MÁSTER INTERUNIVERSITARIO EN MEJORA GENÉTICA ANIMAL Y BIOTECNOLOGÍA DE LA REPRODUCCIÓN

Inclusion of crossbred data and genomic information in the Iberian pig genetic evaluations.

Trabajo de Fin de Máster

Julio, 2020

María Victoria Topayan

Directora:

Dra. Noelia Ibáñez-Escriche

Universitat Politècnica de València

Departamento de Ciencia Animal

Content

Index of tables iii
Index of figures iv
Dedicated to vii
Acknowledgements viii
Abstract. ix
Resumen xi
Resum xiii
1. Introduction. 15
1.1. Iberian pig: brief history and general description. 15
1.2. Current situation: production systems and breeding programs. 16
1.2.1. Within line/breed selection scheme. 17
1.2.2. Pyramidal Scheme 18
1.3. Selection: methods to estimate breeding values. 19
1.3.1. Traditional selection: BLUP 20
1.3.2. Genomic Selection: from SNP-BLUP to single step Genomic BLUP. 21
1.3.2.1. Genomic BLUP 23
1.3.2.2. Single step GBLUP. 24
1.4. Impact of genomics on crossbreeding schemes 26
1.5. Outline. 27
1.6. Objective. 28
2. Materials and methods. 29
2.1. Ethics statements. 29
2.2. Phenotypic data. 29
2.2.1. Exploratory data analysis. 30
2.3. Genomic data. 30
2.3.1. Quality control of genomic data. 30
2.4. Statistical analysis. 31
2.4.1. Models 31
2.4.1.1. Multi trait model (MT). 31
2.4.1.2. Multi breed model (MB). 32
2.4.2. Bayesian analysis. 32
2.5. Cross validation. 34
3. Results and discussion. 37
4. Conclusions. 46

5. Cited bibliography. 47
6. Supplementary material. 56

Index of tables

Table 1. Selection objectives, selection criteria, traits under study and statistical models used in the breeding program in Iberian pig (Ministerio de Medio Ambiente y Medio Rural y Marino & INIA, 2011). 17

Table 2. Models for implementation of genomic selection according how single nucleotide polymorphisms (SNPs) effects are considered and prior assumptions for SNPs effects. 23

Table 3. Number of litter size records according to sow line and farms. 29

Table 4. Number of litter size records according to breed of service boar 30

Table 5. Information of the number of markers and animals genotyped by sex and line. 31

Table 6. Models, data sets, traits and number of records analyzed for Bayesian BLUP (Best Linear Unbiased Prediction) and ssGBLUP (single step Genomic Best Linear Unbiased Prediction) genetic evaluations. 34

Table 7. Number of animals with records, genotyped animals (in brackets) and descriptive statistics of total born (TB), number of born alive (NBA) and stillborn (SB) for pure lines and crossbred Iberian pigs. 38

Table 8. Descriptive statistics of diagonal and off-diagonal elements of Genomic matrix (G) and additive genetic relationship matrix for genotyped animals only (A_{22}) for Entrepelado and Retinto lines. 40

Table 9. Main features of the posterior distribution of the heritability $(h2)$ and additive genetic correlation ($r_{a1,a2}$) of total born (TB), number of born alive (NBA) and stillborn (SB) estimated for Entrepelado (EE) and Retinto (RR) lines with BLUP (Best Linear Unbiased Prediction) and ssGBLUP (Single Step Best Linear Unbiased Prediction) Bayesian multi trait models. 41

Table 10. Main features of the posterior distribution of the heritability $(h2)$ and additive genetic correlation ($r_{a1,a2}$) of total born (TB), number of born alive (NBA) and stillborn (SB) estimated for Entrepelado (EE) and Retinto (RR) lines and crossbred animals (ER-RE) with BLUP (Best Linear Unbiased Prediction) and ssGBLUP (Single Step Genomic Best Linear Unbiased Prediction) Bayesian multi breed models. 43

Table 11. Variance components estimation of total born (TB), number of born alive (NBA) and stillborn (SB) with BLUP (Best Linear Unbiased Prediction) Bayesian multi trait models for Entrepelado (EE) and Retinto (RR) lines. 58

Table 12. Variance components estimation of total born (TB), number of born alive (NBA), stillborn (SB) with ssGBLUP (Single Step Genomic Best Linear Unbiased Prediction) Bayesian multi trait models for Entrepelado (EE) and Retinto (RR) lines. 58

Table 13. Variance components estimation for number of total born (TB), number of born alive (NBA), stillborn (SB) with BLUP (Best Linear Unbiased Prediction) Bayesian multi breed models for Entrepelado (EE) and Retinto (RR) lines and crossbred animals (ER-RE). 58

Table 14. Variance components estimation for number of total born (TB), number of born alive (NBA), stillborn (SB) with ssGBLUP (Single Step Best Linear Unbiased Prediction) Bayesian multi breed models for Entrepelado (EE) and Retinto (RR) lines and crossbred animals (ER-RE).....59

Table 15. Estimates for heritability (standard deviation) in different Iberian pig populations by different authors applying different models. 60

Index of figures

Figure 1: Typical pyramidal scheme of white pig selection programs adapted to Iberian pig (Noguera & Ibáñez-Escriche, 2017)**.**.....19

Figure 2. Schematic representation of the cross-validation pipeline to assess predictive ability. EE: Entrepelado line; RR: Retinto line; ER-RE: crossbred animals; (G)EBV: (genomic) estimated breeding value; $GEBVr$: (genomic) estimated breeding value on training set for validation animals; BV: breeding value; yc: corrected phenotype of validation animals according BLUP (Best Linear Unbiased Prediction) or ssGBLUP (single step Genomic Best Linear Unbiased Prediction) multi trait/multi breed models.....36

Figure 3. Distribution of litter size records for all animals after quality control. TB: total born; NBA: number of born alive; SB: stillborn. 37

Figure 4. Population structure identified by principal component analysis of the genotyped animals. PC1: first principal component (percentage of variance explained); PC2: second principal component (percentage of variance explained); EE: Entrepelado line; RR: Retinto line; ER: Entrepelado x Retinto cross; RE: Retinto x Entrepelado cross. 39

Figure 5. Histograms and plots for diagonal and off-diagonal elements in G (genomic relationship matrix) and A_{22} (numerator relationship matrix for genotyped animals) of all lines

and crosses. a) Frequency distribution for G diagonal elements; b) A_{22} diagonal elements and G diagonal elements plot; c) A_{22} off-diagonal elements and G off-diagonal elements plot.....40

Figure 6. Mean of the accuracy and standardized accuracy for the 10 replicates of the 5-fold cross validation of Entrepelado line for each model and trait by BLUP and ssGBLUP methods. Best Linear Unbiased Prediction; ssGBLUP: single step Genomic Best Linear Unbiased Prediction; MT: multi trait model; MB: multi breed model; TB: total piglets; NBA: number of born alive; SB: stillborn. 44

Figure 7. Mean of the accuracy and standardized accuracy for the 10 replicates of the 5-fold cross validation of Retinto line for each model and trait by BLUP and ssGBLUP methods. Best Linear Unbiased Prediction; ssGBLUP: single step Genomic Best Linear Unbiased Prediction; MT: multi trait model; MB: multi breed model; TB: total piglets; NBA: number of born alive; SB: stillborn. 44

Figure 8. Root mean squared error (RMSE) between the corrected phenotypes and the estimated breeding values of the 10 replicates of the 5-fold cross validation sets of Entrepelado for each model by Best Linear Unbiased Prediction (BLUP) and single step Genomic Best Linear Unbiased Prediction (ssGBLUP) methods. TB: total born; NBA: number of born alive; SB: stillborn; MT: multi trait model; MB: multi breed model. 45

Figure 9. Root mean squared error (RMSE) between the corrected phenotypes and the estimated breeding values of the 10 replicates of the 5-fold cross validation sets of Retinto line for each model by Best Linear Unbiased Prediction (BLUP) and single step Genomic Best Linear Unbiased Prediction ssGBLUP) methods. TB: total born; NBA: number of born alive; SB: stillborn; MT: multi trait model; MB: multi breed model. 45

Figure 10. Distribution for litter size records in Iberian pig population before quality control. TB: total born; NBA: number of born alive; SB: stillborn. 56

Figure 11. Percentage of variance explained by each principal component. PC: principal components of Iberian pig population under study (n=1114) obtained by principal component analysis. 56

Figure 12. Cumulative explained variance by each principal component. PC: principal components of Iberian pig population under study (n=1114) obtained by principal component analysis. 57

Figure 13. Histograms and plots for diagonal and off-diagonal elements in G (genomic relationship matrix) and A_{22} (numerator relationship matrix for genotyped animals) for Entrepelado line population. a) Frequency distribution for G diagonal elements; b) A_{22} diagonal

elements and G diagonal elements plot; c) A_{22} off-diagonal elements and G off-diagonal elements plot.....57

Figure 14. Histograms and plots for diagonal and off-diagonal elements in G (genomic relationship matrix) and A_{22} (numerator relationship matrix for genotyped animals) for Retinto line population. a) Frequency distribution for G diagonal elements; b) A_{22} diagonal elements and G diagonal elements plot; c) A_{22} off-diagonal elements and G off-diagonal elements plot. 57

Dedicated to

Mis padres y hermanas, sostén y motor incondicionales;

mi tío Daniel;

y especialmente a mi sobrino Vito.

Acknowledgements.

A todos los que hicieron posible la ejecución de este Máster: mi familia, que toda la vida me acompañó en mis decisiones; maestros y profesores presentes a lo largo de toda mi formación académica; mis amigos, que a la distancia siempre se hicieron presentes; mis jefes, colegas y personal docente de FCA-UNLZ.

Al equipo de Mejora Genética del Departamento de Ciencia Animal, por el acompañamiento en el trabajo diario y especialmente a mi directora, Noelia Ibáñez-Escriche, por su enseñanza a lo largo de estos meses de trabajo.

Abstract.

Litter size and mortality are the reproductive traits with most important effects on the sows productivity (number of piglets weaned alive per sow and per year) which is one the main factors affecting the cost-effective in pig farms. This especially important in Iberian pig with a lower prolificacy than other commercial pigs and with an increase of the substitution of many traditional producers by intensive management farms. In white pig, maternal breeding schemes are based on a pyramidal scheme. Nucleus herds supply genetically improved breeding stock to multiplier herds, which disseminate genetic gain by supplying hybrid (crosses) stock to commercial herds. Conventionally, data from crosses (CB) were not used in the genetic evaluations of purebreds (CCPS), but in the last years, some white pig breeding programs have included it for increasing the purebred (PB) breeding values accuracy. On another hand, reproductive traits have traditionally been genetic selected by the Best Linear Unbiased Prediction method (BLUP) based on two sources of information: phenotypic records and pedigree information. Nowadays, DNA information such as a large number of single nucleotide polymorphisms (SNPs) along genome are available (genotypes). Recently, this information has been jointly used with the phenotypes and the pedigree through the single step Genomic BLUP (ssGBLUP) method allowing an increase on the accuracy of the breeding value compared to the traditional BLUP. In Iberian pigs, neither of this, CCPS or SSGBLUP have been applied. The aim of this study was to evaluate the inclusion of crossbred information as well as inclusion of genomic data in the genetic evaluation of a pyramidal Iberian pig maternal breeding scheme. A total of 20468 phenotypic records for litter size (total born –TB, number of born alive –NBA and stillborn –SB) from 4753 sows of five farms in Badajoz (Spain), owned by Inga Food S.A., were used. Sows were purebred Iberian pigs from two pure lines, Entrepelado (EE) and Retinto (RR), as well as its reciprocal crosses Entrepelado x Retinto (ER) and Retinto x Entrepelado (RE). The pedigree included 5533 animals and went back 3 generations. From the population under study, 1.435 animals were genotyped for 63072 markers with Illumina Porcine HD Array 70K from samples of blood and tail tissue. Exploratory data analyses were performed on phenotypic records as well as on genomic data in order to detect outliers. Finally, 20152 phenotypic records from 4717 sows and 1114 genotypes remained to be used for the posterior analysis. Two different models using the BLUP and SSBLUP methods were fitted; (1) A trivariate multi trait model with TB, NBA and SB traits analyzed for each pure breed line (MT); and (2) a Multi breed model (MB) based on a trivariate model in which the records of a trait of each pure breed line and its crosses were assumed as different traits. So, MT model were separately performed for TB, NBA and SB traits,

respectively. Genetic parameters estimation was carried out by Bayesian Analysis with Gibbs Sampling. Both evaluations and their approaches were validated by implementing five-fold cross validation (CV). Prediction ability was measured by computing the correlation between the corrected phenotype and the estimated breeding value of the validation population and as the root mean squared error (RMSE). Owing to computational time, the CV was computed using the restricted maximum likelihood method (REML). The estimates of heritability for the three traits by BLUP and ssGBLUP were similar between models. The heritability of TB and NBA in EE was always greater than RR. There were no differences for SB. The additive genetic correlation between TB and NBA was high and superior than the correlations between SB and TB and between SB and NBA. The correlations for TB and NBA between PB and CB were higher compared to SB. Predictive ability measured as Pearson's correlation was higher for SB followed by TB and NBA. MT models and especially ssGBLUP showed better predictive ability for the three traits in EE, while MB models did so for RR. No major differences in terms of RMSE were observed between characters using BLUP and ssGBLUP MT, although a small advantage was found for genomic models. EE presented lower RMSE than RR. The inclusion of genomic data allows to slightly reduce the bias in the prediction of (G)EBV, except for NBA in EE and SB in RR under MB models.

Keywords: Iberian pig, genomic selection, BLUP, Single Step, GBLUP, litter size.

Resumen.

El tamaño de camada y la mortalidad son los caracteres reproductivos con mayor efecto sobre la productividad de las cerdas (número de lechones destetados vivos por cerda y año), siendo este uno de los principales factores que condicionan la rentabilidad en las granjas porcinas. Esto es especialmente importante en el cerdo ibérico, con una prolificidad menor que otros cerdos comerciales y con un aumento de la sustitución de productores tradicionales a granjas de manejo intensivo. En el cerdo blanco, la mejora de líneas maternas se basa en un esquema piramidal. Los núcleos de selección suministran reproductores genéticamente mejorados a las granjas multiplicadores, que difunden la ganancia genética a través del suministro de híbridas (cruces) a granjas comerciales. Clásicamente, los datos de cruces (CB) no se han utilizado en las evaluaciones genéticas de las líneas puras (CCPS). Sin embargo, en los últimos años, algunos esquemas de mejora de cerdo blanco lo han incluido para aumentar la precisión de los valores de mejora de las líneas puras (PB). Por otro lado, los caracteres reproductivos han sido tradicionalmente seleccionados genéticamente usando el método BLUP (mejor predictor lineal insesgado) basado en dos fuentes de información: registros fenotípicos e información de pedigrí. No obstante, hoy en día está disponible la información genómica, como por ejemplo los polimorfismos de un solo nucleótido (SNP) a lo largo del genoma (genotipos). Recientemente, esta información se ha utilizado juntamente con los fenotipos y el pedigrí a través del método BLUP genómico de un solo paso (ssGBLUP) que permite un aumento en la precisión del valor de mejora en comparación con el BLUP tradicional. En cerdo ibérico, ninguno de estos métodos, CCPS o SSGBLUP se han aplicado. El objetivo de este estudio fue evaluar la inclusión de datos de ibéricos cruzados, así como la inclusión de datos genómicos en la evaluación genética de un esquema piramidal materno de cerdo ibérico. Un total de 20.468 registros fenotípicos para el tamaño de camada (nacidos totales –TB, numero de nacidos vivos –NBA y nacidos muertos –TB) de 4.753 cerdas de cinco granjas en Badajoz (España), propiedad de Inga Food S.A., fueron usados. Las cerdas ibéricas correspondieron a dos líneas puras, Entrepelado (EE) y Retinto (RR), así como a sus cruces recíprocos Entrepelado x Retinto (ER) y Retinto x Entrepelado (RE). El pedigrí incluyó 5.533 animales y se remontó a 3 generaciones. De la población en estudio, 1.435 animales fueron genotipados para 63.072 marcadores con Illumina Porcine HD Array 70K a partir de muestras de sangre y tejido de la cola. Los análisis de datos exploratorios se realizaron en registros fenotípicos, así como en datos genómicos para detectar posibles “outliers”. Finalmente, 20.152 registros fenotípicos de 4.717 cerdas y 1.114 genotipos quedaron para el análisis posterior. Se evaluaron los métodos BLUP y ssGBLUP ajustando en cada uno de ellos dos

modelos diferentes. (1) Un modelo multi caracter trivariante con los caracteres TB, NBA y SB analizados para cada línea pura (MT); y (2) un modelo multi raza (MB) basado en un modelo trivariado en el que los registros de un caracter de cada línea de raza pura y sus cruces se asumieron como caracteres diferentes. Por tanto, el modelo MB se analizó por separado para los caracteres TB, NBA y SB, respectivamente. La estimación de los parámetros genéticos se realizó mediante análisis bayesiano con muestreo de Gibbs. Ambas evaluaciones y modelos se validaron mediante la implementación de cinco grupos de validación cruzada (CV). La capacidad de predicción se midió calculando la correlación entre el fenotipo corregido y el valor genético estimado de la población de validación y como la raíz del error cuadrático medio (RMSE). Debido al tiempo de cálculo, la CV se calculó utilizando el método de máxima verosimilitud restringida (REML). Las estimaciones de heredabilidad para los tres caracteres por BLUP y ssGBLUP fueron similares entre modelos. La heredabilidad de TB y NBA en EE fue siempre mayor que RR. No hubo diferencias para SB. La correlación genética aditiva entre TB y NBA fue alta y superior a las correlaciones entre SB y TB y entre SB y NBA. Las correlaciones de TB y NBA entre PB y CB fueron mayores respecto a SB. La capacidad predictiva medida como la correlación de Pearson fue mayor para SB seguida de TB y NBA. Los modelos MT y especialmente ssGBLUP mostraron una mejor capacidad de predicción para los tres caracteres en EE, mientras que los modelos MB lo hicieron para RR. No se observaron grandes diferencias en términos de RMSE entre caracteres mediante BLUP y ssGBLUP MT, aunque se encontró una pequeña ventaja de los modelos genómicos. EE presento RMSE menores que RR. La inclusión de datos genómicos permite reducir ligeramente el sesgo en la predicción de (G)EBV, excepto para NBA en EE y SB en RR bajo los modelos MB.

Palabras clave: cerdo ibérico, selección genómica, BLUP, Single Step GBLUP, tamaño de camada.

Resum.

