouse. Development 130, 369–378. https://doi.org/10.1242/dev.00221.

Islam, R., Liu, X., Gebreselassie, G., Abied, A., Ma, Q., Ma, Y., 2020. Genome-wide association analysis reveals the genetic locus for high reproduction trait in Chinese Arbas Cashmere goat. Genes & Genomics 42, 893–899. https://doi.org/10.1007/s13258-020-00937-5.

Kacperczyk, M., Kmieciak, A., Kratz, E.M., 2021. The Role of ApoE Expression and Variability of Its Glycosylation in Human Reproductive Health in the Light of Current Information. International Journal of Molecular Sciences 22, 7197. https://doi.org/10.3390/ijms22137197.

Keogh, K., Carthy, T.R., McClure, M.C., Waters, S.M., Kenny, D.A., 2021. Genome-wide association study of economically important traits in Charolais and Limousin beef cows. Animal 15,. https://doi.org/10.1016/j.animal.2020.100011 100011.

Kim, H., Kumar, S., Lee, K., 2021. FOXA3, a Negative Regulator of Nur77 Expression and Activity in Testicular Steroidogenesis. International Journal of Endocrinology 2021, 8. https://doi.org/10.1155/2021/6619447.

Laseca, N., Anaya, G., Peña, Z., Pirosanto, Y., Molina, A., Demyda Peyrás, S., 2021a. Impaired Reproductive Function in Equines: From Genetics to Genomics. Animals 11, 393. https://doi.org/10.3390/ani11020393.

Laseca, N., Demyda-Peyrás, S., Goszczynski, D., Ramón, M., Escribano, B., Encina, A., Valera, M., Molina, A., 2021b. Exploratory whole genome association with fertility in PRE horse breed using chromosome X NGS data. In: Proceedings of the 72nd Annual Meeting of the European Federation of Animal Science, 30 August-3 September 2021, Davos, Switzerland, p. 178.

Laseca, N., Molina, A., Ramón, M., Valera, M., Azcona, F., Encina, A., Demyda-Peyrás, S., 2022. Fine-scale analysis of runs of homozygosity islands affecting fertility in mares. Frontiers in Veterinary Science 9, 754028. https://doi.org/10.3389/fvets.2022.754028.

Mahon, G., Cunningham, E., 1982. Inbreeding and the inheritance of fertility in the thoroughbred mare. Livestock Production Science 9, 743–754.

Mantovani, R., Folla, F., Pigozzi, G., Tsuruta, S., Sartori, C., 2020. Genetics of lifetime reproductive performance in italian heavy draught horse mares. Animals 10 (1–14), 1085. https://doi.org/10.3390/ani10061085.

Misztal, I., Tsuruta, S., Lourenco, D., Masuda, Y., Aguilar, I., Legarra, A., Vitezica, Z., 2016. Manual for BLUPF90 family of programs. University of Georgia, Athens, GA, USA.

Moazemi, I., Mohammadabadi, M.R., Mostafavi, A., Esmailizadeh, A.K., Babenko, O.I., Bushtruk, M.V., Tkachenko, S.V., Stavetska, R.V., Klopenko, N.I., 2020. Polymorphism of DMRT3 Gene and Its Association with Body Measurements in Horse Breeds. Russian Journal of Genetics 56, 1232–1240. https://doi.org/10.1134/S1022795420100087.

Nie, G., Li, Y., He, H., Findlay, J.K., Salamonsen, L.A., 2006. HtrA3, a Serine Protease Possessing an IGF-binding Domain, is Selectively Expressed at the Maternal-Fetal Interface During Placentation in the Mouse. Placenta 27, 491–501. https://doi.org/10.1016/j.placenta.2005.03.009.

Onteru, S.K., Fan, B., Du, Z.Q., Garrick, D.J., Stalder, K.J., Rothschild, M.F., 2012. A whole-genome association study for pig reproductive traits. Animal Genetics 43, 18–26. https://doi.org/10.1111/j.1365-2052.2011.02213.x.

