TESIS DE LA UNIVERSIDAD DE ZARAGOZA
2024
277

Houssemeddine Srihi

Genetic and genomic analysis of litter size in a crossbred between two varieties of Iberian pig

Director/es
Varona Aguado, Luis

ISSN 2254-7606

Prensas de la Universidad
Universidad Zaragoza

Universidad de Zaragoza
Servicio de Publicaciones

ISSN 2254-7606

Universidad Zaragoza
1542

Tesis Doctoral

GENETIC AND GENOMIC ANALYSIS OF LITTER SIZE IN A CROSSBRED BETWEEN TWO VARIETIES OF IBERIAN PIG

Autor

Housseméddine Srihi

Director/es

Varona Aguado, Luis

UNIVERSIDAD DE ZARAGOZA
Escuela de Doctorado

Programa de Doctorado en Producción Animal

2024

UNIVERSIDAD DE ZARAGOZA
FACULTAD DE VETERINARIA

TESIS DOCTORAL

Genetic and genomic analysis of litter size in a crossbred between two varieties of Iberian pig

HOUSSEMEDDINE SRIHI

ZARAGOZA, 2024

Universidad Zaragoza
1542

Tesis Doctoral

Genetic and genomic analysis of litter size in a crossbred between two varieties of Iberian pig

Autor

Houssemeddine Srihi

Director

Luis Varona Aguado

Facultad De Veterinaria

Departamento de Anatomía, Embriología y Genética Animal

2024

Departamento de
Anatomía, Embriología
y Genética Animal
Universidad Zaragoza

FACULTAD DE VETERINARIA

Departamento de Anatomía, Embriología y Genética Animal

TESIS DOCTORAL

Genetic and genomic analysis of litter size in a crossbred between two varieties of Iberian pig

Programa de Doctorado en Producción Animal

Housseméddine Srihi

Director

Luis Varona Aguado

2024

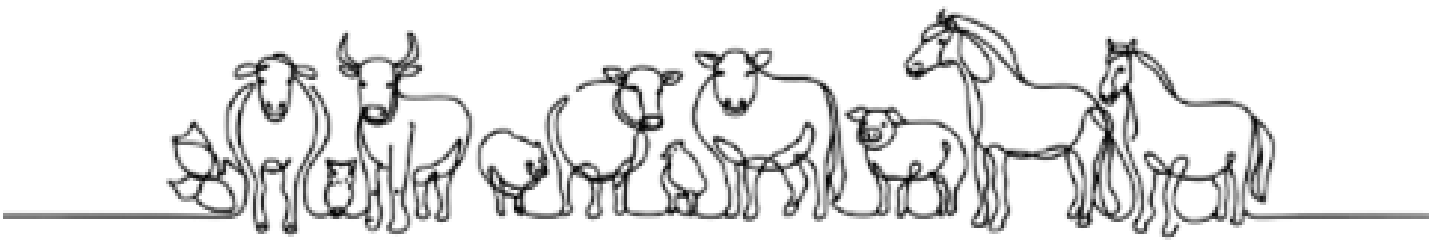
QUANTITATIVE GENETICS LAB
Universidad Zaragoza

Facultad de Veterinaria
Universidad Zaragoza

TESIS DOCTORAL

Genetic and genomic analysis of litter size in a crossbred between two varieties of Iberian pig

Memoria presentada por **Houssemeddine Srihi** para optar al grado de Doctor por la Universidad de Zaragoza
2024



Luis Varona, Doctor en Veterinaria, Catedrático de Universidad del Departamento de Anatomía, Embriología y Genética de la Universidad de Zaragoza e Investigador del Instituto Agroalimentario de Aragón (IA2)

HACE CONSTAR

Que **Houssemeddine Srihi** ha realizado bajo mi dirección los trabajos correspondientes a su Tesis Doctoral titulada "**Genetic and genomic analysis of litter size in a crossbred between two varieties of Iberian pig**", que corresponde con el proyecto de Tesis aprobado por la comisión de Doctorado en Producción Animal, y que cumple con los requisitos exigidos para optar al grado de Doctor por la Universidad de Zaragoza, por lo que autorizo su presentación para que pueda ser juzgada por el Tribunal correspondiente.

Lo que suscribo como director del trabajo, en Zaragoza, a 19 de diciembre 2023

This project has received funding from the European Union's H2020 research and innovation programme under Marie Sklodowska-Curie grant agreement No 801586

This work was possible thanks to the following pre-doctoral scholarship awarded to Houssemeddine Srihi.

funding from the European Union's Horizon 2020 research and innovation program, specifically under the Marie Sklodowska-Curie grant. This support was facilitated through a collaboration with Iberus Talent, with the grant numbered H2020-G.A. number 801586, as part of the research project "PID2020-114705RB-I00".

ACKNOWLEDGEMENTS

As I culminate this significant journey in my academic career, I first and foremost offer my sincerest gratitude to God, whose guidance and blessings have been ever-present.

I am filled with gratitude for the numerous individuals who have supported me along this path. This thesis, a culmination of hard work and dedication, would not have been possible without the encouragement and support of many.

First and foremost, I extend my deepest gratitude to my family. To my parents Sami and Najoua , whose unwavering support and belief in my abilities have been the cornerstone of my strength and perseverance. Your sacrifices and unconditional love have been the guiding light through this journey. To my siblings and extended family, thank you for your endless encouragement and for always believing in me.

My heartfelt appreciation goes to my beloved wife, Feten Chihi. Your love, patience, and unwavering support have been my safe haven. Your understanding and encouragement during this demanding journey have been nothing short of extraordinary. I am endlessly grateful for your companionship and the comfort you have provided me.

I am immensely grateful to my supervisor, Luis Varona, for his invaluable guidance, patience, and expertise. Your mentorship has been instrumental in shaping my academic journey, and your unwavering support has been crucial in bringing this thesis to fruition. I also extend my heartfelt thanks to all the members of the department. Your collective wisdom, support, and kindness have provided a nurturing environment for my growth and learning.

A special word of thanks to my colleagues, David and Carlos. Your presence in the department and partnership during various conferences and events have been invaluable. Our shared experiences and the good times spent together have not only made this journey more enjoyable but also enriched my professional growth and personal experiences.

I also owe a special word of appreciation to Noelia Ibáñez Escriche whose support from Valencia was instrumental in my academic journey. Her willingness to share her extensive knowledge and expertise has been invaluable. Noelia's insights and guidance

have not only enriched my understanding but have also significantly contributed to the depth and quality of my research.

My sincere appreciation goes to Campus Iberus for their support and funding the project, providing the resources and opportunities that have significantly enriched my research experience. Your contribution has been pivotal in my academic journey.

Lastly, I would like to extend my heartfelt gratitude to my friends, who have been pillars of motivation and comfort throughout this journey. Special thanks to Panos, Fabian, and Carmen for your unconditional friendship. In the city of Zaragoza, a place far from my home, you ensured I never felt alone. Together, we have shared countless memorable moments, from adventurous trips to attending various events. These experiences have not only enriched my stay in Zaragoza but also provided a sense of belonging and joy amidst the rigors of academic pursuits.

Along this journey, I have met many people each of whom has played a significant role in my life. While I may not have named everyone here, please know that your impact is no less appreciated. Your presence in my life, whether brief or lasting, has been a valuable part of my journey, helping to shape the person I am today. Thank you to each one of you for crossing my path and for the unique contributions you have made to my story.

Table of Contents

List of Figures

CHAPTER I

CHAPTER II

CHAPTER III

CHAPTER V

CHAPTER VI

CHAPTER VII

CHAPTER VIII

SUPPLEMENTARY MATERIALS

List of Tables

CHAPTER I

CHAPTER II

CHAPTER III

CHAPTER V

CHAPTER VI

CHAPTER VII

SUPPLEMENTARY MATERIALS

List of abbreviations

%IMF	Intramuscular fat percentage
AECERIBER	Asociación Española de Criadores de Cerdo Ibérico
BLUP	Best linear unbiased prediction
CQI	Carcass Quality Index
EE	Entrepelado
ER	Entrepelado x Retinto
HP	Ham percentage on the carcass
L75	piglet Weight at 75 days
LP	Loin percentage
LW	Litter Weight
MAPA	Ministerio de Agricultura, pesca y Alimentación
MT-BLUP	Multi trait animal model Best linear Unbiased Prediction
MUFA	Monounsaturated Fatty Acids
N	Number of phenotypic records
NBA	Number of Piglets Born Alive
NS	Number of recorded sows
NWP	Number of Weaned Piglets
PUFA	Polyunsaturated Fatty Acids
PW45	Piglet Weight at 45 days
QTL	Quantitative trait loci
RE	Retinto x Entrepelado
RR	Retinto
SFA	Saturated Fatty Acids
SNP	Single nucleotide polymorphism
SP	Shoulder percentage
ST-BLUP	Single trait animal model Best Linear Unbiased Prediction
TNB	Total number of piglets born
UPGMA	Unweighted pair group method with arithmetic mean

List of publications

This Doctoral Thesis was carried out at “Departamento de Anatomía, Embriología y Genética Animal”, Facultad veterinaria de Zaragoza within the framework of the following research project: "PID2020-114705RB-I00".

Responsible researcher: Luis Varona.

This thesis is based on the work contained in the following papers.

I. Srihi, H.; Noguera, J.L.; Topayan, V.; Martín de Hijas, M.; Ibañez-Escriche, N.; Casellas, J.; Vázquez-Gómez, M.; Martínez-Castillero, M.; Rosas, J.P.; Varona, L. Additive and Dominance Genomic Analysis for Litter Size in Purebred and Crossbred Iberian Pigs. *Genes* **2022**, *13*, 12. https://doi.org/10.3390/genes13010012.

II. Srihi, H.; López-Carbonell, D.; Ibáñez-Escriche, N.; Casellas, J.; Hernández, P.; Negro, S.; Varona, L. A Bayesian Multivariate Gametic Model in a Reciprocal Cross with Genomic Information: An Example with Two Iberian Varieties. *Animals* **2023**, *13*, 1648. https://doi.org/10.3390/ani13101648.

III. Srihi, H., López-Carbonell, D., Ibáñez-Escriche, N., Casellas, J., Hernández, P., Negro, S., & Varona, L. A multivariate gametic model for the analysis of purebred and crossbred data. An example between two populations of Iberian pigs. *Journal of Animal Breeding and Genetics* **2023**. https://doi.org/10.1111/jbg.12832.

Conference Proceedings

Srihi, H., Ibáñez-Escriche, N., Casellas, J., Noguera, J. L., Hernández, P., Martín de Hijas, M., Vázquez-Gómez, M., Negro, S., Rosas, J. P., & Varona, L. (2022). Bayesian analysis of paternal and maternal gametic effects in a reciprocal cross between two Iberian varieties. In Proceedings of 12th World Congress on Genetics Applied to Livestock Production (WCGALP). Wageningen Academic Publishers. https://doi.org/10.3920/978-90-8686-940-4_772.

ABSTRACT

The Iberian pig breed is widely renowned for their high-quality meat products, particularly the famous Jamón ibérico. However, like any livestock species, Iberian pigs present unique challenges in breeding and genetics, especially when it comes to optimizing their reproductive performance. To maintain the sustainability and profitability of pig farming, INGA FOOD S.A. initiated a crossbreeding program to produce a hybrid sow called CASTÚA by crossing Retinto and Entrepelado varieties of Iberian pigs. The goal is to improve reproductive characteristics and maximize hybrid vigor or heterosis in crossbred population. This thesis focuses on the genetic aspects of a crossbreeding program including additive and dominance genetic effects, genomic imprinting effects, and the correlations between paternal and maternal gametic effects.

The first part of the thesis estimates the additive and dominance variances in both purebred and crossbred populations. The study found that the crossbred population exhibited significantly higher additive genotypic variance for Total Number Born (TNB) and Number Born Alive (NBA) than the purebred population. Additionally, we observed that the genetic correlations between purebred and crossbred performances were higher in the Retinto population, with correlations of 0.663 in Entrepelado and 0.881 in Retinto populations for TNB and NBA. The study also identified four genomic regions on chromosomes 6, 8, and 12, each explaining more than 2% of the additive genetic variance.

Moreover, this thesis expands upon previous studies that initially discovered imprinting effects in this diallelic cross, the second and third parts of the thesis delve deeper into genomic imprinting effects. A multivariate gametic model is employed to estimate gametic correlations between paternal and maternal effects in reciprocal crosses. The results revealed differences in gametic correlations between the two populations, which may explain the distinct performance outcomes observed in reciprocal crosses. We expanded the model to incorporate the correlations between four gametic effects in each parental population. The model, applied to datasets from purebred and crossbred pigs, showed that all posterior estimates of the gametic correlations were positive and confirmed the influence of imprinting effects on the genetic control of litter size. The study also emphasized that maternal genetic contribution plays a more substantial role in determining observed reproductive traits compared to paternal genetic contribution.

RESUMEN

El cerdo ibérico es reconocido por sus productos cárnicos de alta calidad, en particular el famoso Jamón ibérico. Sin embargo, como cualquier especie ganadera, los cerdos ibéricos presentan desafíos particulares en cuanto a su mejora genética, especialmente cuando se trata de optimizar su rendimiento reproductivo. Para mantener la sostenibilidad y rentabilidad de la producción intensiva de cerdo ibérico, la empresa INGA FOOD S.A. inició un programa de cruzamiento para producir una cerda híbrida llamada CASTÚA mediante el cruce de las variedades Retinto y Entrepelado. El objetivo de este cruce es mejorar las características reproductivas y maximizar el vigor híbrido o heterosis en la población cruzada. Esta tesis se centra en los aspectos genéticos de un programa de cruce, incluyendo los efectos genéticos aditivos y dominantes, los efectos de impronta genómica y las correlaciones entre los efectos gaméticos paternos y maternos.

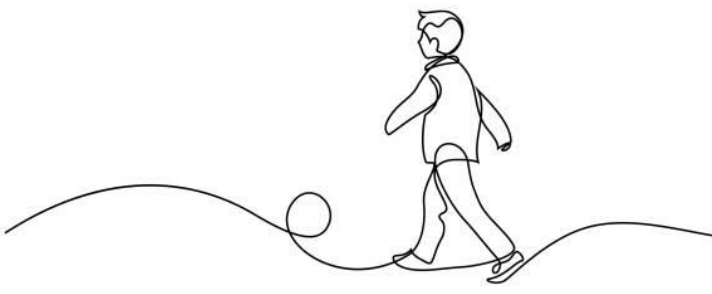
La primera parte de la tesis estima las varianzas aditivas y dominantes en poblaciones tanto puras como cruzadas. El estudio encontró que la población cruzada exhibió una varianza genotípica aditiva significativamente mayor para TNB y NBA que la población pura. Además, se observó que las correlaciones genéticas entre el rendimiento de puras y cruzadas fueron más altas en la población Retinto, con correlaciones de 0.663 en poblaciones Entrepelado y 0.881 en poblaciones Retinto para nacidos totales (TNB) y nacidos vivos (NBA). El estudio también identificó cuatro regiones genómicas en los cromosomas 6, 8 y 12, cada una explicando más del 2% de la varianza genética aditiva.

Adicionalmente, esta investigación amplía estudios anteriores que inicialmente descubrieron efectos de impronta en este cruce dialélico; las segunda y tercera partes de la tesis profundizan en los efectos de impronta genómica. Se empleó un modelo gamético multivariado para estimar las correlaciones gaméticas entre los efectos paternos y maternos en cruces recíprocos. Los resultados revelaron diferencias en las correlaciones gaméticas entre las dos poblaciones, lo que puede explicar los resultados distintos observados en cruces recíprocos. Además, se amplió el modelo para incorporar las correlaciones entre cuatro efectos gaméticos en cada población parental. El modelo propuesto, aplicado a datos de cerdos puros y cruzados, mostró que todas las estimaciones posteriores de las correlaciones gaméticas fueron positivas y confirmaron la influencia de los efectos de impronta en el control genético del tamaño de la camada. El estudio también

enfatizó que la contribución genética materna juega un papel más sustancial en la determinación de las características reproductivas observadas en comparación con la contribución genética paterna.

CHAPTER I

General introduction



1. Introduction

1.1. Pig Sector in the World

The global production of pork meat has witnessed substantial growth since the mid-20th century, and this upward trend has persisted to the present day. According to United States Department of Agriculture (USDA), global pig meat production reached approximately 115 million tonnes in 2022, establishing it as the most extensively produced meat worldwide, surpassing poultry and beef (see Figure 1).

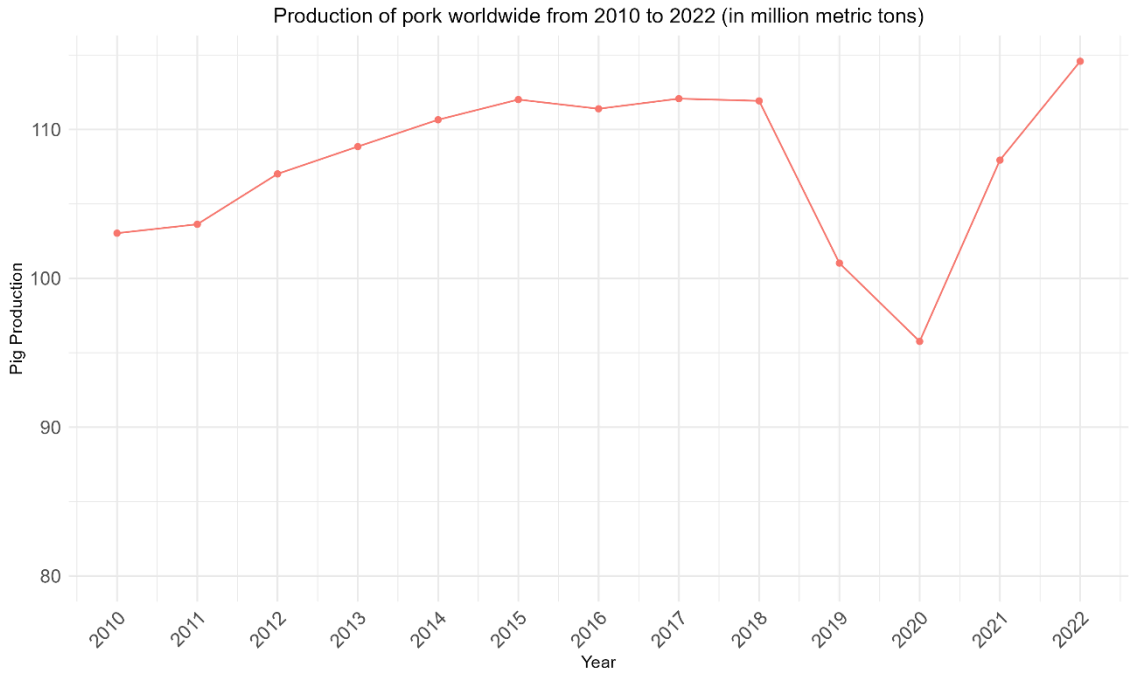


Figure 1: Production of pig Meat in World. (Adapted from USDA , 2023)

China, the European Union, and the United States stand as the foremost contributors to pig meat production, collectively accounting for more than half of the global output (refer to Figure 2). Other notable contributors to the pork industry encompass Brazil, Canada, Russia, and Mexico. The substantial growth in pig meat production can be attributed to the increasing demand for pork, advancements in breeding, feeding, and production technologies, as well as enhancements in infrastructure and logistics.

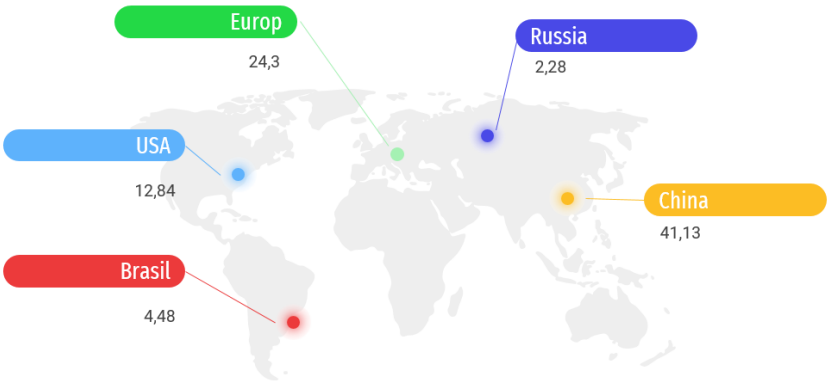


Figure 2: Production of pig meat by country FAO statistics data. (Adapted from FAOSTAT, 2022)

1.2. Pig Sector in Spain

The pig farming sector in Spain stands as one of the nation's most significant agricultural industries, characterized by a rich history and a resilient pork production landscape that has evolved over the years. Distinguished by its adoption of advanced technologies and specialized breeds, Spain's pork industry relies on intensive production methods. According to data from the Spanish Ministry of Agriculture, Fisheries, and Food (MAPA), the Spanish pig sector holds a pivotal role within the nation's economy, constituting approximately 16% of the Final Agrarian Production. Within the broader spectrum of livestock production, the pig sector takes the lead in terms of its economic significance, contributing a substantial 42% to the Final Livestock Production. In 2022, Spain produced approximately 5 million tonnes of pork, representing a modest decline of about 2.2% compared to the previous year. This production was derived from the processing of 56.6 million pigs, solidifying Spain's leadership within the European Union and securing the third position globally, following only China and the United States in terms of pork output.

Spain's expanding role within the European Union is notable, with the country now commanding a 22.9% share of the bloc's pork production, surpassing Germany, which holds a 20.3% share (see Figure 3). The pig population in Spain exceeds 30 million (see Figure 4). This underscores Spain's increasing influence on the European stage; just half a decade ago, its production accounted for a mere 18% of the EU's total. Over the past five years, while the EU's pork production increased by a marginal 0.3%, Spain's

production surged by nearly 24%, highlighting the sector's remarkable expansion at the national level. The distribution of the Spanish pig population across its Autonomous Communities reveals that Aragón and Cataluña are the leading regions with highest number of pigs. Together, these regions account for over 50% of Spain's pig farming, with each boasting approximately eight million pigs. Following closely behind is Castilla y León, with a count of four million pigs, solidifying its position as the third-largest pig farming region in Spain (constituting 14%). Andalucía and the Region of Murcia follow with 8% and 6%, respectively (see Figure 5).

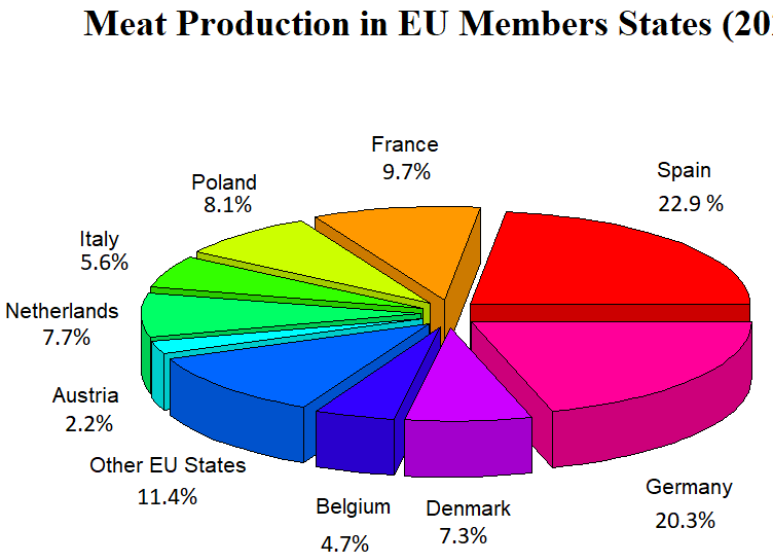


Figure 3: Pig meat in EU Members state. (Adapted MAPA, 2022)

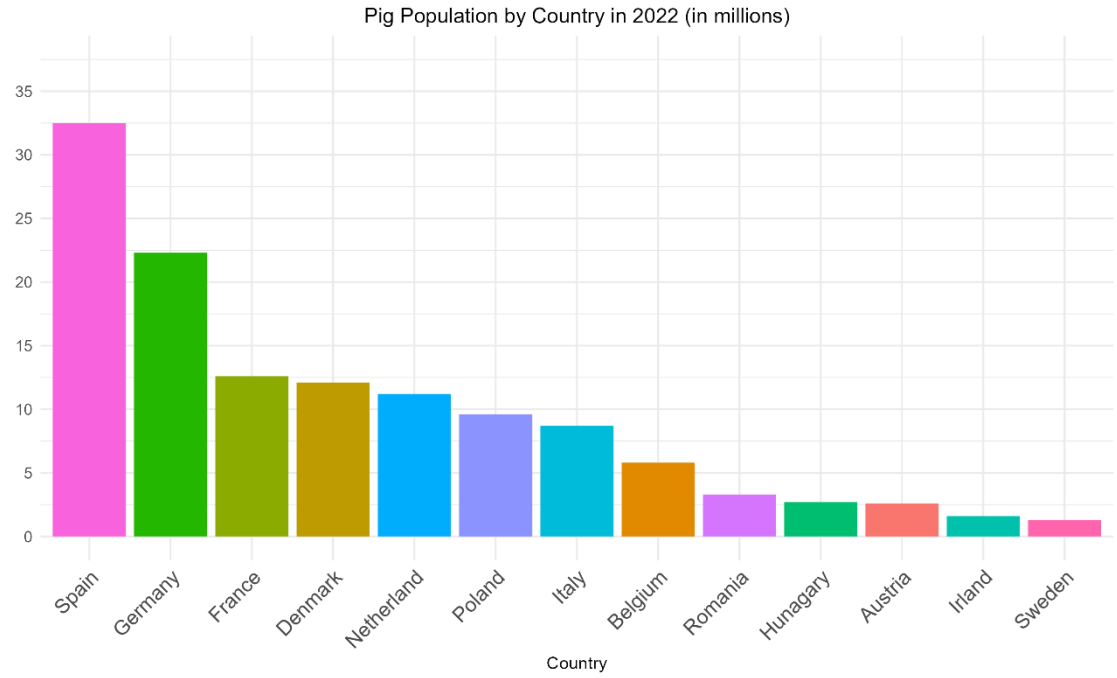


Figure 4: Pig population in EU membres. (Adapted from EUROSAT, 2022)

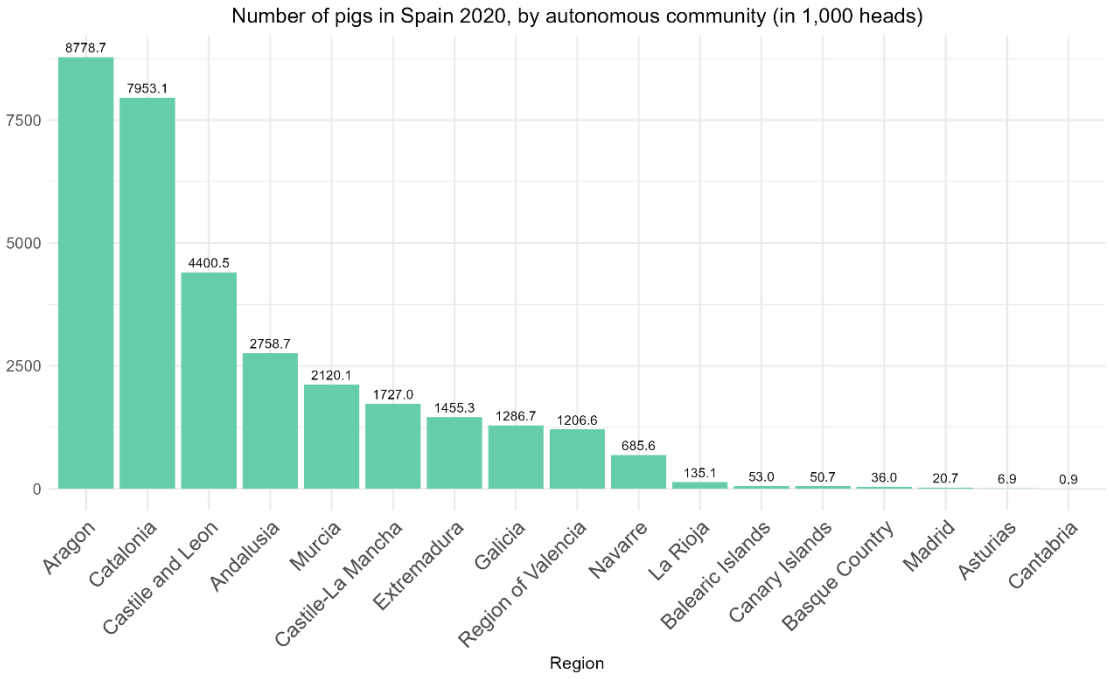


Figure 5: Number of pigs in Spain 2020, by autonomous community. (Adapted from Mercasa, 2022)

The success of the Spanish pig farming industry is exemplified by its status as one of the world's leading pork exporters. Spanish pork products reach numerous countries across the globe, with primary destinations including Italy, France, Germany, Portugal, and China. The widespread global distribution of Spanish pork highlights the nation's proficiency in consistently supplying high-quality meat products to meet international demand.

In Spain, the majority of pig production falls into two distinct categories: intensive production of international white pig breeds and the breeding of Iberian pigs for premium-quality products. Currently, just over one-tenth of the total pig livestock census comprises the Iberian breed. Iberian pigs are selectively raised to produce top-tier products, with breeders adhering to stringent guidelines that require at least 50% Iberian lineage to qualify within the four categories for Iberian pigs. While representing only a fraction of the country's overall pig production, Spain annually raises approximately three million Iberian pigs which represents about 11.3% pigs produced nationwide.

1.3. Iberian pigs

Iberian pigs, native to the Iberian Peninsula encompassing Spain and Portugal, were traditionally raised in a distinctive agroforestry system known as the 'dehesa.' This approach involves extensive or semi-extensive management until the pigs reach a body weight of 95–105 kg. Subsequently, during the finishing period, referred to as the 'montanera,' the pigs graze on acorns and pastures until they reach a body weight of 155–165 kg, typically occurring between 14 and 18 months of age. Purebred pigs produced within this 'montanera' phase fall under the top-quality category known as 'etiqueta negra' (black label).

However, in recent years, most fattened Iberian pigs have resulted from crossbreeding, typically involving Duroc or Crossbred (Duroc x Iberian) boars paired with Iberian sows. This crossbreeding with the Duroc breed enhances precocity, increases lean deposition rates (Serrano et al., 2008). These pigs can be reared in various systems, ranging from intensive methods where concentrates form the primary diet, to mixed and extensive systems that incorporate natural resources such as pastures and acorns.

To illustrate the numerical significance of the various rearing systems, let's consider data from 2017. During that year, a total of 635,000 Iberian pigs were raised using the “montanera” method, consisting of 297,000 purebred and 338,000 crossbred with Duroc. In contrast, 664,000 pigs were raised in extensive or semi-extensive systems without acorn feeding, with the majority being crossbreeds. Additionally, 1,941,000 pigs were raised in intensive systems, all of which were crossbreeds. These statistics highlight that only 20% of the pigs were raised using the traditional “montanera” system, and a mere 10% of the total slaughtered pigs were pure Iberian.

Figure 6: Iberian pigs in "*dehesa"* (Fernando , 2018).

1.3.1. Genetic diversity of Iberian pigs.

There are numerous varieties of Iberian pigs, each characterized by distinct physical traits. Variations in morphological attributes, as depicted in Figure 7, encompass both qualitative factors such as coat color, ear shape, hair density, and frontonasal profile, as well as quantitative factors like snout and ear width, and weight. These differences have been comprehensively observed and documented within various sub-populations, as reported by Giraldo & Tovar (2001).

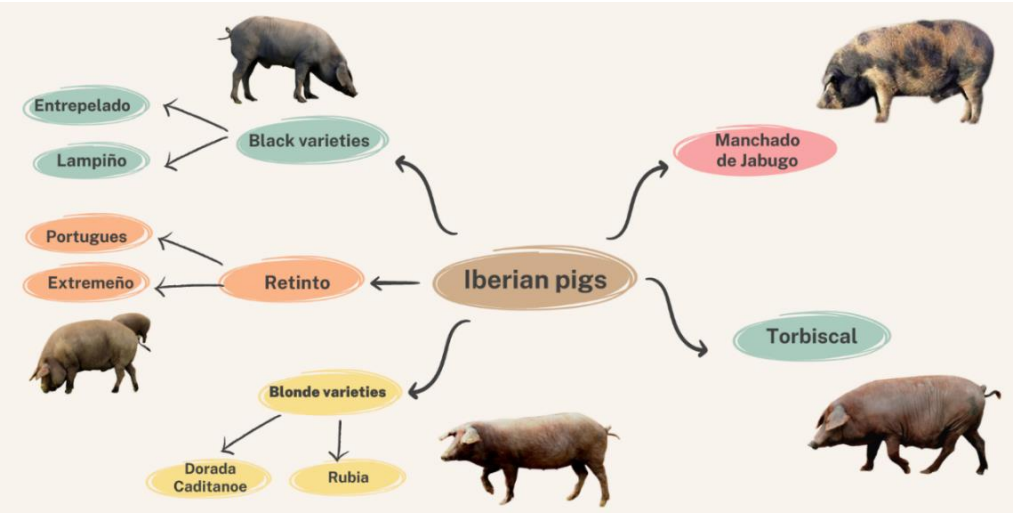


Figure 7: Classification of Iberian Pigs. (Adapted from Rueda & Diéguez, 2004)

These morphological distinctions are particularly pronounced in sub-varieties such as Torbiscal, Retinto, Entrepelado, and Lampiño, as compared to other sub-populations, as emphasized by Benito-Hernández et al. (1997). Moreover, these morphological variations are intricately linked to the pigs' lifestyle and diet, exerting a significant influence on the quality and flavor of their meat.

Over the years, researchers have conducted various studies to characterize the different varieties of Iberian pigs, including black (Mata et al., 1998) and red varieties (Pardo et al., 1998). This study contributed to the morphological differentiation of these varieties and their conservation. Further studies by Cabello et al. (2007), López (2009), Rodrigáñez et al. (1999) and Clemente et al. (2008) have provided specific descriptions for different strains and lines of Iberian pigs. In particular, the Spanish Ministry of Agriculture, Fisheries, and Food (MAPA) now recognizes five varieties based on these characteristics:

Entrepelado: The name Entrepelado is derived from the Spanish word 'pelado,' which translates to 'hairless' or 'bald.' These pigs are called 'Entrepelado' because they have relatively sparse hair compared to other sub-varieties of Iberian pigs. The hairlessness of the Entrepelado pig makes it more manageable and facilitates processing during the slaughtering process, offering a significant advantage in the meat industry. The meat from the Entrepelado pig possesses a slightly distinct flavor profile compared to other Iberian pigs, primarily due to its lower fat content (MAPA, 2016).

Retinto: Retinto is a sub-variety of Iberian pigs that are native to the south of Spain. This sub-variety has a hair color (or coat) that ranges from cinnamon shades to dark or 'retinto', and their distinctive long ears. Retinto pigs are usually medium-sized with a weight range of around 100-160 kg (MAPA, 2016).

Torbiscal: This breed is distinguished by its taller stature, a more extended and straighter back compared to other breeds, and limbs that lack the highly valued fine structure. Additionally, its hooves often exhibit a light streaking that is generally not preferred by breeders. The coat of the Torbiscal pig is a slate-red color, with two variations, light and dark, and is covered in abundant bristles. Its head is uniquely characteristic, marked by the length of its ears and snout (Rodrigáñez et al., 1999).

Lampiño: Lampiño is an early-maturing animal that tends to fatten quickly. It either lacks hair or has very sparse and fine hair, which exposes its characteristic skin folds. The skin itself is thin and comes in two shades: deep black and a lighter slate or pale black. The

head is well-proportioned with a small forehead, featuring characteristic wrinkles in its skin. Its ears are broad and drooping, although some may also have the typical "roof-eave" shape. The snout is elongated, terminating in a knot that may exhibit depigmentation, another characteristic feature of the Lampiño breed. The tail can also appear straight and drooping (MAPA, 2016).

Manchado de Jabugo: Manchado de Jabugo is a native Andalusian subpopulation that originated in the towns of Jabugo and Cortegana, situated in the heart of the Sierra de Aracena and Picos de Aroche. This group consists of individuals with elongated bodies and coats that are either 'retinta' (a reddish-brown color) or 'jara' (a 'dirty white') with black spots. The spots have defined outlines on the 'retinta' coat but are less distinct on the 'jara' coat (Clemente López et al., 2007).

This traditional classification is supported by several studies of genetic diversity. An earlier study (Martínez et al., 2000) used microsatellite information and involved female pigs from Retinto, Negro Lampiño, Entrepelado, Torbiscal, Manchado de Jabugo, and Dorado Gaditano lines, along with two contrasting breeds, Duroc-Jersey and Chato Murciano. The findings of this research (Martínez et al., 2000) demonstrated significant intraracial diversity within the Iberian pig breed, structured into distinct lines and strains with clear genetic differentiation. The genetic definition of individuals from the Negro Lampiño, Torbiscal, and Manchado de Jabugo strains was particularly remarkable.

Furthermore, the study confirmed genetic distances between these subpopulations (Figure 8), revealing the closest genetic proximity between the Retinto and the Entrepelado strains, which was expected given the origin of the latter through crossing between individuals from the Retinto and Lampiño strains (Clemente et al., 2006).

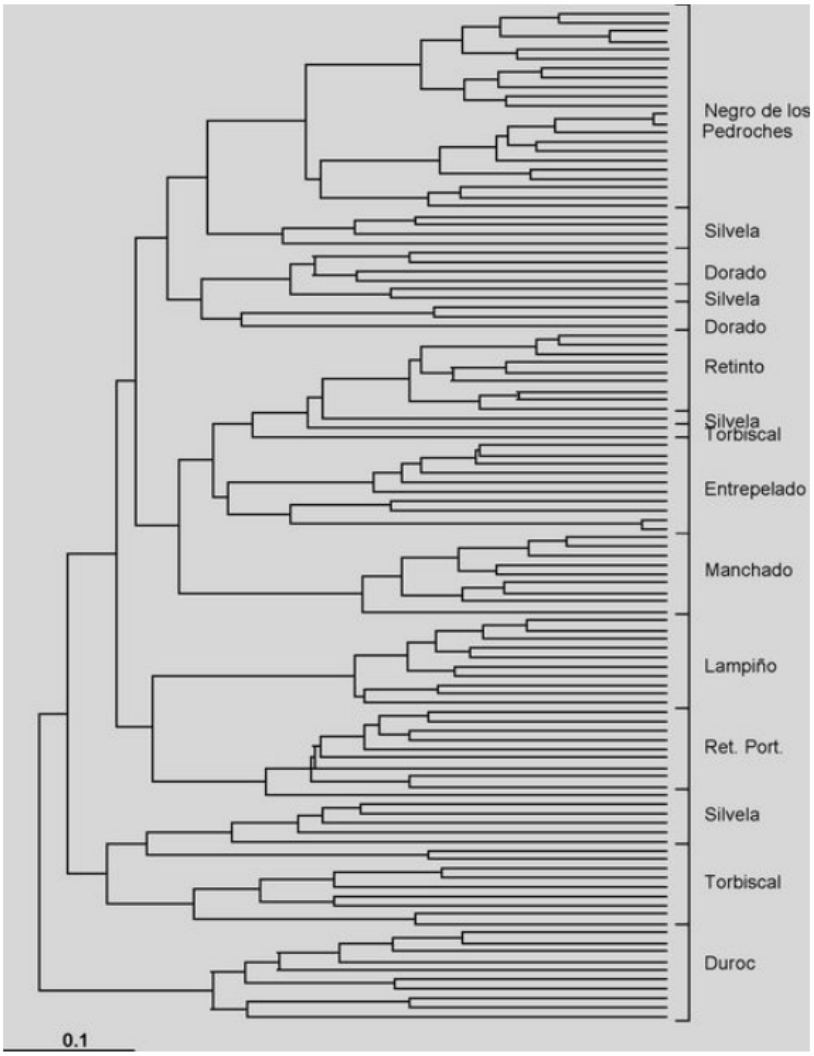


Figure 8: Unweighted pair group method with arithmetic mean (UPGMA) dendrogram for genetic distances of Iberian pig varieties and Duroc (Extracted from Martínez et al., 2000).

These results were confirmed by Fabuel et al. (2004) and by a recent study performed with high-density genotyping conducted by Alonso et al. (2020), they delve into the genomic differentiation among three primary strains of Iberian pigs, namely Entrepelado (EE), Retinto (RR), and Torbiscal (TT), using high-density genotyping methods. The study confirms the genetic closeness between Entrepelado and Retinto strains, corroborating earlier findings by Fabuel et al. (2004).

1.3.2. Phenotypic characterization of the Iberian pig

The genetic diversity across the different strains of the Iberian pig implies a great heterogeneity in the performance in reproductive, productive and carcass and meat quality traits.

Reproductive Traits

In general, the reproductive performance of the Iberian pig is substantially worse than the performance achieved by selected white pig populations (Silio et al., 2001). A summary of the reported performance in reproductive traits is presented in Table 1.

Table 1: Reproductive traits of the Iberian pig breed.

Trait	**Range or Average**	**References**
Age at first parturition	10.0–16.5 months	(Rodriguez et al., 1994)
Litters per year	2.2	(Leenhouwers & Merks, 2013)
Litter size	7.5 (6.0 to 8.3)	(Fernández et al., 2008; García-Casco et al., 2012)
Piglet birth weight	1.1–1.4 kg	(Gómez-Carballar et al., 2009, 2013)
Stillborn percentage	1.7–20.6%	(Perez-Enciso & Gianola, 1992; Saura et al., 2015)
Mortality rate until weaning	2.5–22.9%	(Cebrián et al., 2009; Leenhouwers & Merks, 2013)
Duration of lactation	Up to 60 days (39 days on average)	(Benito-Hernández et al., 1997; Gómez-Carballar et al., 2009)
Farrowing interval	Approximately 173 days	(Cebrián et al., 2009; Leenhouwers & Merks, 2013)
Weaning weight	6.9–20.8 kg	(Aguinaga et al., 2011; Gómez-Carballar et al., 2009)
Trend in lactation duration	Reducing to 25–26 days	(Romero et al., 2016)

Nevertheless, it must be noted that the data presented so far predominantly involves purebred varieties of the Iberian pig. However, scientific literature also provides

compelling evidence for heterosis, or hybrid vigor, when such purebred lines are crossed (Falconer & Mackay, 1996). The influence of heterosis in reproductive traits of the Iberian pig population has been demonstrated by García-Casco et al. (2012). Moreover, a recent study (Noguera et al., 2019) analyzed a diallel cross (Eisen et al., 1983) between three Iberian pig populations (Retinto, Entrepelado and Torbiscal). The results of this study indicated that the effect of heterosis in litter size was over 0.5 piglets in both Total Number Born and Number Born Alive.

Growth traits

The growth performance stages are categorized as lactation, growing stage (from weaning to approximately 30 kg), and early, middle, and late fattening stages, corresponding to approximately 30-60 kg, 60-100 kg, and above 100 kg of live body weight, respectively. A summary of the reported performance in growth traits in the literature is presented in Table 2.

Table 2: Growth performance traits of the Iberian pig breed.

Growth Stage	**Average Daily Gain (g/day)**	**Weight Range (kg)**	**References**
Lactation	257 (168-371)	-	(Aguinaga et al., 2011)
Growing	185-524	Up to 30	(Conde-Aguilera et al., 2011)
Early Fattening	228-566	30-60	(Fernández-Fígares et al., 2008)
Middle Fattening	181-800	60-100	(Barea et al., 2006)
Late Fattening	387-1018	Above 100	(Rey & López-Bote, 2001)
Overall Fattening	181-800	-	(Seiquer et al., 2018)

Table 2 also delineates the distinct growth stages of Iberian pigs. During the lactation stage, the average daily gain is between 168 to 371 g/day, comparable to modern white sows, though the lactation period in Iberian pigs is significantly elongated, averaging around 39 days. In the subsequent growing and fattening stages, there is evident heterogeneity in the daily gain, ranging from 185 to 524 g/day in the growing stage, 228 to 566 g/day in the early fattening stage, 181 to 800 g/day in the middle fattening stage,

and 387 to 1018 g/day in the late fattening stage, with the overall fattening stage also ranging between 181 to 800 g/day. These numbers indicate the maximum growth potential when animals are fed freely or almost freely, marking 524 g/day in the growing stage, 800 g/day in the overall fattening stage, and 1018 g/day post 128 kg. The extensive variability in daily gains through these stages can be attributed to diverse production systems and nutritional regimens, depicting an encompassing overview of the Iberian pig breed's growth performance traits.

Carcass and Meat Quality.

Carcass traits are central to the understanding of the Iberian pig's production characteristics. Table 3 offers insights into various carcass traits based on a compilation of numerous studies. These studies have been diverse in nature, with some emulating the practical conditions prevalent in the Iberian pig rearing systems, while others focus on specific performance and carcass composition parameters throughout different growth stages. A notable observation from the data is the deviation from the standard commercial slaughter weight typical for this breed. Additionally, the variability in traits such as back fat thickness, muscularity, and muscle thickness indicate the Iberian pig's unique propensity for fat deposition and its contrast with lean meat content when compared to conventional pig breeds (Almeida et al., 2019). The emphasis on premium cuts in the studies further highlights this inclination.

Table 3: Carcass traits of the Iberian pig breed

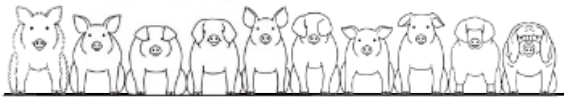
Parameter	Average Value	Range	References
Final Body Weight (kg)	170	160 to 190	(Dobao et al., 1987; Benito et al., 2000 ; Daza et al., 2005; Barea et al., 2006)
Slaughter Age (days)	407	14 to 15 months	(Ayuso et al., 2013; Daza et al., 2006; Dunker et al., 2007)
Back Fat Thickness (mm) - Withers	85	35-90	(García-Valverde et al., 2008; Nieto et al., 2015; Tejerina & García-Torres, 2010)
Back Fat Thickness (mm) - Last Rib	58	10-90	(Ayuso et al., 2015; Ibáñez-Escriche et al., 2016; Mayoral et al., 1999)
Muscularity - Loin Eye Area (cm2)	7,9	5.8 – 9.6	(Daza et al., 2006; Nieto et al., 2003)
Muscle Thickness - Cranial Edge of Gluteus Medius (mm)	14	11.6-16.5	(Martinez-Macipe et al., 2016; Serrano et al., 2008)

Meat quality is a complex and multifaceted parameter that encompasses various traits that define the taste, nutritional value, and sensory characteristics of the final product. In Table 4, an in-depth overview of meat and fat quality characteristics is provided, with a specific focus on the longissimus muscle and back fat tissue. The studies collated in this table highlight the pronounced variability in meat quality, with intramuscular fat content showing a wide range, indicative of the diversity present in meat attributes. The fatty acid profile of the meat is significantly influenced by factors such as diet and management practices, which further emphasize the importance of proper management strategies for optimal meat quality (Parrini et al., 2023; Serra et al., 1998).

Table 4: Meat quality of the Iberian pig breed.

Parameter	**Average Value**	**Range**	**References**
Postmortem pH - 45 min (longissimus muscle)	6.455	6.29–6.62	(Martinez-Macipe et al., 2016; Serra et al., 1998)
Postmortem pH - 24 hours (longissimus muscle)	5.68	5.61–5.75	(Martinez-Macipe et al., 2016; García-Torres, & Cava, 2012)
Intramuscular Fat Content (%)	6.9	3.0 – 9.8	(Daza et al., 2007, 2008; Dunker et al., 2007; Tejerina et al., 2012)
Color (L, a, b)	L: 44 a*: 11.5 b*: 6.6	L: 34 –54, a*: 7.5 –14.8 b*: 1.7 –13.6	(Cava et al., 2003; Estévez et al., 2003; Muriel et al., 2004)
Intramuscular Fat (SFA, MUFA, PUFA % in longissimus muscle)	SFA: 38 MUFA: 56 PUFA :7	31.3% to 46.3% 42.8% to 58.4% 3.5% to 8.7%	(Cava et al., 2003; Muriel et al., 2004; Tejerina et al., 2012)

SFA: Saturated Fatty Acids; MUFA: Monounsaturated Fatty Acids; PUFA: Polyunsaturated Fatty Acids.



CHAPTER II

Genetic improvement of Iberian pigs

The Iberian pig is well-adapted to the "Dehesa" environment in southwestern Spain, characterized by a savannah landscape composed of grass, cork, and holm oaks with seasonal production. Traditionally, Iberian pig production was dominated by purebred varieties and extensive management practices. The genetic improvement was performed empirically, leading to the varieties or subpopulations described in the previous chapter. However, in the recent decades, the Iberian pig production has been split in several productive strategies that ranges between purebred production with traditional extensive methods to crossbreeding with Duroc sires and intensive farming practices.

1. Genetic Improvement in Purebred Iberian Pigs

The heterogeneity of production methods strongly conditioned the genetic improvement program conducted by AECERIBER (*Asociación Española de Criadores de Cerdo Ibérico*). The overall objective of the breeding program for the Iberian Pig should take into account the existence of highly differentiated productive modalities. The specific objectives of selection are:

- Reproductive performance (prolificacy and maternal aptitude -milk production-)
- Early growth (weaning weight and weight at 75 days)
- Carcass quality (percentage of premium cuts: ham, loin and shoulder)
- Meat Quality (percentage of intramuscular fat)

To achieve these objectives, the breeding program designed by AECERIBER used three indexes: maternal, piglet and full cycle. The "Maternal Index" is focused on improving the reproductive performance and it implies the genetic evaluation for three traits: Number of Piglet Born Alive (NBA), Number of Weaned Piglets (NWP) and Litter Weight (LW). The model for genetic evaluation for the three traits is:

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Zu} + \boldsymbol{Wc} + \boldsymbol{e}$$

where $\boldsymbol{y}$ is the vector of phenotypic records, $\boldsymbol{b}$ is the vector of systematic effects that includes the general mean plus age of the dam and herd-year-season, $\boldsymbol{u}$ is the vector of the additive genetic effects, $\boldsymbol{c}$ is the vector of litter effects and $\boldsymbol{e}$ is the vector of residuals. Further, $\boldsymbol{X}$, $\boldsymbol{Z}$ and $\boldsymbol{W}$ are the corresponding incidence matrices. The assumed heritabilities were 0.07, 0.08 and 0.16 for NBA, NWP and LW, respectively. Diving deeper into the

heritability for NBA in Iberian pig populations, García-Casco et al. (2012) provided a heritability estimate of 0.06. In alignment, Fernández et al. (2008) reported a value of 0.09 for the same trait. Concurrently, the same research determined the heritability for the Number of Weaned Piglets (NWP) to range between 0.10 and 0.20. As for the Litter Weight (LW), the ascertained heritability spanned between 0.12 and 0.25 (Fernández-Fígares et al., 2008).

The genetic evaluation of the three traits is performed by a univariate BLUP and the prediction of genetic indexes is weighted as $0.5 \times NBA + 0.25 \times NWP + 0.25 \times LW$. The LW trait is not recorded in all farms. Therefore, the "maternal index" for farms without that index is $0.5 \times NBA + 0.25 \times NWP$.

Regarding the "Piglet Index" the AECERIBER breeding program focuses on enhancing early growth by emphasizing the genetic evaluation of two primary traits: Piglet Weight (PW45) and Piglet Weight at 75 days (PL75). Fernández et al. (2008) provided insights into the heritability of litter weight at 21 days, revealing a range of 0.10 to 0.15. Delving deeper into the growth trajectory, both PW45 and PL75 exhibit heritability estimates ranging from 0.15 to 0.30. This suggests a moderate to high genetic influence, which becomes increasingly evident as piglets age.

The model for this genetic evaluation used by AECERIBER is:

$$\boldsymbol{y = Xb + Zu + Wc + e}$$

Where $\boldsymbol{y}$ represents the vector of phenotypic records for Piglet Weight (PW), $\boldsymbol{b}$ denotes systematic effects that include the general mean, the gender of the piglet, and a fixed effect accounting for seasonal influences. The vector $\boldsymbol{u}$ accounts for the additive genetic effects of the piglet, while $\boldsymbol{c}$ captures the environmental effects shared among piglets from the same litter, and $\boldsymbol{e}$ signifies the vector of residuals. As before, $\boldsymbol{X}$, $\boldsymbol{Z}$ and $\boldsymbol{W}$ are the corresponding incidence matrices.

The genetic evaluation is performed through a univariate BLUP approach. Finally, the piglet indexes are formulated as $0.5 \times PW45 + 0.5 \times PW75$, providing breeders with insights into early growth potential.

The "Full Cycle Index", is also a part of the breeding program for the Iberian pig breed, prioritizes the enhancement of premium cuts (hams, shoulders, and loins) which account

for 65% of the carcass value. The selection criterion focuses on a combined index that encompasses the percentages of the weights of hams and shoulders relative to the carcass weight (post fat trimming and prior to salting) as well as the percentage of loins free from fat. Studies, notably by Fernández et al. (2003), underscore the high heritability of attributes such as weight and yield of noble cuts, pinpointing a heritability range of 0.40 to 0.50. This observation is consistent with findings by Xie et al. (2023), for other pig breeds including Landrace, Yorkshire, and their crossbreeds where they find moderate to high heritability's for 12 distinct carcass traits and six meat quality traits across these four pig populations.

AECERIBER's extensive genetic parameter records, particularly those related to premium cut yield, will be used to perform genetic evaluations using a BLUP-Animal Model based on a multitrait model. The four traits evaluated are Ham Percentage on the Carcass (HP), Shoulder Percentage (SP), Loin Percentage (LP), and Intramuscular Fat Percentage (%IMF).

$$\boldsymbol{y_i = X_i b_i + Z_i u_i + e_i}$$

In this model, $\boldsymbol{y_i}$ (with **i** ranging from 1 to 4) represents records for characteristics: Ham Percentage on the Carcass (HP), Shoulder Percentage (SP), Loin Percentage (LP), and Intramuscular Fat Percentage (%IMF), respectively. Further, $\boldsymbol{b_i}$ are the vectors of systematic effects, that include the mean, sex effects, covariables of carcass weight (for HP, SP, LP) and age at sacrifice (for %IMF). Furthermore, $\boldsymbol{u_i}$ correspond to the vectors of additive genetic effects and $\boldsymbol{e_i}$ are the vectors of random residuals.

The predictions of breeding values for HP, SP, and LP are subsequently integrated to formulate the "Carcass Yield Index" (CYI_i):

$$CYI_i = 1/3HP_i + 1/3SP_i + 1/3LP_i$$

Finally, the prediction of breeding values for the percentage of Intramuscular Fat (%IMF) are directly interpreted as the "Carcass Quality Index" (CQI_i).

A summary of the genetic indexes using for the genetic improvement of purebred Iberian pig in the AECERIBER breeding program are summarized in the Table 5.

Table 5: Selection objectives, selection criteria, traits under study and statistical models used in the breeding program in Iberian pig.

Selection objectives	Selection criteria	Traits	Statistical models
Prolificacy and maternal ability	Maternal Index	NBA, NWP, LW	ST-BLUP
Growth rate at early months of age	Piglet Index	PW45, PL75	ST-BLUP
Carcass quality as noble cuts yield	Complete Cycle Index	HP, SP, LP, %IMF	MT-BLUP

NBA: the Number of Piglets Born Alive; NWP: Number of Weaned Piglets; LW: Litter Weight; PW45: piglet Weight at 45 days; L75: piglet Weight at 75 days; HP: Ham percentage on the carcass; SP; shoulder percentage; LP: loin percentage; %IMF: intramuscular fat percentage; ST-BLUP: single trait animal model Best Linear Unbiased Prediction; MT-BLUP: multi trait animal model Best linear Unbiased Prediction.

2. Genetic Improvement in Crossbred Iberian Pigs

The breeding program implemented by AECERIBER focuses on purebred Iberian pigs. However, the breeding programs of the white pigs are mainly center around a three-way cross, involving three populations are involved (see Figure 9).