La Prolificitat i la mortalitat són els caràcters reproductius amb major efecte sobre la productivitat de les truges (nombre de garrins deslletats vius per truja i any), sent aquest un dels principals factors que condicionen la rendibilitat en les granges porcines. Això és especialment important en el porc ibèric amb una prolificidad menor que altres porcs comercials i amb un augment de la substitució de productors tradicionals a granges de maneig intensiu. En el porc blanc, la millora de línies maternes es basa en un esquema piramidal. Els nuclis de selecció subministren reproductors genèticament millorats a les granges multiplicadors, que difonen el guany genètic a través del subministrament d'híbrides (creus) a granges comercials. Clàssicament, les dades de creus (CB) no s'han utilitzat en les avaluacions genètiques de les línies pures (CCPS). No obstant això, en els últims anys, alguns esquemes de millora de porc blanc ho han inclòs per a augmentar la precisió dels valors de millora de les línies pura (PB). D'altra banda, els caràcters reproductius han sigut tradicionalment seleccionats genèticament usant el mètode BLUP (més ben predictor lineal insesgado) basat en dues fonts d'informació: registres fenotípics i informació de pedigrí. No obstant això, hui dia està disponible informació genòmica, com per exemple polimorfismes d'un sol nucleòtid (SNP) al llarg del genoma (genotips). Recentment, aquesta informació s'ha utilitzat juntament amb els fenotips i el pedigrí a través del mètode BLUP genómico d'un sol pas (ssGBLUP) que permet un augment en la precisió del valor de reproducció en comparació amb el BLUP tradicional. En porc ibèric, cap d'aquests nous mètodes, CCPS o SSGBLUP s'han aplicat. L'objectiu d'aquest estudi va ser avaluar la inclusió de dades d'ibèrics croats, així com la inclusió de dades genómicos en l'avaluació genètica d'un esquema piramidal matern de porc ibèric. Un total de 20.468 registres fenotípics per a la grandària de la ventrada (nascuts totals –TB, numere de nascuts vius –NBA i nascuts morts –TB) de 4.753 truges de cinc granges a Badajoz (Espanya), propietat de Inga Food S. a., van ser usats. Les truges ibèriques van correspondre a dues línies pures, Entrepelado (EE) i Retinto (RR), així com als seus encreuaments recíprocs Entrepelado x Retinto (ER) i Retinto x Entrepelado (RE). El pedigrí va incloure 5.533 animals i es va remuntar a 3 generacions. De la població en estudi, 1.435 animals van ser genotipats per a 63.072 marcadors amb Illumina Porcine HD Array 70K de mostres de sang i teixit de la cua. Les anàlisis de dades exploratòries es van realitzar en registres fenotípics, així com en dades genómicos per a detectar possibles “outliers”. Finalment, 20.152 registres fenotípics de 4.717 truges i 1.114 genotips van quedar per a l'anàlisi posterior. Es van avaluar els mètodes BLUP i SSBLUP ajustant en cadascun d'ells dos models diferents. (1) Un model multicaracter trivariante amb els caràcters TB, NBA i SB analitzats per a cada línia de raça pura

(MT); i (2) un model multi raça (MB) basat en un model trivariado en el qual els registres d'un caràcter de cada línia de raça pura i les seues creus es van assumir com a caràcters diferents. Per tant, el model MB es va analitzar per separat per als caràcters TB, NBA i SB, respectivament. L'estimació dels paràmetres genètics es va realitzar mitjançant anàlisi bayesiana amb mostreig de Gibbs. Totes dues avaluacions i models es van validar mitjançant la implementació de cinc grups de validació creuada (CV). La capacitat de predicció es va mesurar calculant la correlació entre el fenotip corregit i el valor genètic estimat de la població de validació i com l'arrel de l'error mig quadràtic (RMSE). A causa del temps de càlcul, el CV es va calcular utilitzant el mètode de màxima versemblança restringida (REML). Les estimacions de heredabilidad per als tres caràcters per BLUP i ssGBLUP van ser similars entre models. La heredabilidad de TB i NBA en EE va anar sempre major que RR. No va haver-hi diferències per a SB. La correlació genètica additiva entre TB i NBA va ser alta i superior a les correlacions entre SB i TB i entre SB i NBA. Les correlaciones TB i NBA entre PB i CB van ser majors respecte a SB. La capacitat predictiva mesura com la correlació de Pearson va ser major per a SB seguida de TB i NBA. Els models MT i especialment ssGBLUP van mostrar una millor capacitat de predicció per als tres caràcters en EE, mentre que els models MB ho van fer per a RR. No es van observar grans diferències en termes de RMSE entre caràcters mitjançant BLUP i ssGBLUP MT, encara que es va trobar un xicotet avantatge dels models genómicos. EE presente RMSE menors que RR. La inclusió de dades genómicos permet reduir lleugerament el biaix en la predicció de (G)EBV, excepte per a NBA en EE i SB en RR sota els models MB.

Paraules clau: porc ibèric, selecció genòmica, BLUP, Single Step GBLUP, prolificitat.

1. Introduction.

1.1. Iberian pig: brief history and general description.

The Iberian pig *(sus scrofa)* is a native breed derived from one of the three ancestral domestic pig populations of the Iberian Peninsula, and is considered as one of the most important of the Mediterranean-type (Silió, 2000). Iberian pig production has always been deeply bounded to the Mediterranean ecosystem (Lopez-Bote, 1998) and is widely recognized as one of the porcine populations with the highest meat quality (Noguera *et al*., 2019). In fact, the most typical and well-known product that represents Iberian pigs is the bellota ham, which acts as flagships of the increasing export market (Nieto *et al*. 2019).

For centuries and until 1960s, the population was characterized for a large effective size, recording 567,000 sows in the official census of 1955 (Silió, 2000). But during the second half of the 20th century various factors such as urban development, intensification of animal production, outbreak of the African swine fever, lowered value of animal fats and the massive introduction of more efficient foreign breeds have drastically reduced the Iberian pig population (Lopez-Bote, 1998; Silió, 2000; Fabuel *et al*., 2004). Nevertheless, during late 1980s started the Iberian pig breeding recovery based on the revalorization of its products as well as on the social awareness for preservation of the genetic heritage and the natural habitat associated to this breed (Nieto *et al*., 2019). At the same time, the Spanish Association of Iberian Pig Breeders (AECERIBER) requested a normative to protect pure Iberian pig production, leading to the creation of the Iberian Pig Herd Book in 1987 (Ministerio de Medio Ambiente y Medio Rural y Marino, 2011; Ministerio de Medio Ambiente y Medio Rural y Marino & Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, 2011). As a result, in November 2019, 11,938 sires and 347,664 were registered in the censuses of the Iberian breed (Ministerio de Agricultura Pesca y Alimentación, 2019).

Nowadays, the Iberian Pig Herd Book recognizes five strains (Entrepelado, Retinto, Torbiscal, Lampiño and Manchado de Jabugo) with remarkable phenotypic and genetic differences among them (Fabuel *et al*., 2004; Ibáñez-Escriche *et al*., 2014; Noguera & Ibáñez-Escriche, 2017). However, all these strains are characterized by their adaptation to tough environmental conditions in the South West of the Iberian Peninsula, their adipogenic nature (Lopez-Bote, 1998; Ibáñez-Escriche *et al*., 2014) and its low productivity (prolificacy and growth) in comparison with commercial (white) pigs (Silió *et al*., 2001; Barea *et al*., 2011).

1.2. Current situation: production systems and breeding programs.

The traditional system of Iberian pig production takes place in the *dehesas* in southwestern Spain, producing heavy pigs destined as high quality dry-cured meat products (Silió, 2000). Pigs are bred under extensive (or semi-extensive) management up to 95-105 kg of body weight and 8-20 months of age, followed by a finishing period or *montanera* in which are fed with acorn and pastures up to 140-160 kg of body weight and 12 and 24 months of age (Lopez-Bote, 1998). However, this production system coexists with the intensive management farms whose implementation has been greatly increased in the last years (Noguera *et al*., 2019). This transformation involved the use of foreign and more efficient breeds such as Duroc Jersey (DU) in order to improve traits related to feed efficiency and noble pieces yield (Noguera & Ibáñez-Escriche, 2017). As regards this, a normative that regulates the Iberian pig products forces the sow to be Iberian, whereas the boar could be either Iberian, DU or hybrid between these two breeds (BOE-A-2014-318, 2014).

Since 1992, AECERIBER is also responsible for the breeding program in Iberian Pig populations based on the records of the farms involved. For this purpose a Commission on Genetic Improvement with animals geneticist, representatives of farmers, industry and national public institutions related to pig production was established (Asociación Española de Criadores de Cerdo Ibérico, 2011). The main goal of the selection program is to improve rentability of farms accounting on different production systems (Ministerio de Medio Ambiente y Medio Rural y Marino & INIA, 2011). To reach this aim, Iberian pigs are being evaluated for productive and reproductive traits by different indexes of selection (Table 1). Involved farms have to share their records to build the Maternal Index whereas participation in Piglet Index and Complete Cycle Index is optional, and for the last one depends on the available infrastructure (Ministerio de Medio Ambiente y Medio Rural y Marino & INIA, 2011). Nevertheless, those establishments which are interested in participating in the Complete Cycle Index are able to provide their destined-to-slaughter animals (males without reproductive function) to the Complete Cycle Test carried out by AECERIBER (Ministerio de Medio Ambiente y Medio Rural y Marino & INIA, 2011). These tests are performed in a private farm or in public testing centers such as The Animal Selection and Reproduction Center of Badajoz (CENSYRA) of the Extremadura joint, the Hontalbilla Pig Testing Center of the Agrarian Technological Institute from Castilla y León and the Pig Selection Center of the Institute of Agrifood Research and Technology (IRTA) in Monells (Ministerio de Medio Ambiente y Medio Rural y Marino & INIA, 2011). The National Institute of Agrifood Research (INIA) is responsible for program supervising and technical counselling as well

as accomplishment and test supervision (Ministerio de Medio Ambiente y Medio Rural y Marino & INIA, 2011).

Table 1. Selection objectives, selection criteria, traits under study and statistical models used in the breeding program in Iberian pig (Ministerio de Medio Ambiente y Medio Rural y Marino & INIA, 2011).

Selection objectives	Selection criteria	Traits under study	Statistical models
General reproductive ability through prolificacy and maternal ability.	Maternal Index	NBA[1]; NPW[2]; LWW[3].	MT-BLUP[8].
Growth rate at early months of age.	Piglet Index	$BW_{70\text{-}90}$[4]	ST-BLUP[9].
Carcass quality as noble cuts yield	Complete Cycle Index	HW%[5]; SW%[6]; %loin[7].	MT-BLUP[8].

[1]NBA: number of born alive; [2]NPW: number of piglets weaned per dam; [3]LWW: litter weight at weaning: [4]BW70-90: individual bodyweight between 70-90 days; [5]HW%: ham weight relative to carcass weight; [6]SW%: shoulder weight relative to carcass weight; [7]%loin: fat-free loin proportion; [8]MT: multi trait animal model Best linear Unbiased Prediction; [9]ST: single trait animal model Best Linear Unbiased Prediction.

The genetic improvement of litter size (LS) in Iberian pig has been limited in comparison to maternal white pig populations in which has been widely improved through the past years (Fernández *et al*., 2008; Noguera and Ibáñez-Escriche, 2017). Until 2011, the reproductive traits were not included in the Iberian pig breeding program carried out by AECERIBER (Corral *et al*., 2010; Ministerio de Medio Ambiente y Medio Rural y Marino & INIA, 2011). Nevertheless, this system has also some flaws since within line selective breeding is not systematically performed for the farmers and is difficult to separate the herd effect from the genetic effect in the evaluations due the low connections between herds. Another point is that the potential advantages of crossbreeding, such as heterosis and line complementarity, have not been exploited either (Ibáñez-Escriche *et al*., 2016). The introduction of a maternal crossbreeding system like that used for modern pig production in commercial populations (Visscher *et al*., 2000) could significantly improve the efficiency of Iberian pig production. In consequence, two non-exclusive strategies leading the genetic improvement of LS could be considered: (1) within-line selection and (2) appropriate crossbreeding scheme between lines to exploit heterosis (Noguera & Ibáñez-Escriche, 2017; Noguera *et al*., 2019).

1.2.1. Within line/breed selection scheme.

Although genetic evaluations provided for AECERIBER are used for some Iberian pig farms to select individuals, other farms perform its own genetic evaluations. The most common

strategy applied in Iberian pigs breeding programs for reproductive traits by big companies as well as by small farms is the within line/breed selection scheme without using crosses between Iberian strains (Noguera & Ibáñez-Escriche, 2017).

1.2.2. Pyramidal Scheme.

Nowadays the 85% of the total Iberian products has its origin on the intensive production system (Noguera & Ibáñez-Escriche, 2017). Its products are mainly 50% Iberian which means animals are the result of crosses between Iberian sows and DU boars. In this context, reproductive efficiency is a limiting factor in the production of Iberian pigs and its improvement is essential for its economic profitability (Noguera *et al*., 2019). A complementary strategy of the within line selection to increase the improvement on reproductive traits is using crossbred sows in the commercial farms. This pyramidal structure, based on the crossbreeding scheme, has the aim of taking advantage of genetic variability within lines and to exploit heterosis and line complementarity (Figure 1) (Ibáñez-Escriche *et al*., 2014; Noguera & Ibáñez-Escriche, 2017). Nevertheless, one of the main limitations of this strategy is that selection is carried out on purebred nucleus lines with high health conditions and their performance can be a poor predictor of the crossbred offspring raised in commercial farms with different environmental conditions (Dekkers, 2007; Ibáñez-Escriche *et al*., 2009; Wientjes & Calus, 2017).

A simple way of combining performance from purebred animals (PB) with information from crossbred relatives (CB) would be considering PB and CB performance as the expression of different traits with a genetic correlation among them (r_{pc}) (Wei & Van Der Werf, 1994). The genetic regression of CB progeny performance on PB performance is influenced by gene frequency and dominance, genotype-environmental interaction (GxE) and different management of PB and CB (Wei & Van Der Werf, 1994; Lutaaya *et al*., 2001). A theory that accounts for all additive and dominance (co)variances among all crosses of two pure lines does exist (Lo *et al*., 1995) but is too complex for practical applications (Lutaaya *et al*., 2001). Thus, to overcome these limitations the most advance approximation is a multiple trait approach in which records from each line are treated as separated traits (Spilke *et al*., 1998) leading to the implementation of combined crossbred and purebred selection (CCPS) (Wei & Van Der Werf, 1994):

$$\begin{aligned} \boldsymbol{y}_A &= \boldsymbol{X}_A\boldsymbol{\beta}_A + \boldsymbol{Z}_A\boldsymbol{a}_A + \boldsymbol{e}_A \\ \boldsymbol{y}_B &= \boldsymbol{X}_B\boldsymbol{\beta}_B + \boldsymbol{Z}_B\boldsymbol{a}_B + \boldsymbol{e}_B \\ \boldsymbol{y}_{AB} &= \boldsymbol{X}_{AB}\boldsymbol{\beta}_{AB} + \boldsymbol{c}_{AB} + \boldsymbol{e}_{AB} \end{aligned} \qquad [1]$$

where the vectors $\boldsymbol{y}_A, \boldsymbol{y}_B$ and $\boldsymbol{y}_{AB}$ contain phenotypic records on breeds A, B and crossbred AB animals, respectively; $\boldsymbol{\beta}_A$, $\boldsymbol{\beta}_B$ and $\boldsymbol{\beta}_{AB}$ are vectors of fixed effects for A and B breeds and CB respectively and $\boldsymbol{e}_A$, $\boldsymbol{e}_B$ and $\boldsymbol{e}_{AB}$ are vectors of residuals for A breed, B breed and CB respectively; $\boldsymbol{a}_A$ and $\boldsymbol{a}_B$ are vectors of random effects (breeding value) for breed A and B respectively; $\boldsymbol{c}_{AB}$ is the vector of random effects (breeding value) for CB; $\boldsymbol{X}$ and $\boldsymbol{Z}$ are incidence matrices that relates phenotypic records with fixed and random effects, respectively. This methodology has not been applied in Iberian pig because until now a pyramidal breeding system were not implemented. Moreover, its traditional extensive production system makes difficult to control the pedigree with precision as well to register reproductive data (Fernández *et al*., 2008). However, its implementation in Iberian pig breeding schemes has the potential to meaningfully improve the Iberian performance (Ibáñez-Escriche *et al*., 2014).

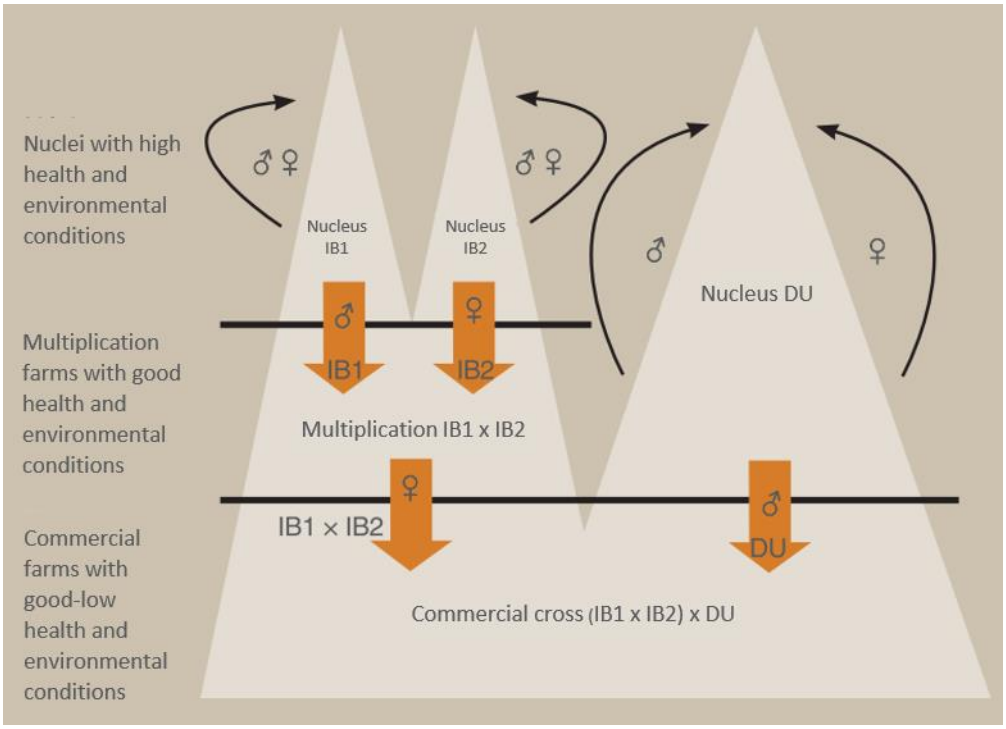

Figure 1: Typical pyramidal scheme of white pig selection programs adapted to Iberian pig (Noguera & Ibáñez-Escriche, 2017)**.**

1.3. Selection: methods to estimate breeding values.

The breeding value (BV) of an individual can be defined as the sum of the average effects of alleles of the genes it carries (at each *locus* for all *loci*) which are transmitted to offspring (Falconer, 1976). If an animal i is randomly mated to a number n of individuals from the

population, then its BV is twice the mean deviation of the progeny from the population mean (Falconer, 1976):

$$BV_i = 2 * EPD \quad [2]$$

where BV_i is the breeding value for the ith animal and EPD is the expected progeny difference (Falconer, 1976). Since only additive effects are inherited from parents to offspring, BV can be also expressed as additive genetic value (a).