Oriá, R.B., de Almeida, J.Z., Moreira, C.N., Guerrant, R.L., Figueiredo, J.R., 2020. Apolipoprotein E Effects on Mammalian Ovarian Steroidogenesis and Human Fertility. Trends in Endocrinology & Metabolism 31, 872–883. https://doi.org/10.1016/j.tem.2020.06.003.

Orlando, L., 2020. Ancient Genomes Reveal Unexpected Horse Domestication and Management Dynamics. BioEssays 42, 1900164. https://doi.org/10.1002/bies.201900164.

Palmer, E., Chavatte-Palmer, P., 2020. Contribution of Reproduction Management and Technologies to Genetic Progress in Horse Breeding. Journal of Equine Veterinary Science 89,. https://doi.org/10.1016/j.jevs.2020.103016 103016.

Paudel, B., Gervasi, M.G., Porambo, J., Caraballo, D.A., Tourzani, D.A., Mager, J., Platt, M.D., Salicioni, A.M., Visconti, P.E., 2019. Sperm capacitation is associated with phosphorylation of the testis-specific radial spoke protein Rsph6a. Biology of Reproduction 100, 440–454. https://doi.org/10.1093/biolre/ioy202.

Perdomo-González, D.I., Molina, A., Sánchez-Guerrero, M.J., Bartolomé, E., Varona, L., Valera, M., 2021. Genetic inbreeding depression load for fertility traits in Pura Raza Española mares. Journal of Animal Science 99, 1–10. https://doi.org/10.1093/jas/skab316.

Perdomo-Gonzalez, D.I., Sanchez-Guerrero, M., Molina, A., Arrebola, F., Valera, M., 2020. Estimation of genetic parameters for fertility criteria in Pura Raza Española mares. In: Proceedings of the 71st Annual Meeting of the European Federation of Animal Science, 1–4 December 2020, Virtual Meeting, p. 345.

Perdomo-González, D.I., Sánchez Guerrero, M., Molina, A., Valera, M., 2020. Genetic Structure Analysis of the Pura Raza Español Horse Population through Partial Inbreeding Coefficient Estimation. Animals 10, 1360. https://doi.org/10.3390/ani10081360.

Pereira, G.L., Chardulo, L.A., Silva, J.A.I., Faria, R., Curi, R.A., 2018. Genomic regions associated with performance in racing line of Quarter Horses. Livestock Science 211, 42–51. https://doi.org/10.1016/j.livsci.2018.02.015.

Pfender, S., Kuznetsov, V., Pleiser, S., Kerkhoff, E., Schuh, M., 2011. Spire-Type Actin Nucleators Cooperate with Formin-2 to Drive Asymmetric Oocyte Division. Current Biology 21, 955–960. https://doi.org/10.1016/j.cub.2011.04.029.

Poyato-Bonilla, J., Laseca, N., Demyda-Peyrás, S., Molina, A., Valera, M., 2022. 500 years of breeding in the Carthusian Strain of Pura Raza Español horse: An evolutional analysis using genealogical and genomic data. Journal of Animal Breeding and Genetics 139, 84–99. https://doi.org/10.1111/jbg.12641.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. American journal of human genetics 81, 559–575. https://doi.org/10.1086/519795.

R-Core-Team, 2021. R: A language and environment for statistical computing V4.1.0 "Camp Pontanezen". R Foundation for Statistical Computing, Vienna, Austria.

Reddy, P., Adhikari, D., Zheng, W., Liang, S., Hämäläinen, T., Tohonen, V., Ogawa, W., Noda, T., Volarevic, S., Huhtaniemi, I., Liu, K., 2009. PDK1 signaling in oocytes controls reproductive aging and lifespan by manipulating the survival of primordial follicles. Human Molecular Genetics 18, 2813–2824. https://doi.org/10.1093/hmg/ddp217.