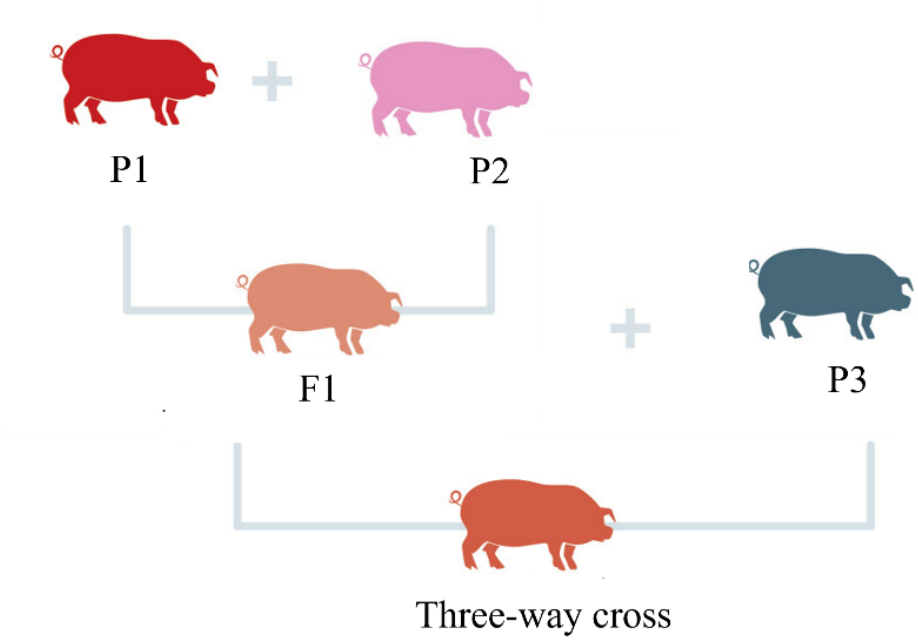


Figure 9: Three Way Cross for pigs.

Two of these populations generate the hybrid sow with the aim of achieving heterosis, and the sows are later mated with sires from the third or paternal population that provided complementarity. Following the same strategy, the INGA FOOD S. A. company and IRTA started a research program to identify the best hybrid sow between the strains of Iberian pig. Over the years, the core of this research has been anchored on two fundamental pillars. The initial step was to perform a diallel cross between three strains of Iberian pig (Retinto, Entrepelado and Torbiscal), focusing on their genetic-productive attributes (see Figure 10). The main aim was to harness the benefits of heterosis (hybrid vigor) and their mutual complementarity across various traits, encompassing reproductive, productive, carcass quality, and meat quality to identify the best cross in terms of productive and reproductive characteristics.

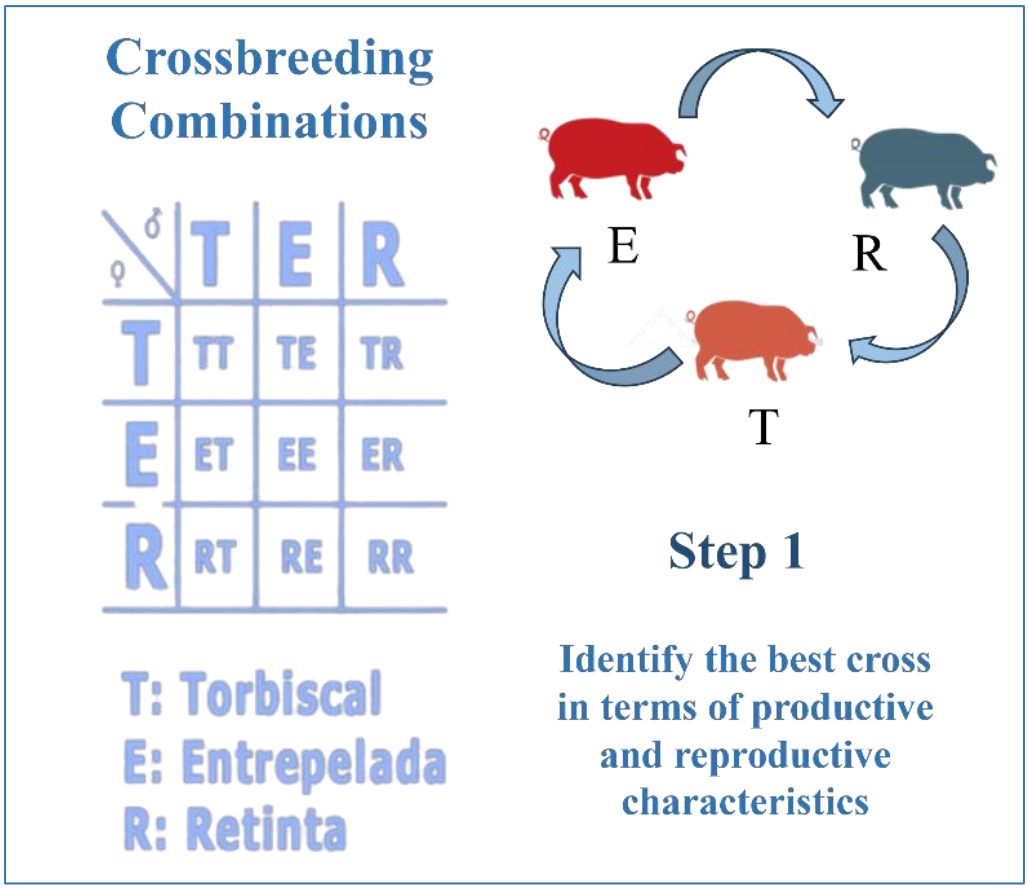


Figure 10: Diallel Crossbreeding Analysis of Iberian Pig Strains.

The results of the study suggest that the cross between Retinto as sire and Entrepelado as dam provided a clear advantage in terms of litter size (Noguera et al., 2019; Varona et al., 2020) and intramuscular fat (Ibañez-Escriche et al., 2016). Based on these findings, INGA FOOD S.A. initiated a breeding program within their private Retinto and Entrepelado populations, aiming to enhance the performance of their hybrid sow, known as CASTUA, that will be mated later with Duroc boars (Noguera & Ibáñez-Escriche, 2017) (refer to figure 11).

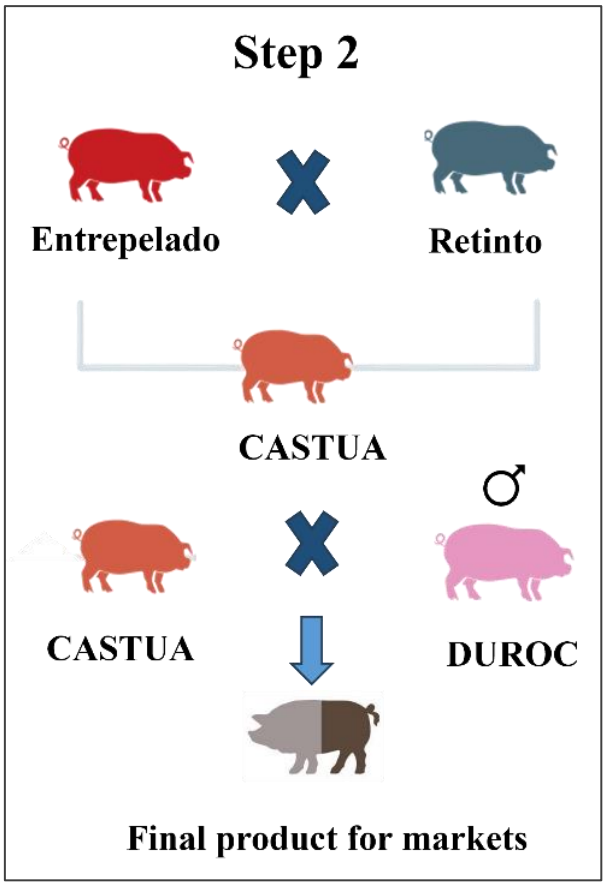
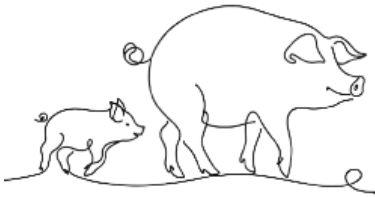


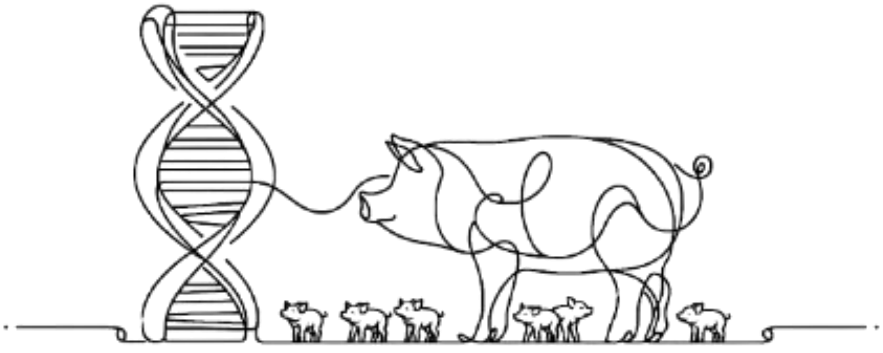
Figure 11: Crossbreeding design with the CASTUA sows.

Once the crossbreeding (Retinto x Entrepelado) was determined, the CASTUA breeding program was designed to improve the reproductive and productive traits in the crossbred sows and final products. Therefore, the parental populations, Retinto and Entrepelado, are evaluated using two primary genetics indices. The "Maternal Index" assesses Litter Size, Variability of Litter Size, and Litter Weight, while the "Paternal Index" focuses on Backfat Thickness, Weight at 180 days, and Feed Efficiency for boars. The statistical models were equivalent to the ones described for the purebred breeding programs of AECERIBER. It is important to note that, nowadays, these genetics evaluations are exclusively carried out within pure lines using purebred phenotypic information.



CHAPTER III

New challenges in the genetic improvement in litter size in crossbred Iberian sows



The breeding program for crossbred Iberian sows developed by the company INGA FOOD S. A. involves the selection of purebred individuals and crossbreeding to generate an 100% Iberian hybrid sow known as CASTUA. The CASTUA sows exhibit a significant increase in litter size thanks to the heterosis effects, as reported by Noguera et al. (2019). However, this increase is still considerably lower when compared to other commercial pig breeds (Silio et al., 2001). Additionally, the study of Noguera et al. (2019) reveals a noteworthy difference in litter size between the two reciprocal crosses (Entrepelado x Retinto and Retinto x Entrepelado).

The breeding program places particular emphasis on litter size due to its crucial impact of the economic efficiency in intensive farms when crossing with Duroc sires. However, there are several challenges that need to be addressed to achieve a higher response to selection in litter size in the crossbred populations. Among then, the studies developed in this thesis will focus on:

- Assessing the predictive ability of breeding values in purebred populations to enhance crossbred performance through the estimation of the genetic correlation between purebred and crossbred performance.
- Exploring the causes of the performance differences between the reciprocal crosses Entrepelado x Retinto and Retinto x Entrepelado and their consequences in the expected response to selection.

1. Genetic correlation between purebred and crossbred performance

The breeding program implemented by INGA FOOD S. A. is centered around the genetic prediction of the breeding values in purebred individuals to enhance the crossbred performance in the CASTUA sow. This strategy is built on the assumption of a high and positive additive genetic correlation (r_{pc}) between purebred and crossbred performance. The genetic correlation (r_{pc}) between purebred and crossbred performance is an important parameter, as the response to selection in crossbred performance depends on its value when selection is based on purebred performance. However, if the genetic correlation between purebred and crossbred performance is low, relying on purebred performance might not accurately reflect the crossbred performance. The reasons of a low genetic correlation may stem from differences in the allele frequencies between parental lines

when non-additive genetic effects are relevant (Duenk et al., 2021), or from genotype x environmental interactions under varying environmental conditions (Dekkers, 2007; Esfandyari et al., 2015; Ibáñez-Escriche et al., 2009). The literature offers mixed results of the estimation of r_{pc} (Wientjes & Calus, 2017), but it has been never explored within the context of the Iberian pig or within the cross involved in the generation of the CASTUA sow.

A straightforward method for integrating performance from purebred animals (PB) with information from crossbred relatives (CB), involves treating PB and CB performance as expressions of distinct traits with a genetic correlation between them (r_{pc}). This approach is based on the idea that the genetic correlation between PB and CB performance can be estimated and used to predict the performance of CB animals from the performance of PB animals (Esfandyari et al., 2015; González-Diéguez et al., 2020; Jiang & Groen, 1999). The statistical model, developed by Wei & Van Der Werf (1994), describes this approach. To illustrate, consider a cross (C) between two purebred populations (A and B). Therefore, the multiple trait model of analysis would be as follows:

$$\begin{pmatrix} \boldsymbol{y_A} \\ \boldsymbol{y_B} \\ \boldsymbol{y_C} \end{pmatrix} = \begin{pmatrix} \boldsymbol{X_A} & 0 & 0 \\ 0 & \boldsymbol{X_B} & 0 \\ 0 & 0 & \boldsymbol{X_C} \end{pmatrix} \begin{pmatrix} \boldsymbol{b_A} \\ \boldsymbol{b_B} \\ \boldsymbol{b_C} \end{pmatrix} + \begin{pmatrix} \boldsymbol{Z_A} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{Z_B} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Z_{AC}} & \boldsymbol{0} & \boldsymbol{Z_{BC}} \end{pmatrix} \begin{pmatrix} \boldsymbol{u_A} \\ \boldsymbol{c_{AC}} \\ \boldsymbol{u_B} \\ \boldsymbol{c_{BC}} \end{pmatrix} + \begin{pmatrix} \boldsymbol{e_A} \\ \boldsymbol{e_B} \\ \boldsymbol{e_C} \end{pmatrix}$$

where the vectors $\boldsymbol{y_A}$, $\boldsymbol{y_B}$ and $\boldsymbol{y_C}$ contain phenotypic records from breeds A, B and crossbred (C) animals, respectively; $\boldsymbol{b_A}$, $\boldsymbol{b_B}$ and $\boldsymbol{b_C}$ are vectors of fixed effects and $\boldsymbol{e_A}$, $\boldsymbol{e_B}$ and $\boldsymbol{e_C}$ are the vectors of residuals for A, B and C populations, respectively; $\boldsymbol{u_A}$ and $\boldsymbol{u_B}$ are vectors of additive genetic effects for the A and B populations; $\boldsymbol{c_{AC}}$ and $\boldsymbol{c_{BC}}$ are the vectors of additive genetic effects of animals in the C population originating from populations A and B, respectively. Further, $\boldsymbol{X_A}, \boldsymbol{X_B}, \boldsymbol{X_C}, \boldsymbol{Z_A}, \boldsymbol{Z_B}, \boldsymbol{Z_{AC}}$ and $\boldsymbol{Z_{BC}}$ are incidence matrices that relates phenotypic records with fixed and random effects, respectively. The total breeding value of crossbred $\boldsymbol{u_C}$ includes $\boldsymbol{c_{AC}}$ and $\boldsymbol{c_{BC}}$ plus the mendelian sampling $\boldsymbol{(\phi_C)}$. In this model, the mendelian samplings are included in the residuals ($\boldsymbol{e_C}$).

The variances of the additive genetic effects are:

$$Var \begin{pmatrix} \boldsymbol{u_A} \\ \boldsymbol{c_{AC}} \\ \boldsymbol{u_B} \\ \boldsymbol{c_{BC}} \end{pmatrix} = \begin{pmatrix} g_{AA}\boldsymbol{A_A} & g_{AC}\boldsymbol{A_A} & 0 & 0 \\ g_{AC}\boldsymbol{A_A} & g_{CCA}\boldsymbol{A_A} & 0 & 0 \\ 0 & 0 & g_{BB}\boldsymbol{A_B} & g_{BC}\boldsymbol{A_B} \\ 0 & 0 & g_{BC}\boldsymbol{A_B} & g_{CCB}\boldsymbol{A_B} \end{pmatrix}$$

where $\boldsymbol{A_A}$ and $\boldsymbol{A_B}$ are the numerator relationship matrices for populations A and B, respectively; g_{AA} and g_{BB} are the additive genetic variances for populations A and B; g_{CCA} and g_{CCB} are the additive genetic variances of population A and B alelles in the crossbred individuals. Finally, g_{AC} and g_{BC} are the genetic covariances between the purebred (A and B) and the crossbred (C).

Note that purebred individuals of populations A and B are not genetically related, and that the genetic connections between young candidates in the purebred with the crossbred populations are only collateral. These constrains of the pedigree information can be partially addressed by the advent of genomic selection procedures pioneered by Meuwissen et al. (2001), facilitated by the development of SNP chips. The initial approach for the genomic evaluation incorporating information from purebred and crossbred individuals was formulated by Ibáñez-Escriche et al. (2009), who defined an SNP based model assuming population specific substitution effects.

Later, Christensen et al. (2014) reformulated the Wei & Van Der Werf (1994) model by incorporating breed-specific partial relationships matrices (García-Cortés & Toro, 2006). The statistical model of analysis remains consistent with the description above. However, in this approach, the genetic covariances for populations A and B were defined as follows:

$$Var\begin{pmatrix}\boldsymbol{u_A}\\ \boldsymbol{u_{AC}}\\ \boldsymbol{c_A}\\ \boldsymbol{c_{AC}}\end{pmatrix} = \boldsymbol{\Sigma}^{(A)} \otimes \mathbf{A}^{(A)} \quad and \quad Var\begin{pmatrix}\boldsymbol{u_B}\\ \boldsymbol{u_{BC}}\\ \boldsymbol{c_B}\\ \boldsymbol{c_{BC}}\end{pmatrix} = \boldsymbol{\Sigma}^{(B)} \otimes \mathbf{A}^{(B)}$$

where $\boldsymbol{u_{AC}}$ and $\boldsymbol{u_{BC}}$ are artificial random vectors that represent the additive genetic effects of the crossbred individuals in the purebred populations. Further, $\boldsymbol{c_A}$ and $\boldsymbol{c_B}$ are the additive genetic effects of purebred individuals for crossbreeding, $\boldsymbol{\Sigma}^{(A)}$ and $\boldsymbol{\Sigma}^{(B)}$ are the 2 x 2 (co)variances matrices for the purebred and crossbred performance in the A and B populations. Finally, $\mathbf{A}^{(A)}$ and $\mathbf{A}^{(B)}$ are the partial numerator relationship matrices (García-Cortés & Toro, 2006), that are calculated as:

$$\mathbf{A}^{(A)} = \begin{pmatrix}\boldsymbol{A_A} & 0.5\boldsymbol{A_A T'_{C,A}}\\ 0.5\boldsymbol{T_{C,A} A_A} & \boldsymbol{A_C^{(A)}}\end{pmatrix} \quad \mathbf{A}^{(B)} = \begin{pmatrix}\boldsymbol{A_B} & 0.5\boldsymbol{A_B T'_{C,B}}\\ 0.5\boldsymbol{T_{C,B} A_B} & \boldsymbol{A_C^{(B)}}\end{pmatrix}$$

where $\boldsymbol{A_A}$ and $\boldsymbol{A_B}$ are the numerator relationship matrix between the individuals of the populations A and B, respectively, $\boldsymbol{A_C^{(A)}}$ and $\boldsymbol{A_C^{(B)}}$ are the partial relationship matrices

between the crossbred individuals (C) with origin in the A and B populations, respectively. Finally, $\boldsymbol{T_{C,A}}$ and $\boldsymbol{T_{C,B}}$ are the matrices assign purebred parents to crossbred origin.

The genealogical partial relationship matrices ($\mathbf{A}^{(A)}$ and $\mathbf{A}^{(B)}$) can be replaced by to the marker based partial relationship matrix ($\mathbf{G}^{(A)}$ and $\mathbf{G}^{(B)}$) that are computed with an algotirhm similar to the one described by VanRaden (2008) for the genomic relationship matrix. In particular

$$\mathbf{G}^{(A)} = \begin{pmatrix} \boldsymbol{G}_{A,A}^{(A)} & \boldsymbol{G}_{A,C}^{(A)} \\ \boldsymbol{G}_{C,A}^{(A)} & \boldsymbol{G}_{C,C}^{(A)} \end{pmatrix} \text{ and } \mathbf{G}^{(B)} = \begin{pmatrix} \boldsymbol{G}_{B,B}^{(B)} & \boldsymbol{G}_{B,C}^{(B)} \\ \boldsymbol{G}_{C,B}^{(B)} & \boldsymbol{G}_{C,C}^{(B)} \end{pmatrix}$$

which are calculated from the SNP marker information as:

$$\boldsymbol{G}_{A,A}^{(A)} = \frac{(\boldsymbol{m^A} - (2\boldsymbol{p^A} - 1)\mathbf{1}')(\boldsymbol{m^A} - (2\boldsymbol{p^A} - 1)\mathbf{1}')'}{s^A}$$

$$\boldsymbol{G}_{A,C}^{(A)} = \frac{(\boldsymbol{m^A} - (2\boldsymbol{p^A} - 1)\mathbf{1}')(\boldsymbol{q^A} - (\boldsymbol{p^A} - 0.5)\mathbf{1}')'}{s^A}$$

$$\boldsymbol{G}_{C,C}^{(A)} = \frac{(\boldsymbol{q^A} - (\boldsymbol{p^A} - 0.5)\mathbf{1}')(\boldsymbol{q^A} - (\boldsymbol{p^A} - 0.5)\mathbf{1}')'}{s^A}$$

$$\boldsymbol{G}_{B,B}^{(B)} = \frac{(\boldsymbol{m^B} - (2\boldsymbol{p^B} - 1)\mathbf{1}')(\boldsymbol{m^B} - (2\boldsymbol{p^B} - 1)\mathbf{1}')'}{s^B}$$

$$\boldsymbol{G}_{B,C}^{(B)} = \frac{(\boldsymbol{m^B} - (2\boldsymbol{p^B} - 1)\mathbf{1}')(\boldsymbol{q^B} - (\boldsymbol{p^B} - 0.5)\mathbf{1}')'}{s^B}$$

$$\boldsymbol{G}_{C,C}^{(B)} = \frac{(\boldsymbol{q^B} - (\boldsymbol{p^B} - 0.5)\mathbf{1}')(\boldsymbol{q^B} - (\boldsymbol{p^B} - 0.5)\mathbf{1}')'}{s^B}$$

Where $\boldsymbol{m^A}$ and $\boldsymbol{m^B}$ are the marker genotype matrices for purebred individuals of the A and B population, s^A and s^B is a scaling parameter. They have elements m_{ij}^X= -1, 0 or 1 if the j^{th} SNP of the ith individual is 11, 12, or 22. The $\boldsymbol{q^A}$ and $\boldsymbol{q^B}$ are the marker allele matrices with elements q_{ij}^X= -1,1 if the j^{th} SNP of the i^{th} individual received from the population X is 1 or 2, respectively. Finally, $\boldsymbol{p^A}$ and $\boldsymbol{p^B}$ are the allelic frequencies in the purebred populations A and B. It must be noted that the calculation of $\boldsymbol{q^A}$ and $\boldsymbol{q^B}$ requires haplotype phasing in the crossbred individuals.

Nevertheless, Stuber & Cockerham (1966) demonstrated that gene substitution effects can be defined within populations or across populations. Furthermore, if all the non-additive effects are considered, both approaches are equivalent. The definition of the gene substitution effects across populations implies the use of "biological" effects linked to the

SNP markers rather than "statistical" effects. However, this necessitates transforming these "biological" effects into "statistical" effects for calculating additive genetic variances, covariances, and correlations.

The relationships between the "biological" and "statistical" definition for a bialelic gene or SNP with "biological" effects of a, d and -a for genotypes 11, 12 and 22 are illustrated in Table 6 for one population (A) with allelic frequency p_A for the 1 allele and q_A for the allele 2 and their crossbreds with another population (B) with allelic frequencies p_A and p_B. In the table, $\alpha_A = a + (1 - p_A)d$ and $\alpha_{A \to B} = a + (1 - p_B)d$ are the allelic substitution effects of the A population in the purebred (α_A) and in the crossbred ($\alpha_{A \to B}$) populations, respectively.

Table 6: "Biological" and "statistical" genetics effects for a biallelic gene with for one population (A) with allelic frequency p_A for the 1 allele and q_A for the allele 2 and their crossbreds with another population (B) with allelic frequencies p_A and p_B.

	Genotypes		
	1-1	1-2	2-2
Biological	a	d	-a
Statistical (Purebred)	$2q_A\alpha_A$	$(q_A - p_A)\alpha_A$	$-2p_A\alpha_A$
Statistical (Crossbred)	$2q_A\alpha_{A \to B}$	$(q_A - p_A)\alpha_{A \to B}$	$-2p_A\alpha_{A \to B}$

a: additive genetic effect; d: dominance effect; p: The frequency of allele 1 in the population; q: The frequency of allele 2 in the population.

Computationally, the "biological" approach does not require the knowledge of the allele origin in the crossbred populations, but involves the inclusion of non-additive effects, such as the dominance genetic variation.

2. Dominance Genetic Variation

Pig litter size is a trait characterized by low heritability (Bidanel, 2011).Theoretically, the genetic variability in traits with low heritability traits is expected to be more associated with non-additive mode of inheritance (Falconer & Mackay, 1996). Dominance is a non-additive genetic effect that refers to the interaction between alleles at the same locus (Falconer & Mackay, 1996), where one allele masks the expression of the other. The impact of dominance and other non-additive effects on hybrid performance, such as heterosis or hybrid vigor, cannot be overlooked in animal breeding programs. Moreover, knowledge of non-additive genetic variation can be valuable for the prediction of future mating outcomes (Toro & Varona, 2010; Varona & Misztal, 1999).

The model for estimation of dominance genetic variance includes a dominance effect (***d***) into the standard mixed model equations, as follows:

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Zu} + \boldsymbol{Zd} + \boldsymbol{e}$$

where ***y*** is the vector of phenotypic records, ***b*** is the vector of systematic effects, ***u*** is the vector of breeding values (additive genetic effects), ***d*** are the dominance effects and ***e*** is the vector or residuals. The variances of the random effects under this model are:

$$Var\begin{pmatrix}\boldsymbol{u}\\ \boldsymbol{d}\\ \boldsymbol{e}\end{pmatrix} = \begin{pmatrix}\boldsymbol{A}\sigma_a^2 & 0 & 0\\ 0 & \boldsymbol{D}\sigma_d^2 & 0\\ 0 & 0 & \boldsymbol{I}\sigma_e^2\end{pmatrix}$$

where ***A*** represents the numerators relationship matrix, ***D*** is the dominance relationship matrix, and ***I*** is the identity matrix. Further, σ_a^2, σ_d^2 and σ_e^2 denote the additive genetic variance, the dominance genetic variance and the residual variance, respectively. The estimation of variance component and the prediction of breeding values require the calculation of the inverses of ***A*** and ***D***, achievable through the procedures outlined by Henderson (1976) and Hoeschele & VanRaden (1991), respectively.

Nevertheless, estimations of non-additive effects using genealogical and phenotypic information in livestock population have been limited (Misztal et al., 1998). This scarcity is attributed to the fact that the available information is primary derived from the resemblance between fullsibs or double cousins. Moreover, the calculations involved are complex and, for practical purposes, the statistical additive variance captures biological dominance or higher order interaction effects (Hill, 2010).

This landscape has undergone a profound transformation with the introduction of massive genotypic information provided by SNP chips, addressing some of the previously existing limitations. Following the study of Vitezica et al. (2013), the $\boldsymbol{A}$ and $\boldsymbol{D}$ matrices can be replaced by the additive ($\boldsymbol{A^G}$) and dominance ($\boldsymbol{D^G}$) genomic covariance matrices. The calculation of the $\boldsymbol{A^G}$ matrix follow the rules described by VanRaden (2008) as follows:

$$\boldsymbol{A^G} = \frac{\boldsymbol{TT'}}{\{tr[\boldsymbol{TT'}]/n\}}$$

where the $\boldsymbol{T}$ matrix is the incidence matrix composed by one vector for each individual. These vectors include values $(2\text{-}2p_i)$, $(1\text{-}2p_i)$ and $(\text{-}2p_i)$ for the SNP genotypes 11,12, and 22, respectively. Further, n is the number of SNP and p_i is the allelic frequency of the i^{th} SNP.

The calculation of the $\boldsymbol{D^G}$ linked with the "statistical" dominance deviations was described by Vitezica et al. (2013) and Nishio & Satoh (2015) as:

$$\boldsymbol{D^G} = \frac{\boldsymbol{WW'}}{\{tr[\boldsymbol{WW'}]/n\}}$$

where $\boldsymbol{W}$ is composed by individual vector with values $-2(1\text{-}p_i)^2$, $2p_i\,(1\text{-}p_i)$ and $-2p_i^2$ for the SNP genotypes 11, 12 and 22, respectively. Alternatively, the model can be also parameterized as the "biological" genotypic additive and dominant effects (Su et al., 2012). With this parameterization the $\boldsymbol{D^G}$ matrix is:

$$\boldsymbol{D^G} = \frac{\boldsymbol{HH'}}{\{tr[\boldsymbol{HH'}]/n\}}$$

here, $\boldsymbol{H}$ is composed by $-2p_i\,(1\text{-}p_i)$, $1\text{-}2p_i\,(1\text{-}p_i)$ and $-2p_i\,(1\text{-}p_i)$ for the SNP genotypes 11, 12 and 22. Vitezica et al. (2013) demonstrated the numerical equivalence between the "biological" and the "statistical" models. Moreover, the estimates of the "biological" genotypic model can be transformed to the "statistical" additive and dominant effects by using the vector of the allelic frequencies in the base generation of any population.

Therefore, the "biological" approach for dominance has been postulated to analyze data for purebred (A and B populations) and crossbred (C population) individuals (Vitezica et al., 2016) as follows:

$$\begin{pmatrix} \boldsymbol{y_A} \\ \boldsymbol{y_B} \\ \boldsymbol{y_C} \end{pmatrix} = \begin{pmatrix} \boldsymbol{X_A} & 0 & 0 \\ 0 & \boldsymbol{X_B} & 0 \\ 0 & 0 & \boldsymbol{X_C} \end{pmatrix} \begin{pmatrix} \boldsymbol{b_A} \\ \boldsymbol{b_B} \\ \boldsymbol{b_C} \end{pmatrix} + \begin{pmatrix} \boldsymbol{Z_A} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Z_B} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{Z_C} \end{pmatrix} \begin{pmatrix} \boldsymbol{u_A} \\ \boldsymbol{u_B} \\ \boldsymbol{u_C} \end{pmatrix}$$

$$+ \begin{pmatrix} \boldsymbol{Z_A} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Z_B} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{Z_C} \end{pmatrix} \begin{pmatrix} \boldsymbol{v_A} \\ \boldsymbol{v_B} \\ \boldsymbol{v_C} \end{pmatrix} + \begin{pmatrix} \boldsymbol{e_A} \\ \boldsymbol{e_B} \\ \boldsymbol{e_C} \end{pmatrix}$$

with variances:

$$var \begin{pmatrix} \boldsymbol{u_A} \\ \boldsymbol{u_B} \\ \boldsymbol{u_C} \end{pmatrix} = \boldsymbol{G_o} \otimes \boldsymbol{A^G}$$

$$var \begin{pmatrix} \boldsymbol{v_A} \\ \boldsymbol{v_B} \\ \boldsymbol{v_C} \end{pmatrix} = \boldsymbol{D_o} \otimes \boldsymbol{D^G}$$

Being $\boldsymbol{A^G}$ and $\boldsymbol{D^G}$ the "biological" additive ($\boldsymbol{A^G}$) and dominance ($\boldsymbol{D^G}$) genomic covariance matrices calculated with allelic frequencies p=q= 0.5 for all SNP markers and $\boldsymbol{G_o}$ and $\boldsymbol{D_o}$ are the genotypic additive and dominance covariance matrices as:

$$\boldsymbol{G_o} = \begin{bmatrix} \sigma^2_{U_A} & \sigma_{U_A U_B} & \sigma_{U_A U_C} \\ \sigma_{U_A U_B} & \sigma^2_{U_B} & \sigma_{U_B U_C} \\ \sigma_{U_A U_C} & \sigma_{U_B U_C} & \sigma^2_{U_C} \end{bmatrix} \text{ and } \boldsymbol{D_o} = \begin{bmatrix} \sigma^2_{V_A} & \sigma_{V_A V_B} & \sigma_{V_A V_C} \\ \sigma_{V_A V_B} & \sigma^2_{V_B} & \sigma_{V_B V_C} \\ \sigma_{V_A V_C} & \sigma_{V_B V_C} & \sigma^2_{V_C} \end{bmatrix}$$

Being $\sigma^2_{U_X}$ and $\sigma^2_{V_X}$ the additive and dominant genotypic variance for the X={A,B or C} populations, and $\sigma_{U_X U_Y}$ and $\sigma_{V_X V_Y}$ are the additive and dominant genotypic covariance between the X={A,B or C} and Y={A,B or C} populations. Given them, the additive and dominance variance components (σ^2_{aA} and σ^2_{dA}) associated with each SNP are:

$$\begin{bmatrix} \sigma^2_{aA} \\ \sigma^2_{aB} \\ \sigma^2_{aC} \end{bmatrix} = \begin{bmatrix} \frac{\sigma^2_{U_A}}{\{tr[\boldsymbol{TT'}]/n\}} \\ \frac{\sigma^2_{U_B}}{\{tr[\boldsymbol{TT'}]/n\}} \\ \frac{\sigma^2_{U_C}}{\{tr[\boldsymbol{TT'}]/n\}} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \sigma^2_{dA} \\ \sigma^2_{dB} \\ \sigma^2_{dC} \end{bmatrix} = \begin{bmatrix} \frac{\sigma^2_{V_A}}{\{tr[\boldsymbol{HH'}]/n\}} \\ \frac{\sigma^2_{V_B}}{\{tr[\boldsymbol{HH'}]/n\}} \\ \frac{\sigma^2_{V_C}}{\{tr[\boldsymbol{HH'}]/n\}} \end{bmatrix}$$

The additive ($\sigma^2_{A_X}$) and dominance ($\sigma^2_{D_X}$) genetic variances of the X={A, B} purebred population were:

$$\sigma^2_{A_X} = \sum_{i=1}^{n} 2p_{Xi}q_{Xi}\sigma^2_{a_X} + 2\,p_{Xi}q_{Xi}(q_{Xi} - p_{Xi})^2\sigma^2_{dX}$$

$$\sigma^2_{D_X} = \sum_{i=1}^{n} (2p_{Xi}\mathrm{q}_{Xi})^2\sigma^2_{dX}$$

Where p_{Xi} and q_{Xi} are the allelic frequencies for A_1 and A_2 at the i^{th} SNP marker and the X= {A,B} population. The estimates of the contributions to the additive variance in the crossbred population from the A ($\sigma^2_{A_{C(A)}}$) and B ($\sigma^2_{A_{C(B)}}$) populations were:

$$\sigma^2_{A_{C(A)}} = \sum_{i=1}^{n} 2p_{Bi}q_{Bi}\sigma^2_{a_C} + 2\,p_{Bi}q_{Bi}(q_{Ai} - p_{Ai})^2\sigma^2_{dC}$$

$$\sigma^2_{A_{C(B)}} = \sum_{i=1}^{n} 2p_{Ai}q_{Ai}\sigma^2_{a_C} + 2\,p_{Ai}q_{Ai}(q_{Bi} - p_{Bi})^2\sigma^2_{dC}$$

and, following Vitezica et al. (2016), the additive variance in the crossbred population ($\sigma^2_{A_C}$) was the average of these two values:

$$\sigma^2_{A_C} = \frac{1}{2}\sigma^2_{A_{C(A)}} + \frac{1}{2}\sigma^2_{A_{C(B)}}$$

the estimate of the dominance variance of the crossbred population ($\sigma^2_{D_C}$) was calculated as follows:

$$\sigma^2_{D_C} = \sum_{i=1}^{n} 4p_{Ai}\mathrm{q}_{Ai}p_{Bi}\mathrm{q}_{Bi}\sigma^2_{d_C}$$

and the covariance between purebred-crossbred additive genetic effects in the A ($\sigma_{A_A A_{C(A)}}$) and B ($\sigma_{A_B A_{C(B)}}$) populations were as follows:

$$\sigma_{A_A A_{C(A)}} = \sum_{i=1}^{n} 2p_{Bi}q_{Bi}\sigma_{a_A a_C} + 2\,p_{Bi}q_{Bi}(q_{Ai} - p_{Ai})^2\sigma_{d_A d_C}$$

$$\sigma_{A_B A_{C(B)}} = \sum_{i=1}^{n} 2p_{Ai}q_{Ai}\sigma_{a_B a_C} + 2\,p_{Ai}q_{Ai}(q_{Bi} - p_{Bi})^2\sigma_{d_B d_C}$$

with

$$\begin{bmatrix}\sigma_{a_A a_C}\\ \sigma_{a_B a_C}\end{bmatrix} = \begin{bmatrix}\frac{\sigma_{U_A U_C}}{\{tr[\boldsymbol{TT'}]/n\}}\\ \frac{\sigma_{U_B U_C}}{\{tr[\boldsymbol{TT'}]/n\}}\end{bmatrix} \quad \text{and} \quad \begin{bmatrix}\sigma_{d_A d_C}\\ \sigma_{d_B d_C}\end{bmatrix} = \begin{bmatrix}\frac{\sigma_{V_A V_C}}{\{tr[\boldsymbol{HH'}]/n\}}\\ \frac{\sigma_{V_B V_C}}{\{tr[\boldsymbol{HH'}]/n\}}\end{bmatrix}$$

From these estimates, the genetic correlations between the purebred and crossbreed breeding values in the A and B populations were:

$$r_{A_A A_{C(A)}} = \frac{\sigma_{A_A A_{C(A)}}}{\sqrt{\sigma^2_{A_A}\sigma^2_{A_{C(A)}}}} \quad \text{and} \quad r_{A_B A_{C(A)}} = \frac{\sigma_{A_B A_{C(B)}}}{\sqrt{\sigma^2_{A_B}\sigma^2_{A_{C(B)}}}}$$

3. Genomic imprinting and parent of origin effects.

In previous studies, the performance of the reciprocal crosses (Retinto x Entrepelado and Entrepelado x Retinto) exhibited significant differences in fat deposition traits (Ibáñez-Escriche et al., 2016) and litter size (Noguera et al., 2019). One potential explanation for this difference can be the presence of genomic imprinting (Barlow & Bartolomei, 2014; Reik & Walter, 2001), a phenomenon initially observed approximately 30 years ago through pronuclear transplantation experiments (Barton et al., 1984). Genomic imprinting is and epigenetic process that involved DNA methylation and other epigenetic marks (Kelsey & Feil, 2013), and it refers to the specific expression pattern of genes based on their parental origin (see Figure 12).

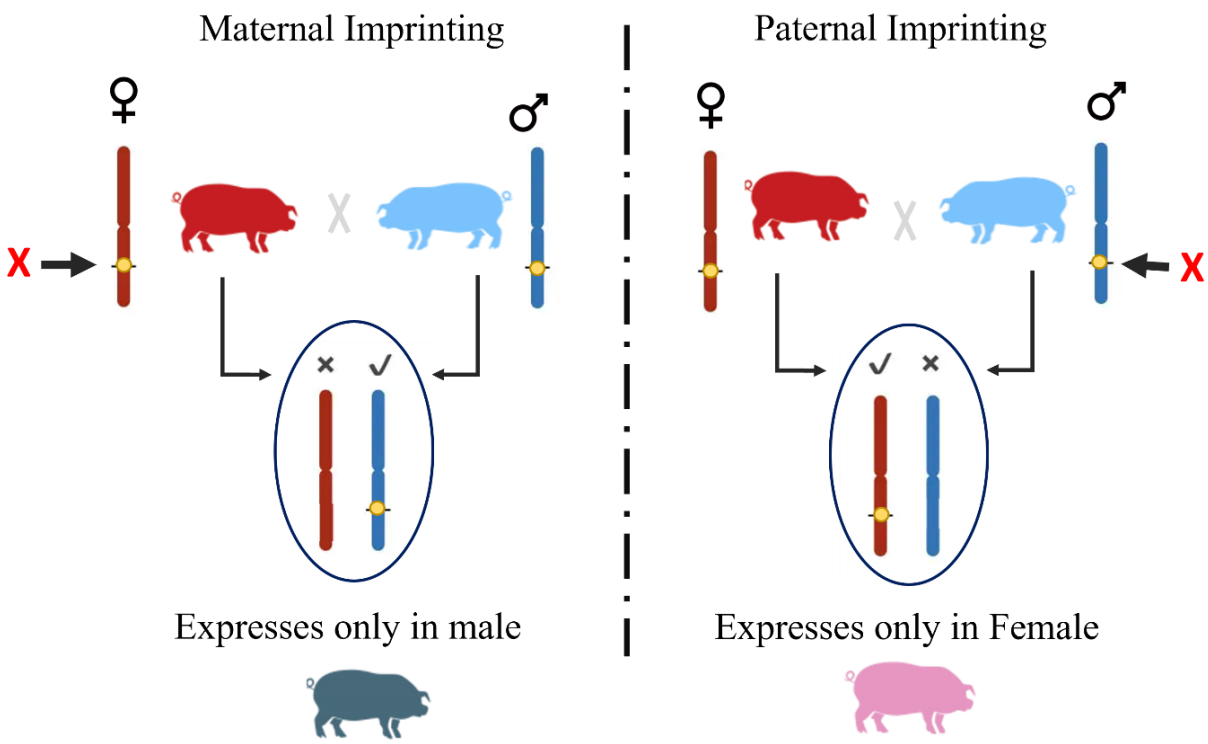


Figure 12: Paternal and Maternal imprinting.

The number of verified imprinted genes is increasing constantly (www.geneimpring.com). In humans, in November 2023, there are 129 verified imprinted genes, while mice have 150 imprinted genes, with an overlap of approximately 70% between the two species. In general, imprinting status is conserved among different species (Thorvaldsen & Bartolomei, 2007). Moreover, in human genetics, there are an increasing evidence of the association of imprinting genes in several disorders, such as Angelman and Prader-Willi syndromes (Mackay & Temple, 2017) and in mammalian development (Thamban et al., 2020).

In pigs, there 44 verified imprinted genes (www.geneimpring.com). In addition, there are several studies that have developed imprinted QTLs (De Koning et al., 2000; Lee et al., 2003; Nezer et al., 1999), and polymorphisms in some imprinted genes have been associated with economically relevant traits, such the IGF2 (*Insuline Growth Factor 2*) (Jungerius et al., 2004; Van Laere et al., 2003). IGF2 is a maternally imprinted gene which promotes growth and its over expression can lead to excessive growth and obesity, while under expression can lead to restricted growth and developmental problems (O'Doherty et al., 2015).

The consequence of genomic imprinting in the quantitative genetic analysis is the appearance of parent-of-origin effects, that has been reported in several traits of pigs (Neugebauer et al., 2010), cattle (Kenny et al., 2022; Varona et al., 2015), poultry (Triantaphyllopoulos et al., 2016) and horses (Perdomo-González et al., 2023).

The quantitative genetic analysis of parent-of-origin effects relies in the gametic model (Gibson, 1988), that was described as follows:

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Za} + \boldsymbol{Wg} + \boldsymbol{e}$$

where $\boldsymbol{y}$ is the vector of phenotypic data, $\boldsymbol{b}$ is the vector of fixed effects, $\boldsymbol{a}$ is the random additive genetic effects, $\boldsymbol{g}$ is the vector of paternal (or maternal) gametic effects and $\boldsymbol{e}$ is the vector of residuals. Moreover, $\boldsymbol{X}$, $\boldsymbol{Z}$ and $\boldsymbol{W}$ are the corresponding incidence matrices. The variance of random effects under this model are:

$$var\begin{pmatrix}\boldsymbol{a}\\ \boldsymbol{g}\\ \boldsymbol{e}\end{pmatrix} = \begin{pmatrix}\boldsymbol{A}\sigma_a^2 & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{P}\sigma_p^2 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}\sigma_e^2\end{pmatrix}$$

Here, $\boldsymbol{A}$ is the numerator relationship matrix, $\boldsymbol{P}$ is the gametic relationship matrix and $\boldsymbol{I}$ is the identity matrix. Further, σ_a^2 is the additive genetic variance, σ_p^2 is the paternal (or maternal) gametic variance and σ_e^2 is the residual variance. The mixed model equations associated with this model are:

$$\begin{pmatrix}\boldsymbol{X'X} & \boldsymbol{X'Z} & \boldsymbol{X'W}\\ \boldsymbol{Z'X} & \boldsymbol{Z'Z} + \boldsymbol{A^{-1}}\alpha_A & \boldsymbol{Z'W}\\ \boldsymbol{W'X} & \boldsymbol{W'Z} & \boldsymbol{W'W} + \boldsymbol{P^{-1}}\alpha_P\end{pmatrix}\begin{pmatrix}\widehat{\boldsymbol{b}}\\ \widehat{\boldsymbol{a}}\\ \widehat{\boldsymbol{g}}\end{pmatrix} = \begin{pmatrix}\boldsymbol{X'y}\\ \boldsymbol{Z'y}\\ \boldsymbol{W'y}\end{pmatrix}$$

where $\alpha_A = \sigma_e^2/\sigma_a^2$ and $\alpha_P = \sigma_e^2/\sigma_p^2$. The inverse of $\boldsymbol{A}$ is easily calculated by the procedure described by Henderson (1976), and the inverse of $\boldsymbol{P}$ can be obtained with the algorithm proposed by Schaeffer et al. (1989).

The above described model assumes that there is only a paternal (or maternal) gametic effect. However, the genetic variation of a quantitative trait can be due to some genes that act additively plus genes paternally imprinted and genes maternally imprinted. In this sense, Meyer & Tier (2012) develop a model that include two genetic effects (one paternal and one maternal).

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Z_p\alpha} + \boldsymbol{Z_m\beta} + \boldsymbol{e}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are the vector of the paternal and maternal genetic effects and $\boldsymbol{Z_p}$ and $\boldsymbol{Z_m}$ are the corresponding incidence matrices. The variance of random effects under this model are:

$$var\begin{pmatrix}\boldsymbol{\alpha}\\ \boldsymbol{\beta}\\ \boldsymbol{e}\end{pmatrix} = \begin{pmatrix}\boldsymbol{P}\sigma_\alpha^2 & \boldsymbol{P}\sigma_{\alpha\beta} & \boldsymbol{0}\\ \boldsymbol{P}\sigma_{\alpha\beta} & \boldsymbol{P}\sigma_\beta^2 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}\sigma_e^2\end{pmatrix}$$

Here, σ_α^2 and σ_β^2 are the paternal and maternal genetic variance, respectively, $\sigma_{\alpha\beta}$ is the covariance between the paternal and maternal genetic effects and σ_e^2 is the residual variance. It must be noted that the size of the $\boldsymbol{P}$ matrix is 2n x 2n, with n the number of animals in the pedigree. Moreover, the mixed model equations are:

$$\begin{pmatrix}\boldsymbol{X'X} & \boldsymbol{X'Z_p} & \boldsymbol{X'Z_p}\\ \boldsymbol{Z_p}'\boldsymbol{X} & \boldsymbol{Z_p}'\boldsymbol{Z_p} + \boldsymbol{P^{-1}}\gamma_1 & \boldsymbol{Z_p}'\boldsymbol{Z_m} + \boldsymbol{P^{-1}}\gamma_2\\ \boldsymbol{Z_m}'\boldsymbol{X} & \boldsymbol{Z_m}'\boldsymbol{Z_p} + \boldsymbol{P^{-1}}\gamma_2 & \boldsymbol{Z_m}'\boldsymbol{Z_m} + \boldsymbol{P^{-1}}\gamma_3\end{pmatrix}\begin{pmatrix}\widehat{\boldsymbol{b}}\\ \widehat{\boldsymbol{g_p}}\\ \widehat{\boldsymbol{g_m}}\end{pmatrix} = \begin{pmatrix}\boldsymbol{X'y}\\ \boldsymbol{Z_p}'\boldsymbol{y}\\ \boldsymbol{Z_m}'\boldsymbol{y}\end{pmatrix}$$

with

$$\begin{pmatrix}\gamma_1 & \gamma_2\\ \gamma_2 & \gamma_3\end{pmatrix} = \sigma_e^2\begin{pmatrix}\sigma_\alpha^2 & \sigma_{\alpha\beta}\\ \sigma_{\alpha\beta} & \sigma_\beta^2\end{pmatrix}^{-1}$$

Under this model, the total genetic variance is $\sigma_\alpha^2 + \sigma_\beta^2$, the imprinting variance (σ_i^2) is $\sigma_\alpha^2 + \sigma_\beta^2 - 2\sigma_{\alpha\beta}$, the mendelian or additive variance (σ_a^2) is $2\sigma_{\alpha\beta}$ and the paternal (σ_{gp}^2) and maternal $\left(\sigma_{gm}^2\right)$ imprinting variances are $\sigma_\alpha^2 - \sigma_{\alpha\beta}$ and $\sigma_\beta^2 - \sigma_{\alpha\beta}$, respectively.

This model can be also reparametrized as:

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Za} + \boldsymbol{Z_p g_p} + \boldsymbol{Z_m g_m} + \boldsymbol{e}$$

where $\boldsymbol{g_p}$ is the vector of paternal gametic (or imprinting) effects and $\boldsymbol{g_m}$ is the vector of maternal gametic (or imprinting) effects. Under this parameterization, the variances of random effects are:

$$var\begin{pmatrix}\boldsymbol{a}\\ \boldsymbol{g_p}\\ \boldsymbol{g_m}\\ \boldsymbol{e}\end{pmatrix} = \begin{pmatrix}\boldsymbol{A}\sigma_\alpha^2 & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{P}\sigma_{gp}^2 & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{P}\sigma_{gm}^2 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}\sigma_e^2\end{pmatrix}$$

The adaptation of these models to the use of genomic information from the SNP chips under a GBLUP approach were proposed by Nishio & Satoh (2015). These authors described to two alternative models (GBLUP-I1) and (GBLUP-I2) that include additive (**a**), dominance (**d**) and imprinting (**i**) models.

The first model (GBLUP-I1) is

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Z_a a} + \boldsymbol{Z_d d} + \boldsymbol{Z_i i} + \boldsymbol{e}$$

with $var(\boldsymbol{a}) = \boldsymbol{G_a}\sigma_a^2$, $var(\boldsymbol{d}) = \boldsymbol{G_d}\sigma_d^2$ and $var(\boldsymbol{i}) = \boldsymbol{G_i}\sigma_i^2$,

where $\boldsymbol{G_a}$ and $\boldsymbol{G_d}$ are the additive and dominance genomic relationship as described by VanRaden (2008) and Vitezica et al. (2013), respectively. The imprinting genomic relationship is calculated as:

$$\boldsymbol{G_i} = \frac{\boldsymbol{M_i M_i}'}{\sum_{\boldsymbol{j}}^{\boldsymbol{N_{snp}}} 2p_j(1-p_j)}$$

Being p_j the allelic frequency of the "1" allele at the j^{th} SNP and $\boldsymbol{M_i}$ is a matrix of number of individual x number of SNP whose values are 0 is the individual in homozygous (11 and 22), +1 if it has received the "1" allele from its father and the "2" allele from its mother and -1 if is have received the "2" allele from its father and the "1" allele from its mother.

The second model proposed by Nishio & Satoh (2015) (GBLUP-I2) is:

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Z_p \alpha} + \boldsymbol{Z_m \beta} + \boldsymbol{Z_d d} + \boldsymbol{e}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are the vector of paternal and maternal genetic effects following the same notation as Meyer & Tier (2012), and $\boldsymbol{Z_p}$ and $\boldsymbol{Z_m}$ are the corresponding incidence matrices. The variances of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $var(\boldsymbol{\alpha}) = \boldsymbol{G_\alpha}\sigma_\alpha^2$, $var(\boldsymbol{\beta}) = \boldsymbol{G_\beta}\sigma_\beta^2$ with

$$\boldsymbol{G_\alpha} = \frac{\boldsymbol{M_p M_p}'}{\sum_{\boldsymbol{j}}^{\boldsymbol{N_{snp}}} p_j(1-p_j)} \text{ and } \boldsymbol{G_\beta} = \frac{\boldsymbol{M_m M_m}'}{\sum_{\boldsymbol{j}}^{\boldsymbol{N_{snp}}} p_j(1-p_j)}$$

The $\boldsymbol{M_p}$ and $\boldsymbol{M_m}$ are matrices of number of individual x number of SNP whose values are (1-p_j) is the individual have received the "1" from its father (or mother) or (-p_j) if it has received the allele "2". Both models (GBLUP-I1) and (GBLUP-I2) are equivalent, and they require to know the paternal and maternal haplotype phases of the phenotyped individuals.

Therefore, the development of methods for unrevealing the parent of origin of alleles or haplotype phasing is required. Haplotype phasing is the process of inferring haplotypes from genotype data and involves determining the parental origin of alleles within an individual's genome (see Figure 13).

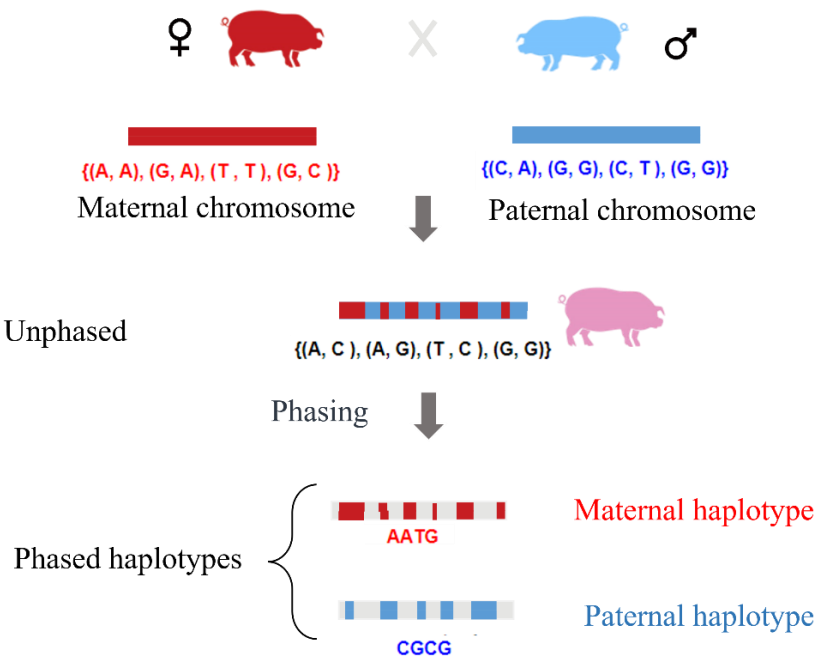


Figure 13: Haplotype-phasing.

Several algorithms and statistical models have been developed to infer haplotype phases from genotypic information, by using within family and across population information. Among them, it is worth to mention Alphaphase (Hickey et al., 2011), Beagle (Browning & Browning, 2008), FImpute (Sargolzaei et al., 2014), Impute2 (Howie et al., 2009), Findhap (VanRaden et al., 2013) and ShapeIt2 (O'Connell et al., 2014). Several studies comparing these software tools (Miar et al., 2017, Srihi et al., 2023) have demonstrated comparable performance among them. However, it is important to note that these tools operate with predefined blocks of SNP genotypes and require the specification of parameters such as "core length" and "tails length" (Figure 14). Adjusting these parameters can significantly impact the results (Srihi et al., 2023).

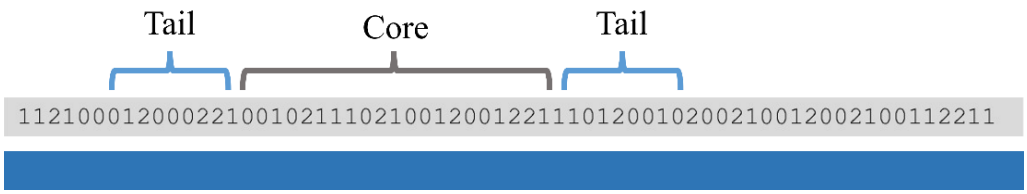
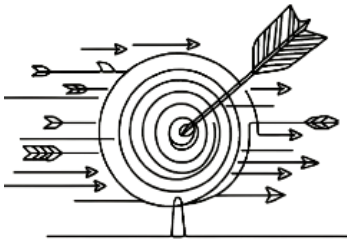


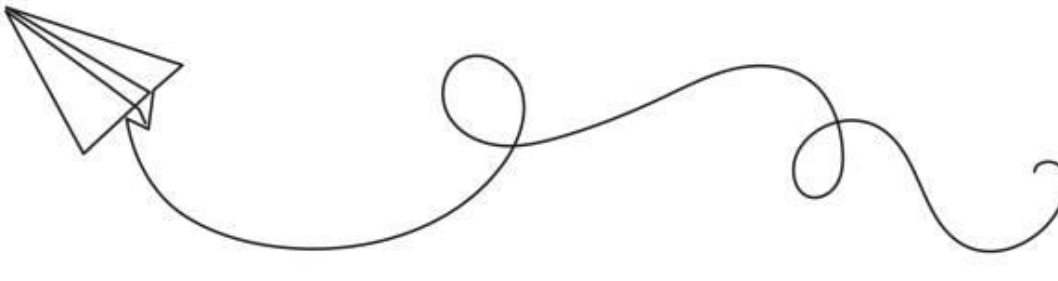
Figure 14: A core and its adjacent tails.