Prediction of the BVs constitutes a key step in breeding programs (Wang *et al.*, 2018) and the type and amount of information available on candidates for selection lead to define the method to be employed for their prediction (Mrode, 2005). In consequence, many methods have been developed during the last decades to predict or estimate BVs. Desirable criteria for evaluations models such as unbiasedness, minimization of variance error of prediction and maximization of correlation between predictor and predictand were described by Henderson (1975) who defined a method capable of fulfill these criteria (the variance parameter have to be the true ones) as Best Linear Unbiased Prediction (BLUP).

1.3.1. Traditional selection: BLUP.

Conventional phenotypic selection (mass selection) in livestock has been practiced successfully for hundreds of years, but in the last 50 years after the BLUP application (Henderson, 1975) the response to selection has increased considerably (Goddard, 2009). The BLUP is a technique for estimating random effects and in animal breeding has been mainly used for estimating BVs. Its acronym stands for "Best Linear Unbiased Prediction" because maximizes the correlation between true and predicted BV or minimized the prediction error variance (PEV), its predictors are a linear functions of observations and the realized values for random variable (e.g. BVS) and the estimable functions of fixed effects are unbiased (Mrode, 2005). Thus, considering a mixed linear model:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Za} + \boldsymbol{e} \quad [3]$$

where $\boldsymbol{y}$ is the vector of records; $\boldsymbol{\beta}$ is the vector of fixed effects; $\boldsymbol{a}$ is the vector of breeding values; $\boldsymbol{e}$ is the vector of the associated errors; $\boldsymbol{X}$ and $\boldsymbol{Z}$ are incidence matrices of $\boldsymbol{\beta}$ and $\boldsymbol{a}$; Henderson (1963) shown that BLUP can be obtained from a solution to the following equations, best known as Mixed Model Equations (MME):

$$\begin{bmatrix} \boldsymbol{XX'} & \boldsymbol{X'Z} \\ \boldsymbol{Z'X} & \boldsymbol{Z'Z + A^{-1}} \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{a}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X'y} \\ \boldsymbol{Z'y} \end{bmatrix} \quad [4]$$

where $\boldsymbol{A^{-1}}$ represents the inverse of numerator relationship matrix for all animals, with or without records; $\widehat{\boldsymbol{\beta}}$ is the vector of solutions for fixed effects and $\widehat{\boldsymbol{a}}$ is the vector of solutions for random effects (BV). Therefore, BVs are estimated (EBVs) using phenotypes and genetic relationships among individuals which are based on pedigree records (Mrode, 2005).

1.3.2. Genomic Selection: from SNP-BLUP to single step Genomic BLUP.

Nowadays a third source of information is available to be considered in a breeding program (Hayes & Goddard, 2010). The rapid development of DNA technologies has allowed having a large number of markers at low cost, being the simplest DNA marker the Single Nucleotide Polymorphism (SNP) (Blasco, 2017). Strictly, SNPs are variations in a base at the same point in the genome among individuals or between individual's chromosome pairs (Hayes & Goddard, 2010) showing the difference between DNA inherited by two individuals (Legarra *et al*., 2014). SNPs can be associated with many of the genes controlling the economically interesting traits, even though the effect of each gene is small (Blasco, 2017). The use of DNA markers to quantitative traits as a tool to map quantitative trait loci (QTL) to chromosome regions was called Marker Assisted Selection (MAS) (Goddard, 2009). Nevertheless, it was not possible to discover enough QTLs to explain a high proportion of the genetic variance in most of the traits and MAS was not as successful as was expected (Goddard, 2009; Wang *et al*., 2018). Owing to this limitation, Meuwissen *et al*. (2001) proposed an approach considering the use of all genome-wide markers at once, so that all genetic variance can be explained by them. It is assumed that there are always some markers in linkage disequilibrium (LD) with any QTL. Hence, sufficient high marker density guarantees near-perfect LD between at least one marker and each QTL (Wang *et al*., 2018). The use of all these genome-wide markers simultaneously considering LD between the marker and the QTL to predict BVs is known as genomic selection (GS) (Meuwissen *et al*., 2001; Ibáñez-Escriche *et al*., 2009; Hayes & Goddard, 2010).

GS implementation is possible if reference population with both phenotype and genotype records exists, which is used to derive prediction equations that predicts genomic estimated breeding values (GEBVs) of animals from markers effects (Hayes & Goddard, 2010):

$$GEBV = \sum_{i}^{n} x_i \hat{g}_i \qquad [5]$$

where n is the number of SNP markers; $\hat{g}_i$ is the effect of the i SNP marker and x is the incidence matrix that relates markers' genotypes to animals. The equations can be used to predict BVs from those selection candidates who have marker genotypes but do not have records (Hayes &

Goddard, 2010). The main advantage of GS relies on decreasing generation interval and increasing accuracy of EBVs (Hayes & Goddard, 2010). However, its accuracy is affected by many factors such as sample size, genetic relationships, marker density, heritability and LD between SNP and QTL (Wang *et al*., 2018). Besides, GS approach does not require pedigree recordings (Meuwissen *et al*., 2016). The impact of GS in terms of selection gain will be greater for those traits which are either sex limited, expensive or difficult to measure, only measurable after slaughter or expressed late in life (Hayes & Goddard, 2010). Hence, in the last years GS has become a main tool in dairy cattle, pig and poultry breeding programs (Blasco, 2017).

In the last years, several methods for deriving prediction equations based on different priors assumptions about distributions of SNPs effects as well as considering SNP markers individually or by a genotype matrix have been defined (Table 2) (Hayes & Goddard, 2010). One of the possible assumptions of the SNP effects is to be normally distributed $\boldsymbol{g} \sim N(0, \boldsymbol{I}\sigma_g^2)$ with the same variance across all SNPs and uncorrelated (Blasco, 2017) as in SNP-BLUP, Genomic BLUP (GBLUP) or single step Genomic BLUP (ssGBLUP). The advantage of considering marker effects normally distributed is its algebraic simplicity (Legarra, 2015). In SNP-BLUP, markers effects are estimated individually (Meuwissen *et al*., 2001; Blasco, 2017) by solving MME (Henderson, 1963). The GBLUP is similar to SNP-BLUP but in this case BV are estimated simultaneously by applying a genomic relationship matrix for genotyped animals (Blasco, 2017). Finally, in ssGBLUP the idea is to replace the pedigree relationships with a matrix that contains the genomic relationships for the genotype animals and the pedigree relationships for those without genotypic data (Meuwissen *et al*., 2016).

Table 2. Models for implementation of genomic selection according how single nucleotide polymorphisms (SNPs) effects are considered and prior assumptions for SNPs effects.

SNP effect	Individually considered	Genomic matrix
Normally distributed	SNP-BLUP[1]	GBLUP[2] ssGBLUP[3]
t-distribution	Bayes A	
Mixed distribution of zero effect and t-distribution	Bayes B	
Mixed distribution of zero effect and normal distribution	Bayes C	
Double exponential distribution	Bayes LASSO[4]	
Mixture of four normal distributions	Bayes R	

[1]BLUP: Best Linear Unbiased Prediction; [2]GBLUP: Genomic Best Linear Unbiased Prediction; [3]ssGBLUP: Single Step Genomic Best Linear Unbiased Prediction; [4]LASSO: Least Absolute Shrinkage and Selection Operator.

There are also methods assuming that only a fraction (π) of the SNPs have an effect on the trait of interest, therefore, a fraction (1 – π) have no effect which is biologically reasonable (Meuwissen *et al*., 2016). These prior assumptions lead to the application of Bayesian regression methods, in which SNPs can have different posterior distributions. Nevertheless, regression coefficients do not reflect the weight of each SNP and these coefficients may vary according to the number of SNPs (Blasco, 2017). Several Bayesian procedures named as Bayes Alphabet have been developed and characterized by different degrees of shrinkage which depend on the prior assumptions (Blasco, 2017). Bayes A assumes prior t-distribution for all SNPs and with the same variance each one. Bayes B is similar to Bayes A, but in this case a large number of SNPs are considered to have zero effect (Meuwissen *et al*., 2001). Bayes C is similar to SNP-BLUP, but in this case a large number of SNPs are considered to have zero effect (Habier *et al*., 2011). Bayes LASSO assumes prior double exponential distribution for all the SNPs (De Los Campos *et al*., 2009). Bayes R consider all SNPs assigned with a mixture of four normal distributions (Erbe *et al*., 2012). Nevertheless, in routinely genomic evaluations the most used are the GBLUP and the ssGBLUP methods.

1.3.2.1. Genomic BLUP.

The Genomic BLUP (GBLUP) method is very similar to BLUP method, except for the replacement of pedigree relationships by the genomic relationships matrix (Meuwissen *et al*., 2016). In BLUP method EBVs are estimated based on phenotype records and pedigree information whereas in

GBLUP method GEBVs are estimated using phenotype records and genomic relationships based on genome-wide dense marker data (Meuwissen *et al*., 2016). GBLUP assumes that SNPs effects are normally distributed with the same variance for every SNP (Meuwissen *et al*., 2016) like in SNP-BLUP. GBLUP results are equivalent to SNP-BLUP by solving MME in which GEBV are predicted instead of marker effects (Legarra *et al*., 2014):

$$\begin{bmatrix} \boldsymbol{X'X} & \boldsymbol{X'Z} \\ \boldsymbol{Z'X} & \boldsymbol{Z'Z + G^{-1}} \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{u}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X'y} \\ \boldsymbol{Z'y} \end{bmatrix} \quad [6]$$

where $\widehat{\boldsymbol{\beta}}$ is the vector of fixed effects, $\widehat{\boldsymbol{u}}$ is the vector of (G)EBV and $\boldsymbol{G}$ is the genomic relationship matrix, with:

$$var(\boldsymbol{u}) = ZD_gZ' \quad [7]$$

where $\boldsymbol{Z} = \boldsymbol{M} - \boldsymbol{P}$, being $\boldsymbol{M}$ the matrix of dimensions number of individuals (*n*) x the number of loci (*m*) that specifies which marker alleles each individual inherited (coded as -1, 0, 1 for the homozygote, heterozygote, and other homozygote, respectively) and $\boldsymbol{P}$ contains allele frequencies expressed as a difference from 0.5 and multiplied by 2, such that column i in $\boldsymbol{P}$ is $2(p_i - 0.5)$ where p_i is the frequency of the second allele at locus i (VanRaden, 2008). Subtraction of $\boldsymbol{P}$ from $\boldsymbol{M}$ sets mean values of the alleles effects to 0. And $D_g = I * \frac{\sigma_a^2}{2\sum p_i q_i}$ where σ_a^2 is the variance for marker effects and p_i and q_i are the allele frequencies at the locus i. Therefore:

$$var(\boldsymbol{u}) = \sigma_{\boldsymbol{a}}^2\boldsymbol{G} \quad [8]$$

where:

$$\boldsymbol{G} = \boldsymbol{ZZ'}/2\sum(1 - p_i) \quad [9]$$

Division by $2\sum(1 - p_i)$ scales $\boldsymbol{G}$ to be analogous to the numerator relationship matrix $\boldsymbol{A}$ (VanRaden, 2008).

Therefore, $\boldsymbol{G}$ represents the genomic relationship matrix which can be seen as an improved estimator of relationships based on markers instead of pedigrees (Legarra *et al*., 2014).

1.3.2.2. Single step GBLUP.

One of the main limitations of the above-mentioned methods (SNP-BLUP; GBLUP and Bayesian regression methods) is that only genotyped individuals are used. However, in the breeding schemes many animals are not genotyped, but their phenotypes can be highly valuable for

estimate their own BV or of its relatives. At first, genomic predictions were obtained by combining traditional genetic evaluation results with genotypic data (VanRaden, 2008; Vitezica *et al*., 2010). This multiple-step procedure (Hayes *et al*., 2009; VanRaden *et al*., 2009) included running a regular BLUP animal model evaluation based on pedigree, extracting pseudo-phenotypes for genotyped animals (e.g., daughter deviations – DD or de-regressed EBV), estimating SNP effects using pseudo-data as records by simple sire models and combining genomic prediction with parent averages (PA) (VanRaden, 2008; Aguilar *et al*., 2010; Misztal *et al*., 2013). This method can lead to a lack of information, inaccuracies and biases (Legarra *et al*., 2014). First of all, key parameters such as pseudo-observations were difficult to obtain (Legarra *et al*., 2014) and its use may inflate genetic evaluation accuracy when they were computed from animals with small progeny numbers (Vitezica *et al*., 2010). Moreover, most genotyped animals have undergone strong selection and the information of a close relative is ignored in the genomic prediction as well as in the creation of pseudo-phenotypes and covariances among pseudo-phenotypes are not corrected modelled (Legarra *et al*., 2014). Later, a model capable of solving multiple-step issues by blending information between genotyped and non-genotyped animals in one step was developed and named Single Step Genomic BLUP (ssGBLUP). The origin of the ssGBLUP method was raised based on two different points of view which were developed almost simultaneously. Legarra *et al*. (2009) developed the method as a Bayesian update of the relationship matrix whereas Christensen and Lund (2010) developed the method as a tool to "impute" phenotypes (Legarra *et al*., 2014). The result of both developments leaded an augmented or extended genomic relationship matrix ($\boldsymbol{H}$) for both genotyped and non-genotyped animals (Legarra *et al*., 2014):

$$\boldsymbol{H} = \begin{pmatrix} \boldsymbol{A_{11}} - \boldsymbol{A_{12}A_{22}^{-1}A_{21}} + \boldsymbol{A_{12}A_{22}^{-1}GA_{22}^{-1}A_{21}} & \boldsymbol{A_{12}A_{22}^{-1}G} \\ \boldsymbol{GA_{22}^{-1}A_{21}} & \boldsymbol{G} \end{pmatrix} \quad [10]$$

where subscripts 1 and 2 indicates ungenotyped and genotyped animals respectively; $\boldsymbol{A}$ is the numerator relationship matrix and $\boldsymbol{G}$ the genomic relationship matrix. The ssGBLUP provides an explicit inverse of the augmented matrix $\boldsymbol{H}$ (Legarra *et al*., 2014):

$$\boldsymbol{H^{-1}} = \boldsymbol{A^{-1}} + \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{G^{-1}} - \boldsymbol{A_{22}^{-1}} \end{bmatrix} \quad [11]$$

thus, its application to genomic evaluation is immediate (Legarra *et al*., 2014) by applying the mixed model:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{e} \quad [12]$$

where $\boldsymbol{y}$ is the vector of records; $\boldsymbol{\beta}$ is the vector of fixed effects; $\boldsymbol{u}$ is the vector of breeding values; $\boldsymbol{e}$ is the vector of the associated errors; $\boldsymbol{X}$ and $\boldsymbol{Z}$ are incidence matrices of the vectors $\boldsymbol{\beta}$ and $\boldsymbol{u}$, respectively, with $Var(u) = \boldsymbol{H}\sigma_u^2$ and $Var(e) = \boldsymbol{I}\sigma_e^2$. Then, solutions are obtained by solving the MME using $\boldsymbol{H}$ matrix instead of $\boldsymbol{A}$ matrix:

$$\begin{bmatrix} \boldsymbol{XX'} & \boldsymbol{X'Z} \\ \boldsymbol{Z'X} & \boldsymbol{Z'Z} + \boldsymbol{H}^{-1} \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{u}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X'y} \\ \boldsymbol{Z'y} \end{bmatrix} \quad [13]$$

Although this method is straightforward applied in routinely genetic evaluation, its implementation has some issues due to the incompatibilities between $\boldsymbol{A_{22}}$ and $\boldsymbol{G}$ matrices that arises in the different amount of information accounting for changes on allele frequencies (Masuda, 2019). To solve this issue, some adjustments should be implemented. First, $\boldsymbol{G}$ should be blending as $\boldsymbol{G} *= \alpha\boldsymbol{G} + \beta\boldsymbol{A_{22}} + \gamma\boldsymbol{I} + \delta\boldsymbol{11'}$ (usually $\gamma = 0$, $\delta = 0$ and $\alpha + \beta = 1$). Second, $\boldsymbol{G}$ should be tuning to be scaled to $\boldsymbol{A_{22}}$, because in real populations most genotypes are from recent generations, but the pedigree comes from long generations. Finally, $\boldsymbol{G}$ and $\boldsymbol{A_{22}}$ need to be scaled in order to achieve the maximum predictive ability in GEBVs for young animals in case of having incomplete pedigrees or/and unqualified genotypes (Masuda, 2019).

1.4. Impact of genomics on crossbreeding schemes

Despite proving CCPS increase response to selection for CB performance in comparison to the classical method of selection on PB performance, its application was limited due to the difficulty and cost of routine collection of pedigree data on the field (Dekkers, 2007; Duenk *et al*., 2019b). However, the availability of genetic markers can overcome this main issue (Ibáñez-Escriche *et al*., 2009). Further, it could reduce the lack of generations between CB and PB and it makes accommodating non-additive gene action easier than the classical CCPS (Dekkers, 2007; Esfandyari *et al*., 2016). Dominance can be an explanation of the heterosis (Falconer & Mackay, 1996; Charlesworth & Willis, 2009). Thus, the inclusion of dominance in GS models may be beneficial for selection of PB for CB performance (Zeng *et al*., 2013).