Saatchi, M., Schnabel, R.D., Rolf, M.M., Taylor, J.F., Garrick, D.J., 2012. Accuracy of direct genomic breeding values for nationally evaluated traits in US Limousin and Simmental beef cattle. Genetics Selection Evolution 44, 38. https://doi.org/10.1186/1297-9686-44-38.

Sairanen, J., Nivola, K., Katila, T., Virtala, A.M., Ojala, M., 2009. Effects of inbreeding and other genetic components on equine fertility. Animal 3, 1662–1672. https://doi.org/10.1017/S1751731109990553.

Schaefer, R.J., Schubert, M., Bailey, E., Bannasch, D.L., Barrey, E., Bar-Gal, G.K., Brem, G., Brooks, S.A., Distl, O., Fries, R., Finno, C.J., Gerber, V., Haase, B., Jagannathan, V., Kalbfleisch, T., Leeb, T., Lindgren, G., Lopes, M.S., Mach, N., da Câmara Machado, A., MacLeod, J.N., McCoy, A., Metzger, J., Penedo, C., Polani, S., Rieder, S., Tammen, I., Tetens, J., Thaller, G., Verini-Supplizi, A., Wade, C.M., Wallner, B., Orlando, L., Mickelson, J.R., McCue, M.E., 2017. Developing a 670k genotyping array to tag ~2M SNPs across 24 horse breeds. BMC Genomics 18, 565. https://doi.org/10.1186/s12864-017-3943-8.

Schrimpf, R., Gottschalk, M., Metzger, J., Martinsson, G., Sieme, H., Distl, O., 2016. Screening of whole genome sequences identified high-impact variants for stallion fertility. BMC Genomics 17, 288. https://doi.org/10.1186/s12864-016-2608-3.

Shrestha, M., Solé, M., Ducro, B.J., Sundquist, M., Thomas, R., Schurink, A., Eriksson, S., Lindgren, G., 2020. Genome-wide association study for insect bite hypersensitivity susceptibility in horses revealed novel associated loci on chromosome 1. Journal of Animal Breeding and Genetics 137, 223–233. https://doi.org/10.1111/jbg.12436.

Smołucha, G., Gurgul, A., Jasielczuk, I., Kawęcka, A., Miksza-Cybulska, A., 2021. A genome-wide association study for prolificacy in three Polish sheep breeds. Journal of Applied Genetics 62, 323–326. https://doi.org/10.1007/s13353-021-00615-6.

Stafuzza, N.B., Costa e Silva, E.V.D., Silva, R.M.D.O., Costa Filho, L.C.C.D., Barbosa, F.B., Macedo, G.G., Lobo, R.B., Baldi, F., 2020. Genome-wide association study for age at puberty in young Nelore bulls. Journal of Animal Breeding and Genetics 137, 234–244. https://doi.org/10.1111/jbg.12438.

Sun, S., Li, C., Liu, S., Luo, J., Chen, Z., Zhang, C., Zhang, T., Huang, J., Xi, L., 2018. RNA sequencing and differential expression reveals the effects of serial oestrus synchronisation on ovarian genes in dairy goats. Reproduction, Fertility and Development 30, 1622–1633. https://doi.org/10.1071/RD17511.

Swegen, A., Smith, N.D., Gibb, Z., Curry, B.J., Aitken, R.J., 2019. The serine protease testisin is present on the surface of capacitated stallion spermatozoa and interacts with key zona pellucida binding proteins. Andrology 7, 199–212. https://doi.org/10.1111/andr.12569.

Todd, E.T., Hamilton, N.A., Velie, B.D., Thomson, P.C., 2020. The effects of inbreeding on covering success, gestation length and foal sex ratio in Australian thoroughbred horses. BMC Genetics 21, 41. https://doi.org/10.1186/s12863-020-00847-1.

van den Berg, S., Vandenplas, J., van Eeuwijk, F.A., Lopes, M.S., Veerkamp, R.F., 2019. Significance testing and genomic inflation factor using high-density genotypes or whole-genome sequence data. Journal of Animal Breeding and Genetics 136, 418–429. https://doi.org/10.1111/jbg.12419.