Determining core and tail lengths relies on factors such as SNP density, expected recombination rates, and the genetic data's structure. Short cores may lead to incomplete phasing, while overly extended cores may encompass multiple recombination events, resulting in inaccurate phasing. Similarly, the tail length should provide sufficient data to accurately identify surrogate parents without unnecessary complexity in the analysis.



CHAPTER IV

Objectives



The main objective of this thesis is to generate new information that may contribute to the improvement of the selection strategies for Iberian crossbreed pigs, with a focus on improving litter size due to its significant impact on economic efficiency in intensive farming.

This main objective aims to address two main challenges in the implementation of these selection strategies:

- The predictive ability of breeding values in the purebred populations to improve the crossbred performance.
- The causes of the performance differences between the reciprocal crosses (Entrepelado x Retinto) and (Retinto x Entrepelado) and its consequence in the expected response to selection.

The first topic (Chapter V) was addressed using the model developed by Vitezica et al. (2016) that allows to infer the genetic correlations between purebred and crossbred populations from a multiple trait analysis, by capturing the information provided by the genomic information from purebred and crossbred individuals.

The second topic was addressed with two different approaches. The first one (Chapter VI) involves the development a multi-trait generalization of the method described by Nishio & Satoh (2015) for the genomic analysis of gametic (or parent-of-origin) effects. However, as the available genomic information is scarce, the second one (Chapter VII) uses only phenotypic and genealogical information and involves the development of a multivariate gametic model that allows covariance between gametic effects within the same genetic origin (Entrepleado and Retinto), by incorporating purebred and crossbred genealogical and phenotypic information.

CHAPTER V

Additive and Dominance Genomic Analysis for Litter Size in Purebred and Crossbred Iberian Pigs

Authors

Srihi, Houssemeddine, José Luis Noguera, Victoria Topayan, Melani Martín de Hijas, Noelia Ibañez-Escriche, Joaquim Casellas, Marta Vázquez-Gómez, María Martínez-Castillero, Juan Pablo Rosas, and Luis Varona

The content of this chapter has been published in Genes 2022.
doi.org/10.3390/genes13010012

Abstract

(1) Background: INGA FOOD S. A. has produced a hybrid Iberian sow called CASTÚA by crossing the Retinto and Entrepelado varieties. Selection of parental populations is based on purebred information under the assumption that the genetic correlation between purebred and crossbred performance is high; however, these correlations can be less than one because of a GxE interaction or the presence of non-additive genetic effects. This study estimated the additive and dominant variances of the purebred and crossbred populations for litter size and calculated the additive genetic correlations between purebred and crossbred performances. (2) Methods: the data set consisted of 2,030 litter size data from the Entrepelado population, 1,977 litters from the Retinto, and 1,958 litters from the crossbred population. Individuals were genotyped with GeneSeek® GGP Porcine70K HD chip. The model of analysis was a 'biological' multivariate mixed model that included additive and dominant SNP effects. (3) Results: the additive genotypic variance of effects for total number born (TNB) ranged from 0.248 (Entrepelado) to 0.546 (Crossbred), and the dominance genotypic variance effects ranged from 0.170 (Retinto) to 0.265 (Crossbred). The genetic correlations between purebred and crossbred performance were 0.663 in Entrepelado and 0.881 in Retinto. After back solving to obtain estimates of the SNP effects, the additive genetic variance associated with genomic regions containing 30 SNPs was estimated and identified four genomic regions in the chromosomes (SSC) no. 6, 8 and 12; one region in SSC6, two regions in SSC8 and *one* region in SSC12 that each explained > 2% of the additive genetic variance.

Keywords: Pig; Iberian; Additive; Dominance; Genetic Correlation; Crossbreeding; Genomic Selection.

1. Introduction

The Iberian pig breed is one of the porcine populations that has the highest meat quality (Serra et al., 1998). Historically, Iberian pig production has been developed extensively with purebred varieties, which has taken advantage of the Dehesa environment in southwestern Spain. In recent decades, however, many traditional breeders have been replaced by intensive production systems that use crossbreeding with Duroc populations to improve growth and efficiency (Serrano et al. 2008). Normative that regulates Iberian pig products (Boletín Oficial del Estado, 2014) obligate the breeders when crossing

Iberian and Duroc varieties, to cross boars from Duroc and sows from Iberian variety. Prolificacy, which is lower than that of white pig populations, is the major limitation in the intensive production of crossbred pigs from Iberian dams (Silio et al., 2001). INGA FOOD, S.A. company has developed a crossbreeding scheme between two Iberian varieties (Retinto and Entrepelado) that has created a hybrid so that has an important heterosis effect in prolificacy (Noguera et al., 2019). In addition, the company has been developing a breeding scheme for increasing litter size through selection in the parental Retinto and Entrepelado populations.

Theoretically, the optimal strategy for selection for crossbreeding is Recurrent Reciprocal Selection (Comstock et al., 1949); however, it has not been routinely used in pig breeding because it involves a delay in the generation interval. Therefore, parental purebred populations are selected based on purebred phenotypic information and under the assumption that the genetic correlation between purebred and crossbred performance is high (Wientjes & Calus, 2017). Those genetic correlations can be imperfect (< 1) because of genotype-by-environment (GxE) interactions and the presence of non-additive genetic effects (Wientjes & Calus, 2017).

Genomic information facilitates the analysis of crossbreeding data, even if genotyped and phenotyped individuals are not directly related (Vitezica et al., 2016), by the definition of an additive-dominant genotypic model that provides estimates of genotype x environmental interactions through genotypic correlations. In addition, the estimates of genotypic and dominance variances can be used to estimate the additive genetic correlation between purebred and crossbred performances. Backsolving, as proposed by Wang et al. (2012), provides an estimate of the SNP effects and allows to calculate the amount of additive genetic variance associated with each genomic region in purebred and crossbred performances.

This study estimated the additive and dominant genotypic and genetic variances and covariances, which were used to calculate the genetic correlations between purebred and crossbred performances in the Retinto and Entrepelado populations. In addition, the distribution of the additive genetic variance within the autosomal genome for purebred and crossbred performance was quantified.

2. Materials And Methods

The phenotypic data included the number of piglets born alive (NBA) and the total number born (TNB) for 306 Entrepelado and 313 Retinto purebred sows, and for 333 crossbred (Entrepelado x Retinto) sows (Table 7).

Table 7: Number of records (and number of sows between brackets) and mean ± standard deviation of the Number Born Alive and Total Number Born in Entrepelado, Crossbred and Retinto populations.

	Entrepelado	Crossbred	Retinto
N (NS)	2,030 (306)	1,958 (333)	1,977 (313)
NBA	7.75 ± 1.85	8.57 ± 2.27	8.07 ± 2.07
TNB	8.02 ± 1.89	8.80 ± 2.29	8.33 ± 2.11

N: Number of records; NS: Number of sows; NBA: Number Born Alive; TNB: Total Number Born.

Genotypes for each sow were identified based on the GeneSeek® GPP Porcine 70K HDchip (Illumina Inc., USA). Filtering excluded genotypes that had a minor allele frequency < 0.05 and a SNP call rate < 0.90 in the overall population. From that, 34,316 SNP markers were used to build the genomic relationship matrices with own developed software. The model of analysis assumed that the phenotypic values of individuals (y) (TNB and NBA) are explained by the (biological) additive (u) and dominant (v) effects of the SNPs, and a covariate (c) with the average homozygosity (f), the systematic effects (b) – order of parity (1, 2, 3, and >3) and sire of service breed (Entreplado, Retinto, or Duroc) and herd-year-season (122 levels), the sow permanent environmental effects (s) and the residuals (e) for the Entrepelado (E), Retinto (R) and Crossbred (ER) populations, as follows:

$$\begin{bmatrix} \boldsymbol{y_E} \\ \boldsymbol{y_R} \\ \boldsymbol{y_{ER}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{f_E} & 0 & 0 \\ 0 & \boldsymbol{f_R} & 0 \\ 0 & 0 & \boldsymbol{f_{RE}} \end{bmatrix} \begin{bmatrix} c_E \\ c_R \\ c_{RE} \end{bmatrix} + \begin{bmatrix} \boldsymbol{X_E} & 0 & 0 \\ 0 & \boldsymbol{X_R} & 0 \\ 0 & 0 & \boldsymbol{X_{RE}} \end{bmatrix} \begin{bmatrix} \boldsymbol{b_E} \\ \boldsymbol{b_R} \\ \boldsymbol{b_{ER}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{T_E} & 0 & 0 \\ 0 & \boldsymbol{T_R} & 0 \\ 0 & 0 & \boldsymbol{T_{RE}} \end{bmatrix} \begin{bmatrix} \boldsymbol{s_E} \\ \boldsymbol{s_R} \\ \boldsymbol{s_{ER}} \end{bmatrix}$$
$$+ \begin{bmatrix} \boldsymbol{T_E} & 0 & 0 \\ 0 & \boldsymbol{T_R} & 0 \\ 0 & 0 & \boldsymbol{T_{RE}} \end{bmatrix} \begin{bmatrix} \boldsymbol{u_E} \\ \boldsymbol{u_R} \\ \boldsymbol{u_{ER}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{T_E} & 0 & 0 \\ 0 & \boldsymbol{T_R} & 0 \\ 0 & 0 & \boldsymbol{T_{RE}} \end{bmatrix} \begin{bmatrix} \boldsymbol{v_E} \\ \boldsymbol{v_R} \\ \boldsymbol{v_{ER}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{e_E} \\ \boldsymbol{e_R} \\ \boldsymbol{e_{ER}} \end{bmatrix}$$

where $\boldsymbol{X}$ and $\boldsymbol{T}$ are the corresponding incidence matrices. Following Vitezica et al., (2016), $\boldsymbol{u}$ and $\boldsymbol{v}$ can be described in terms of the vectors of additive (a) and dominant (d) SNP genotypic effects as follows:

$$\begin{bmatrix} \boldsymbol{u_E} \\ \boldsymbol{u_R} \\ \boldsymbol{u_{RE}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{Za_E} \\ \boldsymbol{Za_R} \\ \boldsymbol{Za_{ER}} \end{bmatrix} \text{and } \begin{bmatrix} \boldsymbol{v_E} \\ \boldsymbol{v_R} \\ \boldsymbol{v_{RE}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{Wd_E} \\ \boldsymbol{Wd_R} \\ \boldsymbol{Wd_{ER}} \end{bmatrix}$$

The matrices $\boldsymbol{Z}$= (z1……zm) and $\boldsymbol{W}$= (w1…….wm) are equal to 1, 0, -1 and 0, 1, 0 for SNP genotypes A1A1, A1A2 and A2A2, respectively.

The covariance across individual genotypic additive ($\boldsymbol{u}$) and dominant ($\boldsymbol{v}$) effects are:

$$cov\begin{bmatrix} \boldsymbol{u_E} \\ \boldsymbol{u_R} \\ \boldsymbol{u_{RE}} \end{bmatrix} = \boldsymbol{G_o} \otimes \boldsymbol{G} \quad \text{and} \quad cov\begin{bmatrix} \boldsymbol{v_E} \\ \boldsymbol{v_R} \\ \boldsymbol{v_{RE}} \end{bmatrix} = \boldsymbol{D_o} \otimes \boldsymbol{D}$$

with

$$\boldsymbol{G_o} = \begin{bmatrix} \sigma^2_{U_E} & \sigma_{U_E U_R} & \sigma_{U_E U_{ER}} \\ \sigma_{U_E U_R} & \sigma^2_{U_R} & \sigma_{U_R U_{ER}} \\ \sigma_{U_E U_{ER}} & \sigma_{U_R U_{ER}} & \sigma^2_{U_{ER}} \end{bmatrix} \quad \text{and} \quad \boldsymbol{D_o} = \begin{bmatrix} \sigma^2_{V_E} & \sigma_{V_E V_R} & \sigma_{V_E V_{ER}} \\ \sigma_{V_E V_R} & \sigma^2_{V_R} & \sigma_{V_R V_{ER}} \\ \sigma_{V_E V_{ER}} & \sigma_{V_R V_{ER}} & \sigma^2_{V_{ER}} \end{bmatrix}$$

and

$$\boldsymbol{G} = \frac{\boldsymbol{ZZ'}}{\{tr[\boldsymbol{ZZ'}]/n\}} \quad \text{and} \quad \boldsymbol{D} = \frac{\boldsymbol{WW'}}{\{tr[\boldsymbol{WW'}]/n\}}$$

The variance components were estimated by REML (Patterson & Thompson, 1971) through the EM-REML algorithm with the remlf90 software (Lourenco et al., 2020) and, to obtain the average information matrix, we used one extra iteration with airemlf90. Additive and dominance variance components were calculated in each of the populations (E, R, and ER) as follows:

$$\begin{bmatrix} \hat{\sigma}^2_{aE} \\ \hat{\sigma}^2_{aR} \\ \hat{\sigma}^2_{aER} \end{bmatrix} = \begin{bmatrix} \frac{\hat{\sigma}^2_{U_E}}{\{tr[\boldsymbol{ZZ'}]/n\}} \\ \frac{\hat{\sigma}^2_{U_R}}{\{tr[\boldsymbol{ZZ'}]/n\}} \\ \frac{\hat{\sigma}^2_{U_{ER}}}{\{tr[\boldsymbol{ZZ'}]/n\}} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \hat{\sigma}^2_{dE} \\ \hat{\sigma}^2_{dR} \\ \hat{\sigma}^2_{dER} \end{bmatrix} = \begin{bmatrix} \frac{\hat{\sigma}^2_{D_E}}{\{tr[\boldsymbol{WW'}]/n\}} \\ \frac{\hat{\sigma}^2_{D_R}}{\{tr[\boldsymbol{WW'}]/n\}} \\ \frac{\hat{\sigma}^2_{D_{ER}}}{\{tr[\boldsymbol{WW'}]/n\}} \end{bmatrix} \quad (1)$$

The additive (σ^2_A)and dominance (σ^2_D) genetic variances of the purebred populations were calculated as follows:

$$\hat{\sigma}^2_{A_E} = \sum_{i=1}^{n} 2\hat{p}_{Ei}\hat{q}_{Ei}\hat{\sigma}^2_{a_E} + 2\,\hat{p}_{Ei}\hat{q}_{Ei}\left(\hat{q}_{Ei} - \hat{p}_{Ei}\right)^2 \hat{\sigma}^2_{d_E} \quad (2)$$

$$\hat{\sigma}^2_{D_E} = \sum_{i=1}^{n} \left(2\hat{p}_{Ei}\hat{q}_{Ei}\right)^2 \hat{\sigma}^2_{d_E} \quad (3)$$

$$\hat{\sigma}^2_{A_R} = \sum_{i=1}^{n} 2\hat{p}_{Ri}\hat{q}_{Ri}\hat{\sigma}^2_{a_R} + 2\,\hat{p}_{Ri}\hat{q}_{Ri}\left(\hat{q}_{Ri} - \hat{p}_{Ri}\right)^2 \hat{\sigma}^2_{d_R} \quad (4)$$

$$\hat{\sigma}^2_{D_R} = \sum_{i=1}^{n} \left(2\hat{p}_{Ri}\hat{q}_{Ri}\right)^2 \hat{\sigma}^2_{d_R} \quad (5)$$

Where $\hat{p}_{Xi}$ and $\hat{q}_{Xi}$ are the raw estimates of the allelic frequencies for A1 and A2 at the i^{th} SNP marker and the X= {E, R or ER} population. The estimates of the contributions to the additive variance in the crossbred population from the Entrepelado ($\sigma^2_{A_{ER(E)}}$) and Retinto ($\sigma^2_{A_{ER(R)}}$) were obtained by (Vitezica et al., 2016) as follows:

$$\hat{\sigma}^2_{A_{ER(E)}} = \sum_{i=1}^{n} 2\hat{p}_{Ri}\hat{q}_{Ri}\hat{\sigma}^2_{a_{ER}} + 2\,\hat{p}_{Ri}\hat{q}_{Ri}\left(\hat{q}_{Ei} - \hat{p}_{Ei}\right)^2 \hat{\sigma}^2_{d_{ER}} \quad (6)$$

$$\hat{\sigma}^2_{A_{ER(R)}} = \sum_{i=1}^{n} 2\hat{p}_{Ei}\hat{q}_{Ei}\hat{\sigma}^2_{a_{ER}} + 2\,\hat{p}_{Ei}\hat{q}_{Ei}\left(\hat{q}_{Ri} - \hat{p}_{Ri}\right)^2 \hat{\sigma}^2_{d_{ER}} \quad (7)$$

And, following Vitezica et al. (2016), the additive variance in the crossbred population was the average of the two values resulted from the equations 6 and 7

$$\hat{\sigma}^2_{A_{ER}} = \frac{1}{2}\hat{\sigma}^2_{A_{ER(E)}} + \frac{1}{2}\hat{\sigma}^2_{A_{ER(R)}} \quad (8)$$

The estimate of the dominance variance of the crossbred population (Vitezica et al., 2016) was calculated as follows:

$$\hat{\sigma}^2_{D_{ER}} = \sum_{i=1}^{n} 4\hat{p}_{Ei}\hat{q}_{Ei}\hat{p}_{Ri}\hat{q}_{Ri}\hat{\sigma}^2_{d_{ER}} \quad (9)$$

and the covariance between purebred-crossbred additive genetic effects in the Entrepelado ($\sigma_{A_E A_{ER(E)}}$) and Retinto ($\sigma_{A_R A_{ER(R)}}$) population were as follows:

$$\hat{\sigma}_{A_E A_{ER(E)}} = \sum_{i=1}^{n} 2\hat{p}_{Ei}\hat{q}_{Ei}\hat{\sigma}_{a_E a_{ER}} + 2\,\hat{p}_{Ei}\hat{q}_{Ei}\left(\hat{q}_{Ei} - \hat{p}_{Ei}\right)\left(\hat{q}_{Ri} - \hat{p}_{Ri}\right)\hat{\sigma}_{d_E d_{ER}} \quad (9)$$

$$\hat{\sigma}_{A_R A_{ER(R)}} = \sum_{i=1}^{n} 2\hat{p}_{Ri}\hat{q}_{Ri}\hat{\sigma}_{a_R a_{ER}} + 2\,\hat{p}_{Ri}\hat{q}_{Ri}\left(\hat{q}_{Ri} - \hat{p}_{Ri}\right)\left(\hat{q}_{Ei} - \hat{p}_{Ei}\right)\hat{\sigma}_{d_R d_{ER}} \quad (10)$$

with

$$\begin{bmatrix}\hat{\sigma}_{a_E a_{ER}} \\ \hat{\sigma}_{a_R a_{ER}}\end{bmatrix} = \begin{bmatrix}\frac{\hat{\sigma}_{U_E U_{ER}}}{\{tr[\boldsymbol{ZZ'}]/n\}} \\ \frac{\hat{\sigma}_{U_R U_{ER}}}{\{tr[\boldsymbol{ZZ'}]/n\}}\end{bmatrix} \quad \text{and} \quad \begin{bmatrix}\hat{\sigma}_{d_E d_{ER}} \\ \hat{\sigma}_{d_R d_{ER}}\end{bmatrix} = \begin{bmatrix}\frac{\hat{\sigma}_{V_E V_{ER}}}{\{tr[\boldsymbol{WW'}]/n\}} \\ \frac{\hat{\sigma}_{V_R V_{ER}}}{\{tr[\boldsymbol{WW'}]/n\}}\end{bmatrix} \quad (11)$$

Therefore, the genetic correlations between the purebred and crossbreed breeding values in the Entrepelado and Retinto populations were computed as follows:

$$\hat{r}_{A_E A_{ER(E)}} = \frac{\hat{\sigma}_{A_E A_{ER(E)}}}{\sqrt{\hat{\sigma}^2_{A_E}\hat{\sigma}^2_{A_{ER(E)}}}} \quad \text{and} \quad \hat{r}_{A_R A_{ER(R)}} = \frac{\hat{\sigma}_{A_R A_{ER(R)}}}{\sqrt{\hat{\sigma}^2_{A_R}\hat{\sigma}^2_{A_{ER(R)}}}} \quad (12)$$

The vector of the SNP additive effects ($\hat{\boldsymbol{a}}_{\boldsymbol{E}}$, $\hat{\boldsymbol{a}}_{\boldsymbol{R}}$ and $\hat{\boldsymbol{a}}_{\boldsymbol{ER}}$) were obtained by backsolving (Wang et al., 2012) as:

$$\hat{\boldsymbol{a}}_{\boldsymbol{E}} = \frac{\hat{\sigma}^2_{a_E}}{\hat{\sigma}^2_{U_E}}\boldsymbol{ZG}^{-1}\hat{\boldsymbol{u}}_{\boldsymbol{E}}, \quad \hat{\boldsymbol{a}}_{\boldsymbol{R}} = \frac{\hat{\sigma}^2_{a_R}}{\hat{\sigma}^2_{U_R}}\boldsymbol{ZG}^{-1}\hat{\boldsymbol{u}}_{\boldsymbol{R}} \quad \text{and} \quad \hat{\boldsymbol{a}}_{\boldsymbol{ER}} = \frac{\hat{\sigma}^2_{a_{ER}}}{\hat{\sigma}^2_{U_{ER}}}\boldsymbol{ZG}^{-1}\hat{\boldsymbol{u}}_{\boldsymbol{ER}} \quad (13)$$

and the vector of SNP dominant effects ($\hat{\boldsymbol{d}}_{\boldsymbol{E}}$, $\hat{\boldsymbol{d}}_{\boldsymbol{R}}$ and $\hat{\boldsymbol{d}}_{\boldsymbol{ER}}$) as follows:

$$\hat{\boldsymbol{d}}_{\boldsymbol{E}} = \frac{\hat{\sigma}^2_{d_E}}{\hat{\sigma}^2_{V_E}}\boldsymbol{WD}^{-1}\hat{\boldsymbol{v}}_{\boldsymbol{E}}, \quad \hat{\boldsymbol{d}}_{\boldsymbol{R}} = \frac{\hat{\sigma}^2_{d_R}}{\hat{\sigma}^2_{V_R}}\boldsymbol{WD}^{-1}\hat{\boldsymbol{v}}_{\boldsymbol{R}} \quad \text{and} \quad \hat{\boldsymbol{d}}_{\boldsymbol{ER}} = \frac{\hat{\sigma}^2_{d_{ER}}}{\hat{\sigma}^2_{V_{ER}}}\boldsymbol{WD}^{-1}\hat{\boldsymbol{v}}_{\boldsymbol{ER}} \quad (14)$$

With those, the genetic additive variances ($\sigma^2_{A_E(k)}, \sigma^2_{A_R(k)}, \sigma^2_{A_{ER(E)}(k)}$ and $\sigma^2_{A_{ER(R)}(k)}$) explained by the kth segment of the genome were calculated as follows:

$$\hat{\sigma}^2_{A_E(k)} = \sum_{i=1}^{n(k)} 2\hat{p}_{Ei}\hat{q}_{Ei}\hat{a}^2_{E_i} + 2\,\hat{p}_{Ei}\hat{q}_{Ei}\left(\hat{q}_{Ei} - \hat{p}_{Ei}\right)^2 \hat{d}^2_{E_i} \quad (15)$$

$$\hat{\sigma}^2_{A_R(k)} = \sum_{i=1}^{n(k)} 2\hat{p}_{Ri}\hat{q}_{Ri}\hat{a}^2_{R_i} + 2\,\hat{p}_{Ri}\hat{q}_{Ri}\left(\hat{q}_{Ri} - \hat{p}_{Ri}\right)^2 \hat{d}^2_{R_i} \quad (16)$$

$$\hat{\sigma}^2_{A_{ER(E)}(k)} = \sum_{i=1}^{n(k)} 2\hat{p}_{Ri}\hat{q}_{Ri}\hat{a}^2_{ER_i} + 2\,\hat{p}_{Ri}\hat{q}_{Ri}\left(\hat{q}_{Ei} - \hat{p}_{Ei}\right)^2 \hat{d}^2_{ER_i} \quad (17)$$

$$\hat{\sigma}^2_{A_{ER(R)}(k)} = \sum_{i=1}^{n(k)} 2\hat{p}_{Ei}\hat{q}_{Ei}\hat{a}^2_{ER_i} + 2\,\hat{p}_{Ei}\hat{q}_{Ei}\left(\hat{q}_{Ri} - \hat{p}_{Ri}\right)^2 \hat{d}^2_{ER_i} \quad (18)$$

Where n(k) is the number of SNP markers within the kth segment, which was set to 30. To identify the genes within the genomic regions that explained > 2.0% of the total genetic variance, we used the biomart tool (www.ensembl.org).

3. Results And Discussion

The results based on TNB and NBA were similar, which was expected because they have a high genetic correlation (Bidanel, 2011); therefore, here we focused on the results with the TNB, and the results for NBA are presented as Supplementary Information (Table S1 to S3 and Figure S1 to S3). The REML estimates of the additive genotypic (co) variances are shown in Table 8.

Table 8: REML estimates of the additive genotypic (co)variances for Total Number Born (TNB).

	Entrepelado	Crossbred	Retinto
Entrepelado	0.248 ± 0.161	0.259 ± 0.178	0.200 ± 0.135
Crossbred	-	0.546 ± 0.268	0.388 ± 0.170
Retinto	-	-	0.282 ± 0.146

The additive genotypic variance was higher in the crossbred than it was in the purebred populations. In addition, the estimates of the genotypic covariances between purebreds (Entrepelado and Retinto) and the crossbred population were all high and positive, and they corresponded to additive genotypic correlations of 0.704 ($0.259/\sqrt{0.248 \times 0.546}$) between Entrepelado and Crossbred, 0.988 ($0.388/\sqrt{0.546 \times 0.282}$) between Retinto and Crossbred, and 0.756 ($0.200/\sqrt{0.248 \times 0.282}$) between the two purebreds. Those results indicated that the genotype x environmental interaction was small, and the additive genotypic correlations were similar to those obtained by (Vitezica et al., 2016) in white pig populations. In addition, they were similar to those obtained in cattle (Karoui et al., 2012) and goat (Carillier et al., 2014) breeds.

The REML estimates of the dominance genotypic (co)variances ranged from 0.170 (Retinto) to 0.265 (Crossbred) (Table 9).

Table 9: REML estimates of the dominance genotypic (co)variances for Total Number Born (TNB).

	Entrepelado	Crossbred	Retinto
Entrepelado	0.177 ± 0.165	0.212 ± 0.171	0.166 ± 0.152
Crossbred	-	0.262 ± 0.210	0.202 ± 0.179
Retinto	-	-	0.172 ± 0.199

The estimates of the dominance genotypic covariances were all positive and reflected genotypic dominance correlations > 0.95. The analysis provided the REML estimates of the sow permanent and residual effects (Table 10).

Table 10: REML estimates of the permanent environmental and residual variances for Total Number Born (TNB) in in Entrepelado, Crossbred and Retinto populations.

Variances	Entrepelado	Crossbred	Retinto
σ_S^2	0.191 ± 0.105	0.009 ± 0.029	0.268 ± 0.120
σ_E^2	2.810 ± 0.099	4.467 ± 0.155	3.534 ± 0.128

σ_S^2: Sow permanent environment variance; σ_E^2: Residual variance.

The additive and dominance genotypic (co) variances were used to calculate the additive and dominance genetic variances in purebred populations based on expressions (1) to (5) (Table 9). The estimates of the additive genetic variances were 0.170 (Entrepelado) and 0.150 (Retinto), and the estimates of the dominance genetic variances were 0.074 (Entrepelado) and 0.056 (Retinto). The heritability estimates were 0.052 (Entrepelado) and 0.037 (Retinto), which were within the range or slightly lower than those of white pig (Bidanel, 2011; Ogawa et al., 2019; Putz et al., 2015) and Iberian (Fernández et al., 2008; García-Casco et al., 2012; Noguera et al., 2019) populations. The dominances ratios (0.023 for Entrepelado and 0.014 for Retinto) were smaller than were the heritability's. Nevertheless, their ratios with respect to the heritabilities were about 40%, higher than were those reported for white pig populations (Culbertson et al., 1998; Vitezica et al., 2016).

We used expressions (6) and (7) to calculate the additive variances for crossbred performance in the purebred populations, which were 0.413 (Entrepelado) and 0.293 (Retinto). Therefore, the additive genetic variance in the crossbred population was the average of the two (0.353), which was higher than the additive genetic variances in the purebred populations, which were similar to the results of Vitezica et al. (2016). Nevertheless, Xiang et al. (2016) found the opposite in a cross between Landrace and Yorkshire breeds. In our study, the dominance genetic variance in the crossbred population (0.079) was calculated based on (8), which was similar to the dominance genetic variance in the purebreds; however, its ratio with the additive genetic variances was lower (22%). Given those variance components, the heritability and dominance ratio estimates in the crossbred population were 0.072 and 0.016, respectively.

In addition, the additive genetic correlations between purebred and crossbred performances in the Entrepelado and Retinto populations were calculated based on expressions (9) to (12), which were 0.663 in Entrepelado and 0.881 in Retinto. Those correlations were within the range of the estimates summarized by Wientjes & Calus (2017), and suggest that the efficiency of selection for increased crossbred performance by selecting for purebred performance will be more effective in Retinto than in Entrepelado.

We used expressions (13) and (14) to calculate the additive and dominance genotypic effects associated with each of the 34,316 SNP markers, which were used in expressions

(15) to (18) to calculate the proportion of the additive genetic variance that was explained by segments of 30 consecutive SNPs (Figure 15).

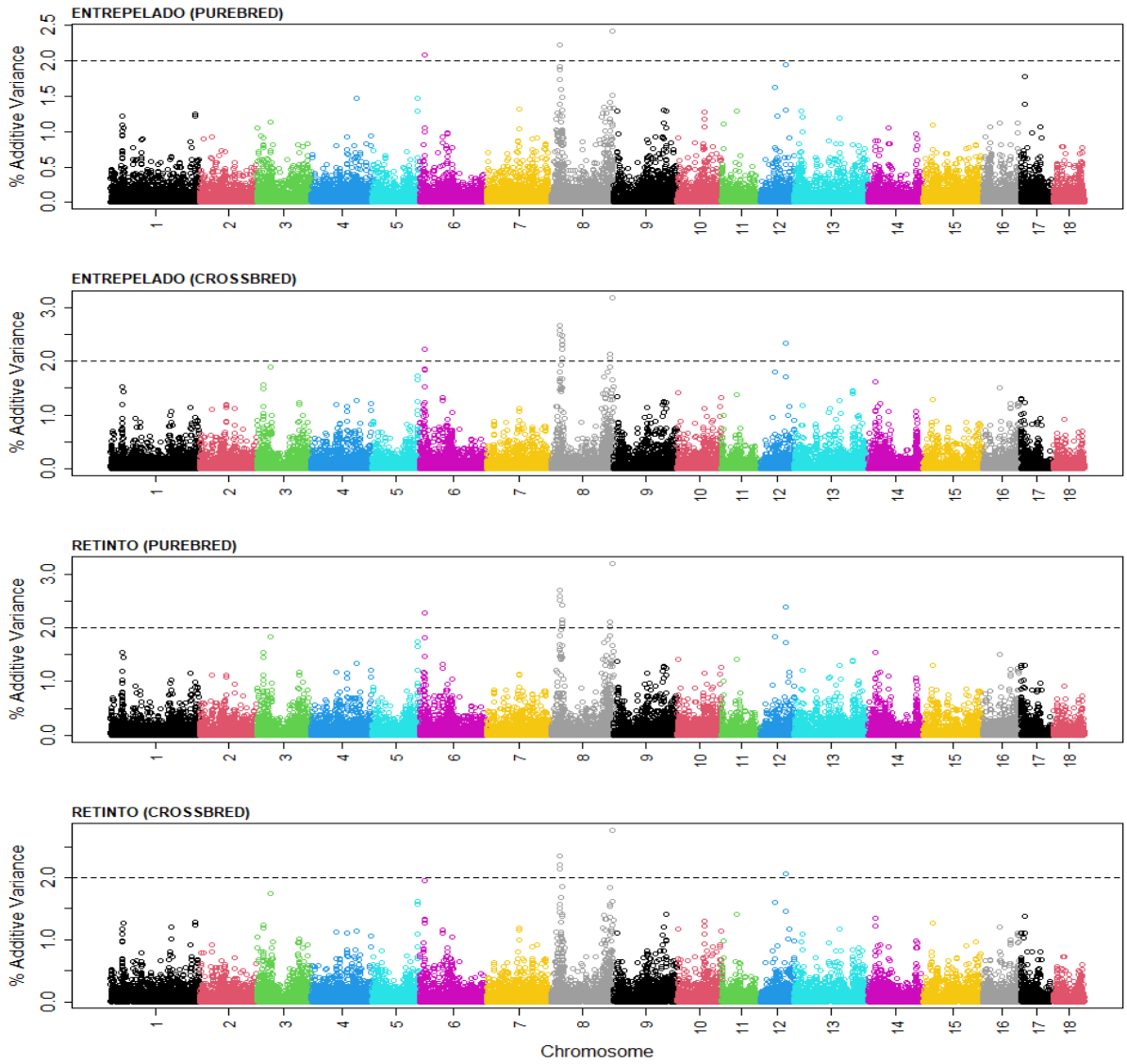


Figure 15: Distribution of the percentage of the additive genetic variance explained by genomic segments of 30 SNPs within the autosomal genome of purebred and crossbred performance for Total Number Born (TNB) in the Entrepelado and Retinto varieties.

Four genomic regions can be highlighted, each explained > 2% of the additive genetic variance in at least one of the populations. The SNPs at the center of each of the genomic regions that explained the highest amount of additive genetic variance, and the genes in the Sus_Scrofa 11.1. genomic map that were within 1 Mb downstream or upstream are presented in Table 11.

Table 11: SNPs at the center of each of the four genomic regions that explained > 2% of the additive genetic variance in at least one of the populations, and genes located within 1 Mb downstream or upstream.

SNP	SSC	bp	Genes
rs326244568	6	7597405	BCO1, PKD1L2, GCSH, ATMIN, CENPN, CDYL2, DYNLRB2
rs81401202	8	11585865	CD38, FGFBP1, PROM1, TAPT1, LDB2
rs81406142	8	137540516	CFAP299, FGF5, PRDM8, ANTXR2
rs345468811	12	46079417	TAOK1, ABHD15, TP53I13, GIT1, ANKRD1, CORO6, EFCAB5, NSRP1, SLC6A4, BLMH, TMIGD1,CPD, GOSR1

SNP: Single Nucleotide Polymorphism; SSC: Sus Scrofa Chromosome; bp: base pair.

Among those genes, several can be proposed as candidate genes for explaining the additive genetic variation. The genomic region surrounding bp 7597405 in SSC6 included BCO1 (Beta-Carotene Oxygenase 1), which encodes an enzyme that catalyzes the breakdown of provitamin A and provides retinoids for embryogenesis (Quadro et al., 2020; Wassef et al., 2013). Furthermore, the GCSH (Glycine Cleavage System H) protein plays an important role in embryonic viability (Leung et al., 2021).

Two genomic regions were identified in SSC8 around bp 11585865 and bp 137540516. Among the genes within those regions, PRDM8 (PR/SET Domain 8), is involved in neurogenesis (Kinameri et al., 2008) of the FGF5 (Fibroblast Growth Factor 5), a member of the fibroblast growth factor family that is involved in several biological processes including embryonic development, cell growth, and morphogenesis (Beer et al., 2005; Haub & Goldfarb, 1991). The genomic region around bp 46079417 in SSC12 contains, among others, the GIT1 (G protein-coupled receptor kinase interactor 1) gene, which plays a role in spine morphogenesis (Segura et al., 2007), the NSRP1 (Nuclear Speckle Splicing Regulatory Protein 1) development process, and in utero embryonic development (Kim et al., 2011), and ANKRD1 (Ankyrin Repeat Domain 1), which is involved in neuron projection development (Stam et al., 2007).

4. Conclusions

1) the additive genetic variance and the heritabilities were higher in the crossbred than those in the purebred populations, 2) the genetic correlation between purebred and crossbreed performances were higher in Retinto than it was in Entrepelado, and 3) the additive genetic variances were heterogeneously distributed throughout the autosomal genome, and several genomic regions in SSC6, SSC8, and SSC12 were identified.

Supplementary Material: The following figure and tables are available in supplementary materials.

Table S1: REML estimates of the additive genotypic (co)variances for Number Born Alive (NBA).

Table S2: REML estimates of the dominance genotypic (co)variances for Number Born Alive (NBA).

Table S3: REML estimates of the permanent environmental and residual variances for Number Born Alive (NBA).

Table S4: GO (Gene Ontology) terms for biological process of the proposed candidate genes.

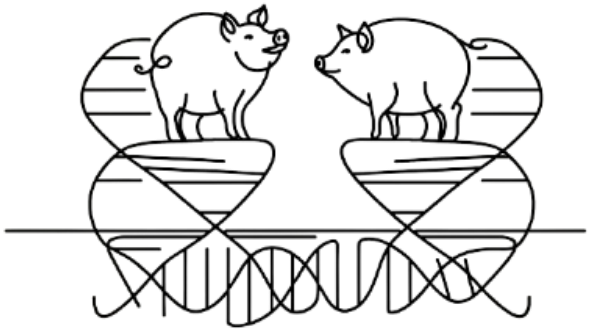
Figure S1: Distribution of the percentage of the additive genetic variance explained by genomic segments of 30 SNPs within the autosomal genome of purebred and crossbred performance for Number Born Alive (NBA) in the Entrepelado and Retinto varieties.

Figure S2: Distribution of the percentage of the additive genetic variance explained by genomic segments of 20 SNPs within the autosomal genome of purebred and crossbred performance for Total Number Born (TNB) in the Entrepelado and Retinto varieties.

Figure S3: Distribution of the percentage of the additive genetic variance explained by genomic segments of 40 SNPs within the autosomal genome of purebred and crossbred performance for Total Number Born (TNB) in the Entrepelado and Retinto varieties.

CHAPTER VI

A Bayesian Multivariate Gametic Model in a Reciprocal Cross with Genomic Information: An Example with Two Iberian Varieties



Authors

Houssemeddine Srihi; David López-Carbonell; Noelia Ibáñez-Escriche; Joaquim Casellas; Pilar Hernández; Sara Negro and Luis Varona.

The content of this chapter has been published in Animals 2023.doi.org/10.3390/ani13101648

Abstract

INGA FOOD, S.A. initiated a crossbreeding program between two Iberian pig varieties, Retinto (R) and Entrepelado (E), with the goal of producing a hybrid sow (F1). Several studies have been conducted to evaluate its productive performance, and these studies have revealed differences in litter size between two reciprocal crosses, suggesting the presence of genomic imprinting effects. To further investigate these effects, this study introduces a multivariate gametic model designed to estimate gametic correlations between paternal and maternal effects originating from both genetic backgrounds involved in the reciprocal crosses. The dataset consisted of 1,258 records (the total number born—TNB and the number born alive—NBA) from 203 crossbred dams for the Entrepelado (sire) × Retinto (dam) cross and 700 records from 125 crossbred dams for the Retinto (sire) × Entrepelado (dam) cross. All animals were genotyped using the GeneSeek® GPP Porcine 70 K HDchip. The results indicated that the posterior distribution of the gametic correlation between paternal and maternal effects was distinctly different between the two populations. Specifically, in the Retinto population, the gametic correlation showed a positive skew with posterior probabilities of 0.78 for the TNB and 0.80 for the NBA. On the other hand, the Entrepelado population showed a posterior probability of a positive gametic correlation between paternal and maternal effects of approximately 0.50. The differences in the shape of the posterior distribution of the gametic correlations between paternal and maternal effects observed in the two varieties may account for the distinct performance observed in the reciprocal crosses.

1. Introduction

The Iberian breed is widely renowned for its ability to produce some of the highest-quality pork (Serra et al., 1998). This breed is particularly well-adapted to the “Dehesa” environment in southwestern Spain, which is characterized by a savannah landscape and is composed of grass, cork, and holm oaks with seasonal production. Traditionally, Iberian pig production was dominated by purebred varieties and extensive management practices. However, in recent decades, there has been a shift toward more intensive farming practices that incorporate crossbreeding with Duroc boars to improve growth and efficiency at commercial stages (Serrano et al., 2008).

The regulatory norms for Iberian pig production allow crossbreeding, as long as the sow is of purebred Iberian stock. The reproductive performance of the Iberian sows is lower than that of white pig populations (Silio et al., 2001), which is a major limitation of its use in intensive farms. Therefore, improvement in the reproductive efficiency of Iberian sows is crucial for their economic efficiency. Several studies have identified genetic variability for prolificacy within and between varieties of Iberian pig (Fernández et al., 2008; García-Casco et al., 2012). To take advantage of this variability, the INGA FOOD, S.A. company has developed a crossbreeding scheme between two Iberian varieties (Retinto and Entrepelado) to generate an F1 hybrid sow, which exhibits significant heterosis for litter size (Noguera et al., 2019). However, this study also found differences in the reproductive performance between the two reciprocal crosses (Entrepelado × Retinto, ER, vs. Retinto × Entrepelado, RE), suggesting that these differences may be attributed to parental imprinting (Reik & Walter, 2001)(i.e., the effects from alleles may differ whether they are transmitted by paternal or maternal gametes). In fact, there is increasing evidence of the importance of imprinting in placenta development (Hanna, 2020), and certain imprinted genes have been proposed as candidates for pig litter size (Coster et al., 2012).

In recent years, some algorithms have been proposed to develop a genomic analysis of imprinting (Nishio & Satoh, 2015) from the genomic information provided by commercial genotyping devices. However, knowledge of the parental haplotype phase of the SNP markers is required to differentiate the paternal or maternal gametic effects, but some approaches have been developed to reconstruct haplotype phases (Hickey et al., 2011).

Phenotypic information from reciprocal crosses offers the opportunity to compare the paternal and maternal effects of each parental population. In the absence of imprinting, a correlation between the paternal and maternal effects from the same population should be performed. Imprinting, on the other hand, results in a lower correlation. Accordingly, the goal of this study was to apply the multivariate gametic model developed in a previous study (Srihi et al., 2022) that utilizes genomic information and is capable of estimating the paternal and maternal gametic contributions of Retinto and Entrepelado varieties in the ER and RE crosses, along with their correlations.

2. Materials And Methods

Phenotypic and Genomic Data. The phenotypic data used in this study consisted of the total number born, TNB, and the number of piglets born alive, NBA, in 203 ER and 125 RE sows. The ER sows were the offspring of 38 purebred Entrepelado boars and 139 Retinto dams, whereas the RE sows were generated from 38 Retinto boars and 92 Entrepelado dams. A summary of the data is presented in Table 12.

Table 12 : The number of records (and sows between brackets), mean (± standard deviation) of the total number born, TNB, and the number born alive, NBA, for Entrepelado × Retinto and Retinto × Entrepelado crosses.

	Entrepelado × Retinto	**Retinto × Entrepelado**
N[1] (NS)[2]	1,258 (203)	700 (125)
TNB[3]	8.78 ± 2.24	8.85 ± 2.37
NBA[4]	8.55 ± 2.23	8.62 ± 2.34

N: number of records; NS: number of sows; TNB: total number born; NBA: number born alive.

Genotyping was performed with the GeneSeek® GPP Porcine 70 K HDchip (Illumina Inc., USA) on all ER and RE crossbred sows, as well as on 341 Retinto and 350 Entrepelado purebred individuals. Due to shared purebred ancestors, there was some degree of relationship between a subset of the ER and RE crossbred sows and the purebred individuals, although not all of them were genotyped. The original genotype data consisted of 60,224 autosomal SNPs, which were filtered by excluding SNP markers with a call rate below 0.90 and a minor allele frequency lower than 0.05 in each population. Among these 4,212 were discarded due to a call rate lower than 0.90, 11,234 were found to be monomorphic, and 9,876 and 11,516 has a minor allele frequency lower than 0.05 in the Entrepelado and Retinto populations, respectively. Finally, a total of 23,386 SNPs were retained.

Haplotype Phasing. AlphaPhase software (Hickey et al., 2011) was used for each chromosome separately, utilizing genotypes of both crossbred and purebred individuals, as well as a pedigree of 1601 individuals. AlphaPhase was executed with a tolerance of 1% of genotype errors and 1% disagreement between genotypes and haplotypes. The

number of surrogates and percentage of surrogate disagreement was set to 10. Nine different scenarios were applied with core lengths of 75, 100, and 125 SNPs and tail lengths of 100, 150, and 200 SNPs (see Table 13). The scenarios were evaluated for concordance, and haplotype assignments that coincided in seven or more scenarios were retained for subsequent analysis.

Table 13: Parameters (core and tail length) in the nine scenarios of haplotype phasing.

Scenario	Core Length	Tail Length
S1	75	100
S2	75	150
S3	75	200
S4	100	100
S5	100	150
S6	100	200
S7	125	100
S8	125	150
S9	125	200

Statistical Model. Once the haplotype phases were calculated, data were analyzed with the model proposed by Srihi et al. (2022)

In this equation, $y_{(ER)}$ and $y_{(RE)}$ refer to the vectors of phenotypic records (TNB or NBA) for the ER and RE crosses, respectively. The terms $\boldsymbol{b}_{(ER)}$ and $\boldsymbol{b}_{(RE)}$ correspond to systematic effects, and $\boldsymbol{s}_{(ER)}$ and $\boldsymbol{s}_{(RE)}$ represent the permanent sow environmental effects. Paternal effects for the Entrepelado (E) and Retinto (R) populations are denoted by $p_{(E)}$ and $p_{(R)}$, respectively. Maternal effects for the Entrepelado (E) and Retinto (R) are represented by $m_{(E)}$ and $m_{(R)}$. Additionally, $e_{(ER)}$ and $e_{(RE)}$ are the residual effects for the ER and RE crosses, respectively. The systematic effects vectors included the order of parity with five levels (first, second, third, fourth, and fifth or more) and herd–year–season with thirty-four levels.

$$\boldsymbol{y}_{(ER)} = \boldsymbol{X}_{(ER)}\boldsymbol{b}_{(ER)} + \boldsymbol{B}_{(ER)}\boldsymbol{s}_{(ER)} + \boldsymbol{Z}_{(ER)}\boldsymbol{p}_{(E)} + \boldsymbol{W}_{(ER)}\boldsymbol{m}_{(R)} + \boldsymbol{e}_{(ER)}$$

$$\boldsymbol{y}_{(RE)} = \boldsymbol{X}_{(RE)}\boldsymbol{b}_{(RE)} + \boldsymbol{B}_{(RE)}\boldsymbol{s}_{(RE)} + \boldsymbol{Z}_{(RE)}\boldsymbol{p}_{(R)} + \boldsymbol{W}_{(RE)}\boldsymbol{m}_{(E)} + \boldsymbol{e}_{(RE)}$$

Further $\boldsymbol{X_{ER}}, \boldsymbol{X_{RE}}, \boldsymbol{B_{ER}}, \boldsymbol{B_{RE}}, \boldsymbol{Z_{ER}}, \boldsymbol{Z_{RE}}, \boldsymbol{W_{ER}},$ and $\boldsymbol{W_{RE}}$ are the corresponding incidence matrices.

Following (Srihi et al., 2022), the prior distribution of the permanent sow environmental effects was:

$$\begin{bmatrix} \boldsymbol{s}_{(ER)} \\ \boldsymbol{s}_{(RE)} \end{bmatrix} \sim N\left(\begin{matrix}0\\0\end{matrix}, \boldsymbol{I} \otimes \boldsymbol{S}\right)$$

where

$$\boldsymbol{S} = \begin{bmatrix} \sigma^2_{s(ER)} & 0 \\ 0 & \sigma^2_{s(RE)} \end{bmatrix}$$

where $\sigma^2_{s(ER)}$ and $\sigma^2_{s(RE)}$ are the variances of the permanent sow environmental effects for ER and RE, respectively. The prior distributions of the gametic effects for the Entrepelado (E) and Retinto (R) populations are:

$$\begin{bmatrix} \boldsymbol{p}_{(E)} \\ \boldsymbol{m}_{(E)} \end{bmatrix} \sim N\left(\begin{matrix}0\\0\end{matrix}, \boldsymbol{G_E} \otimes \boldsymbol{V_E}\right) \begin{bmatrix} \boldsymbol{p}_{(R)} \\ \boldsymbol{m}_{(R)} \end{bmatrix} \sim N\left(\begin{matrix}0\\0\end{matrix}, \boldsymbol{G_R} \otimes \boldsymbol{V_R}\right)$$

where

$$\boldsymbol{V_E} = \begin{bmatrix} \sigma^2_{p(E)} & \sigma_{pm(E)} \\ \sigma_{pm(E)} & \sigma^2_{m(E)} \end{bmatrix}$$

and

$$\boldsymbol{V_R} = \begin{bmatrix} \sigma^2_{p(R)} & \sigma_{pm(R)} \\ \sigma_{pm(R)} & \sigma^2_{m(R)} \end{bmatrix}$$

where $\sigma^2_{p(E)}$, $\sigma^2_{m(E)}$, and $\sigma_{pm(E)}$ refer to the variances of the paternal and maternal gametic effects and the covariance between them for the Entrepelado population. Similarly, $\sigma^2_{p(R)}$, $\sigma^2_{m(R)}$, and $\sigma_{pm(R)}$ represent the variances of the paternal and maternal gametic effects and the covariance between them, respectively, for the Retinto population. Additionally, $\boldsymbol{G_E}$ and $\boldsymbol{G_R}$ are the gametic relationship matrices of the Entrepelado or Retinto gametes, respectively, regardless of whether they are transmitted as paternal or maternal gametes. These matrices describe the relationships among the gametes from Entrepelado and Retinto origins, and they are calculated using the algorithm proposed by Nishio & Satoh (2015):

$$\boldsymbol{G_E} = \frac{\boldsymbol{M_E M_E}'}{\sum_i^{N_{SNP}} q_{(E)i}(1 - q_{(E)i})} \quad \boldsymbol{G_R} = \frac{\boldsymbol{M_R M_R}'}{\sum_i^{N_{SNP}} q_{(R)i}(1 - q_{(R)i})}$$

where $\boldsymbol{M_E}$ and $\boldsymbol{M_R}$ are the matrices of the number of genotyped individuals (n) × the number of SNP (N_{SNP}), whose elements $M_E(i,j)$ (or $M_R(i,j)$ take the value $q(E)_j$) (or $q(R)_j$) or $-\left(1 - q(E)_j\right)$ (or $-\left(1 - q(R)_j\right)$), depending on whether the j^{th} allele of the gametes transmitted for the i^{th} individual is A_1 or A_2 and of Entrepelado (or Retinto) origin. Additionally, $q(E)_j$ and $q(R)_j$ represent the allelic frequencies of the A_2 allele in the Entrepelado (E) and Retinto (R) populations, respectively. The prior distributions for the (co) variance components and the systematic effects were assumed to be flat. The analysis was performed using Bayesian inference with the Gibbs sampler (Gelfand & Smith, 1990) and implemented with Gibbsf90 software (Misztal et al., 2018). The analysis was performed using 10 million iterations after discarding the first million.

At each iteration of the Gibbs sampler, the (co) variances components samples were utilized to compute the samples from the marginal posterior distribution of the correlations between the paternal and maternal gametic effects for Entrepelado ($r_{pm(E)}$) and Retinto ($r_{pm(R)}$):

$$r_{pm(E)} = \frac{\sigma_{pm(E)}}{\sqrt{\sigma^2_{p(E)}\sigma^2_{j(E)}}} \quad \text{and} \quad r_{pm(R)} = \frac{\sigma_{pm(R)}}{\sqrt{\sigma^2_{p(R)}\sigma^2_{j(R)}}}$$

3. Results And Discussion

Haplotype Phasing. The results of comparing haplotype phasing using nine combinations of core length and core tail parameters using Alphaphase software are presented in Figure 16.

Figure 16: Degree of similitude between estimated haplotype phases in the nine scenarios of phasing.

The average degree of similitude was 0.89, and it was consistently above 0.86. Specifically, the predicted haplotype phase was identical across all nine scenarios for only 78.74% of the analysis but had concordance in more than seven scenarios in 92.5% of SNPs. These findings indicated that the output of the phasing algorithm was highly dependent on the specific set of parameters used for its implementation when medium-density SNP chips were used.

Calculation of Gametic Matrices. The diagonal values of the gametic matrices for the Entrepelado population ranged from 0.894 to 1.100, while for the Retinto population, they ranged from 0.901 to 1.179. Table 14 shows the distribution of the gametic relationships observed in the off-diagonal elements of the gametic matrices.

Table 14: Distribution of gametic relationships between the Entrepelado and Retinto gametic effects.

Gametic Relationship	ENTREPELADO	RETINTO
<0.05	92,276 (86.03 %)	94,144 (87.77 %)
0.05–0.10	8130 (7.58 %)	8900 (8.29 %)
0.10–0.20	4670 (4.35 %)	3130 (2.92 %)
0.20–0.30	1076 (1.00 %)	582 (0.54 %)
0.30–0.40	480 (0.44 %)	252 (0.23 %)
0.40–0.50	396 (0.36 %)	188 (0.18 %)
>0.50	228 (0.21 %)	60 (0.05 %)

The calculated gametic matrices yielded results consistent with the familiar relations of the individuals, as gametic relationships around 0.50 indicated that the individuals shared sire (or dam), while gametic relations around 0.25 suggested that the sires (or dams) of the individuals were fullsibs.

Variance Components. The posterior mean and standard deviation estimate of the variance components are presented in Table 15.

Table 15: Posterior mean (and standard deviation) of the variance components for the total number born, TNB, and the number born alive, NBA**.**

Variance Component	TNB	NBA
$\boldsymbol{\sigma^2_{s(ER)}}$	0.144 (0.098)	0.142 (0.097)
$\boldsymbol{\sigma^2_{s(RE)}}$	0.357 (0.187)	0.365 (0.191)
$\boldsymbol{\sigma^2_{p(E)}}$	0.206 (0.103)	0.199 (0.100)
$\boldsymbol{\sigma^2_{m(E)}}$	0.197 (0.114)	0.199 (0.115)
$\boldsymbol{\sigma^2_{p(R)}}$	0.224 (0.132)	0.222 (0.128)
$\boldsymbol{\sigma^2_{m(R)}}$	0.163 (0.087)	0.151 (0.080)
$\boldsymbol{\sigma^2_{e(ER)}}$	4.296 (0.187)	4.251 (0.185)
$\boldsymbol{\sigma^2_{e(RE)}}$	4.795 (0.288)	4.607 (0.278)

Furthermore, Figure 17 shows the posterior distributions of the ratios of gametic variances in the Entrepelado × Retinto (ExR) and Retinto × Entrepelado (RxE) crosses. The posterior mean estimates were similar, ranging between 0.034 for the Retinto maternal gametic effects in the ER cross and 0.043 for the Entrepelado paternal gametic effects in the RE cross.