Several genomic models have been suggested for the prediction of PB individuals BVs for CB performance (Esfandyari *et al*., 2018). These models are the standard additive genomic prediction models (Dekkers, 2007; Christensen *et al*., 2014), models with across-breed effects of SNP genotypes (ASGM) or with breed-specific effects of SNP alleles (BSAM) (Ibáñez-Escriche *et al*., 2009), dominance models (Zeng *et al*., 2013) and breed-specific dominance model (BSDM) (Esfandyari *et al*., 2015). The ASGM assumed SNP effects to be the same across breeds (Ibáñez-Escriche *et al*., 2009). Under this context, r_{pc} can be also estimated when the PB and CB animals

are more distantly related or when pedigree information is not recorded by replacing de $\boldsymbol{A}$ matrix by $\boldsymbol{G}$ (VanRaden, 2008; Legarra *et al*., 2015). Genomic relationships between PB and CB animals should ideally be based on alleles from only one of the PB parental lines (Duenk *et al*., 2019a). However, the ordinary $\boldsymbol{G}$ is based on both alleles of an individual, which in case of CB individuals also include those from the another PB line (Duenk *et al*., 2019a). BSAM assumed breed-specific effects of SNP alleles (Ibáñez-Escriche *et al*., 2009). This model was proposed by Dekkers (2007) based on the fact that substitution effects at the QTL for paternal and maternal alleles would be different if the parental breeds differ in alleles frequencies at the QTL (Zeng *et al*., 2013). Zeng *et al*. (2013) proposed a dominance model that simultaneously fits additive and dominance effects of SNPs enabling the computation of allele substitution effects using appropriate allele frequencies. Finally, Esfandyari *et al*. (2015) extend the dominance model considering the breed specific allele which allows to distinguish between alternate heterozygotes in the CB. These studies have shown that dominance models have a better performance than the additive ones. However, these results are mainly based on simulation studies. In practice, only the additive models have been applied showing that the BSAM models only are superior to ASGM models when r_{pc} values are lower than 0.7. Thus, CCPS can be applied for many traits and breeding schemes using the classical ASGM.

1.5. Outline.

The application of Iberian pig genomic evaluations for reproductive traits such as LS with low heritability and recording in only one sex could be appropriate to improve genetic gain. Furthermore, sows are selected early in life and own records of LS are not available in the evaluations. Nevertheless, the herd nucleus has a small size leading limited phenotypes and genotypes for applying genomic evaluations. Thus, considering commercial sows' records would be a valuable strategy for both genomic and genetic evaluations. Although this strategy has already been studied in white pigs, it has never been evaluated in Iberian pig. In this study we evaluate the implementation of genomic information and the crossbred animal records in the Iberian Pig genetic evaluations of the Inga Food company S.A. This is the only company applying an Iberian maternal pyramidal scheme as in the white pig breeding programs. This study is framed underlying a research line for improvement of reproductive efficiency in Iberian pig by the Institute of Agrifood Research and Technology (IRTA), University of Zaragoza (UNIZAR), Polytechnic University of Valencia (UPV) and the Autonomous University of Barcelona (UAB) in collaboration with Inga Food company S.A.

1.6. Objective.

The aim of this study was to evaluate the inclusion of crossbred information as well as inclusion of genomic data in the genetic evaluation of a pyramidal Iberian pig maternal breeding scheme for improving the accuracy of the estimated breeding values.

2. Materials and methods.

2.1. Ethics statements.

The research ethics committee of Institute of IRTA approved all the management and experimental procedures involving live animals, which were performed in accordance with the Spanish Policy of Animal Protection RD1201/05, which complies with the European Union Directive 86/609 about the protection of animals used in experimentation.

2.2. Phenotypic data.

Phenotypic data of LS come from five farms located in Almadrejo (Badajoz), Spain, belonging to the Iberian pig breeding scheme of the Inga Food company S.A. Data sets compromised two purebred lines recognized in Spain's official Iberian herd-book (AECERIBER), Entrepelado (EE) and Retinto (RR) and their reciprocal crosses (Entrepelado x Retinto – ER and Retino x Entrepelado – RE). LS was quantified as the number of total born (TB), number of born alive (NBA) and stillborn (SB). A total amount of 20468 records for LS of 4753 sows were registered from 2010 to 2019. The pedigree included 5533 animals and went back 3 generations. The distribution of data according location in farms and sow line is presented in Table 3. The purebred sows were located into two selection farms under intensive commercial management, while the purebred boars were kept in an artificial insemination center. The third group, PB and CB, were located in three commercial production farms where both purebred and crossbred sows were mated with DU boars according to usual commercial production system in Iberian pig. Distribution of records according breed of boar of service are shown in Table 4.

Table 3. Number of litter size records according to sow line and farms.

	EE[1]	ER[2]	RE[3]	RR[4]	Total
Farm 1	1683	19	36	2328	4066
Farm 2	2757	37	39	1910	4743
Farm 3	611	4116	1953	1005	7685
Farm 4	2003	0	0	1544	3457
Farm 5	179	30	51	167	427
Total	7233	4202	2079	6954	20468

[1]EE: Entrepelado line; [2]RR: Retinto line; [3]ER: Entrepelado x Retinto cross; [4]RE: Retinto x Entrepelado cross.

Table 4. Number of litter size records according to breed of service boar

Total of records	20468
Duroc Jersey	8174
Entrepelado line	6220
Retinto line	6074

2.2.1. Exploratory data analysis.

Exploratory analysis of TB, NBA and SB were conducted using the R language (R core Team, 2014). Distribution of phenotypic data was checked by histograms and outliers were identified using boxplots analysis. After exploratory data analysis, 20152 records from 4714 sows remained from later analysis. Those records were distributed as follows: 7132 from 1635 EE purebred sows, 6843 records from 1575 RR purebred sows, 4132 records from for 906 ER sows and 2045 records from 598 RE sows.

2.3. Genomic data.

A total number of 1435 animals were genotyped from DNA isolated from samples of blood and tail tissue. DNA from tails' samples was extracted according to phenol/chloroform protocol whereas DNA from blood's samples was extracted following two different methods. First, extraction of DNA on plate was carried out using Invisorb® Blood Mini HTS 96 kit/C (STRATEC Molecular GmbH, 2013) and, second, E. Z. N. A Blood DNA Mini kit d'OMEGA (Omega Bio-tek, 2017) for extraction of DNA in column. Genotyping was performed with Illumina GGP Porcine HD Array 70K (Illumina Inc., San Diego, CA, USA) which contains 63072 Single Nucleotide Polymorphism (SNPs).

2.3.1. Quality control of genomic data.

Quality control (QC) of genomic data was performed with Plink software (Purcell, 2009) by removing SNPs: (i) with minor allele frequency (MAF) below 0.05, (ii) with missing genotype rate per marker > 10%, (iii) with missing genotype rate per individual > 10% and (iv) mapped to sexual chromosomes. Additionally, plink software was used to perform principal components analysis (PCA) to identify outliers. QC was also implemented with PreGSF90 software (Misztal & Tsuruta, 2018) in order to verify and eliminate parent-progeny Mendelian Conflicts. The genomic relationship matrix ($\boldsymbol{G}$) and the additive relationship matrix for genotyped animals ($\boldsymbol{A_{22}}$) were built to check genomic relationships and to identify inconsistencies in the pedigree or mislabeled animals. Analysis of diagonal and off-diagonal elements of $\boldsymbol{G}$ and numerator relationship matrix for genotyped animals $\boldsymbol{A_{22}}$ were conducted in R (R core team, 2014). After

that, 1114 genotyped animals and 34496 SNP markers remained for further analysis. Detailed information of the markers by sex and line after QC is shown Table 5.

Table 5. Information of the number of markers and animals genotyped by sex and line.

	All animals	Entrepelado pure line	Retinto pure line
Markers	34496	34741	26226
Animals	1114	382	357
Females	1040	341	327
Males	74	41	30

2.4. Statistical analysis.

A first step was to identify systematic effects influencing the LS traits. For that, a general linear model for each systematic effect (e.g. parity order, farm-year-season) was fitted to the traits under study using R (R core Team, 2014). Variables with a relevant effect on traits were included in later analysis. To decide whether an effect was relevant for each trait or not confidence intervals were computed. Thus, an effect was considered relevant when zero value was not contained in its 95% confidence interval.

2.4.1. Models.

Two different models were fitted depended on the data set (purebred lines only or purebred lines and its crosses). First, for each purebred line a trivariate multi trait model (MT) for TB, NBA and SB was fitted. Second, for each trait separately, a so-called multi breed model (MB) was fitted. This model is a multivariate model in which the records of each purebred line and crossbred are considered as a different trait (Wei & Van Der Werf, 1994). In this case, diallelic crosses between purebred lines (ER and RE) were considered as only one trait instead of two.

2.4.1.1. Multi trait model (MT).

The assumed model for TB, NBA and SB was:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Za} + \boldsymbol{Wp} + \boldsymbol{e} \qquad [14]$$

where $\boldsymbol{y}$ is the vector of observations (TB, NBA and SB); $\boldsymbol{\beta}$ is the vector of systematic effects including parity order, mating boar line and herd-year-season of parity; $\boldsymbol{a}$ is the vector of additive genetic effects; $\boldsymbol{p}$ is the vector of permanent environmental effects of the dam; $\boldsymbol{e}$ is the vector of residuals; and $\boldsymbol{X}$, $\boldsymbol{Z}$ and $\boldsymbol{W}$ are the incidence matrix associated to the vectors $\boldsymbol{\beta}$, $\boldsymbol{a}$, and $\boldsymbol{p}$, respectively.

In matrix form:

$$\begin{bmatrix} \boldsymbol{y_{TB}} \\ \boldsymbol{y_{NBA}} \\ \boldsymbol{y_{SB}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X_{TB}} & 0 & 0 \\ 0 & \boldsymbol{X_{NBA}} & 0 \\ 0 & 0 & \boldsymbol{X_{SB}} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{b_{TB}} \\ \boldsymbol{b_{NBA}} \\ \boldsymbol{b_{SB}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{Z_{TB}} & 0 & 0 \\ 0 & \boldsymbol{Z_{NBA}} & 0 \\ 0 & 0 & \boldsymbol{Z_{SB}} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{a_{TB}} \\ \boldsymbol{a_{NBA}} \\ \boldsymbol{a_{SB}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{W_{TB}} & 0 & 0 \\ 0 & \boldsymbol{W_{NBA}} & 0 \\ 0 & 0 & \boldsymbol{W_{SB}} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{p_{TB}} \\ \boldsymbol{p_{NBA}} \\ \boldsymbol{p_{SB}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{e_{TB}} \\ \boldsymbol{e_{NBA}} \\ \boldsymbol{e_{SB}} \end{bmatrix} \quad [15]$$

where subscripts $\boldsymbol{TB, NBA}$ and $\boldsymbol{SB}$ indicate the traits.

2.4.1.2. Multi breed model (MB).

The next model was fitted for each trait (TB, NBA and SB):

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Za} + \boldsymbol{Wp} + \boldsymbol{e} \quad [16]$$

where $\boldsymbol{y}$ is the vector of observations of the trait for the EE, RR and ER-RE lines; $\boldsymbol{\beta}$ is the vector of systematic effects including parity order, mating boar line and herd-year-season of parity and TB as covariate when SB was evaluated; $\boldsymbol{a}$ is the vector of additive genetic effects; $\boldsymbol{p}$ is the vector of permanent environmental effects of the dam; $\boldsymbol{e}$ is the vector of residuals; and $\boldsymbol{X}$, $\boldsymbol{Z}$ and $\boldsymbol{W}$ are the incidence matrix associated of the vectors $\boldsymbol{\beta}$, $\boldsymbol{a}$, and $\boldsymbol{p}$, respectively. In matrix form:

$$\begin{bmatrix} \boldsymbol{y_{EE_i}} \\ \boldsymbol{y_{RR_i}} \\ \boldsymbol{y_{ER-RE_i}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X_{EE_i}} & 0 & 0 \\ 0 & \boldsymbol{X_{RR_i}} & 0 \\ 0 & 0 & \boldsymbol{X_{ER-RE_i}} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{b_{EE_i}} \\ \boldsymbol{b_{RR_i}} \\ \boldsymbol{b_{ER-RE_i}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{Z_{EE_i}} & 0 & 0 \\ 0 & \boldsymbol{Z_{RR_i}} & 0 \\ 0 & 0 & \boldsymbol{Z_{ER-RE_i}} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{a_{EE_i}} \\ \boldsymbol{a_{RR_i}} \\ \boldsymbol{a_{ER-RE_i}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{W_{EE_i}} & 0 & 0 \\ 0 & \boldsymbol{W_{RR_i}} & 0 \\ 0 & 0 & \boldsymbol{W_{ER-RE_i}} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{p_{EE_i}} \\ \boldsymbol{p_{RR_i}} \\ \boldsymbol{p_{ER-RE_i}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{e_{EE_i}} \\ \boldsymbol{e_{RR_i}} \\ \boldsymbol{e_{ER-RE_i}} \end{bmatrix} \quad [17]$$

where subscript i indicates the analyzed trait (i= TB, NBA or SB) and EE, RR and ER-RE the corresponding breed data.

2.4.2. Bayesian analysis.

The Bayesian analysis were performed with the BLUPF90 family programs (Misztal & Tsuruta, 2018). All models assumed flat prior distributions for systematic effects ($\boldsymbol{\beta}$) and multivariate Gaussian distributions for additive ($\boldsymbol{a}$), permanent ($\boldsymbol{p}$) and residual effects ($\boldsymbol{e}$). Permanent environmental effects and residual effects were assumed identically and independently distributed, their prior distributions were as follows:

$\boldsymbol{p} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{I} \otimes \boldsymbol{P})$ being $\boldsymbol{I}$ an identity matrix, $\boldsymbol{P}$ the (3x3) permanent environmental variance-covariance matrix of the traits in MT model or the line for trait i in the MB model and $\otimes$ the Kronecker product operator.

$\boldsymbol{e} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{I} \otimes \boldsymbol{R})$ being $\boldsymbol{I}$ an identity matrix, $\boldsymbol{R}$ the (3x3) residual variance-covariance matrix with the residual variances of the trait in MT model or the line for trait i in the MB model and $\otimes$ the Kronecker product operator.

The distribution of the additive genetic effect was assumed as follows:

$\boldsymbol{a} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{K} \otimes \boldsymbol{G_0})$, being $\boldsymbol{K}$ the numerator relationship matrix among animals, $\boldsymbol{G_0}$ the (3x3) additive genetic variance-covariance matrix of the traits in MT model or the line for trait i in the MB model and $\otimes$ the Kronecker product operator. The $\boldsymbol{K}$ matrix was equal to $\boldsymbol{A}$ when the traditional BLUP method with phenotypic and pedigree information was used. However, the $\boldsymbol{K}$ matrix was equal to $\boldsymbol{H}$ when the ssGBLUP method with phenotypic, pedigree and genotyped data was used. $\boldsymbol{H}$ matrix results from blending genomic and pedigree information for both genotyped and non-genotyped animals (Aguilar *et al*., 2010; Legarra *et al*., 2009, 2014). $\boldsymbol{H}$ matrix can be defined as:

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{A_{11}} & \boldsymbol{A_{12}} \\ \boldsymbol{A_{21}} & \boldsymbol{G} \end{bmatrix} = \boldsymbol{A} + \begin{bmatrix} 0 & 0 \\ 0 & \boldsymbol{G} - \boldsymbol{A_{22}} \end{bmatrix} \quad [18]$$

where subscripts 1 and 2 represent ungenotyped and genotyped animals, respectively; $\boldsymbol{A}$ is the numerator relationship matrix and $\boldsymbol{G}$ is a genomic relationship matrix (Aguilar *et al*., 2010). $\boldsymbol{G}$ was constructed as $\boldsymbol{G} = \boldsymbol{Z}\boldsymbol{Z}'/2\sum pq$ (VanRaden, 2008), being $\boldsymbol{Z}$ a matrix for SNP markers and p and q marker allele frequencies. $\boldsymbol{Z}$ can be constructed as: $\boldsymbol{Z} = \boldsymbol{M} - \boldsymbol{P}$ where $\boldsymbol{M}$ is a matrix of dimensions (n) number of individuals x (m) number of *loci,* that specifies which marker alleles each individuals inherited and $\boldsymbol{P}$ a matrix containing allele frequencies expressed as a difference from 0.5 and multiplied by two (VanRaden, 2008).

In order to solve mixed model equations (Henderson, 1963) $\boldsymbol{H^{-1}}$ is needed. Therefore, $\boldsymbol{H^{-1}}$ can be defined as:

$$\boldsymbol{H^{-1}} = \boldsymbol{A^{-1}} + \begin{bmatrix} 0 & 0 \\ 0 & \boldsymbol{G^{-1}} - \boldsymbol{A_{22}}^{-1} \end{bmatrix} \quad [19]$$

Where $\boldsymbol{A_{22}^{-1}}$ is the inverse of numerator relationship matrix for genotyped animals only (Legarra *et al*., 2014). Nevertheless, $\boldsymbol{G}$ cannot be inverted directly due to its singular or close to singular appearance. In order to achieve inversion of $\boldsymbol{G}$, some weights were implemented as described by Aguilar *et al*. (2010) and VanRaden (2008):

$$\mathbf{G} *= 0.95\boldsymbol{G} + 0.05\boldsymbol{A_{22}} \quad [20]$$

Finally, flat priors were also considered for variance-covariances components.

The analysis of each model was performed using Gibbs sampling algorithm (Gelfand *et al*., 1990) with a single chain of 200,000 iterations, a 'burn-in' period of 20000 iterations and a thin interval of 20 samples with THRGIBBS1F90 software (Misztal & Tsuruta, 2018). Posterior marginal inferences were performed with POSTGIBBSF90 software (Misztal & Tsuruta, 2018). The inference was focused on mean, standard deviation and Highest Posterior Density Intervals at 95% of the marginal posterior distributions of the parameters of interest. Convergence of posterior distributions were assessed with graphical analysis by time series plots and histograms using the package boa (Smith, 2007). The models and data used for Bayesian BLUP and ssGBLUP genetic evaluations are shown in Table 6.

Table 6. Models, data sets, traits and number of records analyzed for Bayesian BLUP (Best Linear Unbiased Prediction) and ssGBLUP (single step Genomic Best Linear Unbiased Prediction) genetic evaluations.

Model	Data set	Traits	Number of records
Multi trait	EE[1]	TB[4] NBA[5] SB[6]	7132
	RR[2]	TB[4] NBA[5] SB[6]	6843
Multi breed	EE[1], RR[2], ER-RE[3]	TB_{EE} TB_{RR} $TB_{ER\text{-}RE}$	20152
		NBA_{EE} NBA_{RR} $NBA_{ER\text{-}RE}$	20152
		SB_{EE} SB_{RR} $SB_{ER\text{-}RE}$	20152

[1]EE: Entrepelado line; [2]RR: Retinto line; [3]ER-RE: crossbred animals; [4]TB: total born; [5]NBA: number of born alive; [6]SB: stillborn.