VanRaden, P.M., Tooker, M.E., O'Connell, J.R., Cole, J.B., Bickhart, D.M., 2017. Selecting sequence variants to improve genomic predictions for dairy cattle. Genetics Selection Evolution 49, 32. https://doi.org/10.1186/s12711-017-0307-4.

Velie, B.D., Fegraeus, K.J., Solé, M., Rosengren, M.K., Røed, K.H., Ihler, C.-F., Strand, E., Lindgren, G., 2018. A genome-wide association study for harness racing success in the Norwegian-Swedish coldblooded trotter reveals genes for learning and energy metabolism. BMC Genetics 19, 80. https://doi.org/10.1186/s12863-018-0670-3.

Von Wald, T., Monisova, Y., Hacker, M.R., Yoo, S.W., Penzias, A.S., Reindollar, R.R., Usheva, A., 2010. Age-related variations in follicular apolipoproteins may influence human oocyte maturation and fertility potential. Fertility and Sterility 93, 2354–2361. https://doi.org/10.1016/j.fertnstert.2008.12.129.

Wang, Y., Ding, X., Tan, Z., Xing, K., Yang, T., Wang, Y., Sun, D., Wang, C., 2018. Genome-wide association study for reproductive traits in a Large White pig population. Animal Genetics 49, 127–131. https://doi.org/10.1111/age.12638.

Wolc, A., Torzynski, G., Szwaczkowski, T., 2009. Genetic effects on reproductive traits in Warmblood horses. Canadian Journal of Animal Science 89, 215–218. https://doi.org/10.4141/CJAS08067.

Xu, L., Niu, Q., Chen, Y., Wang, Z., Xu, L., Li, H., Xu, L., Gao, X., Zhang, L., Gao, H., Cai, W., Zhu, B., Li, J., 2021. Validation of the Prediction Accuracy for 13 Traits in Chinese Simmental Beef Cattle Using a Preselected Low-Density SNP Panel. Animals 11, 1890. https://doi.org/10.3390/ani11071890.

Xundong, W., Guiying, Z., Guowen, F., Yonggang, L., 2017. Molecular Cloning of the Porcine HTRA3 Gene and Association of a SNP with Litter Size Traits. Folia Biologica (Praha) 63, 217–221.

Zhang, Y., Liu, H., Li, W., Zhang, Z., Zhang, S., Teves, M.E., Stevens, C., Foster, J.A., Campbell, G.E., Windle, J.J., Hess, R.A., Pazour, G.J., Zhang, Z., 2018. Intraflagellar transporter protein 140 (IFT140), a component of IFT-A complex, is essential for male fertility and spermiogenesis in mice. Cytoskeleton 75, 70–84. https://doi.org/10.1002/cm.21427.

Zhang, Y., Ou, Y., Cheng, M., Shojaei Saadi, H., Thundathil, J.C., van der Hoorn, F.A., 2012. KLC3 is involved in sperm tail midpiece formation and sperm function. Developmental Biology 366, 101–110. https://doi.org/10.1016/j.ydbio.2012.04.026.

Zhou, X., Stephens, M., 2012. Genome-wide efficient mixed-model analysis for association studies. Nature Genetics 44, 821–824. https://doi.org/10.1038/ng.2310.

Ziadi, C., Muñoz-Mejías, E., Sánchez, M., López, M.D., González-Casquet, O., Molina, A., 2021. Genetic analysis of reproductive efficiency in Spanish goat breeds using a random regression model as a strategy for improving female fertility. Italian Journal of Animal Science 20, 1681–1688. https://doi.org/10.1080/1828051X.2021.1979900.