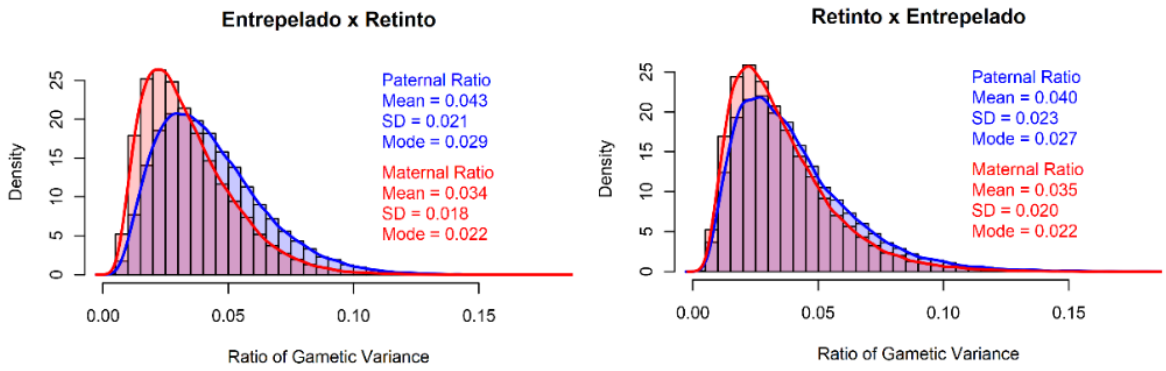


Figure 17: Posterior distributions of the ratio of gametic effect in the Entrepelado × Retinto and in the Retinto × Entrepelado crosses.

These results indicate that there are no relevant differences in the amount of genetic variance contributed by the paternal and maternal origins in either of the two reciprocal crosses, based on the available information.

Gametic Correlations. The posterior distribution of the gametic correlations for the TNB and NBA in the Entrepelado and Retinto populations are presented in Figures 18 and 19, respectively.

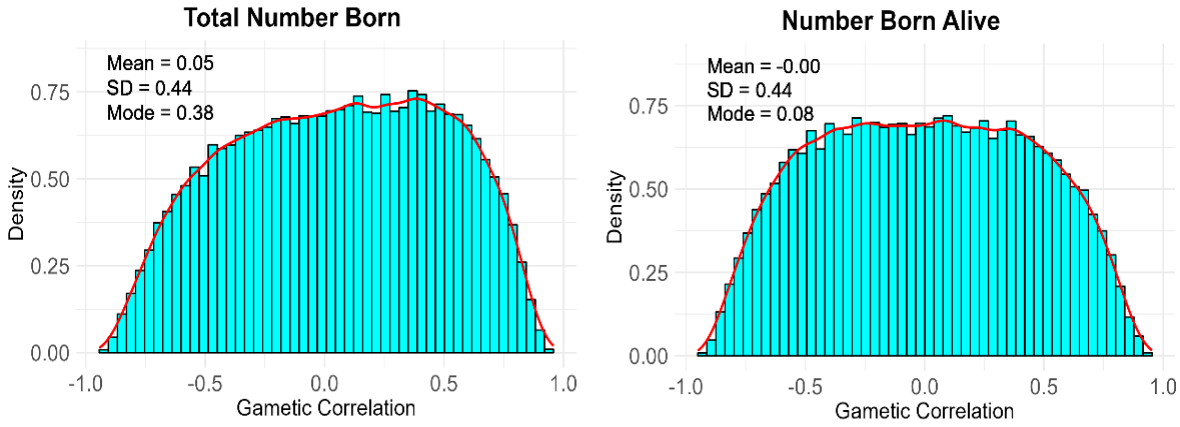


Figure 18: Posterior distributions of the gametic correlation between the paternal and maternal effects in the Entrepelado population.

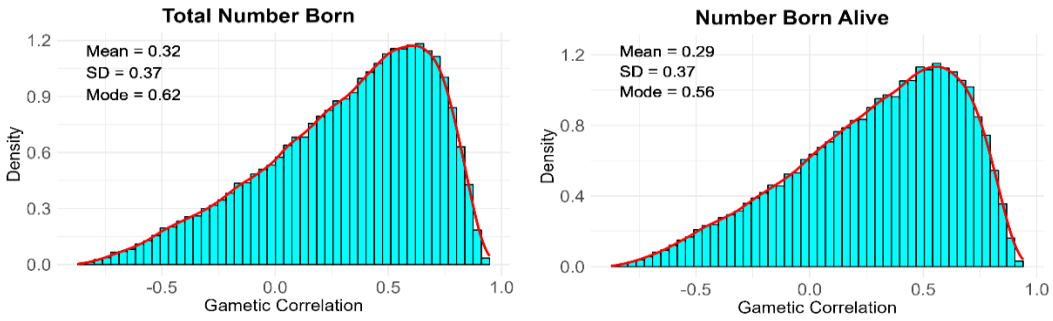


Figure 19: Posterior distributions of the gametic correlation between the paternal and maternal effects in the Retinto population.

The posterior distribution of the correlation between gametic effects in Retinto and Entrepelado showed notable differences in shape. Specifically, the posterior distributions of the gametic correlations in the Retinto population exhibited a higher degree of positive asymmetry compared to those in the Entrepelado population. In fact, the posterior probabilities of a positive gametic correlation in the Retinto population were 0.80 and 0.78 for the TNB and NBA, respectively. In contrast, the posterior probabilities of a positive gametic correlation in the Entrepelado population were 0.50 (TNB) and 0.54 (NBA).

Although caution is needed in interpreting the results due to the limited amount of phenotypic and genotypic information, the shape of the posterior distribution of gametic correlations suggests a potential role of genomic imprinting. This is because a gametic correlation substantially lower than one indicates that the same combination of alleles in a gamete may produce different effects on offspring depending on whether they are transmitted by paternal or maternal gametes, which is consistent with the theory of genomic imprinting. Genomic imprinting is an epigenetic phenomenon that causes genes to be expressed depending on whether they are inherited from the father or mother (Reik & Walter, 2001).

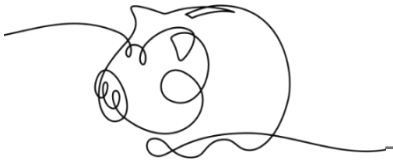
Several theories have been postulated to explain the evolutionary origin of genomic imprinting (Patten et al., 2014), and one of the most popular is the parental investment theory (Moore & Haig, 1991). This theory argues that imprinting is the result of a conflict between the evolutionary success of paternally and maternally derived genes. In mammalian reproduction, the evolutionary success of paternally inherited genes is associated with the increase in fetal growth, while for maternally inherited genes, it is associated with the number of offspring. This theory is reinforced by the discovery of

numerous imprinted genes known to regulate aspects of mammalian development (Thamban et al., 2020), including growth, behavior, and placental function (Fowden et al., 2011), furthermore, there is increasing evidence of imprinted genes in the pig genome (Coster et al., 2012; Wu et al., 2020; Zhang et al., 2012).

From a practical perspective, a low or null gametic correlation between paternal and maternal gametes within the same population indicates that a selection program to improve the performance of the crossbreeding individuals needs to be specifically designed, especially in the Entrepelado population. This is because the selection of purebred animals to increase the performance in the Entrepelado × Retinto cross may not have any noticeable consequences in the performance in the Retinto × Entrepelado cross. Furthermore, this result also may explain the differences in performance among the reciprocal crosses observed by Noguera et al. (Noguera et al., 2019), who proposed using the Retinto variety as a boar and the Entrepelado as a sow, providing better performance than the opposite cross.

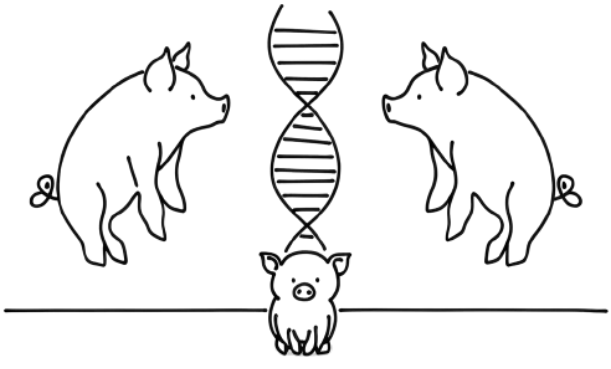
4. Conclusions

The bivariate model proposed in this study provides estimates of the gametic effects of each founder population as either paternal or maternal, as well as their correlation. In the absence of parental imprinting, a perfect correlation of one would be expected. However, our results detect a significant deviation from this ideal scenario, indicating possible differences in the performance of crossbred individuals depending on the paternal or maternal origin of the gametes. These findings provide evidence of the presence of imprinting effects in Iberian pig populations, which could have implications for the design of future breeding programs.



CHAPTER VII

A multivariate gametic model for the analysis of purebred and crossbred data. An example between two populations of Iberian pigs



Authers

Houssemeddine Srihi, David López-Carbonell, Noelia Ibáñez-Escriche, Joaquim Casellas, Pilar Hernández, Sara Negro, Luis Varona

The content of this chapter has been published in ***Journal of Animal Breeding and Genetics 2023***

doi.org/10.1111/jbg.12832

Abstract

Crossbreeding plays a pivotal role within pig breeding programs, aiming to maximize heterosis and improve reproductive traits in crossbred maternal lines. Nevertheless, there is evidence indicating that the performance of reciprocal crosses between two genetic lines might exhibit variability. These variations in performance can be attributed to differences in the correlations between gametic effects, acting as either sire or dam, within purebred and crossbred populations. To address this issue, we propose a multivariate gametic model that incorporates up to four correlated gametic effects for each parental population. The model is employed on a dataset comprising litter size data (total number of piglets born –TNB- and number of piglets born alive –NBA-) derived from a reciprocal cross involving two Iberian pig populations: Entrepelado and Retinto. The dataset comprises 6,933 records from 1,564 purebred Entrepelado (EE) sows, 4,995 records from 1,015 Entrepelado x Retinto (ER) crosses, 2,977 records from 756 Retinto x Entrepelado (RE) crosses, and 7,497 records from 1,577 purebred Retinto (RR) sows. The dataset is further supplemented by a pedigree encompassing 6,007 individual-sire-dam entries. The statistical model also included the order of parity (with six levels), the breed of the service sire (five levels), and the herd-year-season effects (141 levels). Additionally, the model integrates random dominant and permanent environmental sow effects. The analysis employed a Bayesian approach, and the results revealed all the posterior estimates of the gametic correlations to be positive. The range of the posterior mean estimates of the correlations varied across different gametic effects and traits, with a range between 0.04 (gametic correlation between the paternal effects for purebred and the maternal for crossbred in Retinto) and 0.53 (gametic correlation between the paternal effects for purebred and the paternal for crossbred in Entrepelado). Furthermore, the posterior mean variance estimates of the maternal gametic effects were consistently surpassed those for paternal effects within all four populations. The results suggest the possible influence of imprinting effects on the genetic control of litter size, and underscore the importance of incorporating crossbred data into the breeding value predictions for purebred individuals.

Key Words: Crossbreeding, Reciprocal Cross, Gametic Correlation, Parent of Origin.

1. Introduction

The Iberian pig breed is renowned for its adaptability to the natural Mediterranean ecosystem in Southwestern Spain, as well as its product quality (Lopez-Bote, 1998). Despite the Iberian pigs exhibit slower growth, lower feed efficiency, and reduced prolificacy compared to other commercial pig populations, their meat quality (Gilles, 2009) and specialized derivative products contribute to their sustainability. However, in recent decades, the traditional extensive production methods that solely relied on purebred Iberian individuals have been partly replaced by intensive farming practices involving crossbreeding with Duroc pigs. The rationale behind this crossbreeding is to enhance growth and efficiency (Serrano et al., 2008). This shift towards intensive farming has yielded several advantages, such the collection of productive data and the implementation of genetic selection programs. It is important to note that regulatory norms governing the production of Iberian pig products stipulate that the sow must belong to a pure Iberian lineage. This highlights the importance of improving the reproductive efficiency of Iberian sows to ensure the economic sustainability of the breed.

INGA FOOD S.A. is a Spanish company specializing in the production and distribution of premium pig products. As part of their breeding program, they have successfully developed a hybrid Iberian sow called CASTUA. This hybrid sow is the result of crossbreeding between the Retinto and Entrepelado populations of the Iberian breed. Through this crossbreeding, the CASTUA hybrid sow exhibit improved litter size traits due to heterosis, as confirmed by Noguera et al. (2019). Furthermore, INGA FOOD S.A. is actively implementing a breeding program with the aim of further enhancing litter size based on the performance of purebred individuals. The foundation of this program is built upon the assumption of a positive genetic correlation between the performance of purebred and crossbred pigs (Wientjes & Calus, 2017). This hypothesis finds support in a preceding study (Srihi et al., 2022) that quantified the genetic correlation between the performance of purebred and crossbred animals within the Entrepelado and Retinto populations.

Furthermore, as highlighted by Noguera et al. (2019), variations in the performance of reciprocal crosses suggest that the gametic contribution to genetic variance may

diverge when the populations act as either sires or dams in the crossbreeding process (Srihi et al., 2023). While previous literature has employed gametic models to identify differences in parent-of-origin effects (Meyer & Tier, 2012; Varona et al., 2015), to the best of our knowledge, these models have not been previously implemented within the context of crossbreeding. Hence, the goal of this study is to develop a multivariate gametic model encompassing both paternal and maternal gametic effects for purebred and crossbred performance. This framework will enable the estimation of gametic correlations between these effects. Additionally, the model will also incorporate a dominance effect.

2. Material And Methods

The dataset used in this study consisted of a total of 22,402 records for two reproductive traits: total number of piglets born (TNB) and number of piglets born alive (NBA). These records were collected from 4,912 sows that were part of a complete diallelic experiment involving two strains of the Iberian pig breed: Retinto (RR) and Entrepelado (EE). Additionally, the dataset included their reciprocal crosses: Entrepelado × Retinto (ER) and Retinto × Entrepelado (RE). Along with the phenotypic data, a pedigree containing 6,007 individual-sire-dam entries was included for genetic analysis. Table 16 provides a summary of the phenotypic data used in this study.

Table 16: Number of phenotypic records (and number of sows producing them in brackets), mean (± standard deviation) of Total Number Born (TNB) and Number Born Alive (NBA) for Entrepelado (EE), Retinto (RR) and Entrepelado x Retinto (ER) and Retinto × Entrepelado (RE) crosses.

	EE	**RR**	**ER**	**RE**
N (NS)	6,933 (1,564)	7,497 (1,577)	4,995 (1,015)	2,977 (756)
TNB	8.23 ± 2.14	8.44 ± 2.22	8.55 ± 2.27	8.51 ±2.28
NBA	7.87 ± 2.11	8.05 ± 2.18	8.27 ± 2.25	8.18 ± 2.25

N: number of phenotypic records; NS: number of recorded sows.

The data were analyzed with the following models:

$$\boldsymbol{y}_{EE} = \boldsymbol{X}_{EE}\boldsymbol{b}_{EE} + \boldsymbol{Z}_{EE}\boldsymbol{p}_{EE} + \boldsymbol{W}_{EE}\boldsymbol{m}_{EE} + \boldsymbol{Q}_{EE}\boldsymbol{d}_{EE} + \boldsymbol{H}_{EE}\boldsymbol{r}_{EE} + \boldsymbol{e}_{EE}$$

$$\boldsymbol{y}_{ER} = \boldsymbol{X}_{ER}\boldsymbol{b}_{ER} + \boldsymbol{Z}_{ER}\boldsymbol{p}_{ER} + \boldsymbol{W}_{ER}\boldsymbol{m}_{ER} + \boldsymbol{Q}_{ER}\boldsymbol{d}_{ER} + \boldsymbol{H}_{ER}\boldsymbol{r}_{ER} + \boldsymbol{e}_{ER}$$

$$\boldsymbol{y}_{RR} = \boldsymbol{X}_{RR}\boldsymbol{b}_{RR} + \boldsymbol{Z}_{RR}\boldsymbol{p}_{RR} + \boldsymbol{W}_{RR}\boldsymbol{m}_{RR} + \boldsymbol{Q}_{RR}\boldsymbol{d}_{RR} + \boldsymbol{H}_{RR}\boldsymbol{r}_{RR} + \boldsymbol{e}_{RR}$$

$$\boldsymbol{y}_{RE} = \boldsymbol{X}_{RE}\boldsymbol{b}_{RE} + \boldsymbol{Z}_{RE}\boldsymbol{p}_{RE} + \boldsymbol{W}_{RE}\boldsymbol{m}_{RE} + \boldsymbol{Q}_{RE}\boldsymbol{d}_{RE} + \boldsymbol{H}_{\boldsymbol{RE}}\boldsymbol{r}_{RE} + \boldsymbol{e}_{RE}$$

In the given equations, $\boldsymbol{y}_{JK}$ is the vector of phenotypic records (TNB or NBA) for the *JK= {EE, ER, RR, RE}* population. Here, *J= {E, R}* denotes the paternal population, and *K = {E, R}* denotes the maternal. Additionally, $\boldsymbol{b}_{JK}$ is the vector of systematic effects, including order of parity (6 levels), breed of service sire (5 levels) and herd-year-season (141 levels). Moreover, $\boldsymbol{p}_{JK}$, $\boldsymbol{m}_{JK}$, $\boldsymbol{d}_{JK}$, $\boldsymbol{r}_{JK}$ and $\boldsymbol{e}_{JK}$ are the paternal, maternal, dominance, permanent environmental and residual effects of the *JK* population, respectively. It must be noted that $\boldsymbol{p}_{JK}$ is the vector of the paternal gametic effects of the *J={E, R}* population in the *JK* cross, and that $\boldsymbol{m}_{JK}$ is the vector of the *K={E,R}* maternal gametic effects in the *JK={EE, ER, RR, RE}* cross. Besides, $\boldsymbol{X}_{JK}, \boldsymbol{Z}_{JK}, \boldsymbol{W}_{JK}, \boldsymbol{Q}_{JK}$ and $\boldsymbol{H}_{JK}$ are the corresponding incidence matrices involved in the equations.

The statistical model was analyzed by employing a Bayesian approach with a Gibbs sampler (Gelfand & Smith, 1990). In this analysis, bounded uniform distributions were employed as prior distributions for the systematic effects and variance components. The prior distributions for the gametic, dominance, permanent environmental and residual effects were modelled as multivariate Gaussian distributions, characterized by a zero mean and a variance as outlined before:

$$var\begin{pmatrix} p_{EE} \\ p_{RR} \\ p_{ER} \\ p_{RE} \\ m_{EE} \\ m_{RR} \\ m_{ER} \\ m_{RE} \end{pmatrix} = \boldsymbol{T}\otimes\boldsymbol{G}\ ,\ var\begin{pmatrix} d_{EE} \\ d_{RR} \\ d_{ER} \\ d_{RE} \end{pmatrix} = \boldsymbol{Q}\otimes\boldsymbol{D}\ ,\ var\begin{pmatrix} r_{EE} \\ r_{RR} \\ r_{ER} \\ r_{RE} \end{pmatrix} = \boldsymbol{R}\otimes\boldsymbol{I} \text{ and } var\begin{pmatrix} e_{EE} \\ e_{RR} \\ e_{ER} \\ e_{RE} \end{pmatrix} = \boldsymbol{E}\otimes\boldsymbol{I}$$

where **G** and **D** are the gametic and dominance relationship matrix (Smith, 1984), **I** is the identity matrix, and

$$\boldsymbol{T} = \begin{bmatrix} \sigma^2_{P_{EE}} & 0 & \sigma_{P_{EE}P_{ER}} & 0 & \sigma_{P_{EE}M_{EE}} & 0 & 0 & \sigma_{P_{EE}M_{RE}} \\ 0 & \sigma^2_{P_{RR}} & 0 & \sigma_{P_{RR}P_{RE}} & 0 & \sigma_{P_{RR}M_{RR}} & \sigma_{P_{RR}M_{ER}} & 0 \\ \sigma_{P_{EE}P_{ER}} & 0 & \sigma^2_{P_{ER}} & 0 & \sigma_{P_{ER}M_{EE}} & 0 & 0 & \sigma_{P_{ER}M_{RE}} \\ 0 & \sigma_{P_{RR}P_{RE}} & 0 & \sigma^2_{P_{RE}} & 0 & \sigma_{P_{RE}M_{RR}} & \sigma_{P_{RE}M_{ER}} & 0 \\ \sigma_{P_{EE}M_{EE}} & 0 & \sigma_{P_{ER}M_{EE}} & 0 & \sigma^2_{M_{EE}} & 0 & 0 & \sigma_{M_{EE}M_{RE}} \\ 0 & \sigma_{P_{RR}M_{RR}} & 0 & \sigma_{P_{RE}M_{RR}} & 0 & \sigma^2_{M_{RR}} & \sigma_{M_{RR}M_{ER}} & 0 \\ 0 & \sigma_{P_{RR}M_{ER}} & 0 & \sigma_{P_{RE}M_{ER}} & 0 & \sigma_{M_{RR}M_{ER}} & \sigma^2_{M_{ER}} & 0 \\ \sigma_{P_{EE}M_{RE}} & 0 & \sigma_{P_{ER}M_{RE}} & 0 & \sigma_{M_{EE}M_{RE}} & 0 & 0 & \sigma^2_{M_{RE}} \end{bmatrix}$$

where $\sigma^2_{X_{JK}}$ represents the paternal (X=P) or maternal (X=M) gametic variance in the JK population, JK can represent any of the following combinations *{EE, ER, RR, RE}*. Similarly, $\sigma_{X_{JK}Y_{LM}}$ denotes the covariance between the gametic effects between the paternal (if *X=P*) or maternal (if *X=M*) gametic variance in the *JK* cross with the paternal (if *Y=P*) or maternal (if *Y=M*) gametic variance in the LM cross, where *JK* and *LM* can each be *EE,ER,RR* and *RE*, respectively. It is essential to emphasize that in this analysis, the covariances between gametic effects from distinct populations (Retinto or Entrepelado) are explicitly fixed at zero, allowing no correlation between them. Conversely, the model allows for non-null covariance between gametic effects from the same population.

Further,

$$\boldsymbol{Q} = \begin{bmatrix} \sigma^2_{d_{EE}} & 0 & 0 & 0 \\ 0 & \sigma^2_{d_{RR}} & 0 & 0 \\ 0 & 0 & \sigma^2_{d_{ER}} & 0 \\ 0 & 0 & 0 & \sigma^2_{d_{RE}} \end{bmatrix}, \boldsymbol{R} = \begin{bmatrix} \sigma^2_{r_{EE}} & 0 & 0 & 0 \\ 0 & \sigma^2_{r_{RR}} & 0 & 0 \\ 0 & 0 & \sigma^2_{r_{ER}} & 0 \\ 0 & 0 & 0 & \sigma^2_{r_{RE}} \end{bmatrix}, \text{ and}$$

$$\boldsymbol{E} = \begin{bmatrix} \sigma^2_{e_{EE}} & 0 & 0 & 0 \\ 0 & \sigma^2_{e_{RR}} & 0 & 0 \\ 0 & 0 & \sigma^2_{e_{ER}} & 0 \\ 0 & 0 & 0 & \sigma^2_{e_{RE}} \end{bmatrix}$$

with $\sigma^2_{d_{JK}}$, $\sigma^2_{r_{JK}}$ and $\sigma^2_{e_{JK}}$ are the dominance, permanent environmental and residual variances of the *JK* population, respectively.

The inverse of the gametic relationship matrix ($\mathbf{G}^{-1}$) was calculated using a FORTRAN program that follows the algorithm proposed by Meyer & Tier (2012). The gametic relationship matrix (**G**) was itself calculated by direct inversion using R

software (R Core Team., 2021). Subsequently, the elements of **G** were employed to derive the elements of the dominance relationship matrix (**D**) according to the following expression:

$$d_{ij} = g_{ipjp}g_{imjm} + g_{ipjm}g_{imjp}$$

Here, d_{ij} represents the dominance relationship between the i^{th} and j^{th} individuals. Specifically, g_{ixjy} signifies the relationship between the gametes of the i^{th} individual (paternal or maternal, depending on x) and the j^{th} individual (paternal or maternal, depending on y).

The implementation of the Gibbs Sampler was conducted using the BLUPF90 suite of programs, specifically with the gibbsf90+ program (Misztal et al., 2018). The analysis involved a single long chain comprising 1,100,000 iterations, with the initial 100,000 iterations discarded to guarantee convergence towards the stationary distribution.

During each iteration, the following parameters were calculated:

1. Ratios of paternal (p^2_{JK}) or maternal (m^2_{JK}) gametic variances for the JK cross

$$p^2_{JK} = \frac{\sigma^2_{p_{JK}}}{\sigma^2_{p_{JK}} + \sigma^2_{m_{JK}} + 2\sigma_{pJKmJK} + \sigma^2_{d_{JK}} + \sigma^2_{r_{JK}} + \sigma^2_{e_{JK}}}$$

$$m^2_{JK} = \frac{\sigma^2_{m_{JK}}}{\sigma^2_{p_{JK}} + \sigma^2_{m_{JK}} + 2\sigma_{pJKmJK} + \sigma^2_{d_{JK}} + \sigma^2_{r_{JK}} + \sigma^2_{e_{JK}}}$$

2. Ratio of dominance variance (d^2_{JK}) for the JK cross

$$d^2_{JK} = \frac{\sigma^2_{d_{JK}}}{\sigma^2_{p_{JK}} + \sigma^2_{m_{JK}} + 2\sigma_{pJKmJK} + \sigma^2_{d_{JK}} + \sigma^2_{r_{JK}} + \sigma^2_{e_{JK}}}$$

3. Broad sense heritability (H^2_{JK}) for the JK cross.

$$H^2_{JK} = \frac{\sigma^2_{p_{JK}} + \sigma^2_{m_{JK}} + 2\sigma_{pJKmJK} + \sigma^2_{d_{JK}}}{\sigma^2_{p_{JK}} + \sigma^2_{m_{JK}} + 2\sigma_{pJKmJK} + \sigma^2_{d_{JK}} + \sigma^2_{r_{JK}} + \sigma^2_{e_{JK}}}$$

4. Gametic correlations ($r_{X_{JK}Y_{LM}}$) between the X (paternal –P- or maternal –M) gametic effects of the from JK population with the Y ((paternal –P- or maternal –M-) gametic effects of the LM population

$$r_{X_{JK}Y_{LM}} = \frac{\sigma_{X_{JK}Y_{LM}}}{\sigma_{X_{JK}}\sigma_{Y_{LM}}}$$

3. Results And Discussion

Variance Components. The posterior means (and standard deviations) of the variance components for TNB and NBA are presented in Tables 17 and 18, respectively.

Table 17: Posterior means (and standard deviations) of the permanent environmental (σ_r^2), paternal (σ_p^2), maternal (σ_m^2), dominance (σ_d^2) and residual (σ_e^2) variance components for Total Number Born (TNB) in the Entrepelado (EE), Entrepelado x Retinto (ER), Retinto x Entrepelado (RE) and Retinto (RR) populations.

Population	σ_r^2	σ_p^2	σ_m^2	σ_d^2	σ_e^2
EE	0.18 (0.07)	0.14 (0.06)	0.28 (0.07)	0.14 (0.07)	3.50 (0.07)
RR	0.11 (0.04)	0.14 (0.04)	0.20 (0.05)	0.10 (0.04)	4.08 (0.07)
ER	0.14 (0.06)	0.26 (0.09)	0.36 (0.09)	0.13 (0.07)	4.13 (0.09)
RE	0.16 (0.07)	0.15 (0.06)	0.38 (0.10)	0.14 (0.07)	4.17 (0.12)

Table 18: Posterior means (and standard deviations) of the permanent environmental (σ_r^2), paternal (σ_p^2), maternal (σ_m^2), dominance (σ_d^2) and residual (σ_e^2) variance components for Number Born Alive (NBA) in the Entrepelado (EE), Entrepelado x Retinto (ER), Retinto x Entrepelado (RE) and Retinto (RR) populations.

Population	σ_r^2	σ_p^2	σ_m^2	σ_d^2	σ_e^2
EE	0.19 (0.07)	0.13 (0.05)	0.24 (0.06)	0.11 (0.06)	3.47 (0.07)
RR	0.11 (0.04)	0.14 (0.04)	0.19 (0.05)	0.11 (0.04)	3.86 (0.07)
ER	0.13 (0.06)	0.26 (0.08)	0.33 (0.08)	0.12 (0.06)	4.01 (0.09)
RE	0.16 (0.07)	0.16 (0.07)	0.35 (0.09)	0.14 (0.06)	3.96 (0.11)

To begin with, it is important to emphasize that the posterior estimates of paternal and maternal gametic variances in both traits were found to be greater in the crossbred populations (ER and RE) when contrasted with the purebred populations (EE and RR). This discrepancy could potentially be attributed to a scale effect (Falconer & Mackay,

1996), since the phenotypic variance is also higher in the crossbred populations. Another contributing factor to this discrepancy is the model's allowance for covariance between the two gametic effects within the purebred populations, which adds twice the value of $\sigma_{X_{JK}Y_{LM}}$ to the genetic variance.

Nevertheless, the outcomes of our analysis demonstrate that both paternal and maternal gametic effects contribute to the phenotypic variability of litter size traits, as evidenced by their posterior distributions significantly deviating from zero. However, it is important to acknowledge that the variances of maternal gametic variances consistently surpass those of paternal gametic effects. This suggests that the alleles inherited from the mother exert a more pronounced influence on the phenotypic variation of litter size. These findings are further illustrated by the ratios of paternal to maternal gametic variance, visually depicted in Figure 20 for TNB and Supplementary Figure S4 for NBA.

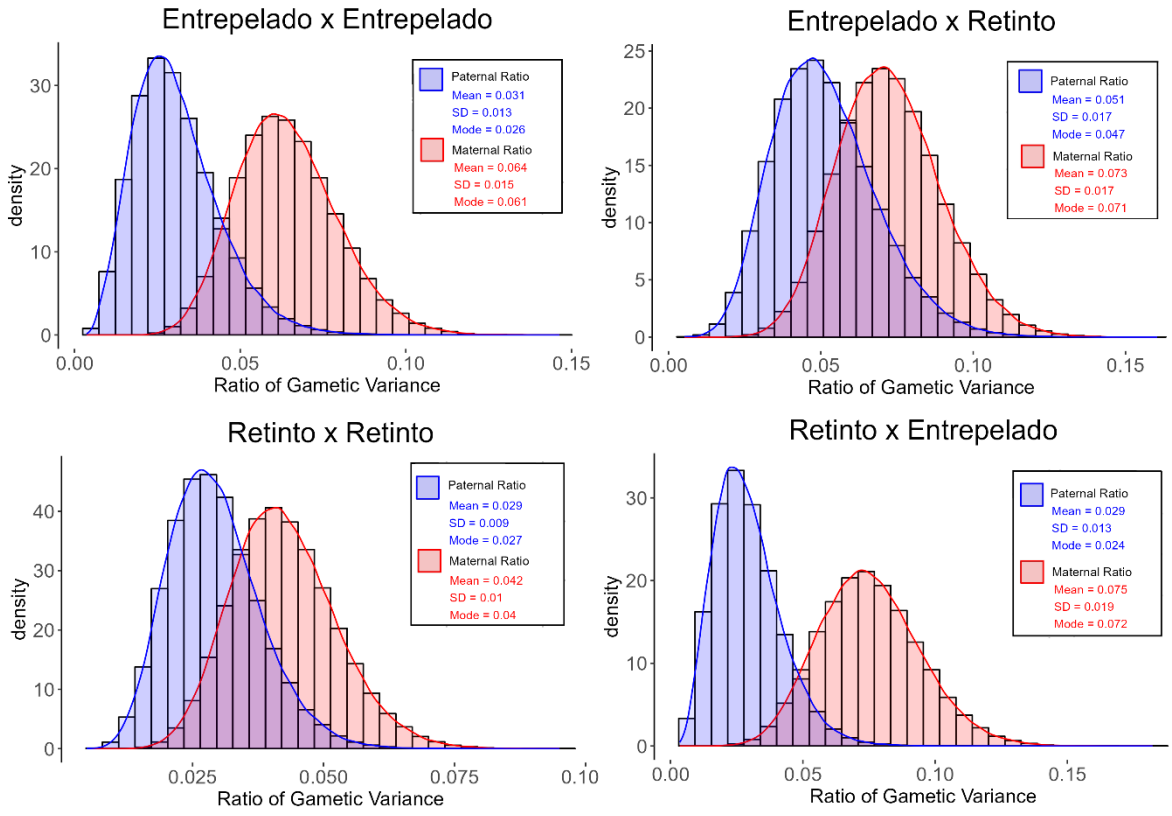


Figure 20: Posterior distributions of the ratios of paternal and maternal gametic variances in the Entrepelado x Entrepelado, Entepelado x Retinto, Retinto x Entrepelado and Retinto x Retinto populations for Total Number Born.

These findings align with those obtained by Stella et al. (2003) in white pigs, suggesting a similar trend. One plausible explanation for this phenomenon is the potential existence of paternal genomic imprinting (Reik & Walter, 2001), a mechanism ensuring certain alleles are only expressed upon heritance from either the mother or father. Numerous theories have been proposed to elucidate the evolutionary origins of genomic imprinting, comprehensive reviewed by Patten et al., (2014). One of the most prominent being the parental investment theory (Moore & Haig, 1991). In accordance with this theory, imprinting arises due to a conflict between the evolutionary interest of the alleles inherited from the father and those inherited from the mother. Within mammalian reproduction, the evolutionary success of the alleles inherited from the father is associated with augmented fetal growth, while the success of the alleles inherited from the mother is linked with offspring number. This theory finds support in the identification of numerous imprinted genes governing diverse facets of mammalian development (Thamban et al., 2020), encompassing growth, behavior, and placental function (Fowden et al., 2011). Additionally, a growing body of evidence indicating the presence of imprinted genes in the pig genome (Coster et al., 2012; Wu et al., 2020; Zhang et al., 2012) is available.

It is important to underline that the paternal and maternal gametic variances exhibited higher values in the crossbred populations when compared to the purebred populations. As mentioned earlier, this divergence could stem from a scale effect or from the inclusion of covariances between the paternal and maternal gametic effects within the purebred populations. Consequently, the covariance between these effects also contributes to the overall genetic variation. Notably, the posterior distributions of the broad-sense heritabilities showed similar patterns across all purebred and crossbred populations, as there are weighted by the increase of phenotypic variation in crossbreds. The posterior mean estimates spanned from 0.125 in the RR population to 0.160 in the EE population for TNB, and from 0.125 (RR) to 0.149 (EE) for NBA, as depicted in Figure 21 for TNB and Supplementary Figure S5 for NBA.

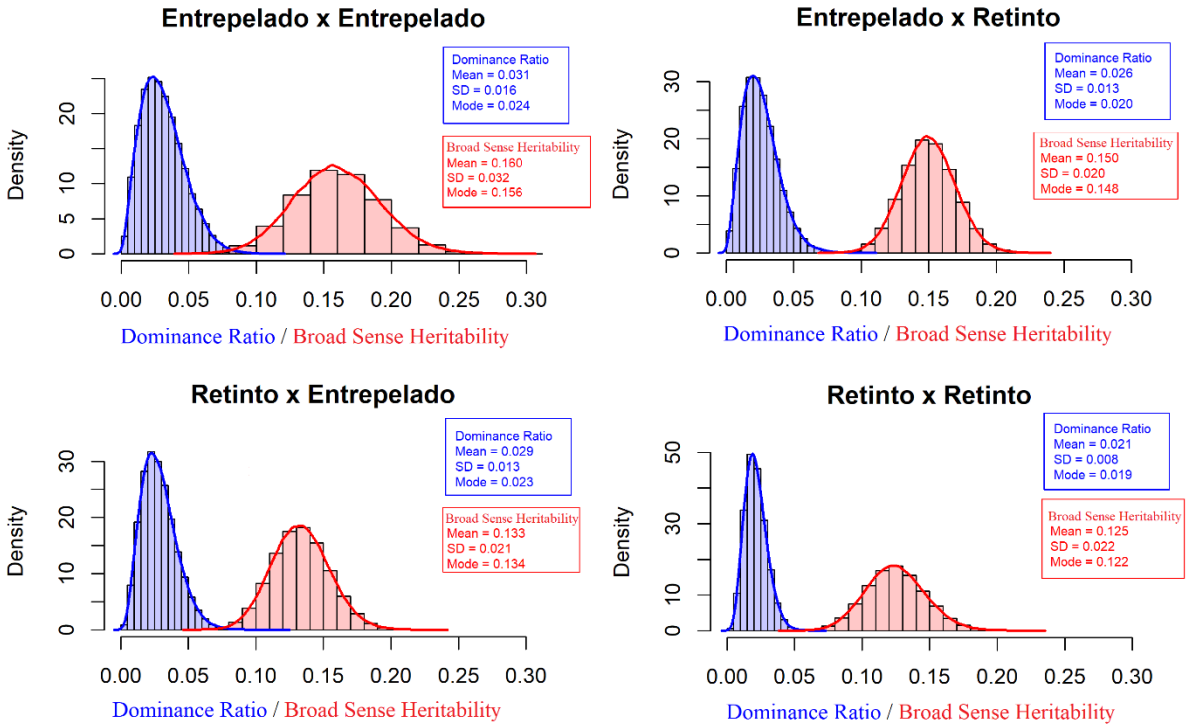


Figure 21: Posterior distributions of the broad-sense heritabilities and ratios of dominance variance in the Entrepelado x Entrepelado, Entrepelado x Retinto, Retinto x Entrepelado and Retinto x Retinto for Total Number Born.

These figures also present the posterior distribution of the ratios of dominance variance, featuring posterior mean estimates ranging from 0.021 in RR to 0.031 in EE for TNB, and from 0.024 in RR to 0.029 in RE for NBA. In general, the ratios of dominance variance were modest, indicating that the implementation of a mate allocation procedure (González-Diéguez et al., 2020; Toro & Varona, 2010) to capture favorable dominance effects may yield limited or insignificant results, even if genotyping information were available.

Gametic Correlations. The proposed model provides the estimates of the gametic covariances and correlations between four gametic effects stemming from each parental population (paternal for purebred, maternal for purebred, paternal for crossbred, and maternal for crossbred). Figures 22 and 23 display the posterior distributions of gametic correlations among these four gametic effects for TNB in the Entrepelado and Retinto populations, respectively. Furthermore, Supplementary Figures S6 and S7 provide the posterior distributions of gametic correlations for NBA.

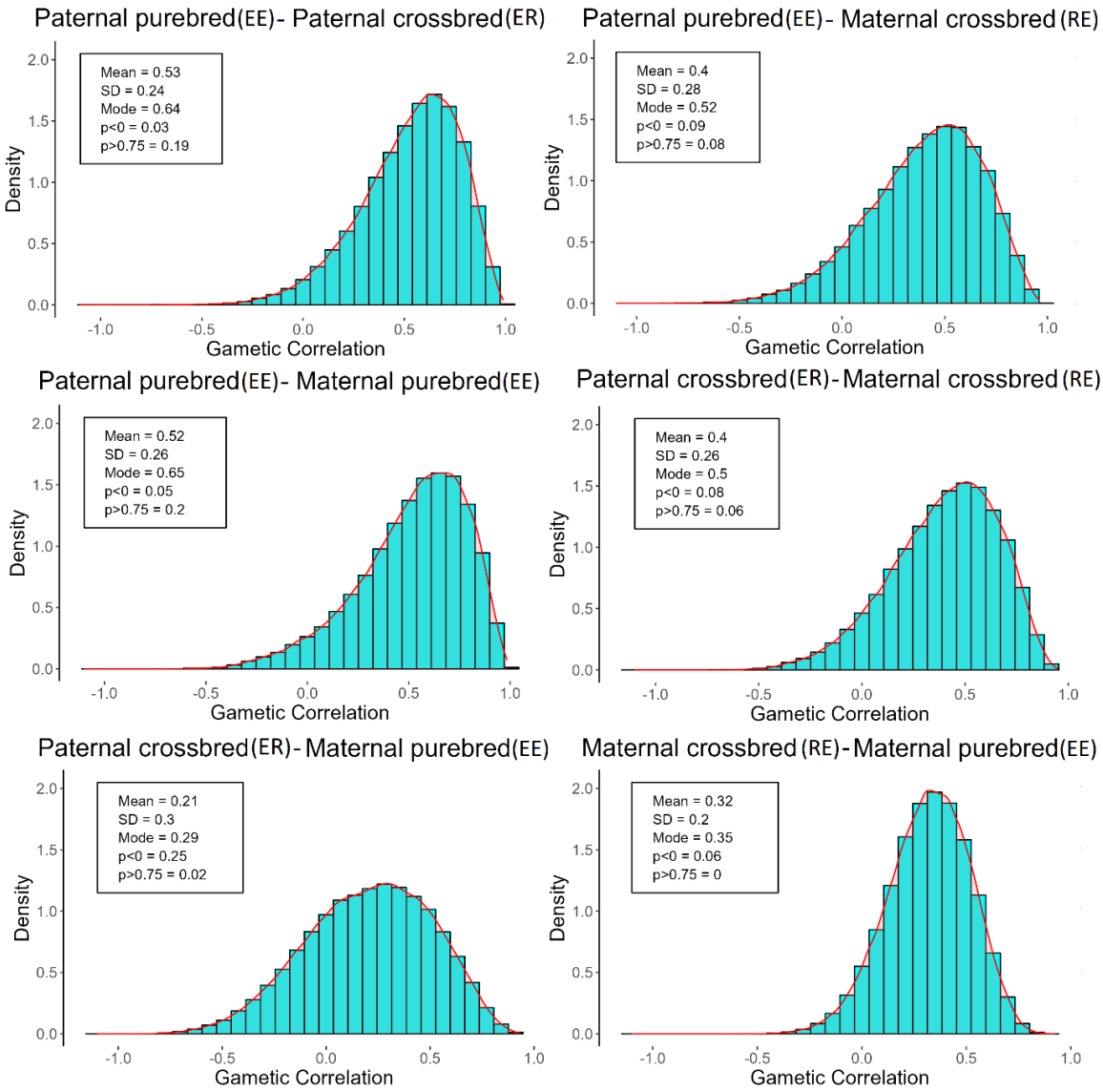


Figure 22: Posterior distribution of the correlations between the four gametic from each parental population (paternal for purebred –EE-, maternal for purebred –EE-, paternal for crossbred –ER-, and maternal for crossbred –RE-) in the Entrepelado population for Total Number Born.

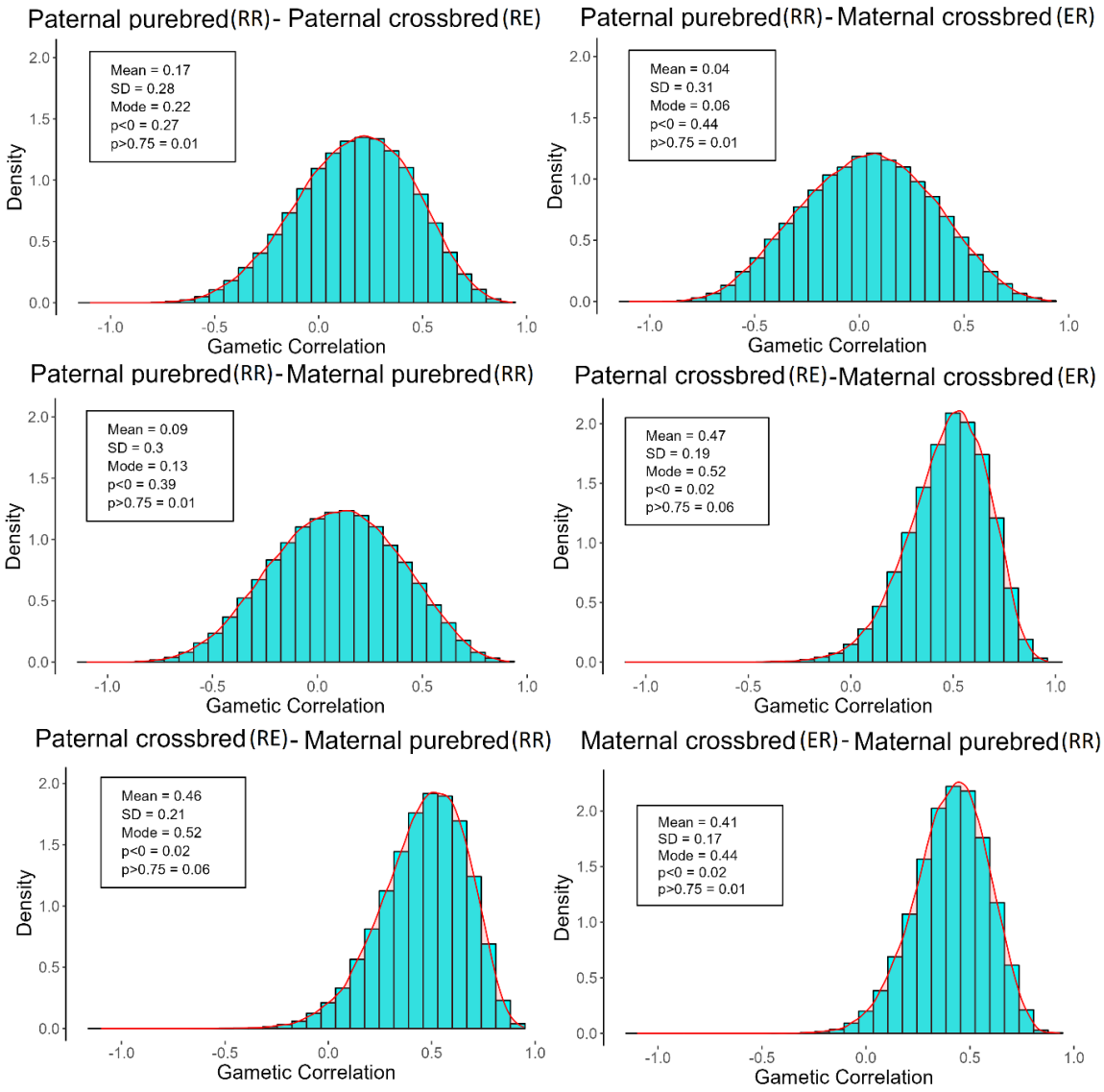


Figure 23: Posterior distribution of the correlations between the four gametic from each parental population (paternal for purebred –RR-, maternal for purebred –RR-, paternal for crossbred –RE-, and maternal for crossbred –ER-) in the Retinto population for Total Number Born.

Every posterior mean estimate of the gametic correlations were positive, spanning from 0.04 (paternal gametic effects for purebred and maternal gametic effects for crossbred in Retinto) to 0.53 (paternal gametic effects for purebred and paternal gametic effects for crossbred in Entrepelado). As far as we know, there are no existing estimates of gametic correlations between purebred and crossbred performance available in the literature. However, our estimates fall within the lower range of the genetic correlation estimates for a wide spectrum of traits available in the literature and reviewed by Wientjes & Calus

(2017). It’s noteworthy that the posterior probability of a gametic correlation surpassing 0.75 was consistently remained below 0.20 for all the gametic correlations. This suggest that selection in the purebred populations might not yield optimal outcomes in the crossbred population. These results differ from those obtained by Srihi et al. (2022); however it's important to recognize that their study was conducted with a notably smaller dataset.

Furthermore, it is interesting to observe that the pattern of gametic correlations differs between purebred and crossbred performance in both populations. Within the Entrepelado population, the gametic effects acting as the sire in the purebred context (paternal for purebred) exhibit relatively high correlations with performance within the crossbred population, whether as a sire (posterior mean of 0.53 with paternal for crossbred) or as a dam (posterior mean of 0.40 with maternal for crossbred). Conversely, the correlations involving the gametic effects when acting as the dam within the purebred population (maternal for purebred) were lower with performance in the crossbred population, both as a sire (posterior mean of 0.21 with paternal for crossbred) and as a dam (posterior mean of 0.32 with maternal for crossbred). In contrast, the scenario in the Retinto population was reversed. The correlations between the gametic effects acting as the sire within the purebred population (paternal for purebred) showed lower correlations with the crossbred population (posterior mean of 0.17 with paternal and 0.04 with maternal gametic effects in the crossbred), whereas the gametic effects acting as the dam (maternal for purebred) displayed elevated correlations (posterior mean of 0.46 with paternal and 0.41 with maternal gametic effects in the crossbred).

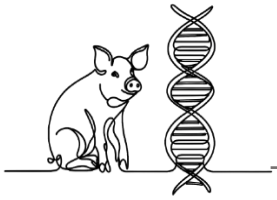
These findings reinforce the importance of using crossbred data to predict the breeding values of purebred individuals and confirm the need to evaluate them for both purebred and crossbred performance. By doing so, balanced selection strategies can be optimized for crossbreeding purposes. Moreover, the observed diversity in the correlation between gametic effects also suggest the potential influence of imprinting effects, which should be considered in genetic evaluation. These results open up alternative strategies for crossbreeding selection and breeding program design.

The populations that contribute as sires or dams in crossbreeding should be selected based on the prediction of their gametic effects when performing as sires or dams in the crossbred population, respectively. In the example provided, the CASTUA population is

commercially produced by crossing Retinto sires with Entrepelado dams. Hence, the selection of the Retinto population should rely on gametic prediction of paternal gametic effects for crossbred, while the Entrepelado population should be selected for maternal gametic effects for crossbred.

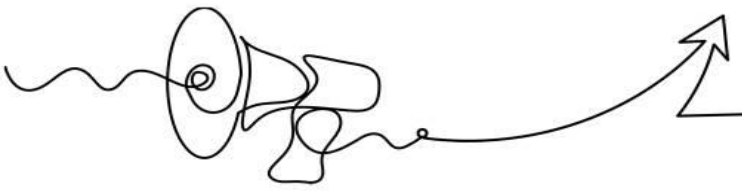
In the proposed model, the performance within the purebred populations contributes to the prediction of gametic effects in crossbreeding via gametic covariances (or correlations). Nonetheless, it should be noted that the covariances (or correlations) between gametic effects from different populations (Retinto or Entrepelado in this example) were assumed to be zero. Future research is required to develop a model that integrates genomic information and accounts for potential covariances among these gametic effects, potentially through adaptations of the metafounders analysis (Xiang et al., 2017).

The findings of this study can be summarized as follows: 1) maternal gametic effects consistently exhibit greater variances than paternal gametic effects, indicating a stronger influence of alleles inherited from the mother on litter size, and 2) distinct patterns of gametic correlations were observed between purebred and crossbred performances within the Entrepelado and Retinto pig populations. These results suggest the potential impact of imprinting effects on the genetic regulation of litter size and underscore the importance of including crossbred data in breeding value predictions for purebred individuals



CHAPTER VIII

General Discussion



Crossbreeding is commonly employed technique in commercial breeding programs spanning diverse animal species, aiming to leverage heterosis and complementarity (Falconer & Mackay, 1996). In white pigs, the most frequent strategy involves a three-way cross, primarily designed to attain heterosis for reproductive traits in the crossbred sow. Simultaneously, this strategy seeks complementarity with the terminal sire concerning growth, carcass traits or meat quality (Cassady et al., 2002).

The Iberian pig presents a distinct scenario. A large proportion of Iberian pig farmers uses Duroc sires to enhance growth rates and efficiency (Serrano et al., 2008). The adoption of intensive management practices has underscored the pivotal role of the reproductive efficiency in ensuring sustainability of Iberian pig farming. Quinton et al. (2006) emphasized that the number of weaned piglets per sow in a year (or numerical productivity) is one of the most important factors in the profitability of pig farms, and litter size is its most important component.

In contrast to other commercial pig populations where efforts to enhance prolificacy in maternal lines through crossbreeding have been commonplace, Iberian pig farmers historically refrained from utilizing crossbred sows. Instead, they traditionally have bred exclusively one of the varieties of the Iberian breed (Martínez et al., 2000). However, several studies have revealed genetic variability for prolificacy within (Fernández et al., 2008; Rodriguez et al., 1994) and between (García-Casco et al., 2012) varieties of Iberian pig. Regulatory norms of the Iberian pig production mandate the use purebred Iberian sows. Consequently, INGA FOOD, S.A. initiated a diallel cross involving three varieties (Retinto, Entrepelado and Torbiscal) to identify the best hybrid sow of complete Iberian genetic origin. This research demonstrated that the cross between Retinto and Entrepelado yielded the highest prolificacy and a notable percentage of heterosis (Noguera et al., 2019). Since then, INGA FOOD, S. A. has started a breeding program to improve the purebred Retinto and Entrepelado populations with a focus on improving litter size. The ongoing breeding program relies solely on purebred phenotypic information to predict the breeding values of the candidates to selection.

The efficacy of this approach hinges on the genetic correlation between purebred and crossbred performance (r_{pc}). This correlation is a key determinant in the response to selection for crossbred performance when the selection is based on purebred (PB) performance metrics. When r_{pc} is low, the integration of crossbred information becomes

increasingly important in the selection process for choosing the best purebred animals. Despite extensive research establishing purebred-crossbred correlation in several species (Calus et al., 2023; Wientjes & Calus, 2017), there exist a notable gap when it comes to Iberian pigs. The absence of specific research in this domain poses a challenge in accurately predicting and enhancing breeding outcomes for Iberian pigs.

One of the primary objectives of the initial study in this thesis is to address this gap. Additionally, the investigation conducted in Chapter V is designed to estimate the additive and dominance genetic variances and covariances within both purebred and crossbred populations. The study also seeks to pinpoint the genomic regions linked to additive genetic variation in both types of performance. To achieve this, the study follows the procedure described by Vitezica et al. (2016), which adopts a genotypic or "biological" model and necessitates the transformation of the results into the "statistical" model to estimate additive and dominance variance components. Further, the procedure was expanded with an adaptation of the method for the calculation of the SNP effects as described by Wang et al. (2012).

The key findings of this study revealed higher additive genetic variances and heritabilities in crossbred populations compared to purebred populations. This suggests that the genetic response in crossbred performance may surpass that of purebred performance. Furthermore, positive purebred-crossbred genetic correlations (0.663 in Entrepelado and 0.881 in Retinto) were observed, falling within the range of correlations estimated in other pig populations (Wientjes & Calus, 2017). otably, several genomic regions on chromosomes 6, 8 (two regions), and 12 were identified, explaining a substantial percentage of the additive genetic variation. The analysis highlighted potential candidate genes, including BCO1 (*β-Carotene Oxygenase* 1), GCSH (*Glycine Cleavage System H*), PRDM8 (*PR/SET Domain 8*), GIT1 (*G protein-coupled receptor kinase interactor 1*), NSRP1 (*Nuclear Speckle Splicing Regulatory Protein 1*), and ANKRD1 (*Ankyrin Repeat Domain 1*). These genes play pivotal roles in essential biological processes such as embryonic development, cell growth, and morphogenesis (Beer et al., 2005; Haub & Goldfarb, 1991; Kim et al., 2011; Segura et al., 2007; Wassef et al., 2013). It is noteworthy that the GCSH (*Glycine Cleavage System H*) protein plays an important role in embryonic viability (Leung et al., 2021). These genomic regions could potentially be given higher weight in the implementation of Weighted GBLUP (Fragomeni et al., 2017; Zhang et al., 2016). However, it's crucial to recognize that these results should be treated

as preliminary due to the constraints of the available genotypic and phenotypic dataset. Additionally, the routine application of genomic selection techniques at INGA FOOD, S.A. is still in its early stages. Furthermore, the practical implementation of Weighted GBLUP is a subject of ongoing discussion (Ren et al., 2021), and its advantages have been demonstrated primarily when specific regions with substantial variance are involved in the genetic determinism of the trait (Ren et al., 2021).

The finding of this first study suggest that the purebred selection should be efficiently for enhancing crossbreeding performance. However, the results must be taken with caution, because the model described by Vitezica et al. (2016) assumes the equivalence of the phenotypic performance in reciprocal crosses, that was not observed in a previous study (Noguera et al., 2019). Motivated by the intriguing results of this study, our research aimed to further investigate parent of origin effects in Iberian pigs. This exploration was conducted using two distinct methodologies. The initial study (Chapter VI) was developed using genotypic information, while the subsequent one (Chapter VII) focused solely on genealogical and phenotypic information.

The presence of parent of origin effects is often associated with genomic imprinting, a phenomenon wherein genes express themselves It is a phenomenon in which genes are expressed differently depending on whether they are transmitted by paternal or maternal gametes (Reik & Walter, 2001). The first approach developed in this thesis involved harnessing molecular information through a novel Bayesian multivariate gametic model. In recent years, some algorithms have been proposed to develop a genomic analysis of imprinting (Nishio & Satoh, 2015) from the genomic information provided by commercial genotyping devices. In this study, we expand upon their approach to a encompass a multivariate scope, considering the performance of Retinto x Entrepelado and Entrepelado x Retinto as two different but correlated traits. As a result, up to four gametic effects were taken into account: Retinto paternal and maternal gametic effects and Entrepelado paternal and maternal effects. The model also allows for the correlation between the paternal and maternal gametic effects within each population.