2.5. Cross validation.

In practice, Cross validation (CV) can be performed in a sample of individuals (validation animals) that are related to those individuals in the training set (training animals) but without being in it (Saatchi *et al*., 2011). In our study the K-fold cross validation (CV) with K equal to five was used in the purebred genotyped animals to compare the prediction ability of the fitted models (Figure 2). For all the scenarios, the five-fold CV was repeated 10 times. Owing to computational time, estimations for BVs were computed based on restricted maximum likelihood (REML) procedure with BLUPF90 software (Misztal & Tsuruta, 2018). Two different criteria were computed on the validation data set for each CV to compare the predictive ability of the models:

- The accuracy of the prediction measured as the Pearson's correlation between true BV and (G)EBV from de validation population divided by the heritability root of the trait estimated

by the classical multi trait animal model (Mehrban *et al*., 2019). This standardization of the Pearson's correlation allows comparison between traits (Saatchi *et al*., 2011):

$$Accuracy = \frac{cor(BV, (G)EBV)}{\sqrt{h^2}} \quad [21]$$

- The root mean squared error (RMSE) between the true BV and (G)EBV from de validation population:

$$RMSE = \sqrt{\sum \frac{(BV, (G)EBV)^2}{N}} \quad [22]$$

Accuracy and RMSE were estimated as the mean of accuracies and RMSE for five-fold cross validation procedures. Due to the true BV is unknown, the corrected phenotype ($\boldsymbol{y_c}$) was used instead (Mehrban *et al*., 2017). Phenotypic records were corrected for fixed effects and permanent effects in all evaluation models. Therefore, the corrected phenotype was defined as:

$$\boldsymbol{y_c = 1\mu + Za + e} \quad [23]$$

where $\boldsymbol{y_c}$, is the corrected phenotype; $\mathbf{1}$, is the vector of ones; $\boldsymbol{\mu}$ is the estimated general mean for each trait; $\boldsymbol{a}$ is the vector of estimated breeding values and $\boldsymbol{e}$ is the vector of estimated associated error; $\boldsymbol{Z}$ was the incidence matrix associated to the vector $\boldsymbol{a}$. $\boldsymbol{y_c}$ was computed with PREDICTF90 software (Misztal & Tsuruta, 2018) based on previous prediction and estimation of fixed and random effects with BLUPF90 software (Misztal & Tsuruta, 2018).

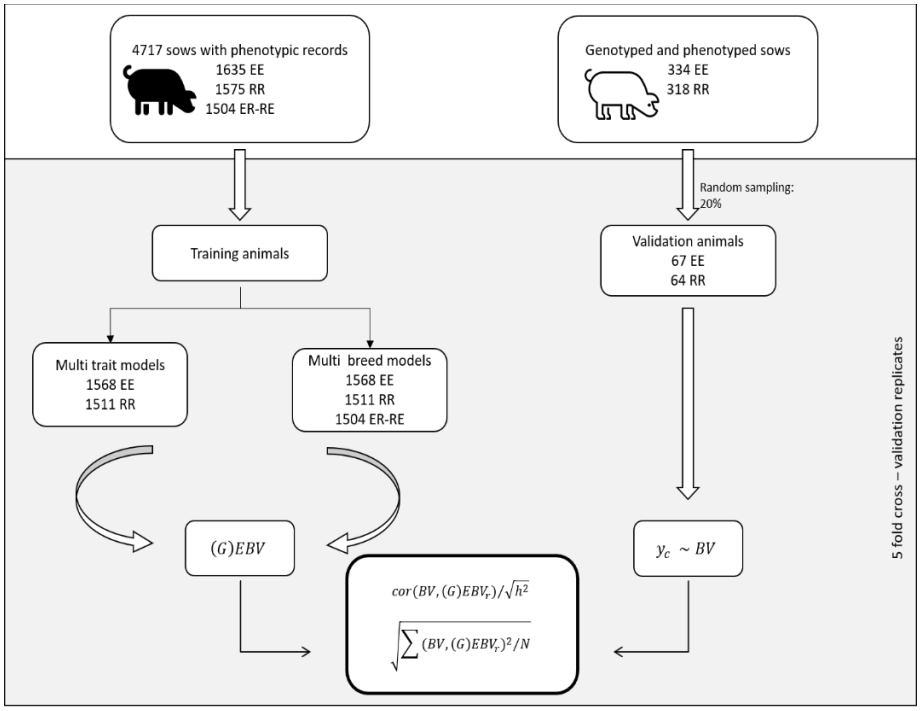

Figure 2. Schematic representation of the cross-validation pipeline to assess predictive ability. EE: Entrepelado line; RR: Retinto line; ER-RE: crossbred animals; (G)EBV: (genomic) estimated breeding value; $(G)EBV_r$: (genomic) estimated breeding value on training set for validation animals; BV: breeding value; y_c: corrected phenotype of validation animals according BLUP (Best Linear Unbiased Prediction) or ssGBLUP (single step Genomic Best Linear Unbiased Prediction) multi trait/multi breed models.

3. Results and discussion.

After the exploratory analysis, 316 phenotypic records (101 for EE, 111 for RR, 70 for ER and 34 for RE) were removed from the data, which represents 1.5% of the total amount of records. The distributions for each trait were not substantially modified after quality control (see Figure 10 in supplementary material). As expected, TB and NBA followed a nearly normal distribution whereas SB followed a Zero-Truncated Poisson distribution (Figure 3).

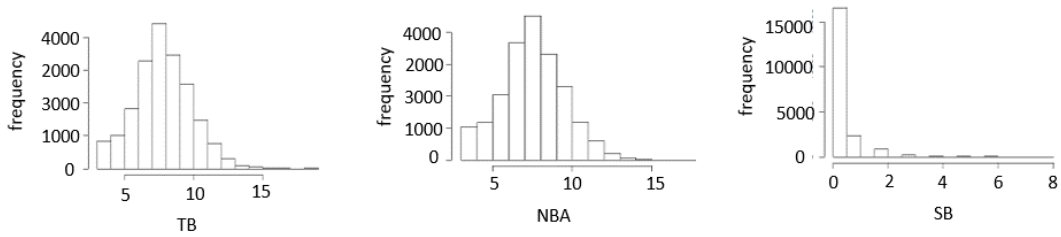

Figure 3. Distribution of litter size records for all animals after quality control. TB: total born; NBA: number of born alive; SB: stillborn.

The number of animals with records, those with genotypes (in brackets) and the basic statistics for LS traits are shown in Table 7. The mean values for all traits in both purebred lines EE and RR were similar and comparable with the values in other Iberian pig populations (Fernández *et al*., 2008; García-Casco *et al*., 2012). The mean values for TB and NBA of CB were around 0.3 piglets greater than those for PB. This advantage of CB over PB can be mainly due to the heterosis as was found by Noguera *et al*. (2019). As expected, the TB and NBA for the Iberian pig lines were clearly lower than those reported for maternal white pigs lines ranging from 11.46 to 15.91 and 10.40 to 15.4, respectively (Arango *et al*., 2005; Freyer, 2018; Lopes *et al*., 2017; Song *et al*., 2017; Zhang *et al*., 2016). The mean values for SB were also into the range than those reported for maternal white pigs lines (Large White) varying from 0.31 to 1.19 (Arango *et al*., 2005; Chu, 2005; Ye *et al*., 2018).

Table 7. Number of animals with records, genotyped animals (in brackets) and descriptive statistics of total born (TB), number of born alive (NBA) and stillborn (SB) for pure lines and crossbred Iberian pigs.

Traits	Records (genotyped animals)	Mean	SD[1]	Min[2]	Max[3]	CV[4]
All lines and crosses						
TB	4714 (1017)	8.25	2.10	3.00	19.00	0.25
NBA	4714 (1017)	7.98	2.04	3.00	18.00	0.26
SB	4714 (1017)	0.27	0.69	0.00	8.00	2.59
Entrepelado						
TB	1635 (333)	7.92	1.95	3.00	19.00	0.25
NBA	1635 (333)	7.67	1.89	3.00	18.00	0.25
SB	1635 (333)	0.24	0.67	0.00	7.00	2.73
Retinto						
TB	1.575 (318)	8.27	2.07	3.00	17.00	0.25
NBA	1575 (318)	7.97	1.99	3.00	15.00	0.25
SB	1575 (318)	0.30	0.74	0.00	8.00	2.44
Crosses						
TB	1504 (366)	8.58	2.23	3.00	19.00	0.25
NBA	1504 (366)	8.34	2.18	3.00	15.00	0.26
SB	1504 (366)	0.24	0.65	0.00	7.00	2.70

[1]SD: standard deviation; [2]Min: minimum; [3]Max: maximum; [4]CV: coefficient of variation.

Population stratification was assessed by a principal component analysis using the genotypic information of 34496 SNPs. The two purebred lines were clearly separated based on the PCA results, while the reciprocal crosses occupied intermediate positions between them, following an expected pattern (Figure 4). Although the Iberian pig has barely been selected, demo-graphic fluctuations and scarce genetic flow among herds led to the development of these lines with phenotypic and genotypic differences (Fabuel *et al*., 2004; García-Casco *et al*., 2012). This result would also indicate that Illumina GGP Porcine HD Array 70K allows proper differentiation among genotypes in Iberian pig lines.

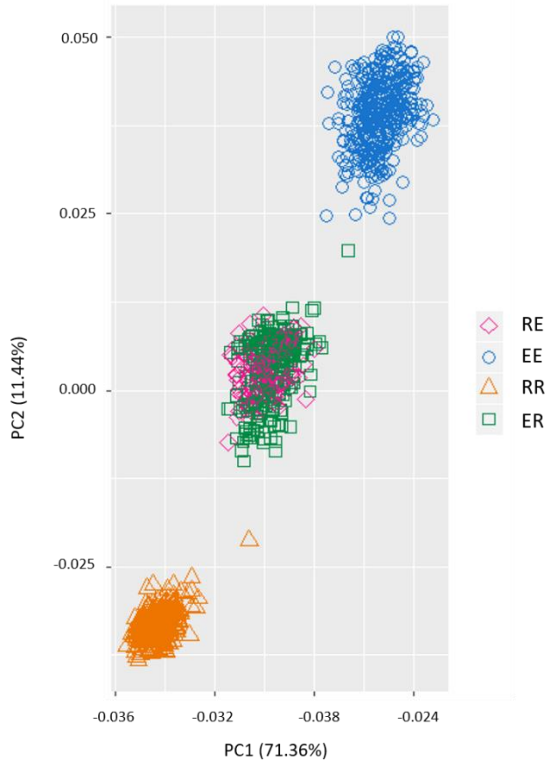

Figure 4. Population structure identified by principal component analysis of the genotyped animals. PC1: first principal component (percentage of variance explained); PC2: second principal component (percentage of variance explained); EE: Entrepelado line; RR: Retinto line; ER: Entrepelado x Retinto cross; RE: Retinto x Entrepelado cross.

Basic statistics for $\boldsymbol{G}$ and $\boldsymbol{A_{22}}$ elements were analyzed (**Table 8** and **Figure 5**) before implementing ssGBLUP. Incompatibilities between $\boldsymbol{G}$ and $\boldsymbol{A_{22}}$ could lead poor convergence rate or/and large reranking (Simeone *et al*., 2011; Misztal *et al*., 2013, Petrini *et al*., 2016). This can be due to problems with the pedigree (i.e. short or incomplete pedigrees, pedigree mistakes and heterogeneous base populations) or genotypes (i.e. incorrect assignment of genotype and poor quality of genotypes). In our study, the maximum difference between $\boldsymbol{G}$ and $\boldsymbol{A_{22}}$ for the diagonal and off-diagonal elements was very small (-0.017) with a SD > 0,04 (**Table 8**). Besides, the distribution of diagonal and off-diagonal elements of $\boldsymbol{G}$ approached to a normal distribution with a single peak (**Figure 5**a) and the correlation between $\boldsymbol{G}$ and $\boldsymbol{A_{22}}$ showed a positive pattern (**Figure 5**b and **Figure 5**c). The same pattern was found when the genotypes were analyzed for each breed separately (see **Figure 13** and Figure 14 in supplementary material). All these results revealed no mismatches between $\boldsymbol{G}$ and $\boldsymbol{A_{22}}$ which indicated absence of problems with either genotypes or pedigree (Simeone *et al*., 2011).

Table 8. Descriptive statistics of diagonal and off-diagonal elements of Genomic matrix (G) and additive genetic relationship matrix for genotyped animals only (A_{22}) for Entrepelado and Retinto lines.

	G			A_{22}			
Diagonal Elements	N[1]	Mean	SD[2]	N[1]	Mean	SD[2]	Difference
All lines and crosses	1114	1.01	0.10	1114	1.01	0.03	0.00
Entrepelado	382	1.02	0.05	382	1.02	0.03	0.00
Retinto	357	1.01	0.05	357	1.01	0.03	0.00
Off-diagonal elements	N[1]	Mean	SD[2]	N[1]	Mean	SD[2]	Difference
All lines and crosses	1,239,882	0.034	0.12	1,239,882	0.039	0.05	-0.005
Entrepelado	145,542	0.083	0.08	145,542	0.100	0.08	-0.017
Retinto	127,092	0.048	0.07	127,092	0.052	0.06	-0.004

[1]N: number of genotyped animals; [2]SD: standard deviation.

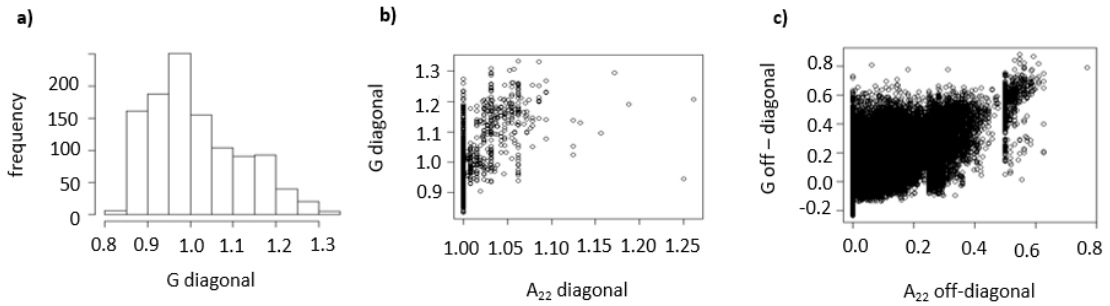

Figure 5. Histograms and plots for diagonal and off-diagonal elements in G (genomic relationship matrix) and A_{22} (numerator relationship matrix for genotyped animals) of all lines and crosses. a) Frequency distribution for G diagonal elements; b) A_{22} diagonal elements and G diagonal elements plot; c) A_{22} off-diagonal elements and G off-diagonal elements plot.

Table 9 shows the heritability and additive genetic correlations of TB, NBA and SB for EE and RR lines under the BLUP and ssGBLUP Bayesian multi trait models. Both methods GBLUP and ssGBLUP provided similar estimates for all traits and lines. The estimates of the heritability by using pedigree based (BLUP) and combined relationship matrices (ssGBLUP) were similar. It can probably be due to the resemblance between these matrices as argued by Misztal *et al*. (2013). The posterior mean of the heritability of EE line by both models BLUP and ssGBLUP for TB (0.09-0.09) and NBA (0.010-0.09) were clearly superior than those of RR line for TB (0.06-0.06) and NBA (0.07-0.06). This scenario was expected based on the results of the estimated variance components (see Table 11 and Table 12 in supplementary material). EE line showed higher values for additive genetic variance component for the two methods and traits than RR line. For SB, no differences were found among lines and models. Moreover, previous analysis with SB as a categorical value showed similar results. The heritability estimates for LS obtained

in this study are in the range of those reported by García-Casco *et al*. (2012) and Fernández *et al*. (2008) for TB and NBA and by Corral *et al*. (2010) for SB in other Iberian pig lines. As expected, the additive genetic correlation between TB and NBA was high for all methods and lines (0.98±0.01 and 0.98±0.01 in EE and RR, respectively). Similar values of additive genetic correlation between both traits were reported by Radojkovic *et al*. (2012) in Swedish Landrace sows (0.938), by Zhang *et al*. (2016) in Landrace sows (0.93±0.04) and by García-Casco *et al*. (2012) in Iberian pigs (0.96±0.02). The posterior mean of additive genetic correlations between SB and TB were positive and moderate low for EE (0.16-0.18) and RR (0.21-0.28) lines, and between SB and NBA positive but close to zero (0.02-0.03 for EE and 0.07-0.11 for RR). It would indicate that although the correlation was low the genetic selection for TB could increase SB. However, the posterior standard deviations were higher and the highest posterior density intervals at 95% clearly included the zero value which made difficult to evaluate the effect of genetic selection for TB or NBA. Several studies in white pigs populations have also reported moderate-low or negative mean values for genetic correlations between SB and TB (Radojkovic *et al*., 2012; Zhang *et al*., 2016) and negative mean values for genetic correlations between SB and NBA (Arango *et al*., 2005; Chu, 2005; Radojkovic *et al*., 2012; Zhang *et al*., 2016).

Table 9. Main features of the posterior distribution of the heritability (h^2) and additive genetic correlation ($r_{a1,a2}$) of total born (TB), number of born alive (NBA) and stillborn (SB) estimated for Entrepelado (EE) and Retinto (RR) lines with BLUP (Best Linear Unbiased Prediction) and ssGBLUP (Single Step Best Linear Unbiased Prediction) Bayesian multi trait models.

		BLUP			ssGBLUP		
		Mean[1]	SD[2]	$HPD_{95\%}$[3]	Mean[1]	SD[2]	$HPD_{95\%}$[3]
Entrepelado							
h^2	TB_{EE}	0.09	0.02	0.05 – 0.14	0.09	0.02	0.04 – 0.13
	NBA_{EE}	0.10	0.02	0.05 – 0.15	0.09	0.02	0.04 – 0.13
	SB_{EE}	0.02	0.01	0.01 – 0.04	0.02	0.01	0.01 – 0.03
$r_{a1,a2}$	TB_{EE}, NBA_{EE}	0.98	0.01	0.97 – 0.99	0.98	0.01	0.97 – 0.99
	TB_{EE}, SB_{EE}	0.18	0.22	-0.41 – 0.51	0.16	0.23	-0.26 – 0.63
	NBA_{EE}, SB_{EE}	0.03	0,22	-0.39 – 0.46	0.02	0.24	-0.41 – 0.51
Retinto							
h^2	TB_{RR}	0.06	0.02	0.03 – 0.10	0.06	0.02	0.03 – 0.10
	NBA_{RR}	0.07	0.02	0.03 – 0.10	0.06	0.02	0.02 – 0.09
	SB_{RR}	0.02	0.01	0.01 – 0.03	0.02	0.01	0.01 – 0.04
$r_{a1,a2}$	TB_{RR}, NBA_{RR}	0.98	0.01	0.96 – 0.99	0.98	0.01	0.95 – 0.99
	TB_{RR}, SB_{RR}	0.21	0.23	-0.21 – 0.65	0.28	0.24	-0.17 – 0.71
	NBA_{RR}, SB_{RR}	0.07	0.25	-0.40 – 0.53	0.11	0.26	-0.37 – 0.60

[1]Mean: posterior mean [2]SD: posterior standard deviation; [3]$HDP_{95\%}$: highest posterior density interval at 95%.