To implement this model, haplotype phasing was necessary. After a preliminary study (Srihi et al., 2023) that compare the performance of AlphaPhase (Hickey et al., 2011), FImpute (Sargolzaei et al., 2014) and Findhap (VanRaden et al., 2013), the AlphaPhase software (Hickey et al., 2011) was utilized separately for each chromosome, using

genotypes from both crossbred and purebred individuals, as well as a pedigree consisting of 1601 individuals. We tested nine combinations of 'CoreLength' (200, 300, and 400) and 'CoreTail' (75, 100, and 125), notable differences in the phasing outcomes with an average similarity of 85.77% ± 4.86% across different settings, and only haplotype assignments that coincided in seven or more scenarios were retained for subsequent analysis.

The model was implemented using a Gibbs Sampler with the blupf90+ software (Misztal et al., 2018). Despite the limited size of the phenotypic and genotypic datasets, the posterior distribution of the gametic correlations between paternal and maternal effects exhibited distinct patterns in the Retinto than in the Entrepelado populations. In the Retinto population, the posterior distribution was clearly skewed with a posterior probability of a positive gametic correlation larger than 0.80, while in Entrepelado, the posterior probability of a positive gametic correlation were only 0.50.

However, the analysis using this molecular approach provided results that were not as informative as anticipated. This limitation was attributed to the small dataset size and the constraints of the available molecular information. To overcome these limitations and obtain a more comprehensive understanding to imprinting effects, we proceeded with the second methodology, which harnessed genealogical information through the application of gametic models (Chapter VII). Gametic models (Meyer & Tier, 2012; Schaeffer et al., 1989; Varona et al., 2015) has been widely used for the identify differences in parent-of-origin effects. Still, to the best of our knowledge, they have not been implemented in the context of crossbreeding. Therefore, the initial challenge of the study was to develop a multivariate model with up to eight gametic effects (including Entrepelado as sire of Entrepelado, Entrepelado as dam of Entrepelado, Entrepelado as sire of the crossbred, Entrepelado as dam of the crossbred, Retinto as sire of Retinto, Retinto as dam of Retinto, Retinto as sire of the crossbred and Retinto as dam of the crossbred). Additionaly, the gametic effects within the same population (Entrepelado and Retinto) are allowed to be covariated, as it is illustrated in Figure 24.

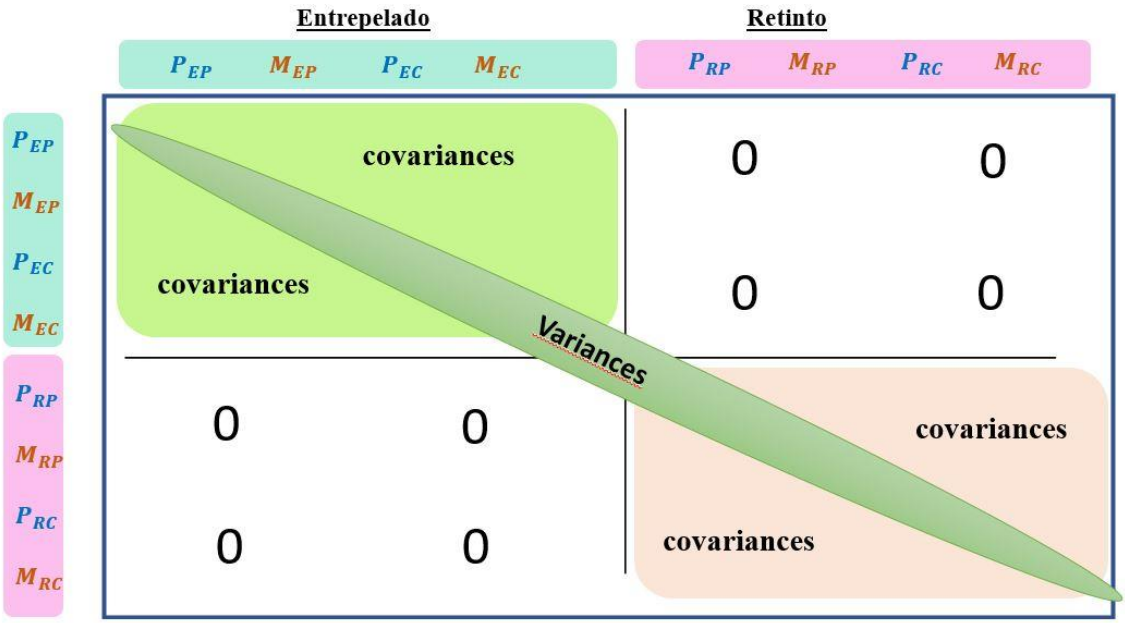


Figure 24: Representation of the structure of the gametic variances and covariances in purebred Retinto and Entrepelado and their reciprocal cross

The model also integrated dominance and the permanent sow environmental effects. The calculation of the gametic and dominance relationship matrices was developed using a self-made program in FORTRAN (see APPENDIX 1), and they were used into the gibbsf90 software (Misztal et al. 1998). The model implementation provides posterior distributions for gametic variances and covariances, as well as the dominance, permanent environmental, and residuals variances. All of them were used to calculate the posterior distributions of the ratios of paternal and maternal gametic variances, the ratios of the dominance variance, the broad-sense heritability and gametic correlations within the gametic effects from the same population.

The results of the study unveiled that the posterior mean variance estimates of the maternal gametic effects consistently surpassed those for paternal effects within all four populations, including both purebred and crossbred. Furthermore, all the posterior mean estimates of the correlations were positive, albeit they ranged between 0.04 (gametic correlation between the paternal effects for purebred and the maternal for crossbred in Retinto) and 0.53 (gametic correlation between the paternal effects for purebred and the paternal for crossbred in Entrepelado).

The imperfect gametic correlations ($r \neq 1$) and the disparities between maternal and paternal gametic variances align consistently with the findings of the preceding study

within this thesis and with those reported by Stella et al. (2003) in with pigs, respectively. Both results align with the hypothesis of the existence of parent-of-origin effects caused by genomic imprinting.

Several theories have been proposed to unravel the evolutionary origins of genomic imprinting (Patten et al., 2014). Among them, one of the most prominent is the parental investment theory (Moore & Haig, 1991), which posits that imprinting arises from a conflict between the evolutionary interest of the alleles inherited from the father and those inherited from the mother. In the case of mammalian reproductive traits, the evolutionary success of the alleles inherited from the father is associated with enhanced pre- and postnatal growth, whereas the success of the alleles inherited from the mother is linked to the number of progenies, with litter size plays a prominent role.

From a practical perspective, the findings of this study may also carry some important consequences. If the gametic correlations within the same population deviate significantly from unity, the selection procedure must have adjusted accordingly. For example, if the final product is a crossbred (CASTUA) between Retinto sires and Entrepelado dams, he selection process for the purebred Retinto population should be based on predicting the paternal gametic effects for the crossbred. Simultaneously, the selection of the Entrepelado population should rely on predicting the maternal gametic effects for the crossbred. The remaining gametic effects should not be factored into the selection process, and their contribution would be beneficial only through their gametic correlation.

Limitations of the study and future perspectives

While this thesis provides valuable insights into crossbreeding between Iberian pig populations, there are certain limitations that should be acknowledged. One significant constraint is the limited availability of reference databases with sufficient size and genetic diversity, particularly in the context of Iberian pigs. This limitation affects the robustness and generalizability of the findings.

The research outcomes presented here should be viewed not as conclusions, However, it serves as a starting point for several interesting research topics.

1. Simulation Studies: It would be valuable to conduct a simulation study that compare the selection response achieved using gametic models for evaluating purebred individuals

for crossbred performance with the selection response achieved using the standard animal model.

2. Integrating Genomic and Genealogical Data: The parent-of-origin studies has been developed separately with genomic (Chapter VI) and genealogical (Chapter VII) information. It will be very interesting to combine the genealogical genomic relationship matrix (Meyer & Tier, 2012; Schaeffer et al., 1989) with the genomic gametic relationship matrix (Nishio & Satoh, 2015) under a single-step (Legarra et al., 2014) based procedure.

3. Exploring Genetic Relationships and Metafounders: Both the genomic and the genealogical analysis of parent-of-origin assumed that the Retinto and the Entrepelado effects were unrelated, as they come from different genetic origins. However, as pointed out in the introduction, there are several studies about the genetic differentiation of the Iberian strains (Alonso et al., 2020; Fabuel et al., 2004), that suggest that some of them can be closely related. Therefore, the development of new procedures that consider the potential relationship between the parental population trough the adaptation of the metafounders theory (Legarra et al., 2015; Xiang et al., 2017) to gametic models .

4. Comparative Model Study : Incorporating a comparative study utilizing the models developed by Christensen et al. (2014), following Wei & Van Der Werf (1994), which use genomic information through marker-based partial relationship matrices.

5. Study of Epistatic interactions: The crossbreeding model developed in Chapter V only included additive and dominance effects. However, heterosis and combination ability between populations may be determined by complex genetic interactions. Therefore, the implementation of models that consider epistasis (Vitezica et al., 2017; Varona et al., 2018;) may provide a better understanding of these processes.

CHAPTER IX

CONCLUSIONS

CONCLUSIONS

The conclusions of this study are:

1. Crossbred population between Retinto and Entrepelado exhibit higher additive genetic variance and heritability than purebred populations, suggesting that crossbreds might be more responsive to selection.

2. There was a higher genetic correlation between purebred and crossbred performances in Retinto pigs compared to Entrepelado pigs.

3. Four genomic regions on chromosomes SSC6, SSC8, and SSC12 were associated with a higher percentage of additive genetic variance, and they harbor several candidate genes, including BCO1 (*β-Carotene Oxygenase 1*), GCSH (*Glycine Cleavage System H*), PRDM8 (*PR/SET Domain 8*), GIT1 (*G protein-coupled receptor kinase interactor 1*), NSRP1 (*Nuclear Speckle Splicing Regulatory Protein 1*), and ANKRD1 (*Ankyrin Repeat Domain 1*).

4. The genomic and genealogical analysis of the data from a reciprocal cross between the Entrepelado and Retinto detected a relevant deviation from the expected perfect correlation between paternal and maternal gametic effects and suggest the presence of imprinting effects in litter size.

5. The maternal gametic effects have shown consistently greater variances than paternal effects, indicating a stronger influence of alleles inherited from the mother on litter size traits

6. The presence of parent of origin effects indicates that selection strategies can be modified to achieve a better selection response by using gametic models.

7. The importance of crossbred phenotypic information for the implementation of selection of crossbred performance is reinforced.

CONCLUSIÓNES

Las conclusiones de este estudio son las siguientes:

1. La población cruzada entre Retinto y Entrepelado presenta una mayor varianza genética aditiva y heredabilidad que las poblaciones de pura raza, lo que sugiere que los cruzados podrían ser más receptivos a la selección.
2. Se ha identificado una mayor correlación genética entre el rendimiento de pura raza y cruzado en cerdos Retinto en comparación con los cerdos Entrepelado.
3. Cuatro regiones genómicas en los cromosomas SSC6, SSC8 y SSC12 se asociaron con un mayor porcentaje de varianza genética aditiva y albergan varios genes candidatos, incluyendo BCO1 (β-Caroteno Oxigenasa 1), GCSH (Sistema de Clivaje de Glicina H), PRDM8 (Dominio PR/SET 8), GIT1 (Interactor 1 de Quinasa Receptora Acoplada a Proteína G), NSRP1 (Proteína Reguladora de Empalme de Espeque Nuclear 1) y ANKRD1 (Dominio de Repetición de Anquirina 1).
4. El análisis genómico y genealógico de los datos de un cruce recíproco entre Entrepelado y Retinto detectó una desviación relevante de la correlación perfecta esperada entre los efectos gaméticos paternos y maternos, lo que sugiere la presencia de efectos de impronta en el tamaño de camada.
5. Los efectos gaméticos maternos han mostrado consistentemente mayores varianzas que los efectos paternos, indicando una influencia más fuerte de los alelos heredados de la madre en las características del tamaño de camada.
6. La presencia de efectos de origen parental indica que las estrategias de selección pueden modificarse para lograr una mejor respuesta a la selección utilizando modelos gaméticos.
7. Se refuerza la importancia de la información fenotípica de los cruzados para la implementación de la selección del rendimiento de los cruzados

REFERENCES

Aguinaga, M. A., Gómez-Carballar, F., Nieto, R., & Aguilera, J. F. (2011). Production and composition of Iberian sow's milk and use of milk nutrients by the suckling Iberian piglet. *Animal*, *5*(9), 1390–1397. https://doi.org/10.1017/S1751731111000474

Almeida, J. M., Bressan, M. C., Amaral, A. J., Bettencourt, C., Santos-Silva, J., Moreira, O., & Gama, L. T. (2019). Body weight and ultrasound measurements over the finishing period in Iberian and F1 Large White × Landrace pigs raised intensively or in free-range conditions. *Livestock Science*, *229*, 170–178. https://doi.org/10.1016/j.livsci.2019.09.020

Alonso, I., Ibáñez-Escriche, N., Noguera, J. L., Casellas, J., Martín de Hijas-Villalba, M., Gracia-Santana, M. J., & Varona, L. (2020). Genomic differentiation among varieties of Iberian pig. *Spanish Journal of Agricultural Research*, *18*(1), 1–20. https://doi.org/10.5424/sjar/2020181-15411

Ayuso, D., González, A., Hernández, F., Corral, J. M., & Izquierdo, M. (2013). Prediction of carcass composition, ham and foreleg weights, and lean meat yields of Iberian pigs using ultrasound measurements in live animals. *Journal of Animal Science*, *91*(4), 1884–1892. https://doi.org/10.2527/jas.2012-5357

Ayuso, M., Óvilo, C., Rodríguez-Bertos, A., Rey, A. I., Daza, A., Fenández, A., González-Bulnes, A., López-Bote, C. J., & Isabel, B. (2015). Dietary vitamin A restriction affects adipocyte differentiation and fatty acid composition of intramuscular fat in Iberian pigs. *Meat Science*, *108*, 9–16. https://doi.org/10.1016/j.meatsci.2015.04.017

Barea, R., Nieto, R., Lara, L., García, M. A., Vílchez, M. A., & Aguilera, J. F. (2006). Effects of dietary protein content and feeding level on carcass characteristics and organ weights of Iberian pigs growing between 50 and 100 kg live weight. *Animal Science*, *82*(3), 405–413. https://doi.org/10.1079/ASC200645

Barlow, D. P., & Bartolomei, M. S. (2014). Genomic Imprinting in Mammals. *Cold Spring Harbor Perspectives in Biology*, *6(2), A018382.*

Barton, S. C., Surani, M. A. H., & Norris, M. L. (1984). Role of paternal and maternal genomes in mouse development. *Nature*, *311*(5984), 374–376. https://doi.org/10.1038/311374a0

Beer, H. D., Bittner, M., Niklaus, G., Munding, C., Max, N., Goppelt, A., & Werner, S. (2005). The fibroblast growth factor binding protein is a novel interaction partner of FGF-7, FGF-10 and FGF-22 and regulates FGF activity: Implications for epithelial repair. *Oncogene*, *24*(34). https://doi.org/10.1038/sj.onc.1208560

Benito-Hernández, J., Vázquez Cisneros, C., García Casco, J., Moreno, M., Ferrera Claramunt, C., & Luis, J. (1997). El cerdo Ibérico. *Revista Computadorizada de Producción Porcina*, *4*, 14.

Benito, J., Vázquez, C., Menaya, C., Ferrera, J. L., García-Casco, J., Silió, L., J., R., & M.C., R. (2000). Evaluation of the productive parameters in different strains of Iberian pig. *Tradition and Innovation in Mediterranean Pig Production, Serie A*(41), 113–121.

Bidanel, J. P. (2011). Biology and genetics of reproduction. In *The genetics of the pig* (pp. 218–241). Wallingford UK:CABI. https://doi.org/10.1079/9781845937560.0218

Boletín Oficial del Estado. (2014). *Real Decreto 4/2014 de 10 enero, por el que se aprueba la norma de calidad para la carne, el jamón, la paleta y la caña de lómo ibérico*. *10*.

Browning, B. L., & Browning, S. R. (2008). A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *American Journal of Human Genetics*, *84*(2), 210–223. https://doi.org/10.1016/j.ajhg.2009.01.005

Cabello, R., León, J., & Barba, C. J. (2007). Contribución a la diferenciación morfológica de las variedades del cerdo Ibérico como base para su conservación. El cerdo Ibérico. Caracterización de sus variedades. *1a Ed. Diputación de Córdoba. Córdoba, España*, 31–87.

Calus, M. P. L., Wientjes, Y. C. J., Bos, J., & Duenk, P. (2023). Animal board invited review: The purebred-crossbred genetic correlation in poultry. In *Animal* (Vol. 17, Issue 11). https://doi.org/10.1016/j.animal.2023.100997

Carillier, C., Larroque, H., & Robert-Granié, C. (2014). Comparison of joint versus purebred genomic evaluation in the French multi-breed dairy goat population. *Genetics Selection Evolution*, *46*(1), 1–10. https://doi.org/10.1186/s12711-014-0067-3

Cassady, J. P., Young, L. D., & Leymaster, K. A. (2002). Heterosis and recombination effects on pig reproductive traits. *Journal of Animal Science*, *80*(9). https://doi.org/10.1093/ansci/80.9.2303

Cava, R., Estévez, M., Ruiz, J., & Morcuende, D. (2003). Physicochemical characteristics of three muscles from free-range reared Iberian pigs slaughtered at 90 kg live weight. *Meat Science*, *63*(4), 533–541. https://doi.org/10.1016/S0309-1740(02)00180-8

Cebrián, M. I., Bazán, J., Ayuso, D., & Paniagua, M. (2009). Cebrián, M. I., Bazán, J., Ayuso, D., & Paniagua, M. (2009). Evaluación del crecimiento y del consumo de pienso de lechones ibéricos destetados a diferentes edades. *XIII Jornadas Sobre Producción Animal.Asociación Interprofesional Para El Desarrollo Agrario, Zaragoza.*, 247–249.

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

Clemente, I., del Pozo, A. M., Ortiz, P. J. A., Dorado, G., & Franganillo, A. R Casanova, A. M. (2006). Algunas consideraciones sobre las diferentes clasificaciones del tronco porcino ibérico: una propuesta integradora. *Solo Cerdo Ibérico*, *(16)*, *7-23*.

Clemente, I., Membrillo, A., Azor, P. J., Polvillo, O., Juárez, M., Santos, E., Jiménez, A. M., Diéguez, E., & Molina, A. (2008). Caracterización de la diversidad genética intrarracial del cerdo Ibérico. *XIV Reunión Nacional de Mejora Genética Animal*, *104*(2).

Clemente López, I., Diéguez Garbayo, E., & Forero Vizcaíno, F. J. (2007). *Las razas porcinas autóctonas andaluzas: el cerdo ibérico y sus estirpes. El Manchado de Jabugo*.

Comstock, R. E., Robinson, H. F., & Harvey, P. H. (1949). A Breeding Procedure Designed To Make Maximum Use of Both General and Specific Combining Ability 1. *Agronomy Journal*, *41*(8), 360–367. https://doi.org/10.2134/agronj1949.00021962004100080006x

Conde-Aguilera, J. A., Aguinaga, M. A., Aguilera, J. F., & Nieto, R. (2011). Nutrient and energy retention in weaned iberian piglets fed diets with different protein concentrations. *Journal of Animal Science*, *89*(3), 754–763. https://doi.org/10.2527/jas.2010-3173

Coster, A., Madsen, O., Heuven, H. C. M., Dibbits, B., Groenen, M. A. M., van Arendonk, J. A. M., & Bovenhuis, H. (2012). The imprinted gene DIO3 is a candidate gene for litter size in pigs. *PLoS ONE*, *7*(2), e31825. https://doi.org/10.1371/journal.pone.0031825

Culbertson, M. S., Mabry, J. W., Misztal, I., Gengler, N., Bertrand, J. K., & Varona, L. (1998). Estimation of Dominance Variance in Purebred Yorkshire Swine. *Journal of Animal Science*, *76*(2), 448–451.

Daza, A., López-Bote, C. J., Tomás Barberán, F. A., Espin, J. C., López Carrasco, C., Olivares, A., & Rey, A. I. (2007). Effect of mediterranean forest parasite with Curculio sp. on nutritional value of acorn for Iberian pig feeding and fat characteristics. *Meat Science*, *76*(2), 316–320. https://doi.org/10.1016/j.meatsci.2006.11.013

Daza, A., López-Bote, C., Rey, A., & Olivares, Á. (2006). Effect of age at the beginning of the free-range fattening period on growth and carcass and fat quality in Iberian pigs. *Archives of Animal Nutrition*, *60*(4), 317–324. https://doi.org/10.1080/17450390600785285

Daza, A., Mateos, A., Carrasco, C. L., Rey, A., Ovejero, I., & López-Bote, C. J. (2006). Effect of feeding system on the growth and carcass characteristics of Iberian pigs, and the use of ultrasound to estimate yields of joints. *Meat Science*, *72*(1), 1–8. https://doi.org/10.1016/j.meatsci.2005.04.031

Daza, A., Mateos, A., Rey, A. I., & López Bote, C. J. (2005). Effect of feeding level during the period previous to free-range fattening on growth and carcass characteristics in Iberian pigs. *Spanish Journal of Agricultural Research*, *3*(4), 387. https://doi.org/10.5424/sjar/2005034-165

Daza, A., Rey, A. I., López-Carrasco, C., & López-Bote, C. J. (2008). Influence of acorn size on growth performance, carcass quality and fatty acid composition of subcutaneous and intramuscular fat from Iberian pigs fattened in confinement. *Spanish Journal of Agricultural Research*, *6*(2), 230–235. https://doi.org/10.5424/sjar/2008062-314

De Koning, D. J., Rattink, A. P., Harlizius, B., Van Arendonk, J. A. M., Brascamp, E. W., & Groenen, M. A. (2000). Genome-wide scan for body composition in pigs reveals important role of imprinting. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(14), 7947-7950. https://doi.org/10.1073/pnas.140216397

Dekkers, J. C. M. (2007). Marker-assisted selection for commercial crossbred performance 1. *Journal of Animal Science*, *85*(9), 2104–2114. https://doi.org/10.2527/jas.2006-683

Dobao, M. T., Rodrigañez, J., Silio, L., Toro, M. A., De Pedro, E., & Garcia de Siles, J. L. (1987). Crecimiento y características de canal en cerdos Ibéricos, Duroc-Jersey x Ibérico y Jiaxing x Ibérico. *Investigación Agraria Producción y Sanidad Animales*, *2*(1), 9–23.

Duenk, P., Bijma, P., Wientjes, Y. C. J., & Calus, M. P. L. (2021). Predicting the purebred-crossbred genetic correlation from the genetic variance components in the parental lines. *Genetics Selection Evolution*, *53*(1), 1–18. https://doi.org/10.1186/s12711-021-00601-w

Dunker, A., Rey, A. I., López-Bote, C. J., & Daza, A. (2007). Effect of the feeding level during the fattening phase on the productive parameters, carcass characteristics and quality of fat in heavy pigs. *Journal of Animal and Feed Sciences*, *16*(4), 624. https://doi.org/10.22358/jafs/66819/2007

Eisen, E. J., Hörstgen-Schwark, G., Saxton, A. M., & Bandy, T. R. (1983). Genetic interpretation and analysis of diallel crosses with animals. *Theoretical and Applied Genetics*, *65*(1), 17–23. https://doi.org/10.1007/BF00276256

Esfandyari, H., Sørensen, A. C., & Bijma, P. (2015). Maximizing crossbred performance through purebred genomic selection. *Genetics Selection Evolution*, *47*(1), 1–16. https://doi.org/10.1186/s12711-015-0099-3

Estévez, M., Morcuende, D., & Cava, R. (2003). Oxidative and colour changes in meat from three lines of free-range reared Iberian pigs slaughtered at 90 kg live weight and from industrial pig during refrigerated storage. *Meat Science*, *65*(3), 1139–1146. https://doi.org/10.1016/S0309-1740(02)00343-1

EUROSAT. (2022). *Pig population - annual data*. https://ec.europa.eu/eurostat/databrowser/view/APRO_MT_LSPIG__custom_3428837/bo

okmark/table?lang=en&bookmarkId=a250e9c6-af2e-49cb-a15a-b8a1e12e9117

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., & Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*. Longman Group, Harlow, UK.

Fernández-Fígares, I., Conde-Aguilera, J. A., Nieto, R., Lachica, M., & Aguilera, J. F. (2008). Synergistic effects of betaine and conjugated linoleic acid on the growth and carcass composition of growing Iberian pigs. *Journal of Animal Science*, *86*(1), 102–111. https://doi.org/10.2527/jas.2006-0230

Fernández, A., De Pedro, E., Núñez, N., Silió, L., García-Casco, J., & Rodríguez, C. (2003). Genetic parameters for meat and fat quality and carcass composition traits in Iberian pigs. *Meat Science*, *64*(4), 405-410. https://doi.org/10.1016/S0309-1740(02)00207-3

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.). https://doi.org/10.5424/sjar/200806s1-378

Fernando adell martí. (2018). *Manifiesto de Asaja Córdoba en defensa del cerdo ibérico extensivo*. Vicepresidente 1º DE Asaja Córdoba 12 Noviembre. https://www.diariodesevilla.es/agr_andalucia/Iberico-Extensivo_0_1299170345.html

Fowden, A. L., Coan, P. M., Angiolini, E., Burton, G. J., & Constancia, M. (2011). Imprinted genes and the epigenetic regulation of placental phenotype. *Progress in Biophysics and Molecular Biology*, *106*(1), 281–288. https://doi.org/10.1016/j.pbiomolbio.2010.11.005

Fragomeni, B. O., Lourenco, D. A. L., Masuda, Y., Legarra, A., & Misztal, I. (2017). Incorporation of causative quantitative trait nucleotides in single-step GBLUP. *Genetics Selection Evolution*, *49*(1). https://doi.org/10.1186/s12711-017-0335-0

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

García-Cortés, L. A., & Toro, M. Á. (2006). Multibreed analysis by splitting the breeding values. *Genetics Selection Evolution*, *38*(6), 1–16. https://doi.org/10.1051/gse:2006024

García-Valverde, R., Barea, R., Lara, L., Nieto, R., & Aguilera, J. F. (2008). The effects of feeding level upon protein and fat deposition in Iberian heavy pigs. *Livestock Science*, *114*(2–3), 263–273. https://doi.org/10.1016/j.livsci.2007.05.005

Gelfand, A. E., & Smith, A. F. M. (1990). Sampling-Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association*, *85*(410), 398–409. https://doi.org/10.2307/2289776

Gibson, J. P. (1988). Gametic models for estimation of autosomally inherited effects that are expressed only when received from eigher male or female parent. *Journal of Dairy Science, 71, 143.*

Gilles, G. (2009). Dry cured ham quality as related to lipid quality of raw material and lipid changes during processing: A review. *Grasas y Aceites*, *60*(3), 297–307. https://doi.org/10.3989/gya.130908

Giraldo, J. D. D. V., & Tovar, M. A. . (2001). Análisis de la evolución de los censos y sistemas de producción del cerdo ibérico. *Revista Española de Estudios Agrosociales y Pesqueros*, *32*(193), 87–118. https://doi.org/10.22004/ag.econ.165081

Gómez-Carballar, F., Aguinaga, M. A., Nieto, R., & Aguilera, J. F. (2009). Effects of intermittent suckling on the performance and digestive efficiency of Iberian piglets weaned at 35 days of age. *Livestock Science*, *124*(1–3), 41–47. https://doi.org/10.1016/j.livsci.2008.12.006

Gómez-Carballar, F., Lara, L., Nieto, R., & Aguilera, J. F. (2013). Effect of increasing lysine supply during last third of gestation on reproductive performance of Iberian sows. *Spanish Journal of Agricultural Research*, *11*(3), 798–807. https://doi.org/10.5424/sjar/2013113-4208

González-Diéguez, D., Tusell, L., Bouquet, A., Legarra, A., & Vitezica, Z. G. (2020). Purebred and crossbred genomic evaluation and mate allocation strategies to exploit dominance in pig crossbreeding schemes. *G3: Genes, Genomes, Genetics*, *10*(8), 2829–2841. https://doi.org/10.1534/g3.120.401376

Hanna, C. W. (2020). Placental imprinting: Emerging mechanisms and functions. *PLoS Genetics*, *16*(4), e1008709. https://doi.org/10.1371/journal.pgen.1008709

Haub, O., & Goldfarb, M. (1991). Expression of FGF5 gene in mouse embryo. *Development*, *406*(112), 397–406.

Henderson, C. R. (1976). A Simple Method for Computing the Inverse of a Numerator Relationship Matrix Used in Prediction of Breeding Values. *Biometrics*, *32*(1), 69–83. https://doi.org/10.2307/2529339

Hickey, J. M., Kinghorn, B. P., Tier, B., Wilson, J. F., Dunstan, N., & Van Der Werf, J. H. J. (2011). A combined long-range phasing and long haplotype imputation method to impute phase for SNP genotypes. *Genetics Selection Evolution*, *43*(1), 1–13. https://doi.org/10.1186/1297-9686-43-12

Hill, W. G. (2010). Understanding and using quantitative genetic variation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1537), 73–85. https://doi.org/10.1098/rstb.2009.0203

Hoeschele, I., & VanRaden, P. M. (1991). Rapid Inversion of Dominance Relationship Matrices for Noninbred Populations by Including Sire by Dam Subclass Effects. *Journal of Dairy Science*, *74*(2), 557–569. https://doi.org/10.3168/jds.S0022-0302(91)78203-9

Howie, B. N., Donnelly, P., & Marchini, J. (2009). A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genetics*, *5*(6), e1000529. https://doi.org/10.1371/journal.pgen.1000529

Ibánẽz-Escriche, N., Fernando, R. L., Toosi, A., & Dekkers, J. C. (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

Jiang, B. X., & Groen, A. F. (1999). Combined crossbred and purebred selection for reproduction traits in a broiler dam line. *Journal of Animal Breeding and Genetics*, *116*(2), 111–125. https://doi.org/10.1046/j.1439-0388.1999.00180

Jungerius, B. J., Van Laere, A. S., Te Pas, M. F. W., Van Oost, B. A., Andersson, L., & Groenen, M. A. M. (2004). The IGF2-intron3-G3072A substitution explains a major imprinted QTL effect on backfat thickness in a Meishan x European white pig intercross. *Genetical Research*, *84*(2), 95–101. https://doi.org/10.1017/S0016672304007098

Karoui, S., Carabaño, M. J., Díaz, C., & Legarra, A. (2012). Joint genomic evaluation of French

dairy cattle breeds using multiple-trait models. *Genetics Selection Evolution*, *44*(1), 1–10. https://doi.org/10.1186/1297-9686-44-39

Kelsey, G., & Feil, R. (2013). New insights into establishment and maintenance of DNA methylation imprints in mammals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1609), 20110336. https://doi.org/10.1098/rstb.2011.0336

Kenny, D., Sleator, R. D., Murphy, C. P., Evans, R. D., & Berry, D. P. (2022). Detection of Genomic Imprinting for Carcass Traits in Cattle Using Imputed High-Density Genotype Data. *Frontiers in Genetics*, *13*, 951087. https://doi.org/10.3389/fgene.2022.951087

Kim, Y. D., Lee, J. Y., Oh, K. M., Araki, M., Araki, K., Yamamura, K. I., & Jun, C. D. (2011). NSrp70 is a novel nuclear speckle-related protein that modulates alternative pre-mRNA splicing in vivo. *Nucleic Acids Research*, *39*(10), 4300-4314. https://doi.org/10.1093/nar/gkq1267

Kinameri, E., Inoue, T., Aruga, J., Imayoshi, I., Kageyama, R., Shimogori, T., & Moore, A. W. (2008). Prdm proto-oncogene transcription factor family expression and interaction with the Notch-Hes pathway in mouse neurogenesis. *PLoS ONE*, *3*(12), e3859. https://doi.org/10.1371/journal.pone.0003859

Lee, H. K., Lee, S. S., Kim, T. H., Jeon, G. J., Jung, H. W., Shin, Y. S., Han, J. Y., Choi, B. H., & Cheong, I. C. (2003). Detection of imprinted quantitative trait loci (QTL) for growth traits in pigs. *Asian-Australasian Journal of Animal Sciences*, *16*(8), 1087–1092. https://doi.org/10.5713/ajas.2003.1087

Leenhouwers, J. I., & Merks, J. W. M. (2013). Suitability of traditional and conventional pig breeds in organic and low-input production systems in Europe: Survey results and a review of literature. *Animal Genetic Resources/Ressources Génétiques Animales/Recursos Genéticos Animales*, *53*, 169–184. https://doi.org/10.1017/s2078633612000446

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

Legarra, A., Christensen, O. F., Vitezica, Z. G., Aguilar, I., & Misztal, I. (2015). Ancestral relationships using metafounders: Finite ancestral populations and across population relationships. *Genetics*, *200*(2). https://doi.org/10.1534/genetics.115.177014

Leung, K. Y., De Castro, S. C. P., Galea, G. L., Copp, A. J., & Greene, N. D. E. (2021). Glycine Cleavage System H Protein Is Essential for Embryonic Viability, Implying Additional Function Beyond the Glycine Cleavage System. *Frontiers in Genetics*, *12*, 625120. https://doi.org/10.3389/fgene.2021.625120

Lopez-Bote, C. J. (1998). Sustained utilization of the Iberian pig breed. *Meat Science*, *49*(98), S17–S27. https://doi.org/10.1016/s0309-1740(98)90036-5

López, I. D. L. C. (2009). El cerdo ibérico autóctono de los Pedroches: La variedad negro de los Pedroches o Pelón de Belalcázar. *Sólo Cerdo Ibérico*, *22*, 19–36.

Lourenco, D., Legarra, A., Tsuruta, S., Masuda, Y., Aguilar, I., & Misztal, I. (2020). Single-step genomic evaluations from theory to practice: using snp chips and sequence data in blupf90. *Genes*, *11*(7), 790. https://doi.org/10.3390/genes11070790

Mackay, D. J. G., & Temple, I. K. (2017). Human imprinting disorders: Principles, practice, problems and progress. *European Journal of Medical Genetics*, *60*(11), 618–626. https://doi.org/10.1016/j.ejmg.2017.08.014

MAPA. (2016). *Raza porcina IBÉRICO*. https://www.mapa.gob.es/es/ganaderia/temas/zootecn

MAPA. (2022). *El sector de la carne de cerdo en cifras:Principales indicadores económicos*. https://www.mapa.gob.es/es/ganaderia/temas/produccion-y-mercados-ganaderos/sectores-ganaderos/porcino/

Martinez-Macipe, M., Rodríguez, P., Izquierdo, M., Gispert, M., Manteca, X., Mainau, E., Hernández, F. I., Claret, A., Guerrero, L., & Dalmau, A. (2016). Comparison of meat quality parameters in surgical castrated versus vaccinated against gonadotrophin-releasing factor male and female Iberian pigs reared in free-ranging conditions. *Meat Science*, *111*, 116–121. https://doi.org/10.1016/j.meatsci.2015.09.002

Martínez, A. M., Delgado, J. V., Rodero, A., & Vega-Pla, J. L. (2000). Genetic structure of the Iberian pig breed using microsatellites. *Animal Genetics*, *31*(5), 295–301. https://doi.org/10.1046/j.1365-2052.2000.00645.x

Mata, C., Pardo, J., Barba, C., Rodero, A., Delgado, J. V, Molina, A., Diéguez, E., & Cañuelo, P. (1998). Estudio morfométrico en las variedades negras del cerdo ibérico. *Archivos de Zootecnia*, *47*(178–179), 287–290.

Mayoral, A. I., Dorado, M., Guillén, M. T., Robina, A., Vivo, J. M., Vázquez, C., & Ruiz, J. (1999). Development of meat and carcass quality characteristics in Iberian pigs reared outdoors. *Meat Science*, *52*(3), 315–324. https://doi.org/10.1016/S0309-1740(99)00008-X

Mercasa. (2022). *Number of pigs in Spain 2020, by autonomous community (in 1,000 heads) [Graph]*. In Statista Retrieved November 09, 2023, from https://www.statista.com/statistics/765054/cabin-spanish-from-porcine-by-dc-aa-in-spain/.

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*(4), 1819–1829. https://doi.org/10.1093/genetics/157.4.1819

Meyer, K., & Tier, B. (2012). Estimates of variances due to parent of origin effects for weights of Australian beef cattle. *Animal Production Science*, *52*(4), 215–224. https://doi.org/10.1071/AN11195

Miar, Y., Sargolzaei, M., & Schenkel, F. S. (2017). A comparison of different algorithms for phasing haplotypes using Holstein cattle genotypes and pedigree data. *Journal of Dairy Science*, *100*(4), 2837–2849. https://doi.org/10.3168/jds.2016-11590

Misztal, I., Tsuruta, S., Lourenco, D., Aguilar, I., Legarra, A., & Vitezica, Z. (2018). Manual for BLUPF90 family of programs. *University of Georgia, Athens, USA*, 142.

Misztal, I., Varona, L., Culbertson, M., Gengler, N., Bertrand, J. K., Mabry, J., Lawlo, T. J., & Tassel, C. P. Van. (1998). Studies on the value of incorporating the effect of dominance in genetic evaluations of dairy cattle, beef cattle and swine. *Biotechnologie, Agronomie, Société et Environnement*, *2*(4).

Moore, T., & Haig, D. (1991). Genomic imprinting in mammalian development: a parental tug-of-war. *Trends in Genetics*, *7*(2), 45–49. https://doi.org/10.1016/0168-9525(91)90230-N

Muriel, E., Ruiz, J., Ventanas, J., Petrón, M. J., & Antequera, T. (2004). Meat quality characteristics in different lines of Iberian pigs. *Meat Science*, *67*(2), 299–307. https://doi.org/10.1016/j.meatsci.2003.11.002

Neugebauer, N., Luther, H., & Reinsch, N. (2010). Parent-of-origin effects cause genetic variation in pig performance traits. *Animal*, *4*(5), 672–681. https://doi.org/10.1017/S1751731109991625

Nezer, C., Moreau, L., Brouwers, B., Coppieters, W., Detilleux, J., Hanset, R., Karim, L., Kvasz, A., Leroy, P., & Georges, M. (1999). An imprinted QTL with a major effect on muscle mass and fat deposition maps to the IGF2 locus in pig. *Nature Genetics*, *21*(2),

155–156. https://doi.org/10.1038/5935

Nieto, R., Lara, L., García, M. A., Vílchez, M. A., & Aguilera, J. F. (2003). Effects of dietary protein content and food intake on carcass characteristics and organ weights of growing Iberian pigs. *Animal Science*, *77*(1), 47–56. https://doi.org/10.1017/s1357729800053637

Nieto, R., Martínez-Pérez, M., Haro, A., Lara, L., & Aguilera, J. F. (2015). Effects of protein intake on rate of growth, protein deposition, and carcass traits of heavy Iberian pigs. *Journal of Animal Science*, *93*(7), 3471–3482. https://doi.org/10.2527/jas.2014-8822

Nishio, M., & Satoh, M. (2015). Genomic best linear unbiased prediction method including imprinting effects for genomic evaluation. *Genetics Selection Evolution*, *47*(1), 1–10. https://doi.org/10.1186/s12711-015-0091-y

Noguera, J. L., & Ibáñez-Escriche, N. (2017). La mejora genómica en porcino: 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*, *13*(12), 2765–2772. https://doi.org/10.1017/S1751731119001125

O'Connell, J., Gurdasani, D., Delaneau, O., Pirastu, N., Ulivi, S., Cocca, M., Traglia, M., Huang, J., Huffman, J. E., Rudan, I., McQuillan, R., Fraser, R. M., Campbell, H., Polasek, O., Asiki, G., Ekoru, K., Hayward, C., Wright, A. F., Vitart, V., … Marchini, J. (2014). A General Approach for Haplotype Phasing across the Full Spectrum of Relatedness. *PLoS Genetics*, *10*(4), e1004234. https://doi.org/10.1371/journal.pgen.1004234

O'Doherty, A. M., MacHugh, D. E., Spillane, C., & Magee, D. A. (2015). Genomic imprinting effects on complex traits in domesticated animal species. *Frontiers in Genetics*, *6*, 156. https://doi.org/10.3389/fgene.2015.00156

Ogawa, S., Konta, A., Kimata, M., Ishii, K., Uemoto, Y., & Satoh, M. (2019). Estimation of genetic parameters for farrowing traits in purebred Landrace and Large White pigs. *Animal Science Journal*, *90*(1), 23–28. https://doi.org/10.1111/asj.13120

Pardo, J., Mata, C., Barba, C., Rodero, A., & Molina, A. (1998). Estudio Morfométrico en las Variedades Rojas del Cerdo Ibérico y Manchado de Jabugo. *Archivos de Zootecnia*, *47*, 287–290.

Parrini, S., Sirtori, F., Čandek-Potokar, M., Charneca, R., Crovetti, A., Kušec, I. D., Sanchez, E. G., Cebrian, M. M. I., Garcia, A. H., Karolyi, D., Lebret, B., Ortiz, A., Panella-Riera, N., Petig, M., Jesus da Costa Pires, P., Tejerina, D., Razmaite, V., Aquilani, C., & Bozzi, R. (2023). Prediction of fatty acid composition in intact and minced fat of European autochthonous pigs breeds by near infrared spectroscopy. *Scientific Reports*, *13*(1), 7874. https://doi.org/10.1038/s41598-023-34996-x

Patten, M. M., Ross, L., Curley, J. P., Queller, D. C., Bonduriansky, R., & Wolf, J. B. (2014). The evolution of genomic imprinting: Theories, predictions and empirical tests. *Heredity*, *113*(2), 119–128. https://doi.org/10.1038/hdy.2014.29

Patterson, H. D., & Thompson, R. (1971). Recovery of Inter-Block Information when Block Sizes are Unequal. *Biometrika*, *58*(3), 545. https://doi.org/10.2307/2334389

Perdomo-González, D. I., Varona, L., Molina, A., Laseca, N., & Valera, M. (2023). Quantitative analysis of parent-of-origin effect in reproductive and morphological selection criteria in the Pura Raza Española horse. *Journal of Animal Breeding and Genetics*, *104*(6), 596–606. https://doi.org/10.1111/jbg.12811

Perez-Enciso, M., & Gianola, D. (1992). Estimates of genetic parameters for litter size in six

strains of Iberian pigs. *Livestock Production Science*, *32*(3), 283–293. https://doi.org/10.1016/S0301-6226(12)80007-8

Putz, A. M., Tiezzi, F., Maltecca, C., Gray, K. A., & Knauer, M. T. (2015). Variance component estimates for alternative litter size traits in swine. *Journal of Animal Science*, *93*(11), 5153–5163. https://doi.org/10.2527/jas.2015-9416

Quadro, L., Giordano, E., Costabile, B. K., Nargis, T., Iqbal, J., Kim, Y., Wassef, L., & Hussain, M. M. (2020). Interplay between β-carotene and lipoprotein metabolism at the maternal-fetal barrier. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, *1865*(11), 158591. https://doi.org/10.1016/j.bbalip.2019.158591

Quinton, V. M., Wilton, J. W., Robinson, J. A., & Mathur, P. K. (2006). Economic weights for sow productivity traits in nucleus pig populations. *Livestock Science*, *99*(1). https://doi.org/10.1016/j.livprodsci.2005.06.002

R Core Team. (2021). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*, http://www.R-project.org/.

Reik, W., & Walter, J. (2001). Genomic imprinting: parental influence on the genome. *Nature Reviews Genetics*, *2*(1), 21–32. https://doi.org/10.1038/35047554

Ren, D., An, L., Li, B., Qiao, L., & Liu, W. (2021). Efficient weighting methods for genomic best linear-unbiased prediction (BLUP) adapted to the genetic architectures of quantitative traits. *Heredity*, *126*(2). https://doi.org/10.1038/s41437-020-00372-y

Rey, A. I., & López-Bote, C. J. (2001). Effect of dietary copper and vitamin E supplementation, and extensive feeding with acorn and grass on longissimus muscle composition and susceptibility to oxidation in Iberian pigs. *Journal of Animal Physiology and Animal Nutrition*, *85*(9–10), 281–292. https://doi.org/10.1046/j.1439-0396.2001.00316.x

Rodrigáñez, J., Ibáñez, M. Á. T., Molina, M. C. R., & López, L. S. (1999). Origen y diversidad genética de la estirpe Torbiscal de cerdo ibérico. *Solo Cerdo Ibérico*, *2*, 11–14.

Rodriguez, C., Rodrigañez, J., & Silio, 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

Romero, P. L., Calabroni, T., & Sáiz, N. A. (2016). Evolución de los principales índices técnicos de las explotaciones porcinas ibéricas españolas. *Solo Cerdo Ibérico*, *36*, 38–41.

Rueda, L., & Diéguez, E. (2004). *Manual del cerdo ibérico. AECERIBER. Badajoz*.

Sargolzaei, M., Chesnais, J. P., & Schenkel, F. S. (2014). A new approach for efficient genotype imputation using information from relatives. *BMC Genomics*, *15*(1), 1-12. https://doi.org/10.1186/1471-2164-15-478

Saura, M., Fernández, A., Varona, L., Fernández, A. I., De Cara, M. Á. R., Barragán, C., & Villanueva, B. (2015). Detecting inbreeding depression for reproductive traits in Iberian pigs using genome-wide data. *Genetics Selection Evolution*, *47*(1), 1–9. https://doi.org/10.1186/s12711-014-0081-5

Schaeffer, L. R., Kennedy, B. W., & Gibson, J. P. (1989). The Inverse of the Gametic Relationship Matrix. *Journal of Dairy Science*, *72*(5), 1266–1272. https://doi.org/10.3168/jds.S0022-0302(89)79231-6

Segura, I., Essmann, C. L., Weinges, S., & Acker-Palmer, A. (2007). Grb4 and GIT1 transduce ephrinB reverse signals modulating spine morphogenesis and synapse formation. *Nature Neuroscience*, *10*(3), 301–310. https://doi.org/10.1038/nn1858

Seiquer, I., Granados, P. P., Lopes, M. L., Lara, L., Fígares, I. F., Haro, A., & Nieto, R. (2018).

Performance and carcass characteristics of immunocastrated and surgically castrated Iberian pigs fed diets of different protein concentration. Preliminary results. *Archivos de Zootecnia*, *66*(253), 141–150.

Serra, X., Gil, F., Pérez-Enciso, M., Oliver, M. A., Vázquez, J. M., Gispert, M., Díaz, I., Moreno, F., Latorre, R., & Noguera, J. L. (1998). A comparison of carcass, meat quality and histochemical characteristics of Iberian (Guadyerbas line) and Landrace pigs. *Livestock Production Science*, *56*(3), 215–223. https://doi.org/10.1016/S0301-6226(98)00151-1

Serrano, M. P., Valencia, D. G., Nieto, M., Lázaro, R., & Mateos, G. G. (2008). Influence of sex and terminal sire line on performance and carcass and meat quality of Iberian pigs reared under intensive production systems. *Meat Science*, *78*(4), 420–428. https://doi.org/10.1016/j.meatsci.2007.07.006

Silio, L., Rodriguez, M. C., Rodrigañez, J., & Toro, M. A. (2001). La selección de cerdos ibéricos. In *Porcino Ibérico: Aspectos Claves* (pp. 125–149). Mundi-Press.

Smith, S. (1984). Dominance relationship matrix and inverse for an inbred population. *Dairy ScUnpublished Mimeo, Dept of Dairy Sci, Ohio State University, Columbus*.

Srihi, H., Ibáñez-Escriche, N., Casellas, J., Noguera, J. L., Hernández, P., Martín de Hijas, M., Vázquez-Gómez, M., Negro, S., Rosas, J. P., & Varona, L. (2022). Bayesian analysis of paternal and maternal gametic effects in a reciprocal cross between two Iberian varieties. *Proceedings of the 12th World Congress on Genetics Applied to Livestock Production*, 3180–3183. https://doi.org/10.3920/978-90-8686-940-4_772

Srihi, H., López-Carbonell, .,D Ibañez-Escriche, N., Casellas, J., Hernández, P., Rosas, J. P., & Varona, L. (2023). Comparación de procedimientos para la asignación de faseshaplotípicas en un cruce dialélico entre estirpes de cerdo ibérico. *XX Jornadas de Producción Anima AIDA,Zaragoza*.

Srihi, H., López-Carbonell, D., Ibáñez-Escriche, N., Casellas, J., Hernández, P., Negro, S., & Varona, L. (2023). A Bayesian Multivariate Gametic Model in a Reciprocal Cross with Genomic Information: An Example with Two Iberian Varieties. *Animals*, *13*(10), 1648. https://doi.org/10.3390/ani13101648

Srihi, H., Noguera, J. L., Topayan, V., de Hijas, M. M., Ibañez-Escriche, N., Casellas, J., Vázquez-Gómez, M., Martínez-Castillero, M., Rosas, J. P., & Varona, L. (2022). Additive and dominance genomic analysis for litter size in purebred and crossbred iberian pigs. *Genes*, *13*(1), 12. https://doi.org/10.3390/genes13010012

Stam, F. J., MacGillavry, H. D., Armstrong, N. J., De Gunst, M. C. M., Zhang, Y., Van Kesteren, R. E., Smit, A. B., & Verhaagen, J. (2007). Identification of candidate transcriptional modulators involved in successful regeneration after nerve injury. *European Journal of Neuroscience*, *25*(12), 3629–3637. https://doi.org/10.1111/j.1460-9568.2007.05597.x

Stella, A., Stalder, K. J., Saxton, A. M., & Boettcher, P. J. (2003). Estimation of variances for gametic effects on litter size in Yorkshire and Landrace swine. *Journal of Animal Science*, *81*(9). https://doi.org/10.2527/2003.8192171x

Stuber, C. W., & Cockerham, C. C. (1966). Gene effects and variances in hybrid populations. *Genetics*, *54*(6), 1279–1286. https://doi.org/10.1093/genetics/54.6.1279

Tejerina, D., & García-Torres, S. (2010). Effect of production system and sex on different carcass traits of Iberian pigs. *Options Méditerranéennes. Séries A. Mediterranean Seminars*, *101*, 401–404. http://om.ciheam.org/om/pdf/a101/00006715.pdf

Tejerina, D., García-Torres, S., Cabeza De Vaca, M., Vázquez, F. M., & Cava, R. (2012). Effect

of production system on physical-chemical, antioxidant and fatty acids composition of Longissimus dorsi and Serratus ventralis muscles from Iberian pig. *Food Chemistry*, *133*(2), 293–299. https://doi.org/10.1016/j.foodchem.2012.01.025

Tejerina, D., García-Torres, S., & Cava, R. (2012). Water-holding capacity and instrumental texture properties of m. Longissimus dorsi and m. Serratus ventralis from Iberian pigs as affected by the production system. *Livestock Science*, *148*(1–2), 46–51. https://doi.org/10.1016/j.livsci.2012.05.005

Thamban, T., Agarwaal, V., & Khosla, S. (2020). Role of genomic imprinting in mammalian development. In *Journal of Biosciences* (Vol. 45, Issue 1, pp. 1–21). https://doi.org/10.1007/s12038-019-9984-1

Thorvaldsen, J. L., & Bartolomei, M. S. (2007). SnapShot: Imprinted Gene Clusters. In *cell* (Vol. 130, Issue 5). https://doi.org/10.1016/j.cell.2007.08.033

Toro, M. A., & Varona, L. (2010). A note on mate allocation for dominance handling in genomic selection. *Genetics Selection Evolution*, *42*(1), 33. https://doi.org/10.1186/1297-9686-42-33

Triantaphyllopoulos, K. A., Ikonomopoulos, I., & Bannister, A. J. (2016). Epigenetics and inheritance of phenotype variation in livestock. *Epigenetics and Chromatin*, *9*(1), 31. https://doi.org/10.1186/s13072-016-0081-5

USDA Foreign Agricultural Service. (2023). *Production of pork worldwide from 2013 to 2023 (in million metric tons)[Graph]. In Statista. Retrieved November 28, 2023, from https://www.statista.com/statistics/237570/pork-production-worldwide*.

Van Laere, A. S., Nguyen, M., Braunschweig, M., Nezer, C., Collette, C., Moreau, L., Archibald, A. L., Haley, C. S., Buys, N., Tally, M., Andersson, G., Georges, M., & Andersson, L. (2003). A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature*, *425*(6960). https://doi.org/10.1038/nature02064

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., Null, D. J., Sargolzaei, M., Wiggans, G. R., Tooker, M. E., Cole, J. B., Sonstegard, T. S., Connor, E. E., Winters, M., van Kaam, J. B. C. H. M., Valentini, A., Van Doormaal, B. J., Faust, M. A., & Doak, G. A. (2013). Genomic imputation and evaluation using high-density Holstein genotypes. *Journal of Dairy Science*, *96*(1), 668–678. https://doi.org/10.3168/jds.2012-5702

Varona, L., & Misztal, I. (1999). Prediction of parental dominance combinations for planned matings, methodology, and simulation results. *Journal of Dairy Science*, *82*(10), 2186–2191. https://doi.org/10.3168/jds.S0022-0302(99)75463-9

Varona, L., Munilla, S., Casellas, J., Moreno, C., & Altarriba, J. (2015). Consequences of paternally inherited effects on the genetic evaluation of maternal effects. *Genetics Selection Evolution*, *47*(1), 1–13. https://doi.org/10.1186/s12711-015-0141-5

Varona, L., Noguera, J., Casellas, J., de Hijas, M., Rosas, J., & Ibáñez-Escriche, N. (2020). A cross-specific multiplicative binomial recursive model for the analysis of perinatal mortality in a diallel cross among three varieties of Iberian pig. *Scientific Reports*, *10*(1), 21190. https://doi.org/10.1038/s41598-020-78346-7

Varona, L., Legarra, A., Toro, M. A., & Vitezica, Z. G. (2018). Non-additive effects in genomic selection. In *Frontiers in Genetics* (Vol. 9, Issue MAR). https://doi.org/10.3389/fgene.2018.00078

Vitezica, Z. G., Legarra, A., Toro, M. A., & Varona, L. (2017). Orthogonal estimates of variances for additive, dominance, and epistatic effects in populations. *Genetics*, *206*(3). https://doi.org/10.1534/genetics.116.199406

Vitezica, Z. G., Varona, L., Elsen, J. M., Misztal, I., Herring, W., & Legarra, A. (2016). Genomic BLUP including additive and dominant variation in purebreds and F1 crossbreds, with an application in pigs. *Genetics Selection Evolution*, *48*(1), 6. https://doi.org/10.1186/s12711-016-0185-1

Vitezica, Z. G., Varona, L., & Legarra, A. (2013). On the additive and dominant variance and covariance of individuals within the genomic selection scope. *Genetics*, *195*(4), 1223–1230. https://doi.org/10.1534/genetics.113.155176

Wang, H., Misztal, I., Aguilar, I., Legarra, A., & Muir, W. M. (2012). Genome-wide association mapping including phenotypes from relatives without genotypes. *Genetics Research*, *94*(2), 73–83. https://doi.org/10.1017/S0016672312000274

Wassef, L., Spiegler, E., & Quadro, L. (2013). Embryonic phenotype, β-carotene and retinoid metabolism upon maternal supplementation of β-carotene in a mouse model of severe vitamin A deficiency. *Archives of Biochemistry and Biophysics*, *539*(2), 223–229. https://doi.org/10.1016/j.abb.2013.07.004

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. C. J., & 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

Wu, Y. Q., Zhao, H., Li, Y. J., Khederzadeh, S., Wei, H. J., Zhou, Z. Y., & Zhang, Y. P. (2020). Genome-wide identification of imprinted genes in pigs and their different imprinting status compared with other mammals. *Zoological Research*, *41*(6), 721. https://doi.org/10.24272/J.ISSN.2095-8137.2020.072

Xiang, T., Christensen, O. F., & Legarra, A. (2017). Technical note: Genomic evaluation for crossbred performance in a single-step approach with metafounders. *Journal of Animal Science*, *95*(4). https://doi.org/10.2527/jas2016.1155

Xiang, T., Christensen, O. F., Vitezica, Z. G., & Legarra, A. (2016). Genomic evaluation by including dominance effects and inbreeding depression for purebred and crossbred performance with an application in pigs. *Genetics Selection Evolution*, *48*(1), 1–14. https://doi.org/10.1186/s12711-016-0271-4

Xie, L., Qin, J., Yao, T., Tang, X., Cui, D., Chen, L., Rao, L., Xiao, S., Zhang, Z., & Huang, L. (2023). Genetic dissection of 26 meat cut, meat quality and carcass traits in four pig populations. *Genetics Selection Evolution*, *55*(1), 43. https://doi.org/10.1186/s12711-023-00817-y

Zhang, F. W., Han, Z. Bin, Deng, C. Y., He, H. J., & Wu, Q. (2012). Conservation of genomic imprinting at the NDN, MAGEL2 and MEST loci in pigs. *Genes and Genetic Systems*, *87*(1), 53–58. https://doi.org/10.1266/ggs.87.53

Zhang, X., Lourenco, D., Aguilar, I., Legarra, A., & Misztal, I. (2016). Weighting strategies for single-step genomic BLUP: An iterative approach for accurate calculation of GEBV and GWAS. *Frontiers in Genetics*, *7*(AUG). https://doi.org/10.3389/fgene.2016.00151

APPENDIX

APPENDIX 1 :

```
INTEGER, PARAMETER :: NANIS=6007,NGAMETES=2*NANIS
    REAL (KIND=8), DIMENSION (0:NANIS)::EFF !INBREEDING
    REAL (KIND=8), DIMENSION (0:NGAMETES):: DIAG,PXP,PXG ! STORAGE OF G
INVERSE
    INTEGER, DIMENSION (2,0:NGAMETES):: GED ! GAMETIC PEDIGREE
    REAL (KIND=8) :: RLOG1
        eff=0
        diag=0
        pxp=0
        pxg=0
    OPEN(11,FILE='genea_def.txt')
    OPEN(88,FILE='gametic_matrix.txt')
    DO I=1,NANIS
        print *,i
      READ(11,*)IA,IS,ID,EFF(IA) ! READ PEDIGREE Y F
      KA2=IA*2
      KA1=KA2-1 ! COMPUTE GAMETE NUMBERS
      IF (IS.NE.0) THEN
        DIS=2/(1.-EFF(IS))
      ELSE
        DIS=1
      ENDIF
      IF (ID.NE.0) THEN
        DID=2/(1.-EFF(ID))
      ELSE
        DID=1
      ENDIF
      RLOG1=RLOG1+LOG(DIS)+LOG(DID) !accumulate log-L
      DIAG(KA1)=DIS
      DIAG(KA2)=DID !Diagonals of inverse
      IF (IS > 0) THEN !SIRE GAMETES
         KS2=IS*2
         KS1=KS2-1 !COMPUTE SIRE GAMETES
         GED(1,KA1)=KS1
         GED(2,KA1)=KS2 !STORE PEDIGREE
         DIS=-0.5*DIS
         PXG(KA1)=DIS ! STORE IN PARENT
         DIS=-0.5*DIS !REDUCE DIS -0.5
         PXP(KS1)=PXP(KS1)+DIS
         PXP(KS2)=PXP(KS2)+DIS ! AND ADD TO THE PARENT X
         DIAG(KS1)=DIAG(KS1)+DIS
         DIAG(KS2)=DIAG(KS2)+DIS
      ENDIF
      IF (ID > 0) THEN
         KD2=ID*2
         KD1=KD2-1
```

```
      GED(1,KA2)=KD1
      GED(2,KA2)=KD2
      DID=-0.5*DID
      PXG(KA2)=DID
      DID=-0.5*DID
      PXP(KD1)=PXP(KD1)+DID
      PXP(KD2)=PXP(KD2)+DID
      DIAG(KD1)=DIAG(KD1)+DID
      DIAG(KD2)=DIAG(KD2)+DID
   ENDIF
ENDDO
PRINT *,-RLOG1,-RLOG1/LOG(10.0) ! LOG-LIKELIHOOD OF DETERMINANT
DO I=1,NGAMETES
   IF (PXG(I).NE.0) THEN
      write(88,*)ged(1,i),i,pxg(i)
      write(88,*)ged(2,i),i,pxg(i)
   ENDIF
   IF ((MOD(I,2).EQ.0).AND.(PXP(I).NE.0)) THEN
      WRITE(88,*)I-1,I,PXP(I)
   ENDIF
   WRITE(88,*)I,I,DIAG(I)
ENDDO
END
```

genes

IMPACT FACTOR 3.5

Indexed in: PubMed

CITESCORE 5.1

Article

Additive and Dominance Genomic Analysis for Litter Size in Purebred and Crossbred Iberian Pigs

Houssemeddine Srihi, José Luis Noguera, Victoria Topayan, Melani Martín de Hijas, Noelia Ibañez-Escriche, Joaquim Casellas, Marta Vázquez-Gómez, María Martínez-Castillero, Juan Pablo Rosas and Luis Varona

Special Issue

Trends and Prospects in Pig Genomics and Genetics

Edited by

Dr. Katarzyna Piórkowska and Dr. Katarzyna Ropka-Molik

MDPI

https://doi.org/10.3390/genes13010012

genes

MDPI

Article

Additive and Dominance Genomic Analysis for Litter Size in Purebred and Crossbred Iberian Pigs

Houssemeddine Srihi [1], José Luis Noguera [2], Victoria Topayan [3], Melani Martín de Hijas [4], Noelia Ibañez-Escriche [3], Joaquim Casellas [4], Marta Vázquez-Gómez [4], María Martínez-Castillero [1], Juan Pablo Rosas [5] and Luis Varona [1,*]

1 Departamento de Anatomía, Embriología y Genética Animal, Facultad de Veterinaria, Instituto Agrolimentario de Aragón (IA2), 50013 Zaragoza, Spain; houssemsrihi@unizar.es (H.S.); mmartinezcastillero@gmail.com (M.M.-C.)
2 Area de Producció Animal, Centre UdL-IRTA, 25198 Lleida, Spain; JoseLuis.Noguera@irta.cat
3 Departamento de Ciència Animal, Universitat Politècnica de València, 46071 Valencia, Spain; victoriatopayan@gmail.com (V.T.); noeibes@dca.upv.es (N.I.-E.)
4 Departament de Ciència Animal i dels Aliments, Universitat Autònoma de Barcelona, 08193 Barcelona, Spain; melani.mhv@gmail.com (M.M.d.H.); joaquim.casellas@uab.cat (J.C.); Marta.Vazquez@uab.cat (M.V.-G.)
5 Programa de Mejora Genética "Castúa", INGA FOOD S.A. (Nutreco), Avda. A Rúa, 2—Bajo Edificio San Marcos, 06200 Almendralejo, Spain; juan.rosas@nutreco.com
* Correspondence: lvarona@unizar.es; Tel.: +34-876-554209

Abstract: INGA FOOD S. A., as a Spanish company that produces and commercializes fattened pigs, has produced a hybrid Iberian sow called CASTÚA by crossing the Retinto and Entrepelado varieties. The selection of the parental populations is based on selection criteria calculated from purebred information, under the assumption that the genetic correlation between purebred and crossbred performance is high; however, these correlations can be less than one because of a GxE interaction or the presence of non-additive genetic effects. This study estimated the additive and dominance variances of the purebred and crossbred populations for litter size, and calculated the additive genetic correlations between the purebred and crossbred performances. The dataset consisted of 2030 litters from the Entrepelado population, 1977 litters from the Retinto population, and 1958 litters from the crossbred population. The individuals were genotyped with a GeneSeek® GGP Porcine70K HDchip. The model of analysis was a 'biological' multivariate mixed model that included additive and dominance SNP effects. The estimates of the additive genotypic variance for the total number born (TNB) were 0.248, 0.282 and 0.546 for the Entrepelado, Retinto and Crossbred populations, respectively. The estimates of the dominance genotypic variances were 0.177, 0.172 and 0.262 for the Entrepelado, Retinto and Crossbred populations. The results for the number born alive (NBA) were similar. The genetic correlations between the purebred and crossbred performance for TNB and NBA—between the brackets—were 0.663 in the Entrepelado and 0.881 in Retinto poplulations. After backsolving to obtain estimates of the SNP effects, the additive genetic variance associated with genomic regions containing 30 SNPs was estimated, and we identified four genomic regions that each explained >2% of the additive genetic variance in chromosomes (SSC) 6, 8 and 12: one region in SSC6, two regions in SSC8, and one region in SSC12.

Keywords: pig; Iberian; additive; dominance; genetic correlation; crossbreeding; genomic selection

Citation: Srihi, H.; Noguera, J.L.; Topayan, V.; Martín de Hijas, M.; Ibañez-Escriche, N.; Casellas, J.; Vázquez-Gómez, M.; Martínez-Castillero, M.; Rosas, J.P.; Varona, L. Additive and Dominance Genomic Analysis for Litter Size in Purebred and Crossbred Iberian Pigs. *Genes* **2022**, *13*, 12. https://doi.org/10.3390/genes13010012

Academic Editor: Katarzyna Piórkowska

Received: 30 November 2021
Accepted: 17 December 2021
Published: 22 December 2021

1. Introduction

The Iberian pig breed is one of the porcine populations that has the highest meat quality [1]. Historically, Iberian pig production was developed extensively with purebred varieties, which took advantage of the *Dehesa* environment in southwestern Spain. In recent decades, however, many traditional production systems have been substituted with intensive production systems that use crossbreeding with Duroc populations to improve growth and efficiency [2]. The norms that regulate Iberian pig production [3] obligate

farmers crossing Iberian and Duroc varieties to cross boars from the Duroc variety and sows from the Iberian variety. Prolificacy, which is lower than that of white pig populations, is the major limitation in the intensive production of crossbred pigs from Iberian dams [4]. The INGA FOOD, S.A. company has developed a crossbreeding scheme between two Iberian varieties (Retinto–R- and Entrepelado–E-) that has created a hybrid sow called CASTUA–ER, which has an important heterosis effect in prolificacy [5]. In addition, the company has been developing a breeding scheme for increasing litter size through selection in the parental Retinto and Entrepelado populations.

Theoretically, the optimal strategy for the selection of purebreds for crossbred performance is Recurrent Reciprocal Selection [6]; however, it has not been routinely used in pig breeding because it involves a delay in the generation interval. In fact, purebred parental populations are selected based on selection criteria calculated from purebred phenotypic information, and under the assumption that the genetic correlation between purebred and crossbred performance is high [7]. Those genetic correlations can be imperfect (<1) because of genotype-by-environment (GxE) interactions and the presence of non-additive genetic effects [7].

Genomic information facilitates the analysis of crossbreeding data, even if genotyped and phenotyped individuals are not directly related [8], by the definition of an additive-dominance genotypic model that provides estimates of genotype x environmental interactions through genotypic correlations. In addition, the estimates of genotypic and dominance variances can be used to estimate the additive genetic correlation between purebred and crossbred performances. Backsolving, as proposed by Wang et al. [9], provides an estimate of the SNP effects and allows us to calculate the amount of additive genetic variance associated with each genomic region in purebred and crossbred performances.

This study estimated the additive and dominance genotypic variances and covariances, which were used to calculate the additive and dominance genetic variances and the genetic correlations between purebred and crossbred performances in the Retinto and Entrepelado populations. In addition, the distribution of the additive genetic variance within the autosomal genome for purebred and crossbred performance was quantified.

2. Materials and Methods

The phenotypic data included the number of piglets born alive (NBA) and the total number born (TNB) for 306 Entrepelado and 313 Retinto purebred sows, and for 333 crossbred (Entrepelado x Retinto) sows when crossed with Entrepelado, Retinto or Duroc boars (Table 1).

Table 1. Number of records (and number of sows between brackets) and the mean ± standard deviation of the number born alive and the total number born in Entrepelado, Crossbred and Retinto populations.

	Entrepelado	Crossbred	Retinto
N [1] (NS) [2]	2030 (306)	1958 (333)	1977 (313)
NBA [3]	7.75 ± 1.85	8.57 ± 2.27	8.07 ± 2.07
TNB [4]	8.02 ± 1.89	8.80 ± 2.29	8.33 ± 2.11

[1] N: number of records. [2] NS: number of sows. [3] NBA: number born alive. [4] TNB: total number born.

All of the sows were genotyped with the GeneSeek® GPP Porcine 70K HDchip (Illumina Inc., San Diego, CA, USA). Filtering excluded genotypes that had a minor allele frequency < 0.05 and an SNP call rate < 0.90 in the overall population. From that, 34,316 SNP markers were used to build the genomic relationship matrices with our own developed software in the R environment [10]. The missing genotypes were replaced with their expectation.

The model of analysis assumed that the phenotypic values of individuals ($\boldsymbol{y}$) (TNB and NBA) are explained by the (biological) additive ($\boldsymbol{u}$) and dominance ($\boldsymbol{v}$) effects of the SNPs, and a covariate (c) with the average homozygosity ($\boldsymbol{f}$), the systematic effects ($\boldsymbol{b}$)

order of parity (1, 2, 3, and >3), the sire of service breed (Entreplado, Retinto, or Duroc) and herd-year-season (122 levels). Phenotypic data were generated in three herds, and herd-year-season effects were defined every 3 months. The sow permanent environmental effects ($\boldsymbol{s}$) with 2030, 1958 and 1977 levels for the Entrepelado, Retinto and Crossbred populations, and the residuals ($\boldsymbol{e}$), were as follows:

$$\begin{bmatrix} y_E \\ y_R \\ y_{ER} \end{bmatrix} = \begin{bmatrix} f_E c_E \\ f_R c_R \\ f_{RE} c_{RE} \end{bmatrix} + \begin{bmatrix} X_E & 0 & 0 \\ 0 & X_R & 0 \\ 0 & 0 & X_{RE} \end{bmatrix} \begin{bmatrix} b_E \\ b_R \\ b_{ER} \end{bmatrix} + \begin{bmatrix} T_E & 0 & 0 \\ 0 & T_R & 0 \\ 0 & 0 & T_{RE} \end{bmatrix} \begin{bmatrix} s_E \\ s_R \\ s_{ER} \end{bmatrix} + \begin{bmatrix} T_E & 0 & 0 \\ 0 & T_R & 0 \\ 0 & 0 & T_{RE} \end{bmatrix} \begin{bmatrix} u_E \\ u_R \\ u_{ER} \end{bmatrix} + \begin{bmatrix} T_E & 0 & 0 \\ 0 & T_R & 0 \\ 0 & 0 & T_{RE} \end{bmatrix} \begin{bmatrix} v_E \\ v_R \\ v_{ER} \end{bmatrix} + \begin{bmatrix} e_E \\ e_R \\ e_{ER} \end{bmatrix}$$

where $\boldsymbol{X}$ and $\boldsymbol{T}$ are the corresponding incidence matrices. Following Vitezica et al. [8], $\boldsymbol{u}$ and $\boldsymbol{v}$ can be described in terms of the vectors of additive ($\boldsymbol{a}$) and dominance ($\boldsymbol{d}$) SNP genotypic effects as follows:

$$\begin{bmatrix} u_E \\ u_R \\ u_{RE} \end{bmatrix} = \begin{bmatrix} Za_E \\ Za_R \\ Za_{ER} \end{bmatrix} \text{ and } \begin{bmatrix} v_E \\ v_R \\ v_{RE} \end{bmatrix} = \begin{bmatrix} Wd_E \\ Wd_R \\ Wd_{ER} \end{bmatrix}$$

The matrices $\boldsymbol{Z} = (\boldsymbol{z_1} \ldots \ldots \boldsymbol{z_m})$ and $\boldsymbol{W} = (\boldsymbol{w_1} \ldots \ldots \boldsymbol{.w_m})$ are equal to 1, 0, −1 and 0, 1, 0 for SNP genotypes A_1A_1, A_1A_2 and A_2A_2, respectively.

The covariance across individual genotypic additive ($\boldsymbol{u}$) and dominance ($\boldsymbol{v}$) effects is

$$cov \begin{bmatrix} u_E \\ u_R \\ u_{RE} \end{bmatrix} = \boldsymbol{G_o} \bigotimes \boldsymbol{G} \text{ and } cov \begin{bmatrix} v_E \\ v_R \\ v_{RE} \end{bmatrix} = \boldsymbol{D_o} \bigotimes \boldsymbol{D}$$

with

$$\boldsymbol{G_o} = \begin{bmatrix} \sigma^2_{U_E} & \sigma_{U_E U_R} & \sigma_{U_E U_{ER}} \\ \sigma_{U_E U_R} & \sigma^2_{U_R} & \sigma_{U_R U_{ER}} \\ \sigma_{U_E U_{ER}} & \sigma_{U_R U_{ER}} & \sigma^2_{U_{ER}} \end{bmatrix} \text{ and } \boldsymbol{D_o} = \begin{bmatrix} \sigma^2_{V_E} & \sigma_{V_E V_R} & \sigma_{V_E V_{ER}} \\ \sigma_{V_E V_R} & \sigma^2_{V_R} & \sigma_{V_R V_{ER}} \\ \sigma_{V_E V_{ER}} & \sigma_{V_R V_{ER}} & \sigma^2_{V_{ER}} \end{bmatrix}$$

and

$$\boldsymbol{G} = \frac{\boldsymbol{ZZ'}}{\{tr[\boldsymbol{ZZ'}]/n\}} \text{ and } \boldsymbol{D} = \frac{\boldsymbol{WW'}}{\{tr[\boldsymbol{WW'}]/n\}}$$

The variance components were estimated by REML [11] through the EM-REML algorithm using *remlf90* software [12] and, in order to obtain the average information matrix, we used one extra iteration with *airemlf90*. Additive and dominance variance components were calculated in each of the populations (*E*, *R*, and *ER*) as follows:

$$\begin{bmatrix} \hat{\sigma}^2_{aE} \\ \hat{\sigma}^2_{aR} \\ \hat{\sigma}^2_{aER} \end{bmatrix} = \begin{bmatrix} \frac{\hat{\sigma}^2_{U_E}}{\{tr[\boldsymbol{ZZ'}]/n\}} \\ \frac{\hat{\sigma}^2_{U_R}}{\{tr[\boldsymbol{ZZ'}]/n\}} \\ \frac{\hat{\sigma}^2_{U_{ER}}}{\{tr[\boldsymbol{ZZ'}]/n\}} \end{bmatrix} \text{ and } \begin{bmatrix} \hat{\sigma}^2_{dE} \\ \hat{\sigma}^2_{dR} \\ \hat{\sigma}^2_{dER} \end{bmatrix} = \begin{bmatrix} \frac{\hat{\sigma}^2_{D_E}}{\{tr[\boldsymbol{WW'}]/n\}} \\ \frac{\hat{\sigma}^2_{D_R}}{\{tr[\boldsymbol{WW'}]/n\}} \\ \frac{\hat{\sigma}^2_{D_{ER}}}{\{tr[\boldsymbol{WW'}]/n\}} \end{bmatrix} \quad (1)$$

The additive (σ^2_A) and dominance (σ^2_D) genetic variances of the purebred populations were calculated as follows:

$$\hat{\sigma}^2_{A_E} = \sum_{i=1}^{n} 2\hat{p}_{Ei}\hat{q}_{Ei}\hat{\sigma}^2_{a_E} + 2\hat{p}_{Ei}\hat{q}_{Ei}(\hat{q}_{Ei} - \hat{p}_{Ei})^2\hat{\sigma}^2_{d_E} \quad (2)$$

$$\hat{\sigma}^2_{D_E} = \sum_{i=1}^{n} (2\hat{p}_{Ei}\hat{q}_{Ei})^2\hat{\sigma}^2_{d_E} \quad (3)$$

$$\hat{\sigma}^2_{A_R} = \sum_{i=1}^{n} 2\hat{p}_{Ri}\hat{q}_{Ri}\hat{\sigma}^2_{a_R} + 2\hat{p}_{Ri}\hat{q}_{Ri}(\hat{q}_{Ri} - \hat{p}_{Ri})^2\hat{\sigma}^2_{d_R} \quad (4)$$

$$\hat{\sigma}^2_{D_R} = \sum_{i=1}^{n} (2\hat{p}_{Ri}\hat{q}_{Ri})^2\hat{\sigma}^2_{d_R} \quad (5)$$

where $\hat{p}_{Xi}$ and $\hat{q}_{Xi}$ are the raw estimates of the allelic frequencies for A_1 and A_2 at the *ith* SNP marker and the X = {*E*,*R* or *ER*} population, respectively. The estimates of the

contributions to the additive variance in the crossbred population from the Entrepelado ($\sigma^2_{A_{ER(E)}}$) and Retinto ($\sigma^2_{A_{ER(R)}}$) were obtained by [8] as follows:

$$\hat{\sigma}^2_{A_{ER(E)}} = \sum_{i=1}^{n} 2\hat{p}_{Ri}\hat{q}_{Ri}\hat{\sigma}^2_{a_{ER}} + 2\hat{p}_{Ri}\hat{q}_{Ri}(\hat{q}_{Ei} - \hat{p}_{Ei})^2\hat{\sigma}^2_{d_{ER}} \quad (6)$$

$$\hat{\sigma}^2_{A_{ER(R)}} = \sum_{i=1}^{n} 2\hat{p}_{Ei}\hat{q}_{Ei}\hat{\sigma}^2_{a_{ER}} + 2\hat{p}_{Ei}\hat{q}_{Ei}(\hat{q}_{Ri} - \hat{p}_{Ri})^2\hat{\sigma}^2_{d_{ER}} \quad (7)$$

Following Vitezica et al. [8], the additive variance in the crossbred population was the average of the two values resulting from Equations (6) and (7), as follows:

$$\hat{\sigma}^2_{A_{ER}} = \frac{1}{2}\hat{\sigma}^2_{A_{ER(E)}} + \frac{1}{2}\hat{\sigma}^2_{A_{ER(R)}} \quad (8)$$

The estimate of the dominance variance of the crossbred population [8] was calculated as follows:

$$\hat{\sigma}^2_{D_{ER}} = \sum_{i=1}^{n} 4\hat{p}_{Ei}\hat{q}_{Ei}\hat{p}_{Ri}\hat{q}_{Ri}\hat{\sigma}^2_{d_{ER}} \quad (9)$$

With these estimates, the heritabilities (h^2_X) and dominance ratios (d^2_X) in the purebred (X = E,R) and crossbred (X = ER) populations were obtained by:

$$\hat{h}^2_X = \hat{\sigma}^2_{A_X} / \left(\hat{\sigma}^2_{A_X} + \hat{\sigma}^2_{D_X} + \hat{\sigma}^2_{S_X} + \hat{\sigma}^2_{E_X}\right) \quad (10)$$

$$\hat{d}^2_X = \hat{\sigma}^2_{D_X} / \left(\hat{\sigma}^2_{A_X} + \hat{\sigma}^2_{D_X} + \hat{\sigma}^2_{S_X} + \hat{\sigma}^2_{E_X}\right) \quad (11)$$

where $\hat{\sigma}^2_{S_X}$ and $\hat{\sigma}^2_{E_X}$ are the estimates of the sow permanent environmental and residual variance in the X = {E,R,ER} population.

The covariance between purebred and crossbred additive genetic effects in the Entrepelado ($\sigma_{A_E A_{ER(E)}}$) and Retinto ($\sigma_{A_R A_{ER(R)}}$) populations were as follows:

$$\hat{\sigma}_{A_E A_{ER(E)}} = \sum_{i=1}^{n} 2\hat{p}_{Ei}\hat{q}_{Ei}\hat{\sigma}_{a_E a_{ER}} + 2\hat{p}_{Ei}\hat{q}_{Ei}(\hat{q}_{Ei} - \hat{p}_{Ei})(\hat{q}_{Ri} - \hat{p}_{Ri})\hat{\sigma}_{d_E d_{ER}} \quad (12)$$

$$\hat{\sigma}_{A_E A_{ER(R)}} = \sum_{i=1}^{n} 2\hat{p}_{Ri}\hat{q}_{Ri}\hat{\sigma}_{a_R a_{ER}} + 2\hat{p}_{Ri}\hat{q}_{Ri}(\hat{q}_{Ri} - \hat{p}_{Ri})(\hat{q}_{Ei} - \hat{p}_{Ei})\hat{\sigma}_{d_R d_{ER}} \quad (13)$$

with

$$\begin{bmatrix} \hat{\sigma}_{a_E a_{ER}} \\ \hat{\sigma}_{a_R a_{ER}} \end{bmatrix} = \begin{bmatrix} \frac{\hat{\sigma}_{U_E U_{ER}}}{\{tr[\boldsymbol{ZZ'}]/n\}} \\ \frac{\hat{\sigma}_{U_R U_{ER}}}{\{tr[\boldsymbol{ZZ'}]/n\}} \end{bmatrix} \text{ and } \begin{bmatrix} \hat{\sigma}_{d_E d_{ER}} \\ \hat{\sigma}_{d_R d_{ER}} \end{bmatrix} = \begin{bmatrix} \frac{\hat{\sigma}_{V_E V_{ER}}}{\{tr[\boldsymbol{WW'}]/n\}} \\ \frac{\hat{\sigma}_{V_R V_{ER}}}{\{tr[\boldsymbol{WW'}]/n\}} \end{bmatrix} \quad (14)$$

Therefore, the genetic correlations between the purebred and crossbreed breeding values in the Entrepelado and Retinto populations were computed as follows:

$$\hat{r}_{A_E A_{ER(E)}} = \frac{\hat{\sigma}_{A_E A_{ER(E)}}}{\sqrt{\hat{\sigma}^2_{A_E}\hat{\sigma}^2_{A_{ER(E)}}}} \text{ and } \hat{r}_{A_R A_{ER(R)}} = \frac{\hat{\sigma}_{A_R A_{ER(R)}}}{\sqrt{\hat{\sigma}^2_{A_R}\hat{\sigma}^2_{A_{ER(R)}}}} \quad (15)$$

The vector of the SNP additive effects ($\hat{\boldsymbol{a}}_{\boldsymbol{E}}$, $\hat{\boldsymbol{a}}_{\boldsymbol{R}}$ and $\hat{\boldsymbol{a}}_{\boldsymbol{ER}}$) was obtained by backsolving [9], as

$$\hat{\boldsymbol{a}}_{\boldsymbol{E}} = \frac{\hat{\sigma}^2_{a_E}}{\hat{\sigma}^2_{U_E}}\boldsymbol{ZG}^{-1}\hat{\boldsymbol{u}}_{\boldsymbol{E}},\ \hat{\boldsymbol{a}}_{\boldsymbol{R}} = \frac{\hat{\sigma}^2_{a_R}}{\hat{\sigma}^2_{U_R}}\boldsymbol{ZG}^{-1}\hat{\boldsymbol{u}}_{\boldsymbol{R}} \text{ and } \hat{\boldsymbol{a}}_{\boldsymbol{ER}} = \frac{\hat{\sigma}^2_{a_{ER}}}{\hat{\sigma}^2_{U_{ER}}}\boldsymbol{ZG}^{-1}\hat{\boldsymbol{u}}_{\boldsymbol{ER}} \quad (16)$$

and the vector of the SNP dominance effects ($\hat{\boldsymbol{d}}_{\boldsymbol{E}}$, $\hat{\boldsymbol{d}}_{\boldsymbol{R}}$ and $\hat{\boldsymbol{d}}_{\boldsymbol{ER}}$) was as follows:

$$\hat{\boldsymbol{d}}_{\boldsymbol{E}} = \frac{\hat{\sigma}^2_{d_E}}{\hat{\sigma}^2_{V_E}}\boldsymbol{WD}^{-1}\hat{\boldsymbol{v}}_{\boldsymbol{E}},\ \hat{\boldsymbol{d}}_{\boldsymbol{R}} = \frac{\hat{\sigma}^2_{d_R}}{\hat{\sigma}^2_{V_R}}\boldsymbol{WD}^{-1}\hat{\boldsymbol{v}}_{\boldsymbol{R}} \text{ and } \hat{\boldsymbol{d}}_{\boldsymbol{ER}} = \frac{\hat{\sigma}^2_{d_{ER}}}{\hat{\sigma}^2_{V_{ER}}}\boldsymbol{WD}^{-1}\hat{\boldsymbol{v}}_{\boldsymbol{ER}} \quad (17)$$

With those, the genetic additive variances ($\sigma^2_{A_E(k)}$, $\sigma^2_{A_R(k)}$, $\sigma^2_{A_{ER(E)}(k)}$ and $\sigma^2_{A_{ER(R)}(k)}$) explained by the *k*th segment of the genome were calculated as follows:

$$\hat{\sigma}^2_{A_E(k)} = \sum_{i=1}^{n(k)} 2\hat{p}_{Ei}\hat{q}_{Ei}\hat{a}^2_{E_i} + 2\hat{p}_{Ei}\hat{q}_{Ei}(\hat{q}_{Ei} - \hat{p}_{Ei})^2\hat{d}^2_{E_i} \quad (18)$$

$$\hat{\sigma}^2_{A_R(k)} = \sum_{i=1}^{n(k)} 2\hat{p}_{Ri}\hat{q}_{Ri}\hat{a}^2_{R_i} + 2\hat{p}_{Ri}\hat{q}_{Ri}(\hat{q}_{Ri} - \hat{p}_{Ri})^2\hat{d}^2_{R_i} \quad (19)$$

$$\hat{\sigma}^2_{A_{ER(E)}(k)} = \sum_{i=1}^{n(k)} 2\hat{p}_{Ri}\hat{q}_{Ri}\hat{a}^2_{ER_i} + 2\hat{p}_{Ri}\hat{q}_{Ri}(\hat{q}_{Ei} - \hat{p}_{Ei})^2\hat{d}^2_{ER_i} \quad (20)$$

$$\hat{\sigma}^2_{A_{ER(R)}(k)} = \sum_{i=1}^{n(k)} 2\hat{p}_{Ei}\hat{q}_{Ei}\hat{a}^2_{ER_i} + 2\hat{p}_{Ei}\hat{q}_{Ei}(\hat{q}_{Ri} - \hat{p}_{Ri})^2\hat{d}^2_{ER_i} \quad (21)$$

where $n(k)$ is the number of SNP markers within the kth segment, which was set to 30 after testing several number of the SNP markers (20, 30 and 40). In order to identify the genes within the genomic regions that explained >2.0% of the total genetic variance, we used the biomart tool (www.ensembl.org (accessed on 10 October 2021)).

3. Results and Discussion

The results based on TNB and NBA were similar, which was expected because these two traits have a high genetic correlation [13], and the raw correlation between them in the analyzed dataset was 0.94; therefore, we focused on the results with the TNB, and the results for NBA are presented as Supplementary Information (Tables S1–S3 and Figure S1). The REML estimates of the additive genotypic (co) variances are shown in Tables 2 and 3 in TNB and NBA.

Table 2. REML estimates ± standard error (SE) of the additive genotypic (co)variances for the total number born (TNB).

	Entrepelado	**Crossbred**	**Retinto**
Entrepelado	0.248 ± 0.161	0.259 ± 0.178	0.200 ± 0.135
Crossbred	-	0.546 ± 0.268	0.388 ± 0.170
Retinto	-	-	0.282 ± 0.146

Table 3. REML estimates ± standard error (SE) of the dominance genotypic (co)variances for the total number born (TNB).

	Entrepelado	**Crossbred**	**Retinto**
Entrepelado	0.177 ± 0.165	0.212 ± 0.171	0.166 ± 0.152
Crossbred	-	0.262 ± 0.210	0.202 ± 0.179
Retinto	-	-	0.172 ± 0.199

The additive genotypic variance was higher in the crossbred populations than it was in the purebred populations. This may be due to scale effects, as the phenotypic variation in also greater. In addition, the estimates of the genotypic covariances between purebreds (Entrepelado and Retinto) and the crossbred population were all high and positive, and they corresponded to additive genotypic correlations of 0.704 ($0.259/\sqrt{0.248 \times 0.546}$) between Entrepelado and Crossbred pigs, 0.988 ($0.388/\sqrt{0.546 \times 0.282}$) between Retinto and Crossbred pigs, and 0.756 ($0.200/\sqrt{0.248 \times 0.282}$) between the two purebreds. These results indicated that the genotype x environmental interaction was small, and the additive genotypic correlations were similar to those obtained by Vitezica et al. [8] in white pig populations. The REML estimates of the dominance genotypic (co)variances ranged from 0.170 (Retinto) to 0.265 (Crossbred) (Table 3).

The estimates of the dominance genotypic covariances were all positive, and reflected genotypic dominance correlations >0.95. The analysis provided the REML estimates of

the sow permanent and residual effects (Table 4). The residual variance (σ^2_E) is greater in the crossbred population than in purebreds, consistently with the greater phenotypic variation. In contrast, the estimate of the sow environmental variance (σ^2_R) was very low in the crossbred population.

Table 4. REML estimates ± standard error (SE) of the permanent environmental and residual variances for the total number born (TNB) in the Entrepelado, Crossbred and Retinto populations.

Variances [1]	Entrepelado	Crossbred	Retinto
σ^2_S	0.191 ± 0.105	0.009 ± 0.029	0.268 ± 0.120
σ^2_E	2.810 ± 0.099	4.467 ± 0.155	3.534 ± 0.128

[1] σ^2_S: Sow permanent environmental variance; σ^2_E: residual variance.

The additive and dominance genotypic (co) variances were used to calculate the additive and dominance genetic variances in the purebred populations based on expressions (1) to (5) (Table 3). The estimates of the additive genetic variances were 0.170 (Entrepelado) and 0.150 (Retinto), and the estimates of the dominance genetic variances were 0.074 (Entrepelado) and 0.056 (Retinto). The heritability estimates were calculated using Equation (10); they were 0.052 (Entrepelado) and 0.037 (Retinto), which were within the range or slightly lower than those of white pigs [13–15] and in the same [5] or other Iberian [16,17] populations. The dominance ratios were obtained from Equation (11), and were 0.023 for Entrepelado and 0.014 for Retinto. They were smaller than the heritabilities, but their ratios with them were approximately 40%, which was higher than those reported for white pig populations [8,18] for litter size and similar to the results of Tusell et al. [19] in other swine traits.

We used Equations (6) and (7) to calculate the additive variances for crossbred performance in the purebred populations, which were 0.413 (Entrepelado) and 0.293 (Retinto). Therefore, the additive genetic variance in the crossbred population was the average of the two (0.353), which was higher than the additive genetic variances in the purebred populations, which were similar to the results of Vitezica et al. [8] with regard to litter size, and to the results of Tusell et al. [19] for other pig traits. Nevertheless, Xiang et al. [20] found the opposite in a cross between Landrace and Yorkshire breeds (0.86 and 0.54 in purebreds and 0.28 in crossbreds). In the present study, the dominance genetic variance in the crossbred population (0.079) was calculated based on the Equation (8), which was similar to the dominance genetic variance in the purebreds; however, its ratio with the additive genetic variances was lower (22%). Given those variance components, the heritability and dominance ratio estimates in the crossbred population were 0.072 and 0.016, respectively.

In addition, the additive genetic correlations between purebred and crossbred performances in the Entrepelado and Retinto populations were calculated based on expressions (12) to (15), which were 0.663 in Entrepelado and 0.881 in Retinto populations. Those correlations were within the range of the estimates summarized by Wientjes and Calus [7], and suggest that the efficiency of the selection for increased crossbred performance by selecting for purebred performance will be more effective in Retinto than in Entrepelado pigs.

We used Equations (16) and (17) to calculate the additive and dominance genotypic effects associated with each of the 34,316 SNP markers, which were used in Equations (18)–(21) to calculate the proportion of the additive genetic variance that was explained by segments of 30 consecutive SNPs (Figure 1). The distribution of the additive variance explained by segments of 20 and 40 SNP markers were similar, and are presented as supplementary information (Figures S2 and S3).

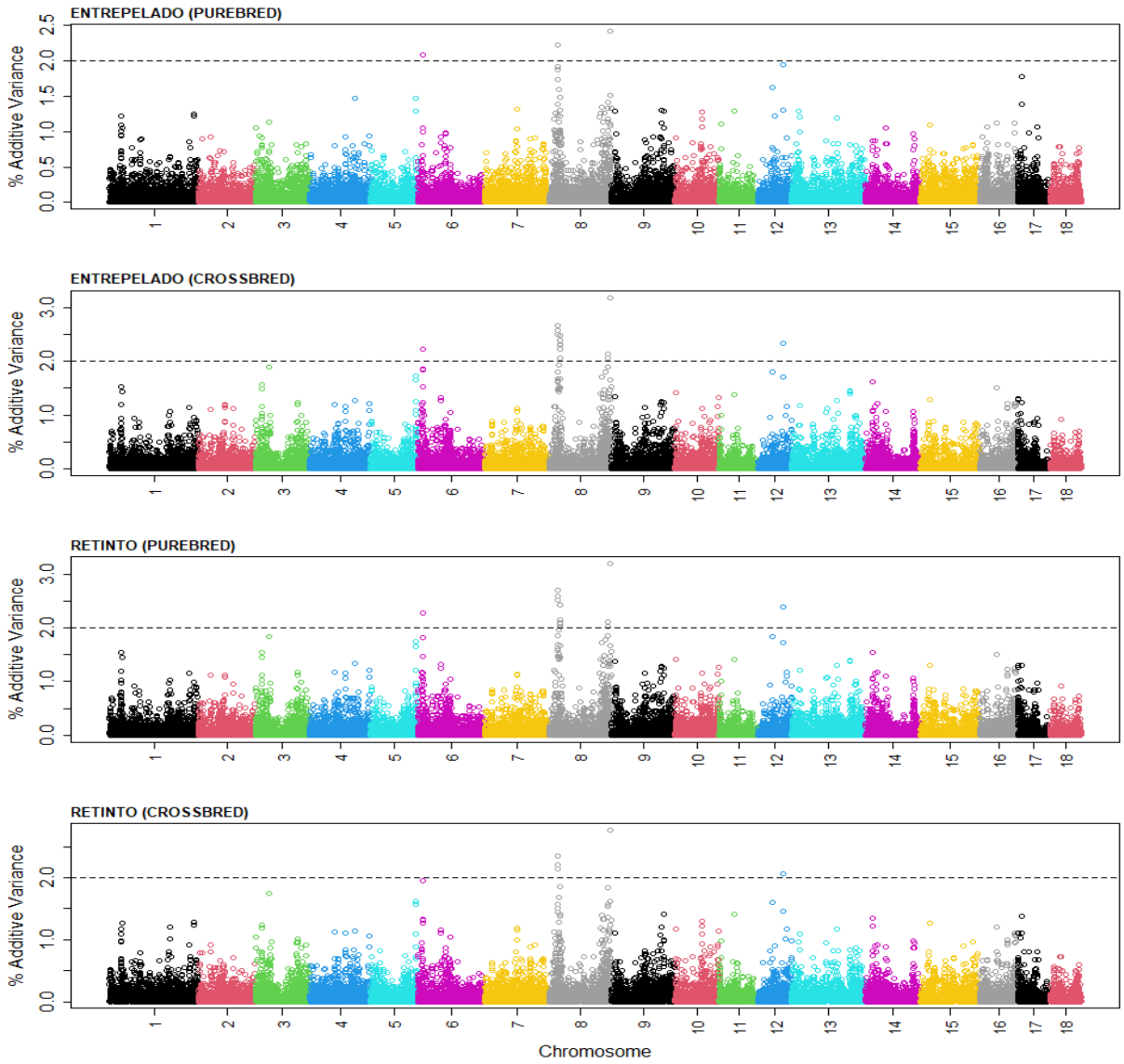


Figure 1. Distribution of the percentage of the additive genetic variance explained by genomic segments of 30 SNPs within the autosomal genome of purebred and crossbred performance for the total number born (TNB) in the Entrepelado and Retinto varieties. Black: chromosomes 1, 9 and 17; red: chromosomes 2, 10 and 18; green: chromosomes 3 and 11; deep blue: chromosomes 4 and 12; blue: chromosomes 5 and 13; purple: chromosomes 6 and 14; yellow: chromosomes 7 and 15; grey: chromosomes 8 and 16.

The figure presents the distribution of the additive variance along the autosomal chromosomes in the Entrepelado and Retinto populations, and for the purebred and crossbred performance. Four genomic regions can be highlighted; each explained >2% of the additive genetic variance in at least one of the populations. The SNPs at the center of each of the genomic regions that explained the highest amount of additive genetic variance, and the genes in the Sus_Scrofa 11.1. genomic map that were within 1 Mb downstream or upstream, are presented in Table 5.

Table 5. SNPs at the center of each of the four genomic regions that explained > 2% of the additive genetic variance in at least one of the populations, and the genes located within 1 Mb downstream or upstream.

SNP [1]	SSC [2]	bp [3]	Genes
rs326244568	6	7,597,405	*BCO1*, *PKD1L2*, *GCSH*, *ATMIN*, *CENPN*, *CDYL2*, *DYNLRB2*
rs81401202	8	11,585,865	*CD38*, *FGFBP1*, *PROM1*, *TAPT1*, *LDB2*
rs81406142	8	137,540,516	*CFAP299*, *FGF5*, *PRDM8*, *ANTXR2*
rs345468811	12	46,079,417	*TAOK1*, *ABHD15*, *TP53I13*, *GIT1*, *ANKRD1*, *CORO6*, *EFCAB5*, *NSRP1*, *SLC6A4*, *BLMH*, *TMIGD1*, *CPD*, *GOSR1*

[1] SNP: single nucleotide polymorphism, [2] SSC: Sus Scrofa chromosome, [3] bp: base pair.

Among those genes, several can be proposed as candidate genes to explain the additive genetic variation. The genomic region surrounding bp 7,597,405 in SSC6 included BCO1 (*β-Carotene Oxygenase 1*), which encodes an enzyme that catalyzes the breakdown of provitamin A and provides retinoids for embryogenesis [21,22]. Furthermore, the GCSH (*Glycine Cleavage System H*) protein plays an important role in embryonic viability [23].

Two genomic regions were identified in SSC8 around bp 11,585,865 and bp 137,540,516. Among the genes within those regions, PRDM8 (*PR/SET Domain 8*) is involved in the neurogenesis [24] of the FGF5 *(Fibroblast Growth Factor 5*), a member of the fibroblast growth factor family that is involved in several biological processes, including embryonic development, cell growth, and morphogenesis [25,26].

The genomic region around bp 46,079,417 in SSC12 contains, among others, the GIT1 (*G protein-coupled receptor kinase interactor 1*) gene, which plays a role in spine morphogenesis [27], the NSRP1 (*Nuclear Speckle Splicing Regulatory Protein 1*) development process, and in utero embryonic development [28], and ANKRD1 (*Ankyrin Repeat Domain 1*), which is involved in neuron projection development [29].

The Gene Ontology (GO) terms for the biological processes for the proposed candidate genes are presented as Supplementary Table S4.

4. Conclusions

(1) The additive genetic variance and the heritabilities were higher in the crossbred than those in the purebred populations, (2) the genetic correlation between purebred and crossbred performances were higher in Retinto than they were in Entrepelado pigs, and (3) the additive genetic variances were heterogeneously distributed throughout the autosomal genome, and four genomic regions in SSC6, SSC8, and SSC12 with several candidate genes were identified.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/genes13010012/s1. Table S1: REML estimates ± the standard error (SE) of the additive genotypic (co)variances for the number born alive (NBA). Table S2: REML estimates ± the standard error (SE) of the dominance genotypic (co)variances for the number born alive (NBA). Table S3: REML estimates ± the standard error (SE) of the permanent environmental and residual variances for the number born alive (NBA). Table S4: GO (Gene Ontology) terms for the biological process of the proposed candidate genes. Figure S1: Distribution of the percentage of the additive genetic variance explained by genomic segments of 30 SNPs within the autosomal genome of the purebred and crossbred performance for the number born alive (NBA) in the Entrepelado and Retinto varieties. Figure S2: Distribution of the percentage of the additive genetic variance explained by genomic segments of 20 SNPs within the autosomal genome of purebred and crossbred performance for the total number born (TNB) in the Entrepelado and Retinto varieties. Figure S3: Distribution of the percentage of the additive genetic variance explained by genomic segments of 40 SNPs within the autosomal genome of the purebred and crossbred performance for the total number born (TNB) in the Entrepelado and Retinto varieties.

Author Contributions: Conceptualization, L.V.; methodology, H.S. and L.V.; software, L.V.; formal analysis, H.S., V.T. and L.V.; data curation, J.L.N., J.P.R., M.M.d.H. and N.I.-E.; writing—original draft preparation, H.S. and L.V.; writing—review and editing M.M.d.H., M.V.-G., N.I.-E., M.M.-C. and J.C.;

supervision, L.V., N.I.-E., J.C. and J.L.N., project administration, J.L.N., J.C. and N.I.-E.; funding acquisition, J.L.N., J.C. and N.I.-E. All authors have read and agreed to the published version of the manuscript.

Funding: The research was partially funded by grants CGL2016-80155-R and PID2020-114705RB-I00 (Ministerio de Ciencia e Innovación), and IDI-20170304 (CDTI). Srihi received funding from the European Union's H2020 research and innovation program under a Marie Sklodowska-Curie grant agreement, No. 801586.

Institutional Review Board Statement: The protocols were approved by the Ethical and Animal Care Committee at IRTA (Institute of Agrifood Research and Technology, Caldes de Montbui, Spain; IRTA-22032012). It was performed according to the Spanish Policy for Animal Protection (RD 53/2013), which meets the European Union Directive 2010/63/UE on the protection of research animals.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset used in this study will be available upon reasonable request to the corresponding author (lvarona@unizar.es).

Acknowledgments: We thank the company Inga Food SA (Almendralejo, Spain) and its technicians (E. Magallón, L. Muñoz, P. Díaz, and D. Iniesta) for their cooperation and technical support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Serra, X.; Gil, F.; Pérez-Enciso, M.; Oliver, M.A.; Vázquez, J.M.; Gispert, M.; Díaz, I.; Moreno, F.; Latorre, R.; Noguera, J.L. A comparison of carcass, meat quality and histochemical characteristics of Iberian (Guadyerbas line) and Landrace pigs. *Livest. Prod. Sci.* **1998**, *56*, 215–223. [CrossRef]
2. Serrano, M.P.; Valencia, D.G.; Nieto, M.; Lázaro, R.; Mateos, G.G. Influence of sex and terminal sire line on performance and carcass and meat quality of Iberian pigs reared under intensive production systems. *Meat Sci.* **2008**, *78*, 420–428. [CrossRef]
3. Ministry of Agriculture, Food and Environment. *Boletín Oficial del Estado Real Decreto 4/2014 de 10 enero, por el que se aprueba la norma de calidad para la carne, el jamón, la paleta y la caña de lómo ibérico*; Ministry of Agriculture, Food and Environment: Madrid, Spain, 2014; pp. 1569–1585.
4. Silio, L.; Rodriguez, M.C.; Rodrigañez, J.; Toro, M.A. La selección de cerdos ibéricos. In *Porcino Ibérico: Aspectos Claves*; Buxade, C., Daza, A., Eds.; Mundi-Press: Madrid, Spain, 2001; pp. 125–149.
5. Noguera, J.L.; Ibáñez-Escriche, N.; Casellas, J.; Rosas, J.P.; Varona, L. Genetic parameters and direct, maternal and heterosis effects on litter size in a diallel cross among three commercial varieties of Iberian pig. *Animal* **2019**, *13*, 2765–2772. [CrossRef] [PubMed]
6. Comstock, R.E.; Robinson, H.F.; Harvey, P.H. A Breeding Procedure Designed to Make Maximum Use of Both General and Specific Combining Ability. *Agron. J.* **1949**, *41*, 360–367. [CrossRef]
7. Wientjes, Y.C.J.; Calus, M.P.L. BOARD INVITED REVIEW: The purebred-crossbred correlation in pigs: A review of theory, estimates, and implications. *J. Anim. Sci.* **2017**, *95*, 3467–3478. [CrossRef]
8. Vitezica, Z.G.; Varona, L.; Elsen, J.-M.; Misztal, I.; Herring, W.; Legarra, A. Genomic BLUP including additive and dominant variation in purebreds and F1 crossbreds, with an application in pigs. *Genet. Sel. Evol.* **2016**, *48*, 6. [CrossRef] [PubMed]
9. Wang, H.; Misztal, I.; Aguilar, I.; Legarra, A.; Muir, W.M. Genome-wide association mapping including phenotypes from relatives without genotypes. *Genet. Res.* **2012**, *94*, 73–83. [CrossRef] [PubMed]
10. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021; Available online: http://www.R-project.org/ (accessed on 1 November 2021).
11. Patterson, H.D.; Thompson, R. Recovery of Inter-Block Information when Block Sizes are Unequal. *Biometrika* **1971**, *58*, 545–554. [CrossRef]
12. Lourenco, D.; Legarra, A.; Tsuruta, S.; Masuda, Y.; Aguilar, I.; Misztal, I. Single-step genomic evaluations from theory to practice: Using snp chips and sequence data in blupf90. *Genes* **2020**, *11*, 790. [CrossRef]
13. Bidanel, J.P. Biology and genetics of reproduction. In *The Genetics of the Pig*; CABI: Wallingford, UK, 2011; pp. 218–241.
14. Putz, A.M.; Tiezzi, F.; Maltecca, C.; Gray, K.A.; Knauer, M.T. Variance component estimates for alternative litter size traits in swine. *J. Anim. Sci.* **2015**, *93*, 5153–5163. [CrossRef]
15. Ogawa, S.; Konta, A.; Kimata, M.; Ishii, K.; Uemoto, Y.; Satoh, M. Estimation of genetic parameters for farrowing traits in purebred Landrace and Large White pigs. *Anim. Sci. J.* **2019**, *90*, 23–28. [CrossRef]
16. Fernández, A.; Rodrigáñez, J.; Zuzúarregui, J.; Rodríguez, M.C.; Silió, L. Genetic parameters for litter size and weight at different parities in Iberian pigs. *Span. J. Agric. Res.* **2008**, *6*, 98–106. [CrossRef]
17. García-Casco, J.M.; Fernández, A.; Rodríguez, M.C.; Silió, L. Heterosis for litter size and growth in crosses of four strains of Iberian pig. *Livest. Sci.* **2012**, *147*, 1–8. [CrossRef]
18. Culbertson, M.S.; Mabry, J.W.; Misztal, I.; Gengler, N.; Bertrand, J.K.; Varona, L. Estimation of Dominance Variance in Purebred Yorkshire Swine. *J. Anim. Sci.* **1998**, *76*, 448–451. [CrossRef]

19. Tusell, L.; Gilbert, H.; Vitezica, Z.G.; Mercat, M.J.; Legarra, A.; Larzul, C. Dissecting total genetic variance into additive and dominance components of purebred and crossbred pig traits. *Animal* **2019**, *13*, 2429–2439. [CrossRef] [PubMed]
20. Xiang, T.; Christensen, O.F.; Vitezica, Z.G.; Legarra, A. Genomic evaluation by including dominance effects and inbreeding depression for purebred and crossbred performance with an application in pigs. *Genet. Sel. Evol.* **2016**, *48*, 92. [CrossRef] [PubMed]
21. Quadro, L.; Giordano, E.; Costabile, B.K.; Nargis, T.; Iqbal, J.; Kim, Y.; Wassef, L.; Hussain, M.M. Interplay between β-carotene and lipoprotein metabolism at the maternal-fetal barrier. *Biochim. Biophys. Acta—Mol. Cell Biol. Lipids* **2020**, *1865*, 158591. [CrossRef]
22. Wassef, L.; Spiegler, E.; Quadro, L. Embryonic phenotype, β-carotene and retinoid metabolism upon maternal supplementation of β-carotene in a mouse model of severe vitamin A deficiency. *Arch. Biochem. Biophys.* **2013**, *539*, 223–229. [CrossRef]
23. Leung, K.Y.; De Castro, S.C.P.; Galea, G.L.; Copp, A.J.; Greene, N.D.E. Glycine Cleavage System H Protein Is Essential for Embryonic Viability, Implying Additional Function Beyond the Glycine Cleavage System. *Front. Genet.* **2021**, *12*, 625120. [CrossRef] [PubMed]
24. Kinameri, E.; Inoue, T.; Aruga, J.; Imayoshi, I.; Kageyama, R.; Shimogori, T.; Moore, A.W. Prdm proto-oncogene transcription factor family expression and interaction with the Notch-Hes pathway in mouse neurogenesis. *PLoS ONE* **2008**, *3*, e3859. [CrossRef]
25. Haub, O.; Goldfarb, M. Expression of FGF5 gene in mouse embryo. *Development* **1991**, *406*, 397–406. [CrossRef]
26. Beer, H.D.; Bittner, M.; Niklaus, G.; Munding, C.; Max, N.; Goppelt, A.; Werner, S. The fibroblast growth factor binding protein is a novel interaction partner of FGF-7, FGF-10 and FGF-22 and regulates FGF activity: Implications for epithelial repair. *Oncogene* **2005**, *24*, 5269–5277. [CrossRef] [PubMed]
27. Segura, I.; Essmann, C.L.; Weinges, S.; Acker-Palmer, A. Grb4 and GIT1 transduce ephrinB reverse signals modulating spine morphogenesis and synapse formation. *Nat. Neurosci.* **2007**, *10*, 301–310. [CrossRef] [PubMed]
28. Kim, Y.D.; Lee, J.Y.; Oh, K.M.; Araki, M.; Araki, K.; Yamamura, K.I.; Jun, C.D. NSrp70 is a novel nuclear speckle-related protein that modulates alternative pre-mRNA splicing in vivo. *Nucleic Acids Res.* **2011**, *39*, 4300–4314. [CrossRef] [PubMed]
29. Stam, F.J.; MacGillavry, H.D.; Armstrong, N.J.; De Gunst, M.C.M.; Zhang, Y.; Van Kesteren, R.E.; Smit, A.B.; Verhaagen, J. Identification of candidate transcriptional modulators involved in successful regeneration after nerve injury. *Eur. J. Neurosci.* **2007**, *25*, 3629–3637. [CrossRef] [PubMed]

animals

IMPACT FACTOR 3.0

Indexed in: PubMed

CITESCORE 4.2

Communication

A Bayesian Multivariate Gametic Model in a Reciprocal Cross with Genomic Information: An Example with Two Iberian Varieties

Houssemeddine Srihi, David López-Carbonell, Noelia Ibáñez-Escriche, Joaquim Casellas, Pilar Hernández, Sara Negro and Luis Varona

Topic Collection

Applications of Quantitative Genetics in Livestock Production

Edited by

Prof. Dr. Michael E. Davis

MDPI

https://doi.org/10.3390/ani13101648

animals

MDPI

Communication

A Bayesian Multivariate Gametic Model in a Reciprocal Cross with Genomic Information: An Example with Two Iberian Varieties

Houssemeddine Srihi [1], David López-Carbonell [1], Noelia Ibáñez-Escriche [2], Joaquim Casellas [3], Pilar Hernández [2], Sara Negro [4] and Luis Varona [1,*]

[1] Facultad de Veterinaria, Instituto Agrolimentario de Aragón (IA2), Universidad de Zaragoza, 50013 Zaragoza, Spain; houssemsrihi@unizar.es (H.S.)
[2] Institute for Animal Science and Technology, Universitat Politècnica de València, 46022 Valencia, Spain
[3] Department of Animal and Food Science, Universitat Autònoma de Barcelona, 08193 Barcelona, Spain
[4] Programa de Mejora Genética "Castúa", INGA FOOD S.A. (Nutreco), 06200 Almendralejo, Spain
* Correspondence: lvarona@unizar.es

Simple Summary: INGA FOOD, S.A. initiated a crossbreeding program involving two Iberian pig varieties: Retinto and Entrepelado. The primary objective of this program is to produce an F1 hybrid sow that exhibits enhanced reproductive performance. In a previous investigation, variations in the reproductive performance of sows, specifically litter size, were observed among the reciprocal crosses. These variations indicate the presence of genomic imprinting effects. To assess the influence of genetic origin, we developed a multivariate gametic model to estimate the gametic correlations between paternal and maternal effects. Gametic correlations lower than one could potentially explain the performance differences observed across the reciprocal crosses. Despite having limited data, the study's findings suggest that the gametic correlation estimate between paternal and maternal effects on litter size is lower in the Entrepelado population compared to the Retinto population.