Heritability and additive genetic correlations under BLUP and ssGBLUP Bayesian multi breed models for TB, NBA and SB are shown in Table 10. The posterior mean of the heritability

for all traits estimated with the multi breed models were similar to those with the multi trait analysis (see Table 9) in accordance with Lutaaya *et al*. (2001). In a CCPS context phenotypic data recorded in CB relatives are used for selection of PB (Wei & Van Der Werf, 1994). Thus, the selection is indirect and its accuracy depends on genetic correlation between PB and CB performance (r_{pc}) (Wientjes & Calus, 2017). This r_{pc} value is key in the success of CCPS and could reflect differences in genetic backgrounds (differences in allele frequencies and the non-additive genetic effects between populations), environments or trait measurements. This study showed that r_{pc} values for TB and NBA in Iberian pig populations were high (around 0.80) which means that the non-additive effects and the differences in allele frequencies between PB and CB lines were low (Duenk *et al*., 2020). It would suggest that the PB's BV can be a good predictor of CB performance for TB and NBA and the benefits of using CB data would relay in the increase of data (Wei & Van Der Werf, 1994; Wientjes & Calus 2017). Similar results were obtained in recent studies of TB by Hidalgo *et al*. (2015) and by Lopes *et al*. (2017) for Large White purebred and Large White x Landrace populations (0.88±0.04 and 0.91±0.04). Conversely, the r_{pc} for SB between RR and the reciprocal crosses was 0.41 (BLUP) and 0.61 (ssGBLUP) and for EE and the crosses was -0.68 (BLUP) and -0.47 (ssGBLUP). It would suggest high non-additives effects as well as high differences between line allele frequencies. Note that the environment between the farms was similar as well as the same method across farms was used to measure the SB trait. These results (r_{pc}< 0.8) would support the implementation of CCPS for SB since it represents an advantage of combined PB and CB selection over pure line selection (Wei & Van Der Werf, 1994; Dekkers, 2007; Wientjes & Calus 2017). The r_{pc} lower than 0.8 have been reported in traits with very low heritability such as fertility which is related with the epistatic interactions (Wientjes & Calus 2017). However, not negative r_{pc} correlations have been previously reported. Hence, further studies are needed to confirm this estimation.

Table 10. Main features of the posterior distribution of the heritability (h^2) and additive genetic correlation ($r_{a1,a2}$) of total born (TB), number of born alive (NBA) and stillborn (SB) estimated for Entrepelado (EE) and Retinto (RR) lines and crossbred animals (ER-RE) with BLUP (Best Linear Unbiased Prediction) and ssGBLUP (Single Step Genomic Best Linear Unbiased Prediction) Bayesian multi breed models.

		BLUP			ssGBLUP		
		Mean[1]	SD[2]	$HPD_{95\%}$[3]	Mean[1]	SD[2]	$HPD_{95\%}$[3]
h^2	TB_{EE}	0.09	0.01	0.06 – 0.12	0.08	0.01	0.05 -0.10
	TB_{RR}	0.05	0.01	0.02 – 0.07	0.05	0.01	0.03 – 0.08
	$TB_{ER\text{-}RE}$	0.1	0.01	0.06 – 0.12	0.09	0.01	0.06 – 0.11
$r_{a1,a2}$	TB_{EE}, TB_{RR}	0.68	0.23	0.24 – 0.95	0.32	0.20	0.05 – 0.75
	TB_{EE}, $TB_{ER\text{-}RE}$	0.91	0.06	0.78 – 0.99	0.70	0.08	0.54 – 0.82
	TB_{RR}, $TB_{ER\text{-}RE}$	0.90	0.06	0.77 – 0.98	0.80	0.10	0.62 – 0.94
h^2	NBA_{EE}	0.09	0.01	0.06 – 0.12	0.08	0.02	0.05 – 0.12
	NBA_{RR}	0.05	0.01	0.02 – 0.08	0.05	0.01	0.03 – 0.07
	$NBA_{ER\text{-}RE}$	0.09	0.01	0.06 – 0.12	0.08	0.01	0.05 – 0.11
$r_{a1,a2}$	NBA_{EE}, NBA_{RR}	0.78	0.18	0.39 – 0.97	0.39	0.20	0.06 – 0.78
	NBA_{EE}, $NBA_{ER\text{-}RE}$	0.94	0.04	0.86 – 0.99	0.69	0.10	0.28 – 0.83
	NBA_{RR}, $NBA_{ER\text{-}RE}$	0.92	0.06	0.78 – 0.98	0.83	0.08	0.67 – 0.95
h^2	SB_{EE}	0.02	0.007	0.008 – 0.035	0.02	0.004	0.009 – 0.025
	SB_{RR}	0.02	0.007	0.011 – 0.034	0.02	0.007	0.012 – 0.038
	$SB_{ER\text{-}RE}$	0.02	0.005	0.012 – 0.033	0.03	0.007	0.017 – 0.047
$r_{a1,a2}$	SB_{EE}, SB_{RR}	-0.78	0.11	-0.97 – -0.58	-0.68	0.14	-0.91 – -0.38
	SB_{EE}, $SB_{ER\text{-}RE}$	-0.68	0.18	-0.91 – -0.31	-0.47	0.13	-0.78 – -0.27
	SB_{RR}, $SB_{ER\text{-}RE}$	0.41	0.23	-0.019 – 0.75	0.61	0.10	0.40 – 0.83

[1]Mean: posterior mean [2]SD: posterior standard deviation; [3]$HDP_{95\%}$: highest posterior density interval at 95%.

Results of the CV for predictive ability measured as the (standardized) Pearson's correlation between corrected phenotype and (G)EBV and as the RMSE for each model and method are shown on Figure 6 and Figure 7 and in Figure 8 and Figure 9, respectively. Results in terms of accuracy measured as Pearson's correlations are in agreement with those found in other studies in pigs (Hidalgo *et al*., 2015; Song *et al*., 2017; Lopez *et al*., 2017). If we compare between traits, the standardized Pearson's correlations showed higher predictive ability (accuracy) for SB (from 0.49 to 0.71) followed by TB (from 0.41 to 0.58) and NBA (from 0.38 to 0.58) (Figure 6 and Figure 7). The Pearson's correlation estimates had a small standard error (0.01-0.02) which allowed the comparison between the means. Regarding lines, a pattern was found among models for EE. In all cases, SB shows the highest accuracy value followed by TB and NBA, which showed a small difference between them (0.06 as maximum). In this breed, multi trait models and especially ssGBLUP exhibited a better prediction ability for the three traits. However, due to the small differences in accuracy between models is not clear that this model leads an important reranking of the candidates of selection. This study also shows that including CB data in the genetic evaluations had different effects on the predictive ability according PB line. For EE line, the inclusion of CB data did not show any advantage whereas for

RR line Pearson's correlations were higher for all traits by both BLUP and ssGBLUP multi breed models (Figure 7). For this breed, the three traits showed the same predictive ability by using BLUP Multi breed models (0.58). The inclusion of genomic data allowed to get the highest predictive ability for SB (0.64) whereas TB and NBA remain a bit lower (0.54). Multi trait models followed the same pattern as was described for EE but with lower values. The advantage of using CB data for RR line compared to EE was consistent with r_{pc} values (see Table 10). The RR-ER r_{pc} values were higher than the EE-ER r_{pc} values for all traits but NBA. In summary, CV results based on the accuracy of GEBV shows that multi trait models and specially ssGBLUP are more appropriate for EE whereas multi breed models are for RR.

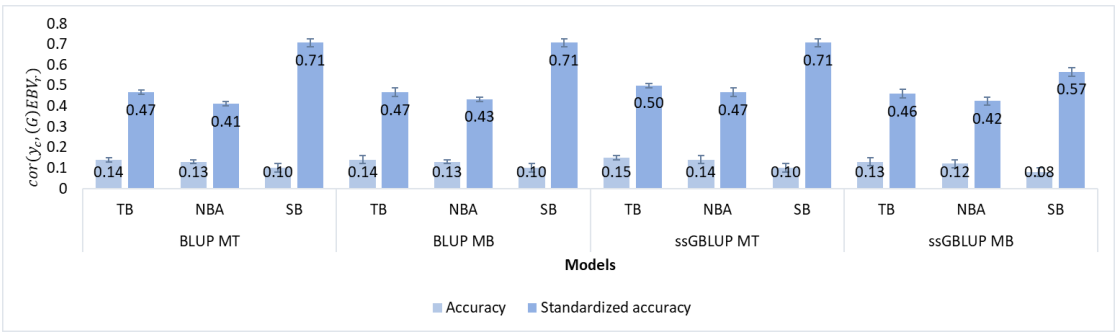

Figure 6. Mean of the accuracy and standardized accuracy for the 10 replicates of the 5-fold cross validation of Entrepelado line for each model and trait by BLUP and ssGBLUP methods. Best Linear Unbiased Prediction; ssGBLUP: single step Genomic Best Linear Unbiased Prediction; MT: multi trait model; MB: multi breed model; TB: total piglets; NBA: number of born alive; SB: stillborn.

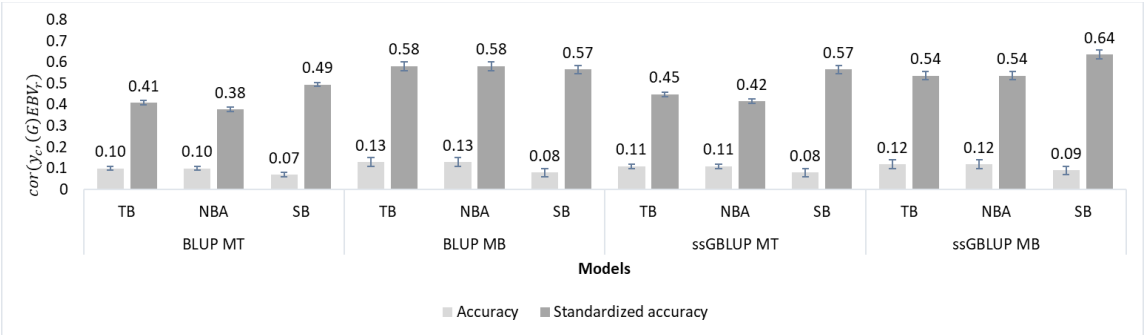

Figure 7. Mean of the accuracy and standardized accuracy for the 10 replicates of the 5-fold cross validation of Retinto line for each model and trait by BLUP and ssGBLUP methods. Best Linear Unbiased Prediction; ssGBLUP: single step Genomic Best Linear Unbiased Prediction; MT: multi trait model; MB: multi breed model; TB: total piglets; NBA: number of born alive; SB: stillborn.

In terms of RMSE, there were no great differences between PB lines for LS traits when BLUP and ssGBLUP multi trait analysis were performed, but in general a little advantage of models including genomic information was found (Figure 8 and Figure 9). RMSE for EE were

lower than for RR. For EE, RMSE values range from 0.58 to 1.55 (Figure 8), whereas for RR values ranged from 0.61 to 1.87 (Figure 9). In line with the accuracy results, these values support the idea that multi trait models are more suitable for EE line. Surprisingly, multi breed models showed higher RMSE values in comparison with multi trait models for RR and specially for TB in ssGBLUP. The inclusion of genomic data may allow a slight reduction in bias in the prediction of (G)EBV except for NBA in EE (Figure 8) and for SB in RR (Figure 9) under MB models. The highest difference of RMSE between BLUP and ssGBLUP methods was found for TB when CB data was included in the RR genetic evaluation (1.87 and 1.81, respectively). This outperformance of ssGBLUP over BLUP could be mainly a result of a better estimation of relationships among individuals by markers (Tusell *et al*., 2013; Hidalgo *et al*., 2015).

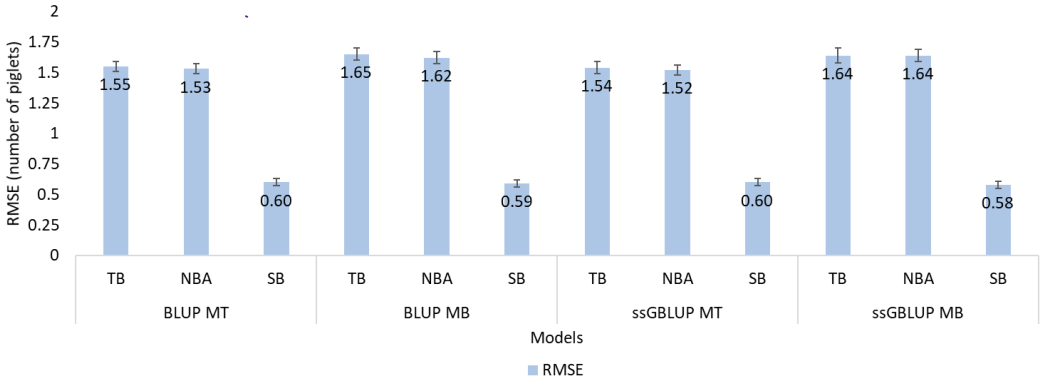

Figure 8. Root mean squared error (RMSE) between the corrected phenotypes and the estimated breeding values of the 10 replicates of the 5-fold cross validation sets of Entrepelado for each model by Best Linear Unbiased Prediction (BLUP) and single step Genomic Best Linear Unbiased Prediction (ssGBLUP) methods. TB: total born; NBA: number of born alive; SB: stillborn; MT: multi trait model; MB: multi breed model.

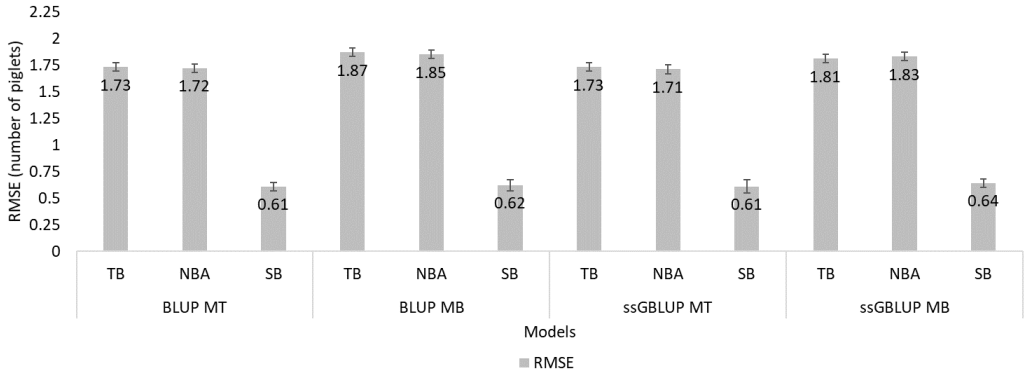

Figure 9. Root mean squared error (RMSE) between the corrected phenotypes and the estimated breeding values of the 10 replicates of the 5-fold cross validation sets of Retinto line for each model by Best Linear Unbiased Prediction (BLUP) and single step Genomic Best Linear Unbiased Prediction ssGBLUP) methods. TB: total born; NBA: number of born alive; SB: stillborn; MT: multi trait model; MB: multi breed model.

4. Conclusions.

- The heritability estimates of TB and NBA for EE were higher than for RR in both multi trait and multi breed models by BLUP and ssGBLUP methods. No differences were found for SB.

- Genetic correlations between CB and PB (r_{pc} values) were higher than 0.8 for TB and NBA and lower than 0.7 and even negative for SB, indicating not relevant differences in genetic backgrounds as well as low GxE for TB and NBA whereas these seems to be relevant for SB.

- Multi trait models are more suitable for EE line in terms of predictive ability of LS traits whereas multi breed models are for RR line. Therefore, within line selection would represent the best choice for EE line whilst accounting on CB data may be useful for RR line.

- The inclusion of genomic data increased the accuracy on (G)EBVs, specially for TB in EE and for SB in RR line and also reduced the (G)EBVs bias except for NBA in EE and SB in RR under MB models. This support the positive impact of including genomic information in the genetic evaluations.

- Although the models have provided differences in the predictive ability of BV for LS traits, these seems to be not enough for relevant changes on the response to selection.