Abstract: INGA FOOD, S.A. initiated a crossbreeding program between two Iberian pig varieties, Retinto (R) and Entrepelado (E), with the goal of producing a hybrid sow (F1). Several studies have been conducted to evaluate its productive performance, and these studies have revealed differences in litter size between the two reciprocal crosses, suggesting the presence of genomic imprinting effects. To further investigate these effects, this study introduces a multivariate gametic model designed to estimate gametic correlations between paternal and maternal effects originating from both genetic backgrounds involved in the reciprocal crosses. The dataset consisted of 1258 records (the total number born—TNB and the number born alive—NBA) from 203 crossbred dams for the Entrepelado (sire) × Retinto (dam) cross and 700 records from 125 crossbred dams for the Retinto (sire) × Entrepelado (dam) cross. All animals were genotyped using the GeneSeek® GPP Porcine 70 K HDchip (Illumina Inc., San Diego, CA, USA). The results indicated that the posterior distribution of the gametic correlation between paternal and maternal effects was distinctly different between the two populations. Specifically, in the Retinto population, the gametic correlation showed a positive skew with posterior probabilities of 0.78 for the TNB and 0.80 for the NBA. On the other hand, the Entrepelado population showed a posterior probability of a positive gametic correlation between paternal and maternal effects of approximately 0.50. The differences in the shape of the posterior distribution of the gametic correlations between paternal and maternal effects observed in the two varieties may account for the distinct performance outcomes observed in the reciprocal crosses.

Keywords: gametic model; Iberian pig; crossbreeding; Retinto; Entrepelado

Citation: Srihi, H.; López-Carbonell, D.; Ibáñez-Escriche, N.; Casellas, J.; Hernández, P.; Negro, S.; Varona, L. A Bayesian Multivariate Gametic Model in a Reciprocal Cross with Genomic Information: An Example with Two Iberian Varieties. *Animals* **2023**, *13*, 1648. https://doi.org/10.3390/ani13101648

Academic Editor: Michael E. Davis

Received: 19 April 2023
Revised: 11 May 2023
Accepted: 13 May 2023
Published: 16 May 2023

1. Introduction

The Iberian breed is widely renowned for its ability to produce some of the highest-quality pork [1]. This breed is particularly well-adapted to the "Dehesa" environment in southwestern Spain, which is characterized by a savannah landscape and is composed of grass, cork, and holm oaks with seasonal production. Traditionally, Iberian pig production was dominated by purebred varieties and extensive management practices. However, in recent decades, there has been a shift toward more intensive farming practices that incorporate crossbreeding with Duroc boars to improve growth and efficiency at commercial stages [2].

The regulatory norms for Iberian pig production allow crossbreeding, as long as the sow is of purebred Iberian stock. The reproductive performance of the Iberian sows is lower than that of white pig populations [3], which is a major limitation of its use in intensive farms. Therefore, improvement in the reproductive efficiency of Iberian sows is crucial for their economic sustainability. Several studies have identified genetic variability for prolificacy within and between varieties of Iberian pig [4,5]. To take advantage of this variability, the INGA FOOD, S.A. company has developed a crossbreeding scheme between two Iberian varieties (Retinto and Entrepelado) to generate an F1 hybrid sow, which exhibits significant heterosis for litter size [6]. However, this study also found differences in the reproductive performance between the two reciprocal crosses (Entrepelado × Retinto, ER, vs. Retinto × Entrepelado, RE), suggesting that these differences may be attributed to parental imprinting [7] (i.e., the effects from alleles may differ whether they are transmitted by paternal or maternal gametes). In fact, there is increasing evidence of the importance of imprinting in placenta development [8], and certain imprinted genes have been proposed as candidates for pig litter size [9].

In recent years, some algorithms have been proposed to develop a genomic analysis of imprinting [10] from the genomic information provided by commercial genotyping devices. However, knowledge of the parental haplotype phase of the SNP markers is required to differentiate the paternal or maternal gametic effects. Some approaches have been developed to reconstruct haplotype phases [11].

Phenotypic information from reciprocal crosses offers the opportunity to compare the paternal and maternal effects of each parental population. In the absence of imprinting, the correlation between the paternal and maternal effects from the same population should be one. Imprinting, on the other hand, results in a lower correlation. Accordingly, the goal of this study was to apply the multivariate gametic model developed in a previous study [12] that utilizes genomic information and is capable of estimating the paternal and maternal gametic contributions of Retinto and Entrepelado varieties in the ER and RE crosses, along with their correlations.

2. Materials and Methods

Phenotypic and Genomic Data. The phenotypic data used in this study consisted of the total number born, TNB, and the number of piglets born alive, NBA, in 203 ER and 125 RE sows. The ER sows were the offspring of 38 purebred Entrepelado boars and 139 Retinto dams, whereas the RE sows were generated from 38 Retinto boars and 92 Entrepelado dams. A summary of the data is presented in Table 1.

Table 1. The number of records (and sows between brackets), mean (±standard deviation) of the total number born and the number born alive, for Entrepelado × Retinto and Retinto × Entrepelado crosses.

	Entrepelado × Retinto	**Retinto × Entrepelado**
N [1] (NS) [2]	1258 (203)	700 (125)
TNB [3]	8.78 ± 2.24	8.85 ± 2.37
NBA [4]	8.55 ± 2.23	8.62 ± 2.34

[1] N: number of records. [2] NS: number of sows. [3] TNB: total number born. [4] NBA: number born alive.

Genotyping was performed with the GeneSeek® GPP Porcine 70 K HDchip (Illumina Inc., San Diego, CA, USA) on all ER and RE crossbred sows, as well as on 341 Retinto and 350 Entrepelado purebred individuals. Due to shared purebred ancestors, there was some degree of relationship between a subset of the ER and RE crossbred sows and the purebred individuals, although not all of them were genotyped. The original genotype data consisted of 60,224 autosomal SNPs, which were filtered by excluding SNP markers with a call rate below 0.90 and a minor allele frequency lower than 0.05 in each population. Among these, 4212 were discarded due to a call rate lower than 0.90, 11,234 were found to be monomorphic, and 9876 and 11,516 had a minor allele frequency lower than 0.05 in the Entrepelado and Retinto populations, respectively. Finally, a total of 23,386 SNPs were retained.

Haplotype Phasing. AlphaPhase software [11] was used for each chromosome separately, utilizing genotypes of both crossbred and purebred individuals, as well as a pedigree of 1601 individuals. AlphaPhase was executed with a tolerance of 1% of genotype errors and 1% disagreement between genotypes and haplotypes. The number of surrogates and percentage of surrogate disagreement was set to 10. Nine different scenarios were applied with core lengths of 75, 100, and 125 SNPs and tail lengths of 100, 150, and 200 SNPs (see Table 2). The scenarios were evaluated for concordance, and haplotype assignments that coincided in seven or more scenarios were retained for subsequent analysis.

Table 2. Parameters (core and tail length) in the nine scenarios of haplotype phasing.

Scenario	Core Length	Tail Length
S1	75	100
S2	75	150
S3	75	200
S4	100	100
S5	100	150
S6	100	200
S7	125	100
S8	125	150
S9	125	200

Statistical Model. Once the haplotype phases were calculated, data were analyzed with the model proposed by Shiri et al. [12]:

$$\boldsymbol{y}_{(ER)} = \boldsymbol{X}_{(ER)}\boldsymbol{b}_{(ER)} + \boldsymbol{B}_{(ER)}\boldsymbol{s}_{(ER)} + \boldsymbol{Z}_{(ER)}\boldsymbol{p}_{(E)} + \boldsymbol{W}_{(ER)}\boldsymbol{m}_{(R)} + \boldsymbol{e}_{(ER)}$$
$$\boldsymbol{y}_{(RE)} = \boldsymbol{X}_{(RE)}\boldsymbol{b}_{(RE)} + \boldsymbol{B}_{(RE)}\boldsymbol{s}_{(RE)} + \boldsymbol{Z}_{(RE)}\boldsymbol{p}_{(R)} + \boldsymbol{W}_{(RE)}\boldsymbol{m}_{(E)} + \boldsymbol{e}_{(RE)}$$

In this equation, $\boldsymbol{y}_{(ER)}$ and $\boldsymbol{y}_{(RE)}$ refer to the vectors of phenotypic records (TNB or NBA) for the ER and RE crosses, respectively. The terms $\boldsymbol{b}_{(ER)}$ and $\boldsymbol{b}_{(RE)}$ correspond to systematic effects, and $\boldsymbol{s}_{(ER)}$ and $\boldsymbol{s}_{(RE)}$ represent the permanent sow environmental effects. Paternal effects for the Entrepelado (E) and Retinto (R) populations are denoted by $\boldsymbol{p}_{(E)}$ and $\boldsymbol{p}_{(R)}$, respectively. Maternal effects for the Entrepelado (E) and Retinto (R) are represented by $\boldsymbol{m}_{(E)}$ and $\boldsymbol{m}_{(R)}$. Additionally, $\boldsymbol{e}_{(ER)}$ and $\boldsymbol{e}_{(RE)}$ are the residual effects for the ER and RE crosses, respectively. The systematic effects vectors included the order of parity with five levels (first, second, third, fourth, and fifth or more) and herd–year–season with thirty-four levels. Further, $\boldsymbol{X}_{ER}, \boldsymbol{X}_{RE}, \boldsymbol{B}_{ER}, \boldsymbol{B}_{RE}, \boldsymbol{Z}_{ER}, \boldsymbol{Z}_{RE}, \boldsymbol{W}_{ER}$, and $\boldsymbol{W}_{RE}$ are the corresponding incidence matrices.

Following [12], the prior distribution of the permanent sow environmental effects was:

$$\begin{bmatrix} \boldsymbol{s}_{(ER)} \\ \boldsymbol{s}_{(RE)} \end{bmatrix} \sim N\begin{pmatrix} 0 \\ 0, \boldsymbol{I} \otimes \boldsymbol{S} \end{pmatrix}$$

where:

$$\boldsymbol{S} = \begin{bmatrix} \sigma^2_{s(ER)} & 0 \\ 0 & \sigma^2_{s(RE)} \end{bmatrix}$$

where $\sigma^2_{s(ER)}$ and $\sigma^2_{s(RE)}$ are the variances of the permanent sow environmental effects for ER and RE, respectively. The prior distributions of the gametic effects for the Entrepelado (E) and Retinto (R) populations are:

$$\begin{bmatrix} \boldsymbol{p}_{(E)} \\ \boldsymbol{m}_{(E)} \end{bmatrix} \sim N\begin{pmatrix} 0 \\ 0' \end{pmatrix}, \boldsymbol{G_E} \otimes \boldsymbol{V_E}) \begin{bmatrix} \boldsymbol{p}_{(R)} \\ \boldsymbol{m}_{(R)} \end{bmatrix} \sim N\begin{pmatrix} 0 \\ 0' \end{pmatrix}, \boldsymbol{G_R} \otimes \boldsymbol{V_R})$$

where:

$$\boldsymbol{V_E} = \begin{bmatrix} \sigma^2_{p(E)} & \sigma_{pm(E)} \\ \sigma_{pm(E)} & \sigma^2_{m(E)} \end{bmatrix}$$

and:

$$\boldsymbol{V_R} = \begin{bmatrix} \sigma^2_{p(R)} & \sigma_{pm(R)} \\ \sigma_{pm(R)} & \sigma^2_{m(R)} \end{bmatrix}$$

where $\sigma^2_{p(E)}$, $\sigma^2_{m(E)}$, and $\sigma_{pm(E)}$ refer to the variances of the paternal and maternal gametic effects and the covariance between them, respectively, for the Entrepelado population. Similarly, $\sigma^2_{p(R)}$, $\sigma^2_{m(R)}$, and $\sigma_{pm(R)}$ represent the variances of the paternal and maternal gametic effects and the covariance between them, respectively, for the Retinto population. Additionally, $\boldsymbol{G_E}$ and $\boldsymbol{G_R}$ are the gametic relationship matrices of the Entrepelado or Retinto gametes, respectively, regardless of whether they are transmitted as paternal or maternal gametes. These matrices describe the relationships among the gametes from Entrepelado and Retinto origins, and they are calculated using the algorithm proposed by Nishio and Satoh [10]:

$$\boldsymbol{G_E} = \frac{\boldsymbol{M_E}\boldsymbol{M_E}'}{\sum_i^{N_{SNP}} q_{(E)i}\left(1 - q_{(E)i}\right)} \; \boldsymbol{G_R} = \frac{\boldsymbol{M_R}\boldsymbol{M_R}'}{\sum_i^{N_{SNP}} q_{(R)i}\left(1 - q_{(R)i}\right)}$$

where $\boldsymbol{M_E}$ and $\boldsymbol{M_R}$ are the matrices of the number of genotyped individuals ($\boldsymbol{n}$) $\times$ the number of SNP (N_{SNP}), whose elements $M_E(i,j)$ (or $M_R(i,j)$ take the value $q(E)_j$ (or $q(R)_j$) or $-\left(1 - q(E)_j\right)$ (or $-\left(1 - q(R)_j\right)$), depending on whether the *jth* allele of the gametes transmitted for the *ith* individual is A_1 or A_2 and of Entrepelado (or Retinto) origin. Additionally, $q(E)_j$ and $q(R)_j$ represent the allelic frequencies of the A_2 allele in the Entrepelado (E) and Retinto (R) populations, respectively. The prior distributions for the (co) variance components and the systematic effects were assumed to be flat. The analysis was performed using Bayesian inference with the Gibbs sampler [13] and implemented with Gibbsf90 software [14]. The analysis was performed using 10 million iterations after discarding the first million.

At each iteration of the Gibbs sampler, the (co) variances components samples were utilized to compute the samples from the marginal posterior distribution of the correlations between the paternal and maternal gametic effects for Entrepelado ($r_{pm(E)}$) and Retinto ($r_{pm(R)}$):

$$r_{pm(E)} = \frac{\sigma_{pm(E)}}{\sqrt{\sigma^2_{p(E)}\sigma^2_{j(E)}}} \text{ and } r_{pm(R)} = \frac{\sigma_{pm(R)}}{\sqrt{\sigma^2_{p(R)}\sigma^2_{j(R)}}}$$

3. Results and Discussion

Haplotype Phasing. The results of comparing haplotype phasing using nine combinations of core length and core tail parameters using Alphaphase software are presented in Figure 1.

Figure 1. Degree of similitude between estimated haplotype phases in the nine scenarios of phasing.

The average degree of similitude was 0.89, and it was consistently above 0.86. Specifically, the predicted haplotype phase was identical across all nine scenarios for only 78.74% of the analyses but had concordance in more than seven scenarios for 92.5% of SNPs. These findings indicated that the output of the phasing algorithm was highly dependent on the specific set of parameters used for its implementation when medium-density SNP chips were used.

Calculation of Gametic Matrices. The diagonal values of the gametic matrices for the Entrepelado population ranged from 0.894 to 1.100, while for the Retinto population, they ranged from 0.901 to 1.179. Table 3 shows the distribution of the gametic relationships observed in the off-diagonal elements of the gametic matrices.

Table 3. Distribution of gametic relationships between the Entrepelado and Retinto gametic effects.

Gametic Relationship	Entrepelado	Retinto
<0.05	92,276 (86.03%)	94,144 (87.77%)
0.05–0.10	8130 (7.58%)	8900 (8.29%)
0.10–0.20	4670 (4.35%)	3130 (2.92%)
0.20–0.30	1076 (1.00%)	582 (0.54%)
0.30–0.40	480 (0.44%)	252 (0.23%)
0.40–0.50	396 (0.36%)	188 (0.18%)
>0.50	228 (0.21%)	60 (0.05%)

The calculated gametic matrices yielded results consistent with the familiar relationships of the individuals, as gametic relationships around 0.50 indicated that the individuals shared sire (or dam), while gametic relationships around 0.25 suggested that the sires (or dams) of the individuals were fullsibs.

Variance Components. The posterior mean and standard deviation estimate of the variance components are presented in Table 4.

Table 4. Posterior mean (and standard deviation) of the variance components for the total number born, TNB, and the number born alive, NBA.

Variance Component	TNB	NBA
$\sigma^2_{s(ER)}$	0.144 (0.098)	0.142 (0.097)
$\sigma^2_{s(RE)}$	0.357 (0.187)	0.365 (0.191)
$\sigma^2_{p(E)}$	0.206 (0.103)	0.199 (0.100)
$\sigma^2_{m(E)}$	0.197 (0.114)	0.199 (0.115)
$\sigma^2_{p(R)}$	0.224 (0.132)	0.222 (0.128)
$\sigma^2_{m(R)}$	0.163 (0.087)	0.151 (0.080)
$\sigma^2_{e(ER)}$	4.296 (0.187)	4.251 (0.185)
$\sigma^2_{e(RE)}$	4.795 (0.288)	4.607 (0.278)

Furthermore, Figure 2 shows the posterior distributions of the ratios of gametic variances in the Entrepelado × Retinto (E × R) and Retinto × Entrepelado (R × E) crosses. The posterior mean estimates were similar, ranging between 0.034 for the Retinto maternal gametic effects in the ER cross and 0.043 for the Entrepelado paternal gametic effects in the RE cross.

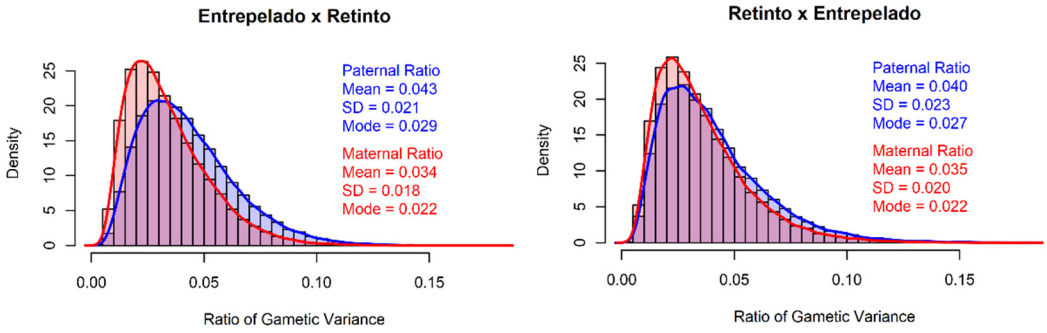


Figure 2. Posterior distributions of the ratio of gametic effect in the Entrepelado × Retinto and in the Retinto × Entrepelado crosses.

These results indicate that there are no relevant differences in the amount of genetic variance contributed by the paternal and maternal origins in either of the two reciprocal crosses, based on the available information.

Gametic Correlations. The posterior distribution of the gametic correlations for the TNB and NBA in the Entrepelado and Retinto populations are presented in Figures 3 and 4, respectively.

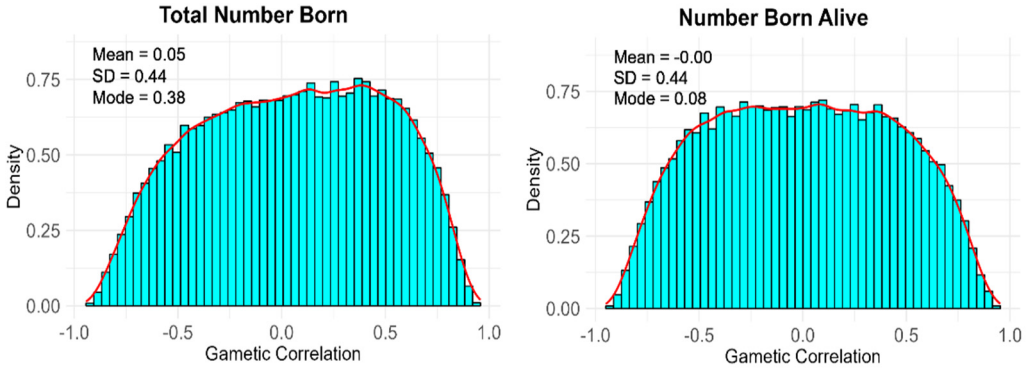


Figure 3. Posterior distributions of the gametic correlation between the paternal and maternal effects in the Entrepelado population.

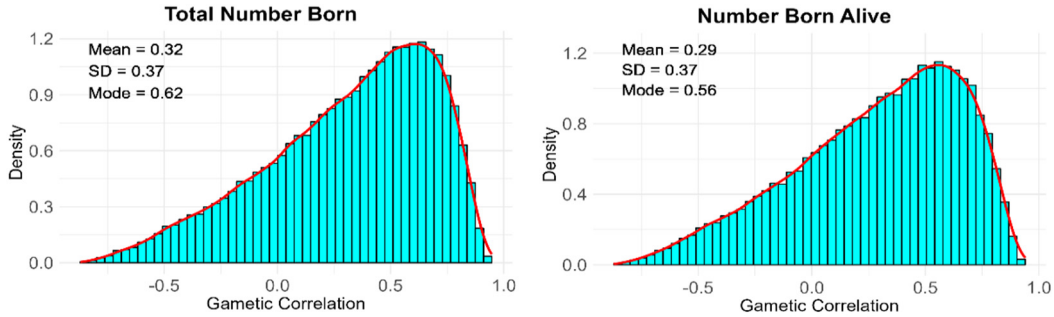


Figure 4. Posterior distributions of the gametic correlation between the paternal and maternal effects in the Retinto population.

The posterior distribution of the correlation between gametic effects in Retinto and Entrepelado showed notable differences in shape. Specifically, the posterior distributions of the gametic correlations in the Retinto population exhibited a higher degree of positive asymmetry compared to those in the Entrepelado population. In fact, the posterior probabilities of a positive gametic correlation in the Retinto population were 0.80 and 0.78 for the TNB and NBA, respectively. In contrast, the posterior probabilities of a positive gametic correlation in the Entrepelado population were 0.50 (TNB) and 0.54 (NBA).

Although caution is needed in interpreting the results due to the limited amount of phenotypic and genotypic information, the shape of the posterior distribution of gametic correlations suggests a potential role of genomic imprinting. This is because a gametic correlation substantially lower than one indicates that the same combination of alleles in a gamete may produce different effects on offspring depending on whether they are transmitted by paternal or maternal gametes, which is consistent with the theory of genomic imprinting. Genomic imprinting is an epigenetic phenomenon that causes genes to be expressed depending on whether they are inherited from the father or mother [7].

Several theories have been postulated to explain the evolutionary origin of genomic imprinting [15], and one of the most popular is the parental investment theory [16]. This theory argues that imprinting is the result of a conflict between the evolutionary success of paternally and maternally derived genes. In mammalian reproduction, the evolutionary success of paternally inherited genes is associated with the increase in fetal growth, while for maternally inherited genes, it is associated with the number of offspring. This theory is reinforced by the discovery of numerous imprinted genes known to regulate aspects of mammalian development [17], including growth, behavior, and placental function [18] and, furthermore, there is increasing evidence of imprinted genes in the pig genome [9,19,20].

From a practical perspective, a low or null gametic correlation between paternal and maternal gametes within the same population indicates that a selection program to improve the performance of the crossbreeding individuals needs to be specifically designed, especially in the Entrepelado population. This is because the selection of purebred animals to increase the performance in the Entrepelado × Retinto cross may not have any noticeable consequences in the performance in the Retinto × Entrepelado cross. Furthermore, this result also may explain the differences in performance among the reciprocal crosses observed by Noguera et al. [6], who proposed using the Retinto variety as a boar and the Entrepelado as a sow, providing better performance than the opposite cross.

4. Conclusions

The bivariate model proposed in this study provides estimates of the gametic effects of each founder population as either paternal or maternal, as well as their correlation. In the absence of parental imprinting, a perfect correlation of one would be expected. However, our results detect a significant deviation from this ideal scenario, indicating possible differences in the performance of crossbred individuals depending on the paternal or maternal origin of the gametes. These findings provide evidence of the presence of

imprinting effects in Iberian pig populations, which could have implications for the design of future breeding programs.

Author Contributions: Conceptualization, L.V.; methodology, H.S., D.L.-C. and L.V.; software, H.S. and L.V.; validation, H.S., D.L.-C. and L.V.; formal analysis, H.S. and D.L.-C.; investigation, H.S.; resources, N.I.-E. and J.C.; data curation, S.N., P.H. and N.I.-E.; writing—original draft preparation, H.S. and L.V.; writing—review and editing, N.I.-E., P.H. and J.C.; visualization, D.L.-C. and H.S.; supervision, L.V.; project administration, N.I.-E. and J.C.; funding acquisition, N.I.-E. and J.C. All authors have read and agreed to the published version of the manuscript.

Funding: The research was partially funded by grants CGL2016-80155-R, PID2020-114705RB-I00 (Ministerio de Ciencia e Innovación), and IDI-20170304 (CDTI). Srihi received funding from the European Union's H2020 research and innovation program under a Marie Sklodowska–Curie grant agreement, No. 801586.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset used in this study will be available upon reasonable request to the corresponding author (lvarona@unizar.es).

Acknowledgments: We thank INGA FOOD, S.A. (Almendralejo, Spain) and its technicians (E. Magallón, L. Muñoz, P. Díaz, and D. Iniesta) for their cooperation and technical support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Serra, X.; Gil, F.; Pérez-Enciso, M.; Oliver, M.A.; Vázquez, J.M.; Gispert, M.; Díaz, I.; Moreno, F.; Latorre, R.; Noguera, J.L. A comparison of carcass, meat quality and histochemical characteristics of Iberian (Guadyerbas line) and Landrace pigs. *Livest. Prod. Sci.* **1998**, *56*, 215–223. [CrossRef]
2. Serrano, M.P.; Valencia, D.G.; Nieto, M.; Lázaro, R.; Mateos, G.G. Influence of sex and terminal sire line on performance and carcass and meat quality of Iberian pigs reared under intensive production systems. *Meat Sci.* **2008**, *78*, 420–428. [CrossRef]
3. Silio, L.; Rodriguez, M.C.; Rodrigañez, J.; Toro, M.A. La selección de cerdos ibéricos. In *Porcino Ibérico: Aspectos Claves*; Buxade, C., Daza, A., Eds.; Mundi-Press: Madrid, Spain, 2001; pp. 125–149.
4. Fernández, A.; Rodrigáñez, J.; Zuzúarregui, J.; Rodríguez, M.C.; Silió, L. Genetic parameters for litter size and weight at different parities in Iberian pigs. *Span. J. Agric. Res.* **2008**, *6*, 98–106. [CrossRef]
5. García-Casco, J.M.; Fernández, A.; Rodríguez, M.C.; Silió, L. Heterosis for litter size and growth in crosses of four strains of Iberian pig. *Livest. Sci.* **2012**, *147*, 1–8. [CrossRef]
6. Noguera, J.L.; Ibáñez-Escriche, N.; Casellas, J.; Rosas, J.P.; Varona, L. Genetic parameters and direct, maternal and heterosis effects on litter size in a diallel cross among three commercial varieties of Iberian pig. *Animal* **2019**, *13*, 2765–2772. [CrossRef] [PubMed]
7. Reik, W.; Walter, J. Genomic imprinting: Parental influence on the genome. *Nat. Rev. Genet.* **2001**, *2*, 21–32. [CrossRef] [PubMed]
8. Hanna, C.W. Placental imprinting: Emerging mechanisms and functions. *PLoS Genet.* **2020**, *16*, e1008709. [CrossRef] [PubMed]
9. Coster, A.; Madsen, O.; Heuven, H.C.M.; Dibbits, B.; Groenen, M.A.M.; van Arendonk, J.A.M.; Bovenhuis, H. The imprinted gene DIO3 is a candidate gene for litter size in pigs. *PLoS ONE* **2012**, *7*, e31825. [CrossRef] [PubMed]
10. Nishio, M.; Satoh, M. Genomic best linear unbiased prediction method including imprinting effects for genomic evaluation. *Genet. Sel. Evol.* **2015**, *47*, 32. [CrossRef] [PubMed]
11. Hickey, J.M.; Kinghorn, B.P.; Tier, B.; Wilson, J.F.; Dunstan, N.; Van Der Werf, J.H.J. A combined long-range phasing and long haplotype imputation method to impute phase for SNP genotypes. *Genet. Sel. Evol.* **2011**, *43*, 12. [CrossRef] [PubMed]
12. Srihi, H.; Ibáñez-Escriche, N.; Casellas, J.; Noguera, J.L.; Hernández, P.; Martín de Hijas, M.; Vázquez-Gómez, M.; Negro, S.; Rosas, J.P.; Varona, L. Bayesian analysis of paternal and maternal gametic effects in a reciprocal cross between two Iberian varieties. In Proceedings of the 12th World Congress on Genetics Applied to Livestock Production, Rotterdam, The Netherlands, 3–8 July 2022; pp. 3180–3183.
13. Gelfand, A.E.; Smith, A.F.M. Sampling-Based Approaches to Calculating Marginal Densities. *J. Am. Stat. Assoc.* **1990**, *85*, 398–409. [CrossRef]
14. Misztal, I.; Tsuruta, S.; Lourenco, D.; Aguilar, I.; Legarra, A.; Vitezica, Z. *Manual for BLUPF90 Family of Programs*; University of Georgia: Athens, GA, USA, 2018; p. 142.
15. Patten, M.M.; Ross, L.; Curley, J.P.; Queller, D.C.; Bonduriansky, R.; Wolf, J.B. The evolution of genomic imprinting: Theories, predictions and empirical tests. *Heredity* **2014**, *113*, 119–128. [CrossRef] [PubMed]
16. Moore, T.; Haig, D. Genomic imprinting in mammalian development: A parental tug-of-war. *Trends Genet.* **1991**, *7*, 45–49. [CrossRef] [PubMed]

17. Thamban, T.; Agarwaal, V.; Khosla, S. Role of genomic imprinting in mammalian development. *J. Biosci.* **2020**, *45*, 20. [CrossRef] [PubMed]
18. Fowden, A.L.; Coan, P.M.; Angiolini, E.; Burton, G.J.; Constancia, M. Imprinted genes and the epigenetic regulation of placental phenotype. *Prog. Biophys. Mol. Biol.* **2011**, *106*, 281–288. [CrossRef] [PubMed]
19. Wu, Y.Q.; Zhao, H.; Li, Y.J.; Khederzadeh, S.; Wei, H.J.; Zhou, Z.Y.; Zhang, Y.P. Genome-wide identification of imprinted genes in pigs and their different imprinting status compared with other mammals. *Zool. Res.* **2020**, *41*, 721–725. [CrossRef] [PubMed]
20. Zhang, F.W.; Han, Z.B.; Deng, C.Y.; He, H.J.; Wu, Q. Conservation of genomic imprinting at the NDN, MAGEL2 and MEST loci in pigs. *Genes Genet. Syst.* **2012**, *87*, 53–58. [CrossRef] [PubMed]

Received: 28 June 2023 | Revised: 6 September 2023 | Accepted: 14 October 2023

DOI: 10.1111/jbg.12832

ORIGINAL ARTICLE

Journal of Animal Breeding and Genetics WILEY

A multivariate gametic model for the analysis of purebred and crossbred data. An example between two populations of Iberian pigs

Houssemeddine Srihi[1] | **David López-Carbonell**[1] | **Noelia Ibáñez-Escriche**[2] | **Joaquim Casellas**[3] | **Pilar Hernández**[2] | **Sara Negro**[4] | **Luis Varona**[1]

[1]Facultad de Veterinaria, Instituto Agroalimentario de Aragón (IA2), Universidad de Zaragoza, Zaragoza, Spain

[2]Universitat Politècnica de València, Valencia, Spain

[3]Universitat Autònoma de Barcelona, Barcelona, Spain

[4]INGA FOOD S.A. (Nutreco), Almendralejo, Spain

Correspondence
Luis varona, Facultad de Veterinaria, Instituto Agroalimentario de Aragón (IA2), Universidad de Zaragoza, 50013 Zaragoza, Spain.
Email: lvarona@unizar.es

Funding information
Ministerio de Ciencia e Innovación, Grant/Award Number: PID2020-114705RB-I00 and CGL2016-80155-R; CDTI, Grant/Award Number: IDI-20170304; European Union's H2020 research and innovation program, Grant/Award Number: 801586

Abstract

Crossbreeding plays a pivotal role within pig breeding programmes, aiming to maximize heterosis and improve reproductive traits in crossbred maternal lines. Nevertheless, there is evidence indicating that the performance of reciprocal crosses between two genetic lines might exhibit variability. These variations in performance can be attributed to differences in the correlations between gametic effects, acting as either sire or dam, within purebred and crossbred populations. To address this issue, we propose a multivariate gametic model that incorporates up to four correlated gametic effects for each parental population. The model is employed on a data set comprising litter size data (total number of piglets born—TNB- and number of piglets born alive—NBA-) derived from a reciprocal cross involving two Iberian pig populations: Entrepelado and Retinto. The data set comprises 6933 records from 1564 purebred Entrepelado (EE) sows, 4995 records from 1015 Entrepelado × Retinto (ER) crosses, 2977 records from 756 Retinto × Entrepelado (RE) crosses and 7497 records from 1577 purebred Retinto (RR) sows. The data set is further supplemented by a pedigree encompassing 6007 individual-sire-dam entries. The statistical model also included the order of parity (with six levels), the breed of the service sire (five levels) and the herd-year-season effects (141 levels). Additionally, the model integrates random dominant and permanent environmental sow effects. The analysis employed a Bayesian approach, and the results revealed all the posterior estimates of the gametic correlations to be positive. The range of the posterior mean estimates of the correlations varied across different gametic effects and traits, with a range between 0.04 (gametic correlation between the paternal effects for purebred and the maternal for crossbred in Retinto) and 0.53 (gametic correlation between the paternal effects for purebred and the paternal for crossbred in Entrepelado). Furthermore, the posterior mean variance estimates of the maternal gametic effects were consistently surpassed those for paternal effects within all four

populations. The results suggest the possible influence of imprinting effects on the genetic control of litter size, and underscore the importance of incorporating crossbred data into the breeding value predictions for purebred individuals.

KEYWORDS
crossbreeding, gametic correlation, parent of origin, reciprocal cross

1 | INTRODUCTION

The Iberian pig breed is renowned for its adaptability to the natural Mediterranean ecosystem in Southwestern Spain, as well as its product quality (Lopez-Bote, 1998). Despite the Iberian pigs exhibit slower growth, lower feed efficiency and reduced prolificacy compared to other commercial pig populations, their meat quality (Gilles, 2009) and specialized derivative products contribute to their sustainability. However, in recent decades, the traditional extensive production methods that solely relied on purebred Iberian individuals have been partly replaced by intensive farming practices involving crossbreeding with Duroc pigs. The rationale behind this crossbreeding is to enhance growth and efficiency (Serrano et al., 2008). This shift towards intensive farming has yielded several advantages, such the collection of productive data and the implementation of genetic selection programmes. It is important to note that regulatory norms governing the production of Iberian pig products stipulate that the sow must belong to a pure Iberian lineage. This highlights the importance of improving the reproductive efficiency of Iberian sows to ensure the economic sustainability of the breed.

INGA FOOD S.A. is a Spanish company specializing in the production and distribution of premium pig products. As part of their breeding programme, they have successfully developed a hybrid Iberian sow called CASTUA. This hybrid sow is the result of crossbreeding between the Retinto and Entrepelado populations of the Iberian breed. Through this crossbreeding, the CASTUA hybrid sow exhibit improved litter size traits due to heterosis, as confirmed by Noguera et al. (2019). Furthermore, INGA FOOD S.A. is actively implementing a breeding programme with the aim of further enhancing litter size based on the performance of purebred individuals. The foundation of this programme is built upon the assumption of a positive genetic correlation between the performance of purebred and crossbred pigs (Wientjes & Calus, 2017). This hypothesis finds support in a preceding study (Srihi et al., 2022) that quantified the genetic correlation between the performance of purebred and crossbred animals within the Entrepelado and Retinto populations.

Furthermore, as highlighted by Noguera et al. (2019), variations in the performance of reciprocal crosses suggest that the gametic contribution to genetic variance may diverge when the populations act as either sires or dams in the crossbreeding process (Srihi et al., 2023). While previous literature has employed gametic models to identify differences in parent-of-origin effects (Meyer & Tier, 2012; Varona et al., 2015), to the best of our knowledge, these models have not been previously implemented within the context of crossbreeding. Hence, the goal of this study was to develop a multivariate gametic model encompassing both paternal and maternal gametic effects for purebred and crossbred performance. This framework will enable the estimation of gametic correlations between these effects. Additionally, the model will also incorporate a dominance effect.

2 | MATERIALS AND METHODS

The data set used in this study consisted of a total of 22,402 records for two reproductive traits: total number of piglets born (TNB) and number of piglets born alive (NBA). These records were collected from 4912 sows that were part of a complete diallelic experiment involving two strains of the Iberian pig breed: Retinto (RR) and Entrepelado (EE). Additionally, the data set included their reciprocal crosses: Entrepelado × Retinto (ER) and Retinto × Entrepelado (RE). Along with the phenotypic data, a pedigree containing 6007 individual-sire-dam entries was included for genetic analysis. Table 1 provides a summary of the phenotypic data used in this study.

The data were analysed with the following models:

$$\boldsymbol{y}_{EE} = \boldsymbol{X}_{EE}\boldsymbol{b}_{EE} + \boldsymbol{Z}_{EE}\boldsymbol{p}_{EE} + \boldsymbol{W}_{EE}\boldsymbol{m}_{EE} + \boldsymbol{Q}_{EE}\boldsymbol{d}_{EE} + \boldsymbol{H}_{EE}\boldsymbol{r}_{EE} + \boldsymbol{e}_{EE}$$
$$\boldsymbol{y}_{ER} = \boldsymbol{X}_{ER}\boldsymbol{b}_{ER} + \boldsymbol{Z}_{ER}\boldsymbol{p}_{ER} + \boldsymbol{W}_{ER}\boldsymbol{m}_{ER} + \boldsymbol{Q}_{ER}\boldsymbol{d}_{ER} + \boldsymbol{H}_{ER}\boldsymbol{r}_{ER} + \boldsymbol{e}_{ER}$$
$$\boldsymbol{y}_{RR} = \boldsymbol{X}_{RR}\boldsymbol{b}_{RR} + \boldsymbol{Z}_{RR}\boldsymbol{p}_{RR} + \boldsymbol{W}_{RR}\boldsymbol{m}_{RR} + \boldsymbol{Q}_{RR}\boldsymbol{d}_{RR} + \boldsymbol{H}_{RR}\boldsymbol{r}_{RR} + \boldsymbol{e}_{RR}$$
$$\boldsymbol{y}_{RE} = \boldsymbol{X}_{RE}\boldsymbol{b}_{RE} + \boldsymbol{Z}_{RE}\boldsymbol{p}_{RE} + \boldsymbol{W}_{RE}\boldsymbol{m}_{RE} + \boldsymbol{Q}_{RE}\boldsymbol{d}_{RE} + \boldsymbol{H}_{\boldsymbol{RE}}\boldsymbol{r}_{RE} + \boldsymbol{e}_{RE}$$

In the given equations, $\boldsymbol{y}_{JK}$ is the vector of phenotypic records (TNB or NBA) for the $JK = \{EE, ER, RR, RE\}$ population. Here, $J = \{E, R\}$ denotes the paternal population, and $K = \{E, R\}$ denotes the maternal. Additionally, $\boldsymbol{b}_{JK}$ is the vector of systematic effects, including order of parity six levels), breed of service sire (five levels) and herd-year-season (141 levels). Moreover, $\boldsymbol{p}_{JK}$, $\boldsymbol{m}_{JK}$, $\boldsymbol{d}_{JK}$, $\boldsymbol{r}_{JK}$ and $\boldsymbol{e}_{JK}$ are the paternal, maternal, dominance, permanent environmental and residual effects of the JK population respectively. It must be noted that $\boldsymbol{p}_{JK}$ is the vector of the paternal gametic effects of the $J = \{E, R\}$ population in the

TABLE 1 Number of phenotypic records (and number of sows producing them in brackets), mean (± standard deviation) of Total Number Born (TNB) and Number Born Alive (NBA) for Entrepelado (EE), Retinto (RR) and Entrepelado × Retinto (ER) and Retinto × Entrepelado (RE) crosses.

	EE	RR	ER	RE
N[a] (NS[b])	6933 (1564)	7497 (1577)	4995 (1015)	2977 (756)
TNB	8.23 ± 2.14	8.44 ± 2.22	8.55 ± 2.27	8.51 ± 2.28
NBA	7.87 ± 2.11	8.05 ± 2.18	8.27 ± 2.25	8.18 ± 2.25

[a]N: number of phenotypic records.

[b]NS: number of recorded sows.

JK cross, and that $\boldsymbol{m}_{JK}$ is the vector of the $K=\{E, R\}$ maternal gametic effects in the $JK=\{EE, ER, RR, RE\}$ cross. Besides, $\boldsymbol{X}_{JK}, \boldsymbol{Z}_{JK}, \boldsymbol{W}_{JK}, \boldsymbol{Q}_{JK}$ and $\boldsymbol{H}_{JK}$ are the corresponding incidence matrices involved in the equations.

The statistical model was analysed by employing a Bayesian approach with a Gibbs sampler (Gelfand & Smith, 1990). In this analysis, bounded uniform distributions were employed as prior distributions for the systematic effects and variance components. The prior distributions for the gametic, dominance, permanent environmental and residual effects were modelled as multivariate Gaussian distributions, characterized by a zero mean and the following variances:

$$var\begin{pmatrix} p_{EE} \\ p_{RR} \\ p_{ER} \\ p_{RE} \\ m_{EE} \\ m_{RR} \\ m_{ER} \\ m_{RE} \end{pmatrix} = \boldsymbol{T} \otimes \boldsymbol{G}, var\begin{pmatrix} d_{EE} \\ d_{RR} \\ d_{ER} \\ d_{RE} \end{pmatrix} = \boldsymbol{Q} \otimes \boldsymbol{D},$$

$$var\begin{pmatrix} r_{EE} \\ r_{RR} \\ r_{ER} \\ r_{RE} \end{pmatrix} = \boldsymbol{R} \otimes \boldsymbol{I} \text{ and } var\begin{pmatrix} e_{EE} \\ e_{RR} \\ e_{ER} \\ e_{RE} \end{pmatrix} = \boldsymbol{E} \otimes \boldsymbol{I}$$

where $\boldsymbol{G}$ and $\boldsymbol{D}$ are the gametic and dominance relationship matrix (Smith, 1984), $\boldsymbol{I}$ is the identity matrix and where $\sigma^2_{X_{JK}}$ represents the paternal ($X=P$) or maternal ($X=M$) gametic variance in the JK population, JK can represent any of the following combinations *{EE, ER, RR, RE}*. Similarly, $\sigma_{X_{JK}Y_{LM}}$ denotes the covariance between the gametic effects between the paternal (if $X=P$) or maternal (if $X=M$) gametic variance in the *JK* cross with the paternal (if $Y=P$) or maternal (if $Y=M$) gametic variance in the LM cross, where *JK* and *LM* can each be *EE, ER, RR* and *RE* respectively. It is essential to emphasize that in this analysis, the covariances between gametic effects from distinct populations (Retinto or Entrepelado) are explicitly fixed at zero, allowing no correlation between them. Conversely, the model allows for non-null covariance between gametic effects from the same population.

Further,

$$\boldsymbol{Q} = \begin{bmatrix} \sigma^2_{d_{EE}} & 0 & 0 & 0 \\ 0 & \sigma^2_{d_{RR}} & 0 & 0 \\ 0 & 0 & \sigma^2_{d_{ER}} & 0 \\ 0 & 0 & 0 & \sigma^2_{d_{RE}} \end{bmatrix},$$

$$\boldsymbol{R} = \begin{bmatrix} \sigma^2_{r_{EE}} & 0 & 0 & 0 \\ 0 & \sigma^2_{r_{RR}} & 0 & 0 \\ 0 & 0 & \sigma^2_{r_{ER}} & 0 \\ 0 & 0 & 0 & \sigma^2_{r_{RE}} \end{bmatrix}, \text{and}$$

$$\boldsymbol{T} = \begin{bmatrix} \sigma^2_{P_{EE}} & 0 & \sigma_{P_{EE}P_{ER}} & 0 & \sigma_{P_{EE}M_{EE}} & 0 & 0 & \sigma_{P_{EE}M_{RE}} \\ 0 & \sigma^2_{P_{RR}} & 0 & \sigma_{P_{RR}P_{RE}} & 0 & \sigma_{P_{RR}M_{RR}} & \sigma_{P_{RR}M_{ER}} & 0 \\ \sigma_{P_{EE}P_{ER}} & 0 & \sigma^2_{P_{ER}} & 0 & \sigma_{P_{ER}M_{EE}} & 0 & 0 & \sigma_{P_{ER}M_{RE}} \\ 0 & \sigma_{P_{RR}P_{RE}} & 0 & \sigma^2_{P_{RE}} & 0 & \sigma_{P_{RE}M_{RR}} & \sigma_{P_{RE}M_{ER}} & 0 \\ \sigma_{P_{EE}M_{EE}} & 0 & \sigma_{P_{ER}M_{EE}} & 0 & \sigma^2_{M_{EE}} & 0 & 0 & \sigma_{M_{EE}M_{RE}} \\ 0 & \sigma_{P_{RR}M_{RR}} & 0 & \sigma_{P_{RE}M_{RR}} & 0 & \sigma^2_{M_{RR}} & \sigma_{M_{RR}M_{ER}} & 0 \\ 0 & \sigma_{P_{RR}M_{ER}} & 0 & \sigma_{P_{RE}M_{ER}} & 0 & \sigma_{M_{RR}M_{ER}} & \sigma^2_{M_{ER}} & 0 \\ \sigma_{P_{EE}M_{RE}} & 0 & \sigma_{P_{ER}M_{RE}} & 0 & \sigma_{M_{EE}M_{RE}} & 0 & 0 & \sigma^2_{M_{RE}} \end{bmatrix}$$

$$\boldsymbol{E} = \begin{bmatrix} \sigma^2_{e_{EE}} & 0 & 0 & 0 \\ 0 & \sigma^2_{e_{RR}} & 0 & 0 \\ 0 & 0 & \sigma^2_{e_{ER}} & 0 \\ 0 & 0 & 0 & \sigma^2_{e_{RE}} \end{bmatrix}$$

with $\sigma^2_{d_{JK}}$, $\sigma^2_{r_{JK}}$ and $\sigma^2_{e_{JK}}$ are the dominance, permanent environmental and residual variances of the *JK* population respectively.

The inverse of the gametic relationship matrix ($\boldsymbol{G}^{-1}$) was calculated using a FORTRAN program that follows the algorithm proposed by Meyer and Tier (2012). The gametic relationship matrix ($\boldsymbol{G}$) was itself calculated by direct inversion using R software (R Core Team, 2021). Subsequently, the elements of $\boldsymbol{G}$ were employed to derive the elements of the dominance relationship matrix ($\boldsymbol{D}$) according to the following expression:

$$d_{ij} = g_{ipjp}g_{imjm} + g_{ipjm}g_{imjp}$$

Here, d_{ij} represents the dominance relationship between the *ith* and *jth* individuals. Specifically, g_{ixjy} signifies the relationship between the gametes of the *ith* individual (paternal or maternal, depending on *x*) and the *jth* individual (paternal or maternal, depending on *y*).

The implementation of the Gibbs Sampler was conducted using the BLUPF90 suite of programs, specifically with the gibbsf90+ program (Misztal et al., 2018). The analysis involved a single long chain comprising 1,100,000 iterations, with the initial 100,000 iterations discarded to guarantee convergence towards the stationary distribution.

During each iteration, the following parameters were calculated:

Ratios of paternal (p^2_{JK}) or maternal (m^2_{JK}) gametic variances for the JK cross

$$p^2_{JK} = \frac{\sigma^2_{p_{JK}}}{\sigma^2_{p_{JK}} + \sigma^2_{m_{JK}} + 2\sigma_{pJKmJK} + \sigma^2_{d_{JK}} + \sigma^2_{r_{JK}} + \sigma^2_{e_{JK}}}$$

$$m^2_{JK} = \frac{\sigma^2_{m_{JK}}}{\sigma^2_{p_{JK}} + \sigma^2_{m_{JK}} + 2\sigma_{pJKmJK} + \sigma^2_{d_{JK}} + \sigma^2_{r_{JK}} + \sigma^2_{e_{JK}}}$$

Ratio of dominance variance (d^2_{JK}) for the JK cross

$$d^2_{JK} = \frac{\sigma^2_{d_{JK}}}{\sigma^2_{p_{JK}} + \sigma^2_{m_{JK}} + 2\sigma_{pJKmJK} + \sigma^2_{d_{JK}} + \sigma^2_{r_{JK}} + \sigma^2_{e_{JK}}}$$

Broad sense heritability (H^2_{JK}) for the JK cross

$$H^2_{JK} = \frac{\sigma^2_{p_{JK}} + \sigma^2_{m_{JK}} + 2\sigma_{pJKmJK} + \sigma^2_{d_{JK}}}{\sigma^2_{p_{JK}} + \sigma^2_{m_{JK}} + 2\sigma_{pJKmJK} + \sigma^2_{d_{JK}} + \sigma^2_{r_{JK}} + \sigma^2_{e_{JK}}}$$

Gametic correlations ($r_{X_{JK}Y_{LM}}$) between the X (paternal—P- or maternal—M-) gametic effects of the from JK population with the Y (paternal—P- or maternal—M-) gametic effects of the LM population

$$r_{X_{JK}Y_{LM}} = \frac{\sigma_{X_{JK}Y_{LM}}}{\sigma_{X_{JK}}\sigma_{Y_{LM}}}$$

3 | RESULTS AND DISCUSSION

3.1 | Variance components

The posterior means (and standard deviations) of the variance components for TNB and NBA are presented in Tables 2 and 3 respectively.

To begin with, it is important to emphasize that the posterior estimates of paternal and maternal gametic variances in both traits were found to be greater in the crossbred populations (ER and RE) when contrasted with the purebred populations (EE and RR). This discrepancy could potentially be attributed to a scale effect (Falconer & Mackay, 1996), since the phenotypic variance is also higher in the crossbred populations. Another contributing factor to this discrepancy is the model's allowance for covariance between the two gametic effects within the purebred populations, which adds twice the value of $\sigma_{X_{JK}Y_{LM}}$ to the genetic variance.

Nevertheless, the outcomes of our analysis demonstrate that both paternal and maternal gametic effects contribute to the phenotypic variability of litter size traits, as evidenced by their posterior distributions significantly deviating from zero. However, it is important to acknowledge

TABLE 2 Posterior means (and standard deviations) of the permanent environmental (σ^2_r), paternal (σ^2_p), maternal (σ^2_m), dominance (σ^2_d) and residual (σ^2_e) variance components for Total Number Born (TNB) in the Entrepelado (EE), Entrepelado × Retinto (ER), Retinto × Entrepelado (RE) and Retinto (RR) populations.

Population	σ^2_r	σ^2_p	σ^2_m	σ^2_d	σ^2_e
EE	0.18 (0.07)	0.14 (0.06)	0.28 (0.07)	0.14 (0.07)	3.50 (0.07)
RR	0.11 (0.04)	0.14 (0.04)	0.20 (0.05)	0.10 (0.04)	4.08 (0.07)
ER	0.14 (0.06)	0.26 (0.09)	0.36 (0.09)	0.13 (0.07)	4.13 (0.09)
RE	0.16 (0.07)	0.15 (0.06)	0.38 (0.10)	0.14 (0.07)	4.17 (0.12)

TABLE 3 Posterior means (and standard deviations) of the permanent environmental (σ^2_r), paternal (σ^2_p), maternal (σ^2_m), dominance (σ^2_d) and residual (σ^2_e) variance components for Number Born Alive (NBA) in the Entrepelado (EE), Entrepelado × Retinto (ER), Retinto × Entrepelado (RE) and Retinto (RR) populations.

Population	σ^2_r	σ^2_p	σ^2_m	σ^2_d	σ^2_e
EE	0.19 (0.07)	0.13 (0.05)	0.24 (0.06)	0.11 (0.06)	3.47 (0.07)
RR	0.11 (0.04)	0.14 (0.04)	0.19 (0.05)	0.11 (0.04)	3.86 (0.07)
ER	0.13 (0.06)	0.26 (0.08)	0.33 (0.08)	0.12 (0.06)	4.01 (0.09)
RE	0.16 (0.07)	0.16 (0.07)	0.35 (0.09)	0.14 (0.06)	3.96 (0.11)

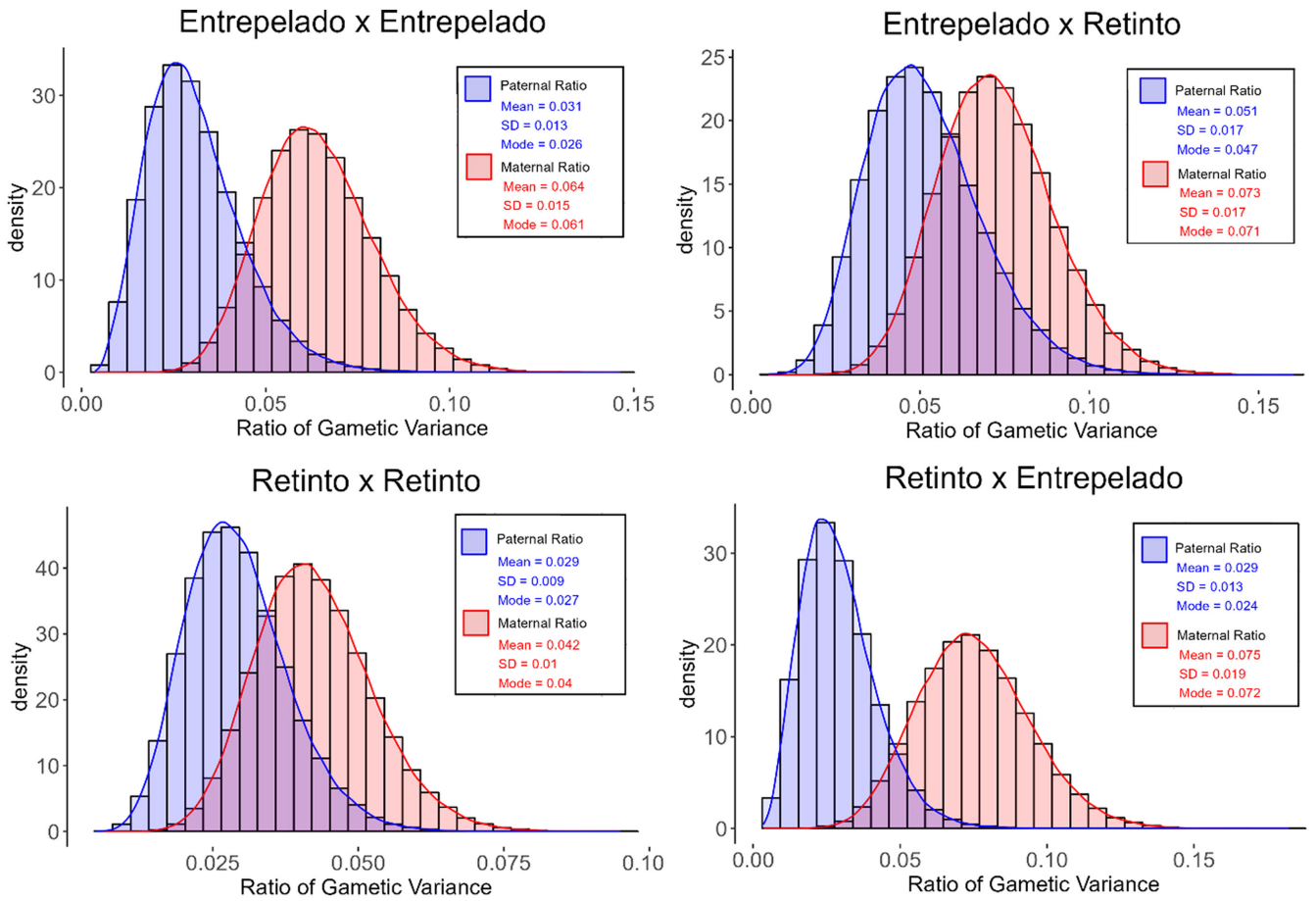


FIGURE 1 Posterior distributions of the ratios of paternal and maternal gametic variances in the Entrepelado × Entrepelado, Entrepelado × Retinto, Retinto × Entrepelado and Retinto × Retinto populations for Total Number Born.

that the variances of maternal gametic variances consistently surpass those of paternal gametic effects. This suggests that the alleles inherited from the mother exert a more pronounced influence on the phenotypic variation of litter size. These findings are further illustrated by the ratios of paternal to maternal gametic variance, visually depicted in Figure 1 for TNB and Figure S1 for NBA.