5. Cited bibliography.

Aguilar, I., Misztal, I., Johnson, D. L., Legarra, A., Tsuruta, S., & Lawlor, T. J. (2010). Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *Journal of Dairy Science*, *93*(2), 743–752. https://doi.org/10.3168/jds.2009-2730

Arango, J., Misztal, I., Tsuruta, S., Culbertson, M., & Herring, W. (2005). Threshold-linear estimation of genetic parameters for farrowing mortality, litter size, and test performance of Large White sows. *Journal of Animal Science*, *83*(3), 499–506. https://doi.org/10.2527/2005.833499x

Asociación Española de Criadores de Cerdo Ibérico. (2011). *Conoce la Asociación Española de Criadores de Ganado Porcino Selecto Ibérico Puro y Tronco Ibérico.* http://www.aeceriber.es/aeceriber-asociacion-espanola-cerdo-iberico.html

Barea, R., Nieto, R., Vitari, F., Domeneghini, C., & Aguilera, J. F. (2011). Effects of pig genotype (Iberian v. Landrace Large White) on nutrient digestibility, relative organ weight and small intestine structure at two stages of growth. *Animal*, *5*(4), 547–557. https://doi.org/10.1017/S1751731110002181

Blasco, A. (2017). *Bayesian data analysis for animal scientists: The basics*. Springer. https://doi.org/10.1007/978-3-319-54274-4

Charlesworth, D., & Willis, J. H. (2009). The genetics of inbreeding depression. *Nature Reviews Genetics*, *10*(11), 783–796. https://doi.org/10.1038/nrg2664

Christensen, O. F., & Lund, M. S. (2010). Genomic prediction when some animals are not genotyped. *Genetics Selection Evolution*, *42*, 1–8. http://www.doaj.org/doaj?func=abstract&id=515339

Christensen, O. F., Madsen, P., Nielsen, B., & Su, G. (2014). Genomic evaluation of both purebred and crossbred performances. *Genetics Selection Evolution*, *46*(1), 1–9. https://doi.org/10.1186/1297-9686-46-23

Chu, M. X. (2005). Statistical analysis of stillbirths in different genotypes of sows. *Asian-Australasian Journal of Animal Sciences*, *18*(10), 1475–1478.

https://doi.org/10.5713/ajas.2005.1475

Corral, J. M., Izquierdo, M., Ayuso, D., Carolino, N., Gama, L. T., & Padilla, J. A. (2010). Genetic parameters for direct and maternal effects on growth traits of sheep. *Proceedings of the World Congress on Genetics Applied to Livestock Production*, 603–611. https://doi.org/10.5194/aab-55-603-2012

De Los Campos, G., Naya, H., Gianola, D., Crossa, J., Legarra, A., Manfredi, E., Weigel, K., & Cotes, J. M. (2009). Predicting quantitative traits with regression models for dense molecular markers and pedigree. *Genetics*, *182*(1), 375–385. https://doi.org/10.1534/genetics.109.101501

Dekkers, J. (2007). Marker-assisted selection for commercial crossbred performance. *Journal of Animal Science*, *85*(9), 2104–2114. https://doi.org/10.2527/jas.2006-683

Duenk, P., Bijma, P., Calus, M. P. L., Wientjes, Y., & Van Der Werf, J. H. J. (2020). The impact of non-additive effects on the genetic correlation between populations. *G3: Genes, Genomes, Genetics*, *10*(2), 783–795. https://doi.org/10.1534/g3.119.400663

Duenk, P., Calus, M. P. L., Wientjes, Y., Breen, V., Henshall, J., Hawken, R., & Bijma, P. (2019a). Estimating the purebred-crossbred genetic correlation of body weight in broiler chickens with pedigree or genomic relationships. *Genetics Selection Evolution*, *51*(1), 1–11. https://doi.org/10.1186/s12711-019-0447-9

Duenk, P., Calus, M. P. L., Wientjes, Y., Breen, V., Henshall, J., Hawken, R., & Bijma, P. (2019b). Validation of genomic predictions for body weight in broilers using crossbred information and considering breed-of-origin of alleles. *Genetics Selection Evolution*, *51*(1), 1–12. https://doi.org/10.1186/s12711-019-0481-7

Erbe, M., Hayes, B., Matukumalli, L. K., Goswami, S., Bowman, P. J., Reich, C. M., Mason, B. A., & Goddard, M. (2012). Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. *Journal of Dairy Science*, *95*(7), 4114–4129. https://doi.org/10.3168/jds.2011-5019

Esfandyari, H., Berg, P., & Sørensen, A. C. (2018). Balanced selection on purebred and crossbred performance increases gain in crossbreds. *Genetics Selection Evolution*, *50*(1), 1–14. https://doi.org/10.1186/s12711-018-0379-9

Esfandyari, H., Bijma, P., Henryon, M., Christensen, O. F., & Sørensen, A. C. (2016). Genomic prediction of crossbred performance based on purebred Landrace and Yorkshire data using a dominance model. *Genetics Selection Evolution*, *48*(1), 1–9. https://doi.org/10.1186/s12711-016-0220-2

Esfandyari, H., Sørensen, A. C., & Bijma, P. (2015). A crossbred reference population can improve the response to genomic selection for crossbred performance. *Genetics Selection Evolution*, *47*(1), 1–12. https://doi.org/10.1186/s12711-015-0155-z

Fabuel, E., Barragán, C., Silió, L., Rodríguez, M. C., & Toro, M. A. (2004). Analysis of genetic diversity and conservation priorities in Iberian pigs based on microsatellite markers. *Heredity*, *93*(1), 104–113. https://doi.org/10.1038/sj.hdy.6800488

Falconer, D. S. (1976). *Introducción a la genética cuantitativa.* (1ª). Compañía Editorial Continental S.A.

Falconer, D.S., & Mackay, T.F.C. (1996). *Introduction to quantitative genetics.* 4th ed. Longmans Green, Harlow, Essex, UK.

Fernández, A., Rodrigáñez, J., Zuzúarregui, J., Rodríguez, M. C., & Silió, L. (2008). Genetic parameters for litter size and weight at different parities in Iberian pigs. *Spanish Journal of Agricultural Research*, *6*(SPEC. ISS.), 98–106. https://doi.org/10.5424/sjar/200806s1-378

Freyer, G. (2018). Maximum number of total born piglets in a parity and individual ranges in litter size expressed as specific characteristics of sows. *Journal of Animal Science and Technology*, *60*(1), 1–7. https://doi.org/10.1186/s40781-018-0172-x

García-Casco, J. M., Fernández, A., Rodríguez, M. C., & Silió, L. (2012). Heterosis for litter size and growth in crosses of four strains of Iberian pig. *Livestock Science*, *147*(1–3), 1–8. https://doi.org/10.1016/j.livsci.2012.03.005

Gelfand, A. E., Hills, S. E., Racine-Poon, A., & Smith, A. F. M. (1990). Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association*, *85*(412), 972–985. https://doi.org/10.1080/01621459.1990.10474968

Goddard, M. (2009). Genomic selection: Prediction of accuracy and maximisation of long term response. *Genetica*, *136*(2), 245–257. https://doi.org/10.1007/s10709-008-9308-0

Habier, D., Fernando, R. L., Kizilkaya, K., & Garrick, D. J. (2011). Extension of the bayesian alphabet for genomic selection. *BMC Bioinformatics*, *12*. https://doi.org/10.1186/1471-2105-12-186

Hayes, B., Bowman, P. J., Chamberlain, A. J., & Goddard, M. (2009). Invited review: Genomic selection in dairy cattle: Progress and challenges. *Journal of Dairy Science*, *92*(2), 433–443. https://doi.org/10.3168/jds.2008-1646

Hayes, B., & Goddard, M. (2010). Genome-wide association and genomic selection in animal breeding. *Genome*, *53*(11), 876–883. https://doi.org/10.1139/G10-076

Henderson, C. R. (1963). Selection index and expected genetic advance. *Statistical Genetics and Plant Breeding*, 141–163.

Henderson, C. R. (1975). Comparison of Alternative Sire Evaluation Methods. *Journal of Animal Science*, *41*(3), 760–770.

Hidalgo, A. M., Bastiaansen, J. W. M., Lopes, M. S., Harlizius, B., Groenen, M. A. M., & de Koning, D. J. (2015). Accuracy of predicted genomic breeding values in purebred and crossbred pigs. *G3: Genes, Genomes, Genetics*, *5*(8), 1575–1583. https://doi.org/10.1534/g3.115.018119

Ibáñez-Escriche, N., Fernando, R. L., Toosi, A., & Dekkers, J. (2009). Genomic selection of purebreds for crossbred performance. *Genetics Selection Evolution*, *41*(12), 1–10. https://doi.org/10.1186/1297-9686-41-12

Ibáñez-Escriche, N., Magallón, E., Gonzalez, E., Tejeda, J. F., & Noguera, J. L. (2016). Genetic parameters and crossbreeding effects of fat deposition and fatty acid profiles in Iberian pig lines. *Journal of Animal Science*, *94*(1), 28–37. https://doi.org/10.2527/jas.2015-9433

Ibáñez-Escriche, N., Varona, L., Magallón, E., & Noguera, J. L. (2014). Crossbreeding effects on pig growth and carcass traits from two Iberian strains. *Animal*, *8*(10), 1569–1576. https://doi.org/10.1017/S1751731114001712

Legarra, A. (2015). *Bases for Genomic Prediction* (pp. 25–34). https://doi.org/10.1016/j.joca.2004.04.008

Legarra, A., Aguilar, I., & Misztal, I. (2009). A relationship matrix including full pedigree and

genomic information. *Journal of Dairy Science*, *92*(9), 4656–4663. https://doi.org/10.3168/jds.2009-2061

Legarra, A., Christensen, O. F., Aguilar, I., & Misztal, I. (2014). Single Step, a general approach for genomic selection. *Livestock Science*, *166*(1), 54–65. https://doi.org/10.1016/j.livsci.2014.04.029

Legarra, A., Lourenco, D., Tsuruta, S., Fragomeni, B., Chen, C. Y., Herring, W., Misztal, I., Legarra, A., Christensen, O. F., Vitezica, Z., Aguilar, I., & Misztal, I. (2015). Multi-population Genomic Relationships for. *Journal of Animal Science*, *200*(October), 503–515. https://doi.org/10.1534/genetics.117.300152/-/DC1.1

Lo, L. L., Fernando, R. L., Cantet, R. J. C., & Grossman, M. (1995). Theory for modelling means and covariances in a two-breed population with dominance inheritance. *Theoretical and Applied Genetics*, *90*(1), 49–62. https://doi.org/10.1007/BF00220995

Lopes, M. S., Bovenhuis, H., Hidalgo, A. M., Van Arendonk, J. A. M., Knol, E. F., & Bastiaansen, J. W. M. (2017). Genomic selection for crossbred performance accounting for breed-specific effects. *Genetics Selection Evolution*, *49*(1), 1–9. https://doi.org/10.1186/s12711-017-0328-z

Lopez-Bote, C. J. (1998). Sustained utilization of the Iberian pig breed. *Meat Science*, *49*(SUPPL. 1). https://doi.org/10.1016/S0309-1740(98)00072-2

Lutaaya, E., Misztal, I., Mabry, J. W., Short, T., Timm, H. H., & Holzbauer, R. (2001). Genetic parameter estimates from joint evaluation of purebreds and crossbreds in swine using the crossbred model. *Journal of Animal Science*, *79*(12), 3002–3007. https://doi.org/10.2527/2001.79123002x

Masuda, Y. (2019). *Introduction to BLUPF90 suite programs Standard Edition* (Issue September). University of Georgia.

Meuwissen, T., Hayes, B., & Goddard, M. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*(4), 1819–1829.

Meuwissen, T., Hayes, B., & Goddard, M. (2016). Genomic selection: A paradigm shift in animal breeding. *Animal Frontiers*, *6*(1), 6–14. https://doi.org/10.2527/af.2016-0002

BOE-A-2014-318, 1 (2014).

Ministerio de Agricultura Pesca y Alimentación. (2019). *Encuestas ganaderas 2019. Ganado porcino.* http://publicacionesoficiales.boe.es/NIPO:003191624

Ministerio de Medio Ambiente y Medio Rural y Marino. (2011). *Resolucion de Aprobación Programa de Mejora Raza Porcina Iberica* (pp. 1–2).

Ministerio de Medio Ambiente y Medio Rural y Marino, & Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria. (2011). *Programa de mejora de la raza porcina Ibérico* (Issue Julio).

Misztal, I., Aggrey, S. E., & Muir, W. M. (2013). Experiences with a single-step genome evaluation. *Poultry Science*, *92*(9), 2530–2534. https://doi.org/10.3382/ps.2012-02739

Misztal, I., & Tsuruta, S. (2018). *Manual for BLUPF90 family of programs - Ignacy Misztal et al*. *May*, 0–142.

Mrode, R. (2005). *Linear Models for the Prediction of Animal Bredding Values* (2nd ed.). CABI Publishing.

Muñoz, M., Sánchez-Esquilche, F. Caraballo, C., Gómez, F., Pariente, J. M., Silió, L., Rodríguez, C., & García-Casco, J. M. (2018). Animal breeding scheme applied to the quality of pure Iberian montanera pigs. *Archivos de Zootecnia*, *67*(260), 596–603. https://www.cabdirect.org/cabdirect/abstract/20183364209

Nieto, R., García-Casco, J. M., Lara, L., Palma-Granados, P., Izquierdo, M., Hernandez, F., Dieguez, E., Duarte, J. L., & Batorek-Lukač, N. (2019). Ibérico (Iberian) Pig. In *European Local Pig Breeds - Diversity and Performance. A Study of Project TREASURE* (pp. 1–25). https://doi.org/10.1016/j.colsurfa.2011.12.014

Noguera, J. L., & Ibáñez-Escriche, N. (2017). Mejora del cerdo Ibérico. *Suis*, *142*, 28–34.

Noguera, J. L., Ibáñez-Escriche, N., Casellas, J., Rosas, J. P., & Varona, L. (2019). Genetic parameters and direct, maternal and heterosis effects on litter size in a diallel cross among three commercial varieties of Iberian pig. *Animal*, 2765–2772. https://doi.org/10.1017/S1751731119001125

Noguera, J. L., Ibáñez-Escriche, N., Magallón, E., & Varona, L. (2018). Genetics parameters and

crossbreeding effects of prolificacy in three commercial lines of Iberian pigs. *Proceedings of the World Congress on Genetics Applied to Livestock Production.*, 238–241.

Omega Bio-tek. (2017). *E.Z.N.A.® Blood DNA Mini Kit*. *January*, D3392-01. https://www.omegabiotek.com/

Radojkovic, D., Petrovic, M., Radovic, C., Parunovic, N., Popovac, M., & Gogic, M. (2012). Heritability and correlation of litter traits in pigs determined by REML method. *Biotechnology in Animal HusbandryBiotehnologija u Stocarstvu*, *28*(4), 771–778. https://doi.org/10.2298/bah1204771r

Rodriguez, C., Rodrigañez, J., & Silió, L. (1994). Genetic analysis of maternal ability in Iberian pigs. *Journal of Animal Breeding and Genetics*, *111*(1–6), 220–227. https://doi.org/10.1111/j.1439-0388.1994.tb00461.x

Saatchi, M., McClure, M. C., McKay, S. D., Rolf, M. M., Kim, J., Decker, J. E., Taxis, T. M., Chapple, R. H., Ramey, H. R., Northcutt, S. L., Bauck, S., Woodward, B., Dekkers, J., Fernando, R. L., Schnabel, R. D., Garrick, D. J., & Taylor, J. F. (2011). Accuracies of genomic breeding values in American Angus beef cattle using K-means clustering for cross-validation. *Genetics Selection Evolution*, *43*(1), 1–16. https://doi.org/10.1186/1297-9686-43-40

Silió, L. (2000). Iberian pig breeding programme. In *ICAR Technical Series. Developing breeding strategies for lower input animal production environments.* ICAR.

Silió, L., Rodríguez, C., Rodrigáñez, J., & Toro, M.A. (2001). La selección de cerdos ibéricos. Porcino ibérico: aspectos claves (Buxadé C., Daza A., eds). Ed. Mundi-Prensa, Madrid, 125-149.

Simeone, R., Misztal, I., Aguilar, I., & Legarra, A. (2011). Evaluation of the utility of diagonal elements of the genomic relationship matrix as a diagnostic tool to detect mislabelled genotyped animals in a broiler chicken population. *Journal of Animal Breeding and Genetics*, *128*(5), 386–393. https://doi.org/10.1111/j.1439-0388.2011.00926.x

Smith, B. J. (2007). boa: An R package for MCMC output convergence assessment and posterior inference. *Journal of Statistical Software*, *21*(11), 1–37. https://doi.org/10.18637/jss.v021.i11

Song, H., Zhang, J., Jiang, Y., Gao, H., Tang, S., Mi, S., Yu, F., Meng, Q., Xiao, W., Zhang, Q., & Ding, X. (2017). Genomic prediction for growth and reproduction traits in pig using an admixed reference population. *Journal of Animal Science*, *95*(8), 3415. https://doi.org/10.2527/jas2017.1656

Spilke, J., Groeneveld, E., & Mielenz, N. (1998). Joint purebred and crossbred (co)variance component estimation with a pseudo multiple trait model: loss in efficiency. *Journal of Animal Breeding and Genetics*, *115*(1–6), 341–350. https://doi.org/10.1111/j.1439-0388.1998.tb00355.x

STRATEC Molecular GmbH. (2013). *User Manual User Manual* (Vol. 3304, Issue January).

VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, *91*(11), 4414–4423. https://doi.org/10.3168/jds.2007-0980

VanRaden, P. M., Van Tassell, C. P., Wiggans, G. R., Sonstegard, T. S., Schnabel, R. D., Taylor, J. F., & Schenkel, F. S. (2009). Invited review: Reliability of genomic predictions for North American Holstein bulls. *Journal of Dairy Science*, *92*(1), 16–24. https://doi.org/10.3168/jds.2008-1514

Vitezica, Z. G., Aguilar, I., & Legarra, A. (2010). One-Step vs . Multi-Step Methods For Genomic Prediction In Presence Of Selection. *Proceedings of 9th World Congress of Genetics Applied to Livestock Production: 1-6 August 2010*, *August*.

Wang, X., Xu, Y., Hu, Z., & Xu, C. (2018). Genomic selection methods for crop improvement: Current status and prospects. *Crop Journal*, *6*(4), 330–340. https://doi.org/10.1016/j.cj.2018.03.001

Wei, M., & Van Der Werf, J. H. J. (1994). Maximizing genetic response in crossbreds using both purebred and crossbred information. *Animal Production*, *59*(3), 401–413. https://doi.org/10.1017/S0003356100007923

Wientjes, Y., & Calus, M. P. L. (2017). BOARD INVITED REVIEW: The purebred-crossbred correlation in pigs: A review of theory, estimates, and implications. *Journal of Animal Science*, *95*(8), 3467. https://doi.org/10.2527/jas2017.1669

Ye, J., Tan, C., Hu, X., Wang, A., & Wu, Z. (2018). Genetic parameters for reproductive traits at different parities in Large White Pigs. *Journal of Animal Science*, *96*(4), 1215–1220.

https://doi.org/10.1093/jas/sky066

Zeng, J., Toosi, A., Fernando, R. L., Dekkers, J., & Garrick, D. J. (2013). Genomic selection of purebred animals for crossbred performance in the presence of dominant gene action. *Genetics Selection Evolution*, *45*(1), 1–17. https://doi.org/10.1186/1297-9686-45-11

Zhang, T., Wang, L. G., Shi, H. B., Yan, H., Zhang, L. C., Liu, X., Pu, L., Liang, J., Zhang, Y. B., Zhao, K. B., & Wang, L. X. (2016). Heritabilities and genetic and phenotypic correlations of litter uniformity and litter size in Large White sows. *Journal of Integrative Agriculture*, *15*(4), 848–854. https://doi.org/10.1016/S2095-3119(15)61155-8

6. Supplementary material.

Figure 10. Distribution for litter size records in Iberian pig population before quality control. TB: total born; NBA: number of born alive; SB: stillborn.

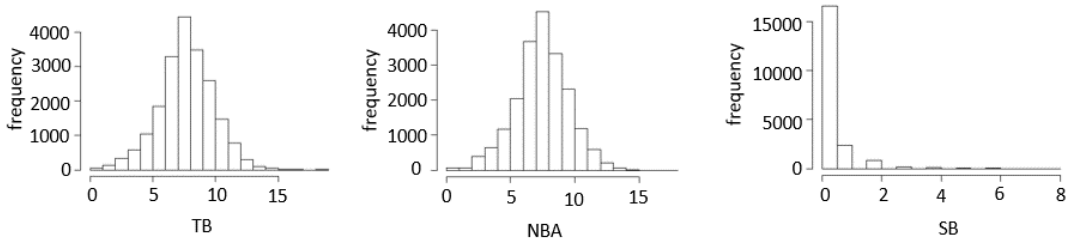

Figure 11. Percentage of variance explained by each principal component. PC: principal components of Iberian pig population under study (n=1114) obtained by principal component analysis.

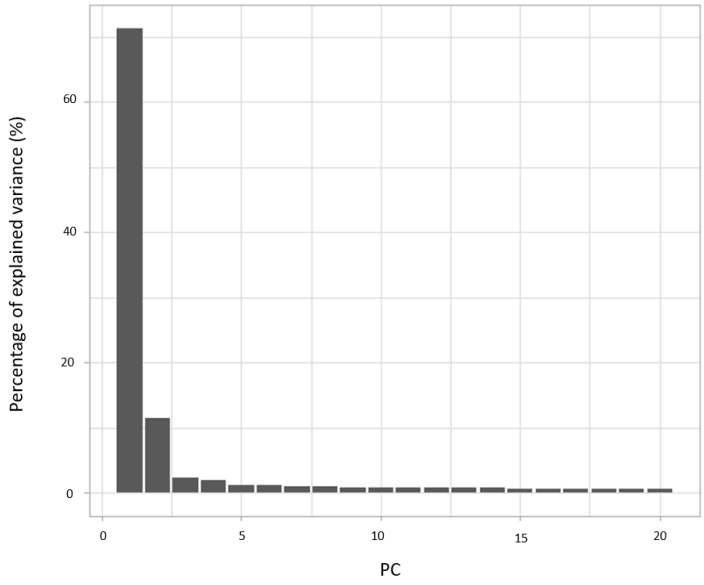

Figure 12. Cumulative explained variance by each principal component. PC: principal components of Iberian pig population under study (n=1114) obtained by principal component analysis.

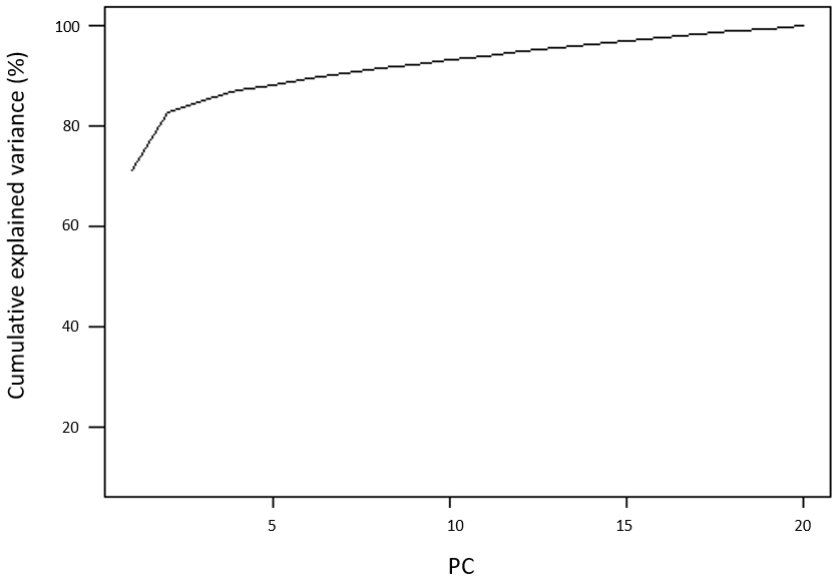

Figure 13. Histograms and plots for diagonal and off-diagonal elements in G (genomic relationship matrix) and A_{22} (numerator relationship matrix for genotyped animals) for Entrepelado line population. a) Frequency distribution for G diagonal elements; b) A_{22} diagonal elements and G diagonal elements plot; c) A_{22} off-diagonal elements and G off-diagonal elements plot.

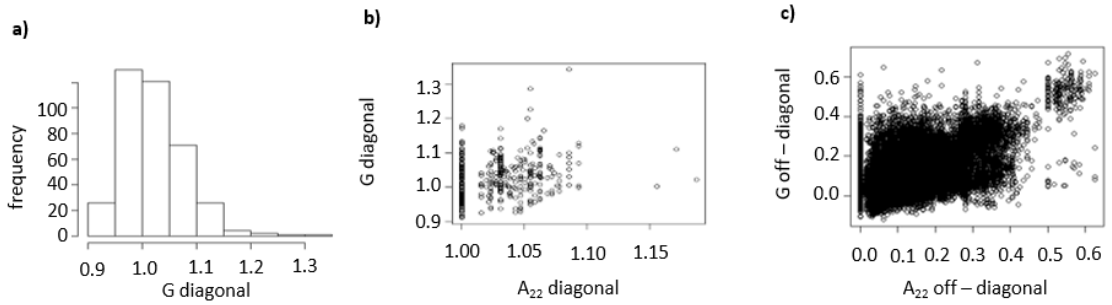

Figure 14. Histograms and plots for diagonal and off-diagonal elements in G (genomic relationship matrix) and A_{22} (numerator relationship matrix for genotyped animals) for Retinto line population. a) Frequency distribution for G diagonal elements; b) A_{22} diagonal elements and G diagonal elements plot; c) A_{22} off-diagonal elements and G off-diagonal elements plot.

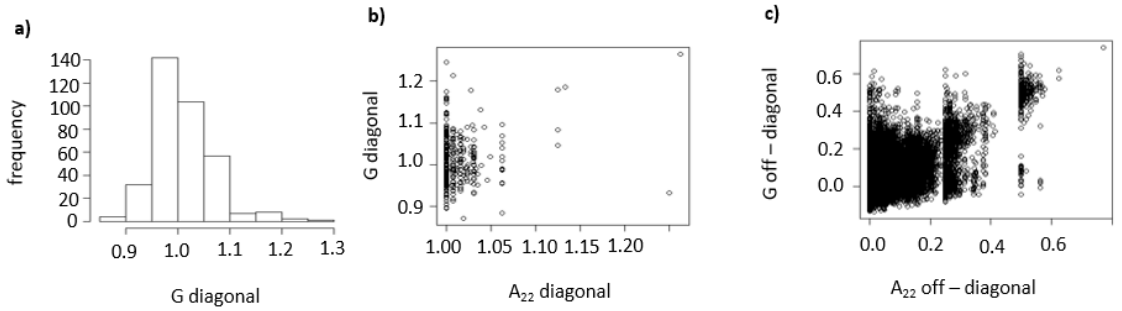

Table 11. Variance components estimation of total born (TB), number of born alive (NBA) and stillborn (SB) with BLUP (Best Linear Unbiased Prediction) Bayesian multi trait models for Entrepelado (EE) and Retinto (RR) lines.

	$\sigma_a^{2\,1}$			$\sigma_p^{2\,2}$			$\sigma_e^{2\,3}$		
Variance	Mean[4]	SD[5]	$HPD_{95\%}$[6]	Mean[4]	SD[5]	$HPD_{95\%}$[6]	Mean[4]	SD[5]	$HPD_{95\%}$[6]
TB_{EE}	0.35	0.09	0.18 – 0.52	0.74	0.07	0.59 – 0.88	2.57	0.04	2.47 – 2.65
NBA_{EE}	0.36	0.09	0.19 – 0.55	0.68	0.07	0.54 – 0.82	2.49	0.04	2.41 – 2.58
SB_{EE}	0.01	0.003	0.003 – 0.016	0.008	0.003	0.0026 – 0.014	0.39	0.007	0.38 – 0.41
TB_{RR}	0.25	0.08	0.11 – 0.40	0.74	0.08	0.60 – 0.89	3.09	0.06	2.98 – 3.20
NBA_{RR}	0.26	0.08	0.11 – 0.42	0.67	0.07	0.54 – 0.81	2.94	0.05	2.84 – 3.04
SB_{RR}	0.01	0.004	0.004 – 0.020	0.01	0.005	0.005 – 0.024	0.49	0.009	0.48 – 0.51

[1]σ_a^2: additive genetic variance; [2]σ_p^2: permanent environmental variance; [3]σ_e^2: residual variance; [4]Mean: posterior mean; [5]SD: posterior standard deviation; [6]HPD95%: highest posterior density interval at 95%.

Table 12. Variance components estimation of total born (TB), number of born alive (NBA), stillborn (SB) with ssGBLUP (Single Step Genomic Best Linear Unbiased Prediction) Bayesian multi trait models for Entrepelado (EE) and Retinto (RR) lines.

	$\sigma_a^{2\,1}$			$\sigma_p^{2\,2}$			$\sigma_e^{2\,3}$		
Variance	Mean[4]	SD[5]	$HPD_{95\%}$[6]	Mean[4]	SD[5]	$HPD_{95\%}$[6]	Mean[4]	SD[5]	$HPD_{95\%}$[6]
TB_{EE}	0.33	0.09	0.17 – 0.51	0.74	0.07	0.60 – 0.87	2.57	0.04	2.49 – 2.66
NBA_{EE}	0.33	0.09	0.15 – 0.50	0.71	0.08	0.55 – 0.86	2.49	0.04	2.41 – 2.58
SB_{EE}	0.009	0.003	0.003 – 0.014	0.009	0.003	0.002 – 0.015	0.39	0.007	0.38 – 0.41
TB_{RR}	0.27	0.08	0.12 -0.44	0.75	0.08	0.61 – 0.91	3.10	0.05	2.99 – 3.21
NBA_{RR}	0.22	0.07	0.08 – 0.36	0.70	0.07	0.56 – 0.83	2.94	0.05	2.83 – 3.04
SB_{RR}	0.01	0.005	0.004 – 0.021	0.01	0.005	0.004 – 0.023	0.49	0.009	0.48 – 0.51

[1]σ_a^2: additive genetic variance; [2]σ_p^2: permanent environmental variance; [3]σ_e^2: residual variance; [4]Mean: posterior mean; [5]SD: posterior standard deviation; [6]HPD95%: highest posterior density interval at 95%.

Table 13. Variance components estimation for number of total born (TB), number of born alive (NBA), stillborn (SB) with BLUP (Best Linear Unbiased Prediction) Bayesian multi breed models for Entrepelado (EE) and Retinto (RR) lines and crossbred animals (ER-RE).

	$\sigma_a^{2\,1}$			$\sigma_p^{2\,2}$			$\sigma_e^{2\,3}$		
Variance	Mean[4]	SD[5]	$HPD_{95\%}$[6]	Mean	SD	$HPD_{95\%}$	Mean	SD	$HPD_{95\%}$
TB_{EE}	0.31	0.06	0.20 – 0.42	0.26	0.05	0.16 – 0.36	2.94	0.06	2.84 – 3.04
TB_{RR}	0.19	0.05	0.11 – 0.31	0.18	0.04	0.09 – 0.26	3.56	0.07	3.42 – 3.70
$TB_{ER\,-\,RE}$	0.46	0.07	0.30 – 0.58	0.17	0.06	0.07 – 0.27	4.06	0.08	3.92 – 4.21
NBA_{EE}	0.31	0.05	0.21 – 0.42	0.23	0.05	0.15 – 0.33	2.84	0.05	2.76 – 2.95
NBA_{RR}	0.18	0.05	0.09 – 0.28	0.17	0.04	0.09 – 0.26	3.36	0.07	3.25 – 3.50
$NBA_{ER\text{-}RE}$	0.43	0.06	0.29 – 0.54	0.15	0.05	0.06 – 0.25	3.92	0.07	3.76 – 4.05
SB_{EE}	0.009	0.003	0.004 – 0.014	0.003	0.001	0.002 – 0.006	0.38	0.007	0.36 – 0.39
SB_{RR}	0.010	0.003	0.0002 – 0.008	0.007	0.002	0.004 – 0.012	0.47	0.008	0.45 – 0.48
$SB_{ER\text{-}RE}$	0.008	0.002	0.004 – 0.13	0.014	0.004	0.008 – 0.022	0.37	0.007	0.36 – 0.38

[1]σ_a^2: additive genetic variance; [2]σ_p^2: permanent environmental variance; [3]σ_e^2: residual variance; [4]Mean: posterior mean; [5]SD: posterior standard deviation; [6]HPD95%: highest posterior density interval at 95%.

Table 14. Variance components estimation for number of total born (TB), number of born alive (NBA), stillborn (SB) with ssGBLUP (Single Step Best Linear Unbiased Prediction) Bayesian multi breed models for Entrepelado (EE) and Retinto (RR) lines and crossbred animals (ER-RE).

	σ_a^2			σ_{pe}^2			σ_e^2		
Variance	Mean[4]	SD[5]	$HPD_{95\%}$[6]	Mean	SD	$HPD_{95\%}$	Mean	SD	$HPD_{95\%}$
TB_{EE}	0.27	0.04	0.20 – 0.36	0.28	0.04	0.21 – 0.35	2.93	0.06	2.83 – 3.04
TB_{RR}	0.21	0.06	0.12 – 0.30	0.17	0.02	0.13 – 0.23	3.54	0.07	3.40 – 3.66
$TB_{ER\text{-}RE}$	0.41	0.07	0.28 – 0.56	0.27	0.09	0.11 – 0.42	4.03	0.08	3.88 – 4.18
NBA_{EE}	0.28	0.04	0.21 – 0.35	0.26	0.03	0.19 – 0.31	2.83	0.05	2.74 – 2.95
NBA_{RR}	0.17	0.03	0.09 – 0.23	0.21	0.03	0.16 – 0.26	3.34	0.07	3.20 – 3.46
$NBA_{ER\text{-}RE}$	0.38	0.07	0.26 – 0.53	0.25	0.09	0.11 – 0.39	3.90	0.08	3.73 – 4.03
SB_{EE}	0.006	0.001	0.003 – 0.008	0.006	0.002	0.003 – 0.01(	0.38	0.006	0.37 – 0.39
SB_{RR}	0.011	0.003	0.005 – 0.017	0.008	0.003	0.003 – 0.01	0.47	0.007	0.45 – 0.48
$SB_{ER\text{-}RE}$	0.011	0.003	0.006 – 0.019	0.009	0.004	0.003 – 0.01(	0.37	0.007	0.36 – 0.38

[1]σ_a^2: additive genetic variance; [2]σ_p^2: permanent environmental variance; [3]σ_e^2: residual variance; [4]Mean: posterior mean; [5]SD: posterior standard deviation; [6]HPD95%: highest posterior density interval at 95%.

Table 15. Estimates for heritability (standard deviation) in different Iberian pig populations by different authors applying different models.

Trait	Population	Estimates (SD[1])	Model	Author(s)
	RE[11], RC[12], BHC[13] and BHP[14]	0.07 (0.02)	Repeatability animal model	García-Casco *et al*., (2012)
TB[2]	Entrepelado	0.140 (0.022)	Multiple population repeatability model	Noguera *et al*. (2019)
	Retinto	0.009 (0.017)	Multiple population repeatability model	Noguera *et al*. (2019)
	Torbiscal	0.086 (0.022)	Multiple population repeatability model	Noguera *et al*. (2019)
$TB_{<2}$[3]	RE, RC, BHC and BHP	0.07 (0.02)	Multi trait animal model	García-Casco *et al*. (2012)
$TB_{>3}$[4]	RE, RC, BHC and BHP	0.10 (0.02)	Multi trait animal model	García-Casco *et al*. (2012)
NBA[5]	RE, RC, BHC and BHP	0.06 (0.02)	Repeatability animal model	García-Casco *et al*. (2012)
	Torbiscal	0.07 (0.01)	Repeatability animal model	Fernández *et al*. (2008)
	Entrepelado	0.131 (0.022)	Multiple population repeatability model	Noguera *et al*. (2019)
	Retinto	0.084 (0.017)	Multiple population repeatability model	Noguera *et al*. (2019)
	Torbiscal	0.078 (0.021)	Multiple population repeatability model	Noguera *et al*. (2019)
		0.06 (0.02)	Bivariate animal model	Muñoz *et al*. (2018)
	Retinto	0.07 (0.07)	Single trait Animal model	Corral *et al*. (2010)
	Torbiscal	0.06 (0.025)	Single trait animal model	Rodriguez *et al*. (1994)
	Torbiscal	0.06 (0.025)	Animal model with maternal effects	Rodriguez *et al*. (1994)
NBA_1[6]	Torbiscal	0.13 (0.02)	Repeatability animal model	Fernández *et al*. (2008)
NBA_2[6]	Torbiscal	0.09 (0.02)	Repeatability animal model	Fernández *et al*. (2008)
NBA_3[6]	Torbiscal	0.10 (0.02)	Repeatability animal model	Fernández *et al*. (2008)
NBA_4[6]	Torbiscal	0.06 (0.02)	Repeatability animal model	Fernández *et al*. (2008)

NBA_5[6]	Torbiscal	0.11 (0.03)	Repeatability animal model	Fernández *et al*. (2008)
$NBA_{>5}$[7]	Torbiscal	0.09 (0.01)	Repeatability animal model	Fernández *et al*. (2008)
$NBA_{<2}$[8]	RE, RC, BHC and BHP	0.06 (0.02)	Multi trait animal model	García-Casco *et al*. (2012)
$NBA_{>3}$[9]	RE, RC, BHC and BHP	0.11 (0.02)	Multi trait animal model	García-Casco *et al*. (2012)
SB[10]	Retinto	0.02 (0.05)	Single trait animal model	Corral *et al*. (2010)

[1]SD: standard deviation; [2]TB: total born; [3]$TB_{<2}$: total born at first and second parities; [4]$TB_{>3}$: total born at third and subsequent parities; [5]NBA: number of born alive; [6]$NBA_{1\text{-}2\text{-}3\text{-}4\text{-}5}$: number of born alive at first, second, third, fourth and fifth parity, respectively; [7]$NBA_{>5}$: number of born alive at fifth and subsequent parities; [8]$NBA_{<2}$: number of born alive at first and second parities; [9]$NBA_{>3}$: number of born alive at third and subsequent parities; [10]SB: stillborn; [11]RE: Portuguese Red Ervideira strain; [12]RC: Portuguese Red Caldeira strain; [13]BHC: Black Hairless Campanar strain; [14]BHP: Black Hairless Puebla strain.