These findings align with those obtained by Stella et al. (2003) in white pigs, suggesting a similar trend. One plausible explanation for this phenomenon is the potential existence of paternal genomic imprinting (Reik & Walter, 2001), a mechanism ensuring certain alleles are only expressed upon heritance from either the mother or father. Numerous theories have been proposed to elucidate the evolutionary origins of genomic imprinting, comprehensive reviewed by Patten et al. (2014). One of the most prominent being the parental investment theory (Moore & Haig, 1991). In accordance with this theory, imprinting arises due to a conflict between the evolutionary interest of the alleles inherited from the father and those inherited from the mother. Within mammalian reproduction, the evolutionary success of the alleles inherited from the father is associated with augmented foetal growth, while the success of the alleles inherited from the mother is linked with offspring number. This theory finds support in the identification of numerous imprinted genes

governing diverse facets of mammalian development (Thamban et al., 2020), encompassing growth, behaviour and placental function (Fowden et al., 2011). Additionally, a growing body of evidence indicating the presence of imprinted genes in the pig genome (Coster et al., 2012; Wu et al., 2020; Zhang et al., 2012) is available.

It is important to underline that the paternal and maternal gametic variances exhibited higher values in the crossbred populations when compared to the purebred populations. As mentioned earlier, this divergence could stem from a scale effect or from the inclusion of covariances between the paternal and maternal gametic effects within the purebred populations. Consequently, the covariance between these effects also contributes to the overall genetic variation. Notably, the posterior distributions of the broad-sense heritabilities showed similar patterns across all purebred and crossbred populations, as there are weighted by the increase of phenotypic variation in crossbreds. The posterior mean estimates spanned from 0.125 in the RR population to 0.160 in the EE population for TNB, and from 0.125 (RR) to 0.149 (EE) for NBA, as depicted in Figure 2 for TNB and Figure S2 for NBA.

These figures also present the posterior distribution of the ratios of dominance variance, featuring posterior mean estimates ranging from 0.021 in RR to 0.031 in EE for TNB, and from 0.024 in RR to 0.029 in RE for NBA. In general, the ratios of dominance variance were modest, indicating that the implementation of a mate allocation procedure (González-Diéguez et al., 2020; Toro & Varona, 2010) to capture favourable dominance effects may yield limited or insignificant results, even if genotyping information were available.

3.2 | Gametic correlations

The proposed model provides the estimates of the gametic covariances and correlations between four gametic effects stemming from each parental population (paternal for purebred, maternal for purebred, paternal for crossbred and maternal for crossbred). Figures 3 and 4 display the posterior distributions of gametic correlations among these four gametic effects for TNB in the Entrepelado and Retinto populations respectively. Furthermore, Figures S3 and S4 provide the posterior distributions of gametic correlations for NBA.

Every posterior mean estimates of the gametic correlations were positive, spanning from 0.04 (paternal gametic effects for purebred and maternal gametic effects for crossbred in Retinto) to 0.53 (paternal gametic effects

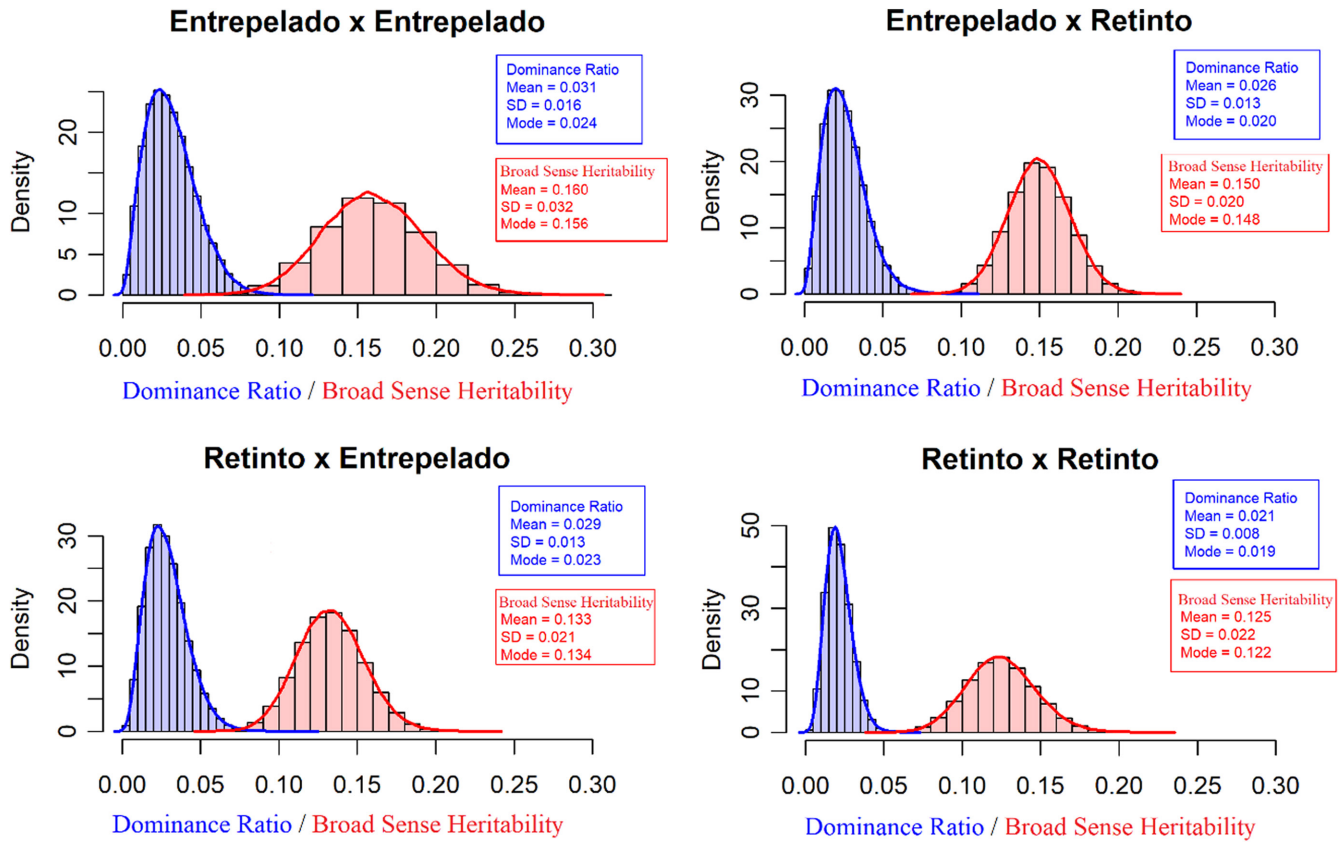


FIGURE 2 Posterior distributions of the broad-sense heritabilities and ratios of dominance variance in the Entrepelado × Entrepelado, Entrepelado × Retinto, Retinto × Entrepelado and Retinto × Retinto for Total Number Born.

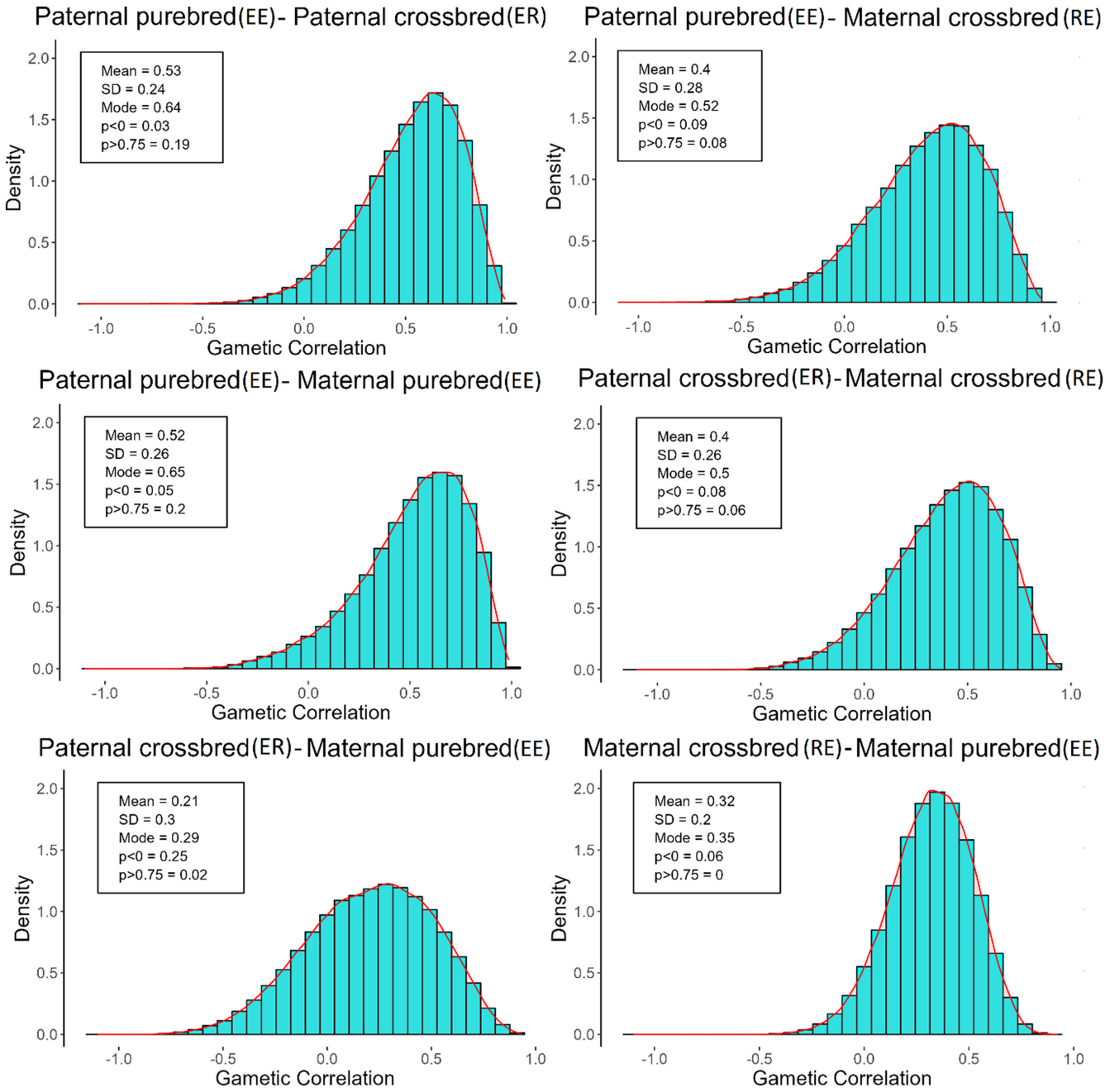


FIGURE 3 Posterior distribution of the correlations between the four gametic from each parental population (paternal for purebred—EE-, maternal for purebred—EE-, paternal for crossbred—ER- and maternal for crossbred—RE-) in the Entrepelado population for Total Number Born.

for purebred and paternal gametic effects for crossbred in Entrepelado). As far as we know, there are no existing estimates of gametic correlations between purebred and crossbred performance available in the literature. However, our estimates fall within the lower range of the genetic correlation estimates for a wide spectrum of traits available in the literature and reviewed by Wientjes and Calus (2017). It is noteworthy that the posterior probability of a gametic correlation surpassing 0.75 was consistently remained below 0.20 for all the gametic correlations. This suggest that selection in the purebred populations might not yield optimal outcomes in the crossbred population. These results differ from those obtained by Srihi et al. (2022); however, it is important to recognize that their study was conducted with a notably smaller data set.

Furthermore, it is interesting to observe that the pattern of gametic correlations differs between purebred and crossbred performance in both populations. Within the Entrepelado population, the gametic effects acting as the sire in the purebred context (paternal for purebred)

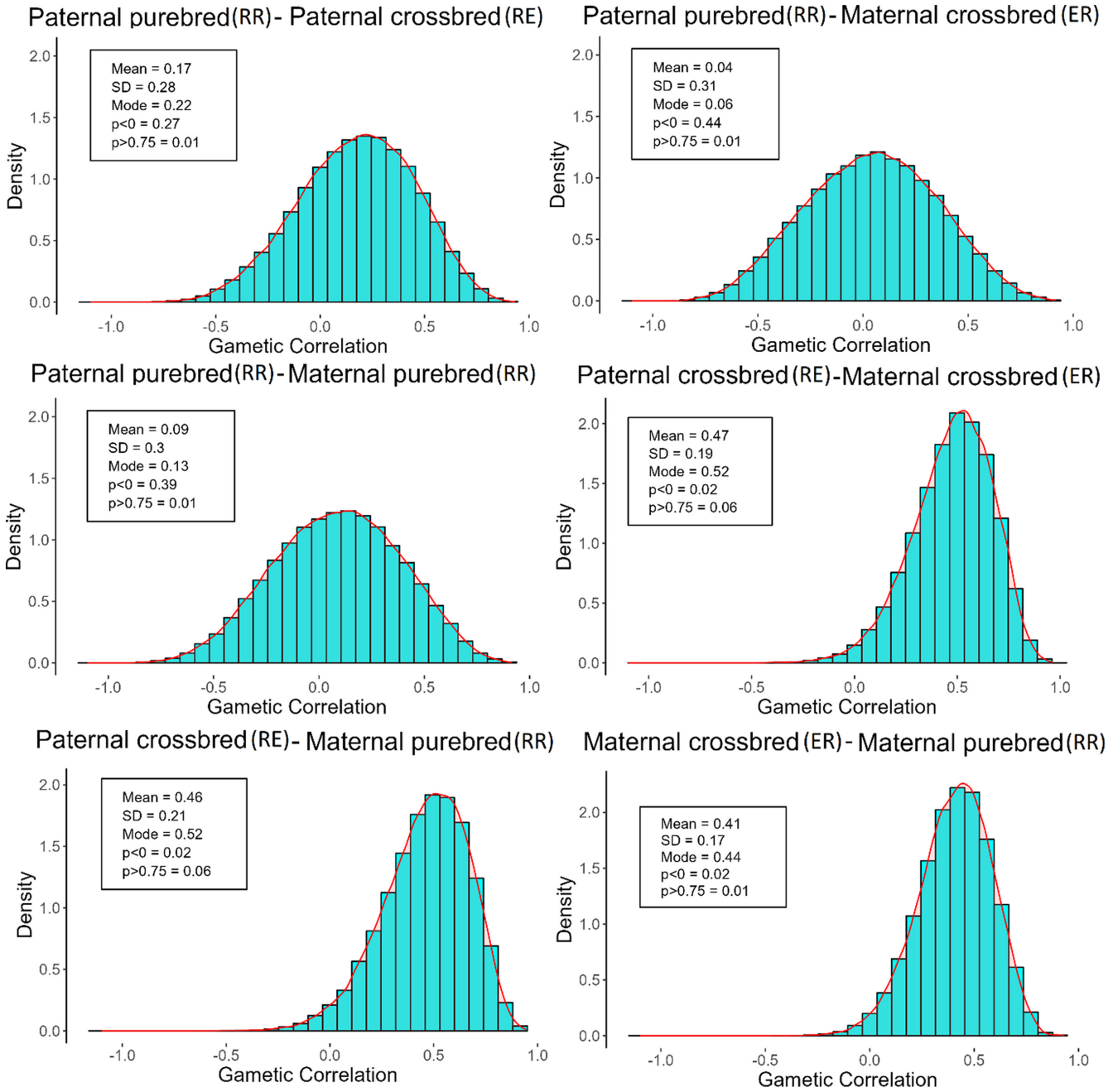


FIGURE 4 Posterior distribution of the correlations between the four gametic from each parental population (paternal for purebred—RR-, maternal for purebred—RR-, paternal for crossbred—RE-, and maternal for crossbred—ER-) in the Retinto population for Total Number Born.

exhibit relatively high correlations with performance within the crossbred population, whether as a sire (posterior mean of 0.53 with paternal for crossbred) or as a dam (posterior mean of 0.40 with maternal for crossbred). Conversely, the correlations involving the gametic effects when acting as the dam within the purebred population (maternal for purebred) were lower with performance in the crossbred population, both as a sire (posterior mean of 0.21 with paternal for crossbred) and as a dam (posterior mean of 0.32 with maternal for crossbred). In contrast, the scenario in the Retinto population was reversed. The correlations between the gametic effects acting as the sire within the purebred population (paternal for purebred) showed lower correlations with the crossbred population (posterior mean of 0.17 with paternal and 0.04 with maternal gametic effects in the crossbred), whereas the gametic effects acting as the dam (maternal for purebred) displayed elevated correlations (posterior mean of 0.46 with paternal and 0.41 with maternal gametic effects in the crossbred).

These findings reinforce the importance of using crossbred data to predict the breeding values of purebred

individuals and confirm the need to evaluate them for both purebred and crossbred performance. By doing so, balanced selection strategies can be optimized for crossbreeding purposes. Moreover, the observed diversity in the correlation between gametic effects also suggest the potential influence of imprinting effects, which should be considered in genetic evaluation. These results open up alternative strategies for crossbreeding selection and breeding programme design.

The populations that contribute as sires or dams in crossbreeding should be selected based on the prediction of their gametic effects when performing as sires or dams in the crossbred population respectively. In the example provided, the CASTUA population is commercially produced by crossing Retinto sires with Entrepelado dams. Hence, the selection of the Retinto population should rely on gametic prediction of paternal gametic effects for crossbred, while the Entrepelado population should be selected for maternal gametic effects for crossbred.

In the proposed model, the performance within the purebred populations contributes to the prediction of gametic effects in crossbreeding via gametic covariances (or correlations). Nonetheless, it should be noted that the covariances (or correlations) between gametic effects from different populations (Retinto or Entrepelado in this example) were assumed to be zero. Future research is required to develop a model that integrates genomic information and accounts for potential covariances among these gametic effects, potentially through adaptations of the metafounders analysis (Xiang et al., 2017).

The findings of this study can be summarized as follows: 1) maternal gametic effects consistently exhibit greater variances that paternal gametic effects, indicating a stronger influence of alleles inherited from the mother on litter size, and 2) distinct patterns of gametic correlations were observed between purebred and crossbred performances within the Entrepelado and the Retinto pig populations. These results suggest the potential impact of imprinting effects on the genetic regulation of litter size and underscore the importance of including crossbred data in breeding value predictions for purebred individuals.

AUTHOR CONTRIBUTIONS

Conceptualization: Luis Varona; methodology: Houssemeddine Srihi, David López-Carbonell and Luis Varona; software: Houssemeddine Srihi and Luis Varona; validation: Houssemeddine Srihi, David López-Carbonell and Luis Varona; formal analysis: Houssemeddine Srihi and David López-Carbonell; investigation: Houssemeddine Srihi; resources: Noelia Ibáñez-Escriche and Joaquim Casellas; data curation: Sara Negro, Pilar Hernández and Noelia Ibáñez-Escriche; writing—original draft preparation: Houssemeddine Srihi and Luis Varona; writing—review and editing: Noelia Ibáñez-Escriche, Pilar Hernández and Joaquim Casellas; visualization: David López-Carbonell and Houssemeddine Srihi; supervision: Luis Varona; project administration: Noelia Ibáñez-Escriche and Joaquim Casellas; funding acquisition: Noelia Ibáñez-Escriche and Joaquim Casellas. All authors have read and agreed to the published version of the article.

ACKNOWLEDGEMENTS

We thank INGA FOOD, S.A. (Almendralejo, Spain) and its technicians (E. Magallón, L. Muñoz, P. Díaz and D. Iniesta) for their cooperation and technical support. The authors also want to thank J. L. Noguera for his work.

FUNDING INFORMATION

The research was partially funded by grants CGL2016-80155-R and PID2020-114705RB-I00 (Ministerio de Ciencia e Innovación), and IDI-20170304 (CDTI). Srihi received funding from the European Union's H2020 research and innovation program under a Marie Sklodowska-Curie grant agreement with Iberus Talent (H2020-G.A.number 801586).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data set used in this study will be available upon reasonable request to the corresponding author (lvarona@unizar.es).

ORCID

Houssemeddine Srihi https://orcid.org/0000-0002-9249-3070
David López-Carbonell https://orcid.org/0000-0003-0964-9212
Noelia Ibáñez-Escriche https://orcid.org/0000-0002-6221-3576
Joaquim Casellas https://orcid.org/0000-0002-4982-3556
Pilar Hernández https://orcid.org/0000-0001-7857-7334
Sara Negro https://orcid.org/0000-0002-2984-873X
Luis Varona https://orcid.org/0000-0001-6256-5478

REFERENCES

Coster, A., Madsen, O., Heuven, H. C. M., Dibbits, B., Groenen, M. A. M., van Arendonk, J. A. M., & Bovenhuis, H. (2012). The imprinted gene DIO3 is a candidate gene for litter size in pigs. *PLoS One*, *7*, e31825. https://doi.org/10.1371/journal.pone.0031825

Falconer, D. S., & Mackay, T. F. C. (1996). *Introduction to quantitative genetics*. Longman Group.

Fowden, A. L., Coan, P. M., Angiolini, E., Burton, G. J., & Constancia, M. (2011). Imprinted genes and the epigenetic regulation of placental phenotype. *Progress in Biophysics and Molecular Biology*, *106*, 281–288. https://doi.org/10.1016/j.pbiomolbio.2010.11.005

Gelfand, A. E., & Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, *85*, 398–409. https://doi.org/10.2307/2289776

Gilles, G. (2009). Dry cured ham quality as related to lipid quality of raw material and lipid changes during processing: A review. *Grasas y Aceites*, *60*, 297–307. https://doi.org/10.3989/gya.130908

González-Diéguez, D., Tusell, L., Bouquet, A., Legarra, A., & Vitezica, Z. G. (2020). Purebred and crossbred genomic evaluation and mate allocation strategies to exploit dominance in pig crossbreeding schemes. *G3 Genes, Genomes, Genetics*, *10*, 2829–2841. https://doi.org/10.1534/g3.120.401376

Lopez-Bote, C. J. (1998). Sustained utilization of the Iberian pig breed. *Meat Science*, *49*, S17–S27. https://doi.org/10.1016/s0309-1740(98)90036-5

Meyer, K., & Tier, B. (2012). Estimates of variances due to parent of origin effects for weights of Australian beef cattle. *Animal Production Science*, *52*, 215–224. https://doi.org/10.1071/AN11195

Misztal, I., Tsuruta, S., Lourenco, D., Aguilar, I., Legarra, A., & Vitezica, Z. (2018). *Manual for BLUPF90 family of programs*. University of Georgia.

Moore, T., & Haig, D. (1991). Genomic imprinting in mammalian development: A parental tug-of-war. *Trends in Genetics*, *7*, 45–49. https://doi.org/10.1016/0168-9525(91)90230-N

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*, *13*, 2765–2772. https://doi.org/10.1017/S1751731119001125

Patten, M. M., Ross, L., Curley, J. P., Queller, D. C., Bonduriansky, R., & Wolf, J. B. (2014). The evolution of genomic imprinting: Theories, predictions and empirical tests. *Heredity (Edinburgh)*, *113*, 119–128. https://doi.org/10.1038/hdy.2014.29

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Reik, W., & Walter, J. (2001). Genomic imprinting: Parental influence on the genome. *Nature Reviews. Genetics*, *2*, 21–32. https://doi.org/10.1038/35047554

Serrano, M. P., Valencia, D. G., Nieto, M., Lázaro, R., & Mateos, G. G. (2008). Influence of sex and terminal sire line on performance and carcass and meat quality of Iberian pigs reared under intensive production systems. *Meat Science*, *78*, 420–428. https://doi.org/10.1016/j.meatsci.2007.07.006

Smith, S. (1984). *Dominance relationship matrix and inverse for an outbred population*. Dairy Sc.

Srihi, H., López-Carbonell, D., Ibáñez-Escriche, N., Casellas, J., Hernández, P., Negro, S., & Varona, L. (2023). A Bayesian multivariate gametic model in a reciprocal cross with genomic information: An example with two Iberian varieties. *Animals*, *13*, 1648. https://doi.org/10.3390/ani13101648

Srihi, H., Noguera, J. L., Topayan, V., de Hijas, M. M., Ibañez-Escriche, N., Casellas, J., Vázquez-Gómez, M., Martínez-Castillero, M., Rosas, J. P., & Varona, L. (2022). Additive and dominance genomic analysis for litter size in purebred and crossbred iberian pigs. *Genes (Basel)*, *13*, 12. https://doi.org/10.3390/genes13010012

Stella, A., Stalder, K. J., Saxton, A. M., & Boettcher, P. J. (2003). Estimation of variances for gametic effects on litter size in Yorkshire and Landrace swine. *Journal of Animal Science*, *81*, 2171–2178. https://doi.org/10.2527/2003.8192171x

Thamban, T., Agarwaal, V., & Khosla, S. (2020). Role of genomic imprinting in mammalian development. *Journal of Biosciences*, *45*, 20. https://doi.org/10.1007/s12038-019-9984-1

Toro, M. A., & Varona, L. (2010). A note on mate allocation for dominance handling in genomic selection. *Genetics, Selection, Evolution*, *42*, 33. https://doi.org/10.1186/1297-9686-42-33

Varona, L., Munilla, S., Casellas, J., Moreno, C., & Altarriba, J. (2015). Consequences of paternally inherited effects on the genetic evaluation of maternal effects. *Genetics, Selection, Evolution*, *47*, 63. https://doi.org/10.1186/s12711-015-0141-5

Wientjes, Y. C. J., & 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*, 3467–3478. https://doi.org/10.2527/jas.2017.1669

Wu, Y. Q., Zhao, H., Li, Y. J., Khederzadeh, S., Wei, H. J., Zhou, Z. Y., & Zhang, Y. P. (2020). Genome-wide identification of imprinted genes in pigs and their different imprinting status compared with other mammals. *Zoological Research*, *41*, 721–725. https://doi.org/10.24272/J.ISSN.2095-8137.2020.072

Xiang, T., Christensen, O. F., & Legarra, A. (2017). Technical note: Genomic evaluation for crossbred performance in a single-step approach with metafounders. *Journal of Animal Science*, *95*, 1472–1480. https://doi.org/10.2527/jas2016.1155

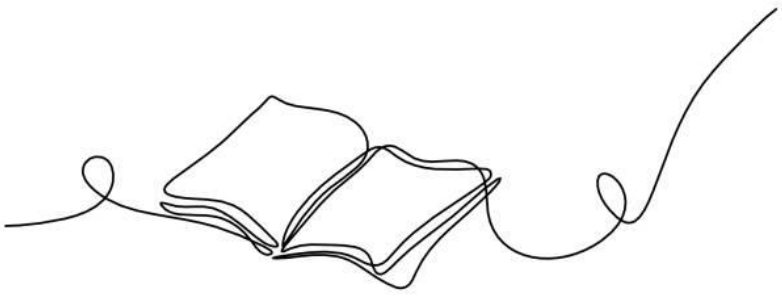
Zhang, F. W., Bin Han, Z., Deng, C. Y., He, H. J., & Wu, Q. (2012). Conservation of genomic imprinting at the NDN, MAGEL2 and MEST loci in pigs. *Genes & Genetic Systems*, *87*, 53–58. https://doi.org/10.1266/ggs.87.53

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Srihi, H., López-Carbonell, D., Ibáñez-Escriche, N., Casellas, J., Hernández, P., Negro, S., & Varona, L. (2023). A multivariate gametic model for the analysis of purebred and crossbred data. An example between two populations of Iberian pigs. *Journal of Animal Breeding and Genetics*, *00*, 1–10. https://doi.org/10.1111/jbg.12832

SUPPLEMENTARY MATERIALS



Suplementary Material:

Additive and Dominant genomic analysis for litter size in pure and crossbred Iberian pigs.

Houssemeddine Srihi[1], José Luis Noguera[2], Victoria Topayan[3], Melani Martín de Hijas[4], Noelia Ibañez-Escriche[3], Joaquim Casellas[4], Marta Vázquez-Gómez[4], María Martínez-Castillero[1], Juan Pablo Rosas[5] and Luis Varona[1]*

Table S1: REML estimates ± standard error (SE) of the additive genotypic (co)variances for Number Born Alive (NBA).

	Entrepelado	**Crossbred**	**Retinto**
Entrepelado	0.147 ± 0.156	0.202 ± 0.169	0.184 ± 0.150
Crossbred	-	0.477 ± 0.241	0.410 ± 0.172
Retinto	-	-	0.359 ± 0.183

Table S2: REML estimates ± standard error (SE) of the dominance genotypic (co)variances for Number Born Alive (NBA).

	Entrepelado	**Crossbred**	**Retinto**
Entrepelado	0.293 ± 0.216	0.286 ± 0.182	0.204 ± 0.173
Crossbred	-	0.282 ± 0.188	0.202 ± 0.177
Retinto	-	-	0.155 ± 0.205

Table S3: REML estimates ± standard error (SE) of the permanent environmental and residual variances for Number Born Alive (NBA).

Variances	**Entrepelado**	**Crossbred**	**Retinto**
σ_S^2	0.111 ± 0.111	0.004 ± 0.022	0.252 ± 0.126
σ_E^2	2.820 ± 0.099	4.385 ± 0.152	3.468 ± 0.125

σ_S^2: Sow permanent environment variance ; σ_E^2: Residual variance.

Table S4: GO (Gene Ontology) terms for biological process of the proposed candidate genes.

Gene	Name	Biological process	
BCO1	*Beta-carotene oxygenase 1*	GO:0001523 GO:0042574 GO:1901810	- retinoid metabolic process - retinal metabolic process - beta-carotene metabolic process
GCSH	*Glycine cleavage system protein H*	GO:0019464	- glycine decarboxylation via glycine cleavage system
ATMIN	*ATM interactor*	GO:0010628 GO:0044458 GO:0045893 GO:0045944 GO:1902857	- positive regulation of gene expression - motile cilium assembly - positive regulation of transcription, DNA-templated - positive regulation of transcription by RNA polymerase II - positive regulation of non-motile cilium assembly
CENPN	*Centromere protein N*	GO:0007059 GO:0051382	- chromosome segregation - kinetochore assembly
DYNLRB2	*Dynein light chain roadblock-type 2*	GO:0007018	- microtubule-based movement
CD38	*CD38 molecule*	GO:0050794	- regulation of cellular process
FGFBP1	*Fibroblast growth factor binding protein 1*	GO:0090050 GO:1903589	- positive regulation of cell migration involved in sprouting angiogénesis - positive regulation of cell migration involved in sprouting angiogenesis
PROM1	*Prominin 1*	GO:0045494 GO:0060042 GO:0060219 GO:0072112 GO:0072139 GO:2000768	- Photoreceptor cell maintenance - retina morphogenesis in camera-type eye - camera-type eye photoreceptor cell differentiation - glomerular visceral epithelial cell differentiation - glomerular parietal epithelial cell differentiation - positive regulation of nephron tubule epithelial cell differentiation
TAPT1	*Transmembrane anterior posterior transformation 1*	GO:0045724 GO:0048706 GO:0048856	- positive regulation of cilium assembly - embryonic skeletal system development - anatomical structure development
LDB2	*LIM domain binding 2*	GO:0045944 GO:0044089 GO:0043549 GO:0035019 GO:0030334 GO:0010669 GO:0006357 GO:0001942	- positive regulation of transcription by RNA polymerase II - positive regulation of cellular component biogenesis - regulation of kinase activity - somatic stem cell population maintenance - regulation of cell migration - epithelial structure maintenance - regulation of transcription by RNA polymerase II - hair follicle development
FGF5	*Fibroblast growth factor 5*	GO:0023019 GO:0010001	- signal transduction involved in regulation of gene expression

		GO:0008543 GO:0008284	- glial cell differentiation - fibroblast growth factor receptor signaling pathway - positive regulation of cell population proliferation
TAOK1	*TAO kinase 1*	GO:1901985 GO:0097194 GO:0070050 GO:0046330 GO:0032874 GO:0016310 GO:0007095 GO:0007026 GO:0006974 GO:0006468 GO:0000226	- positive regulation of protein acetylation - execution phase of apoptosis - neuron cellular homeostasis - positive regulation of JNK cascade - positive regulation of stress-activated MAPK cascade - phosphorylation - mitotic G2 DNA damage checkpoint signaling - negative regulation of microtubule depolymerization - cellular response to DNA damage stimulus - protein phosphorylation - microtubule cytoskeleton organization
GIT1	*GIT ArfGAP 1*	GO:2000300 GO:0106015 GO:0099171 GO:0090063 GO:0071222 GO:0061743 GO:0048666 GO:0048013 GO:0045820 GO:0045454 GO:0032691 GO:0032465 GO:0007626 GO:0007420 GO:0001957	- regulation of synaptic vesicle exocytosis - negative regulation of inflammatory response to wounding - presynaptic modulation of chemical synaptic transmission - positive regulation of microtubule nucleation - cellular response to lipopolysaccharide - motor learning - neuron development - ephrin receptor signaling pathway - negative regulation of glycolytic process - cell redox homeostasis - negative regulation of interleukin-1 beta production - regulation of cytokinesis - locomotory behavior - brain development - intramembranous ossification
ANKRD13B	*ankyrin repeat domain 13B*	GO:0002091	- negative regulation of receptor internalization

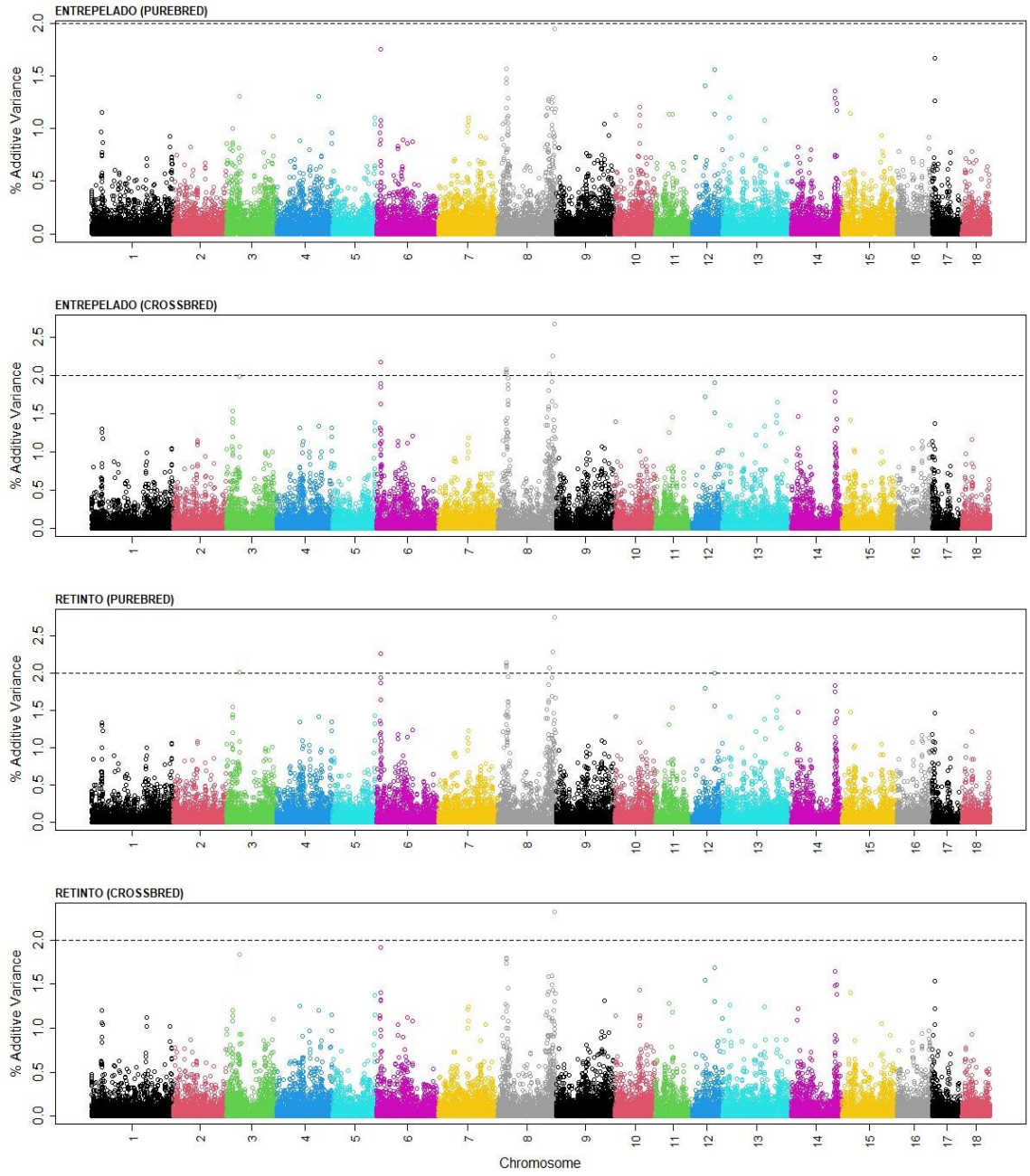


Figure S1: Distribution of the percentage of the additive genetic variance explained by genomic segments of 30 SNPs within the autosomal genome of purebred and crossbred performance for Number Born Alive (NBA) in the Entrepelado and Retinto varieties.

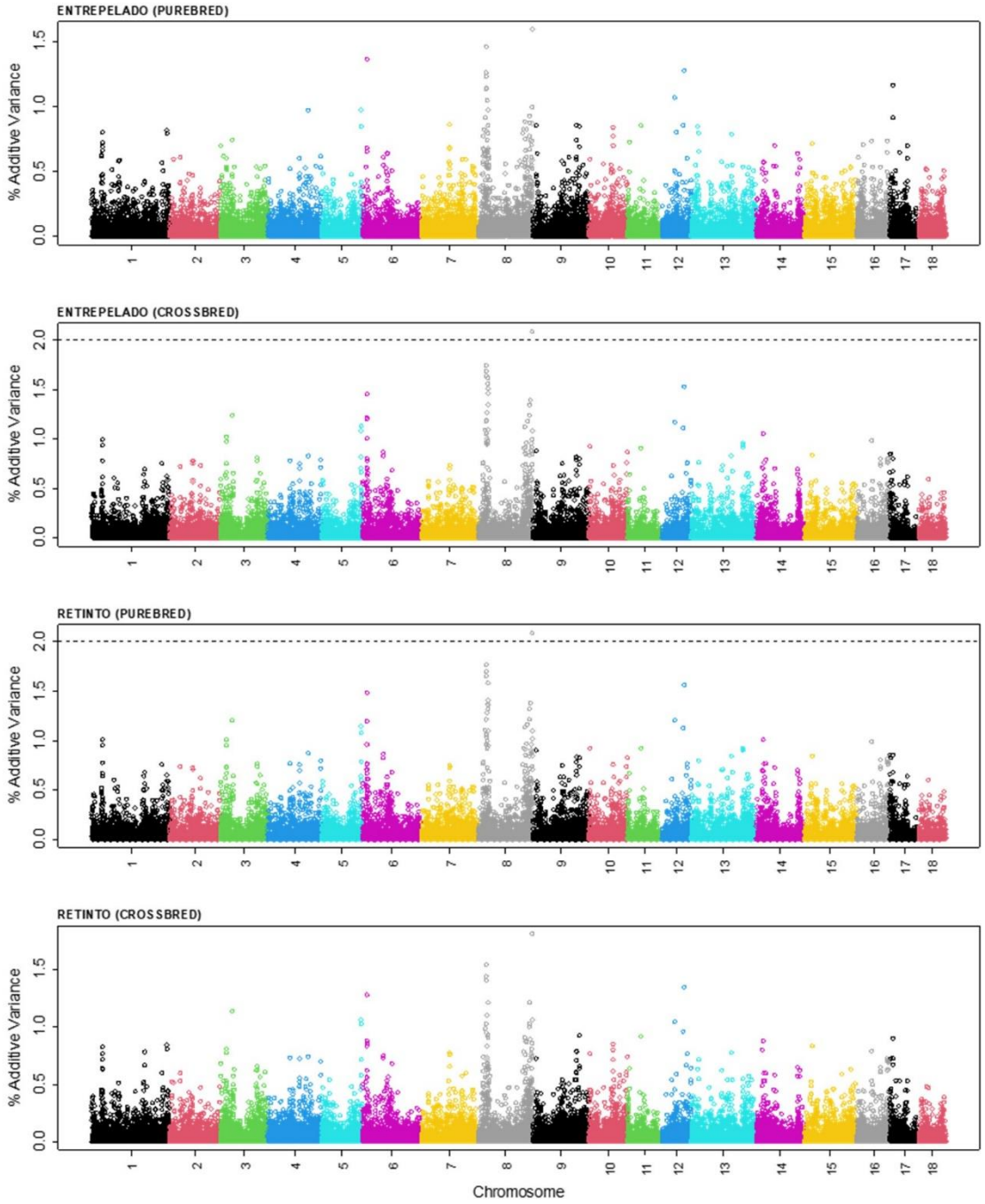


Figure S2: Distribution of the percentage of the additive genetic variance explained by genomic segments of 20 SNPs within the autosomal genome of purebred and crossbred performance for Total Number Born (TNB) in the Entrepelado and Retinto varieties.

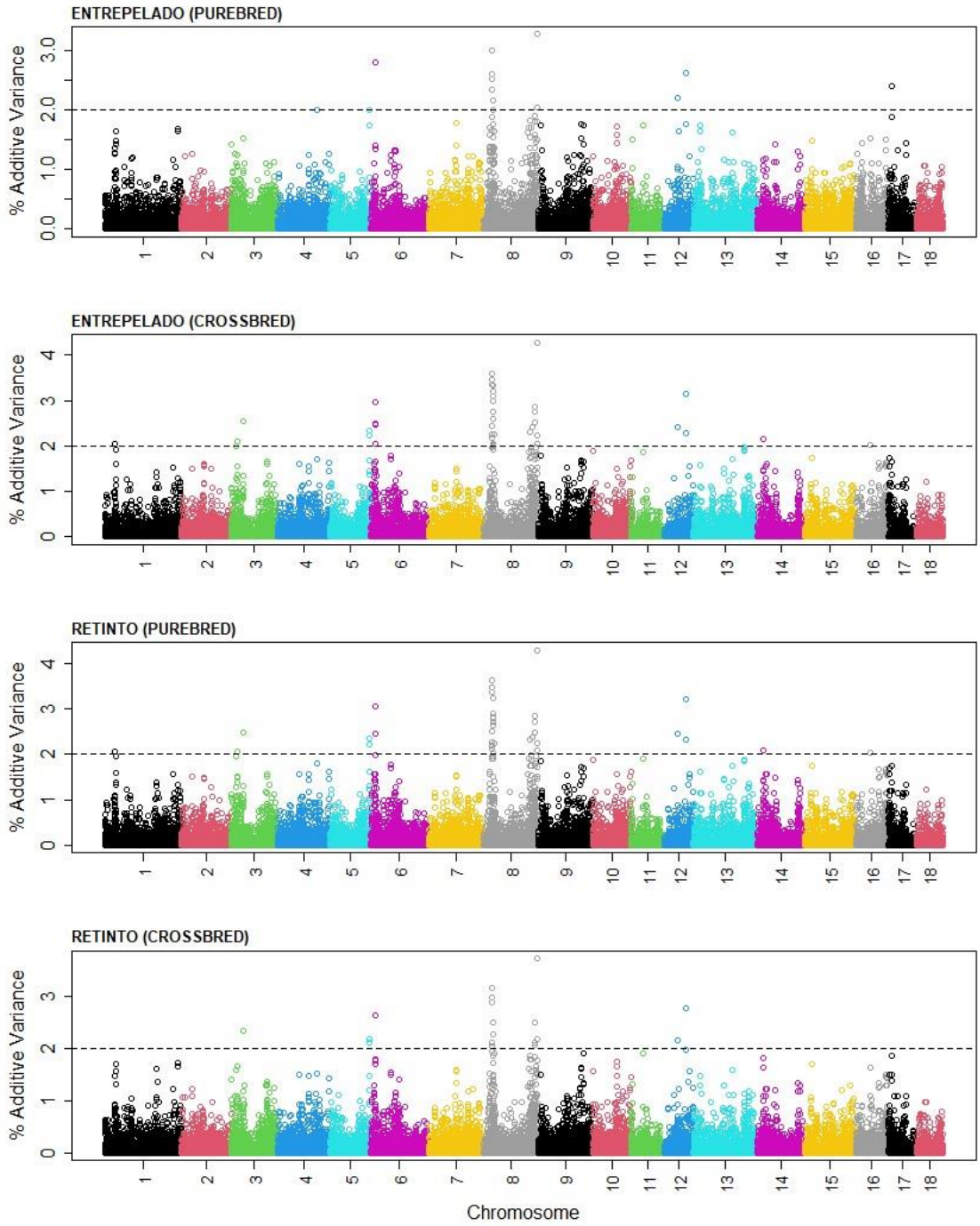


Figure S3: Distribution of the percentage of the additive genetic variance explained by genomic segments of 40 SNPs within the autosomal genome of purebred and crossbred performance for Total Number Born (TNB) in the Entrepelado and Retinto varieties.

A multivariate gametic model for the analysis of purebred and crossbred data.

An example between two varieties of Iberian pigs.

SUPPLEMENTARY INFORMATION

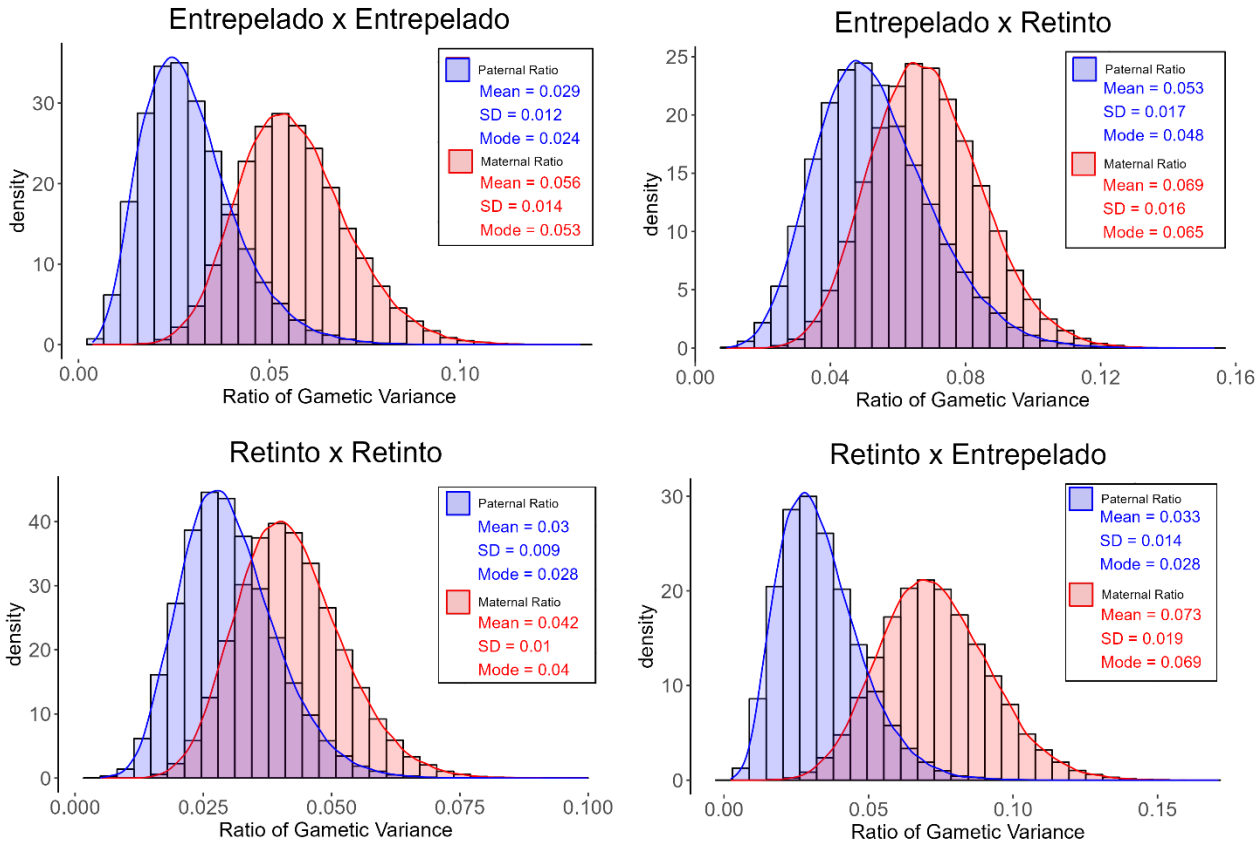


Figure S4: Posterior distributions of the ratios of paternal and maternal gametic variances in the Entrepelado x Entrepelado, Entepelado x Retinto, Retinto x Entrepelado and Retinto x Retinto populations for Number Born Alive.

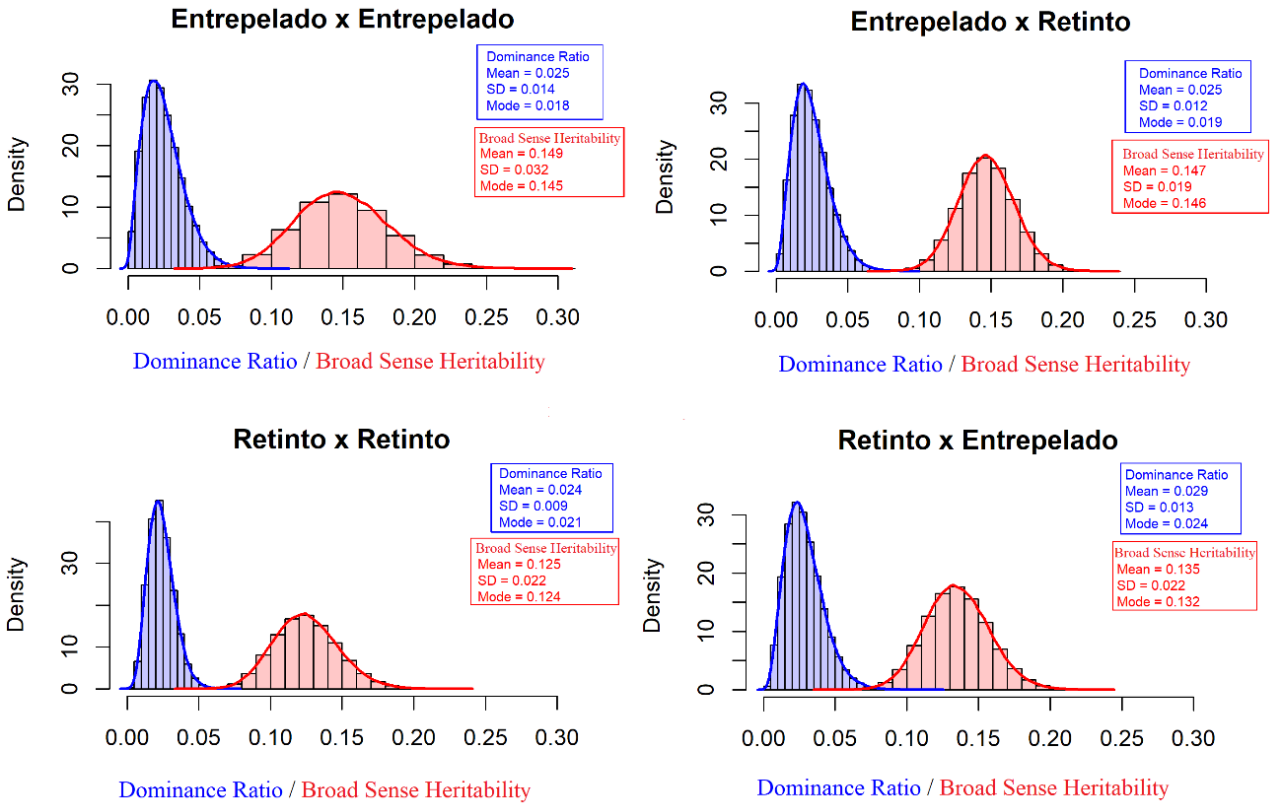


Figure S5: Posterior distributions of the broad-sense heritability and ratios of dominance variance in the Entrepelado x Entrepelado, Entrepelado x Retinto, Retinto x Entrepelado and Retinto x Retinto for Number Born Alive.

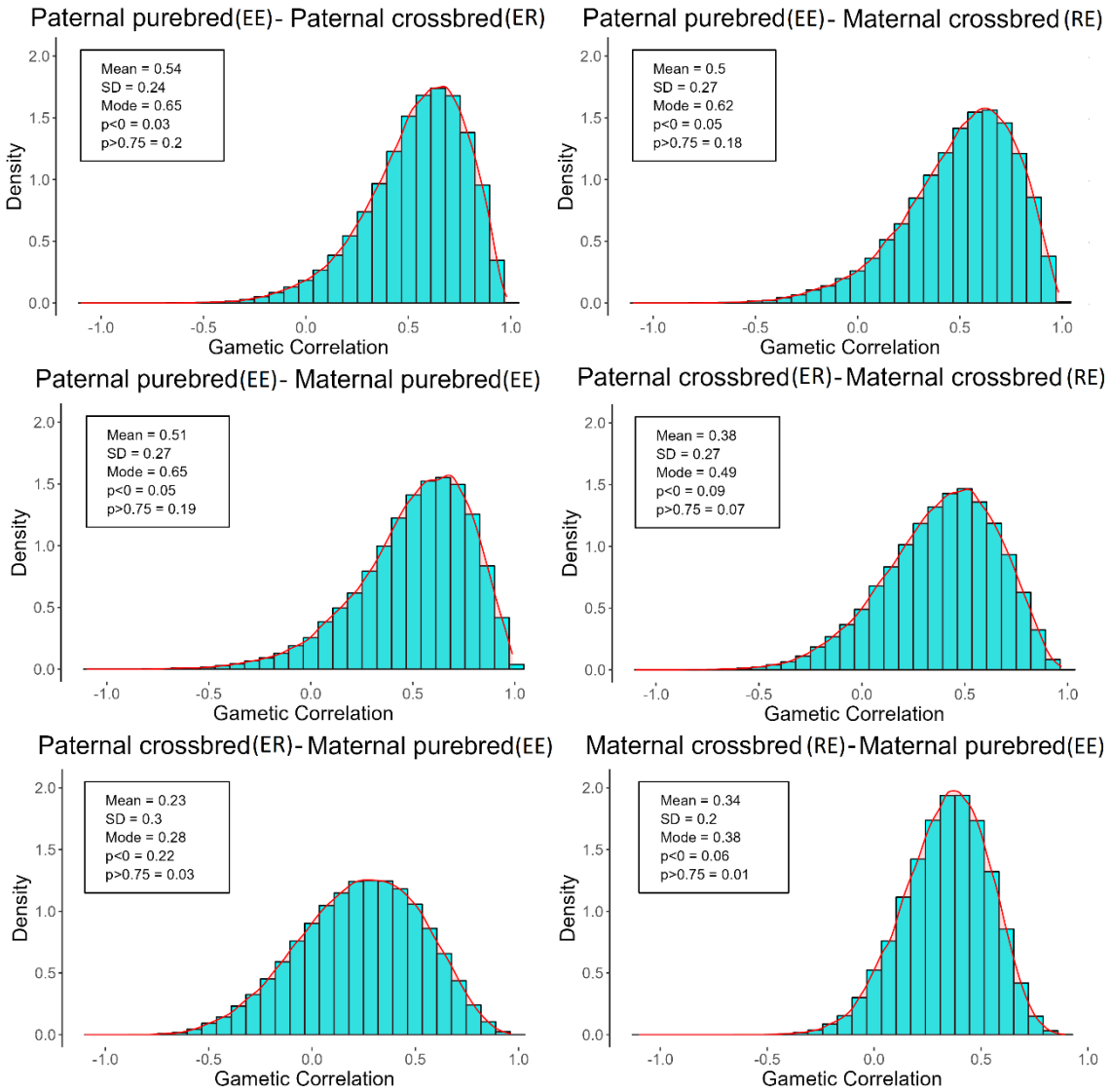


Figure S6: Posterior distribution of the correlations between the four gametic from each parental population (paternal for purebred –EE-, maternal for purebred –EE-, paternal for crossbred –ER-, and maternal for crossbred –RE-) in the Entrepelado population for Number Born Alive.

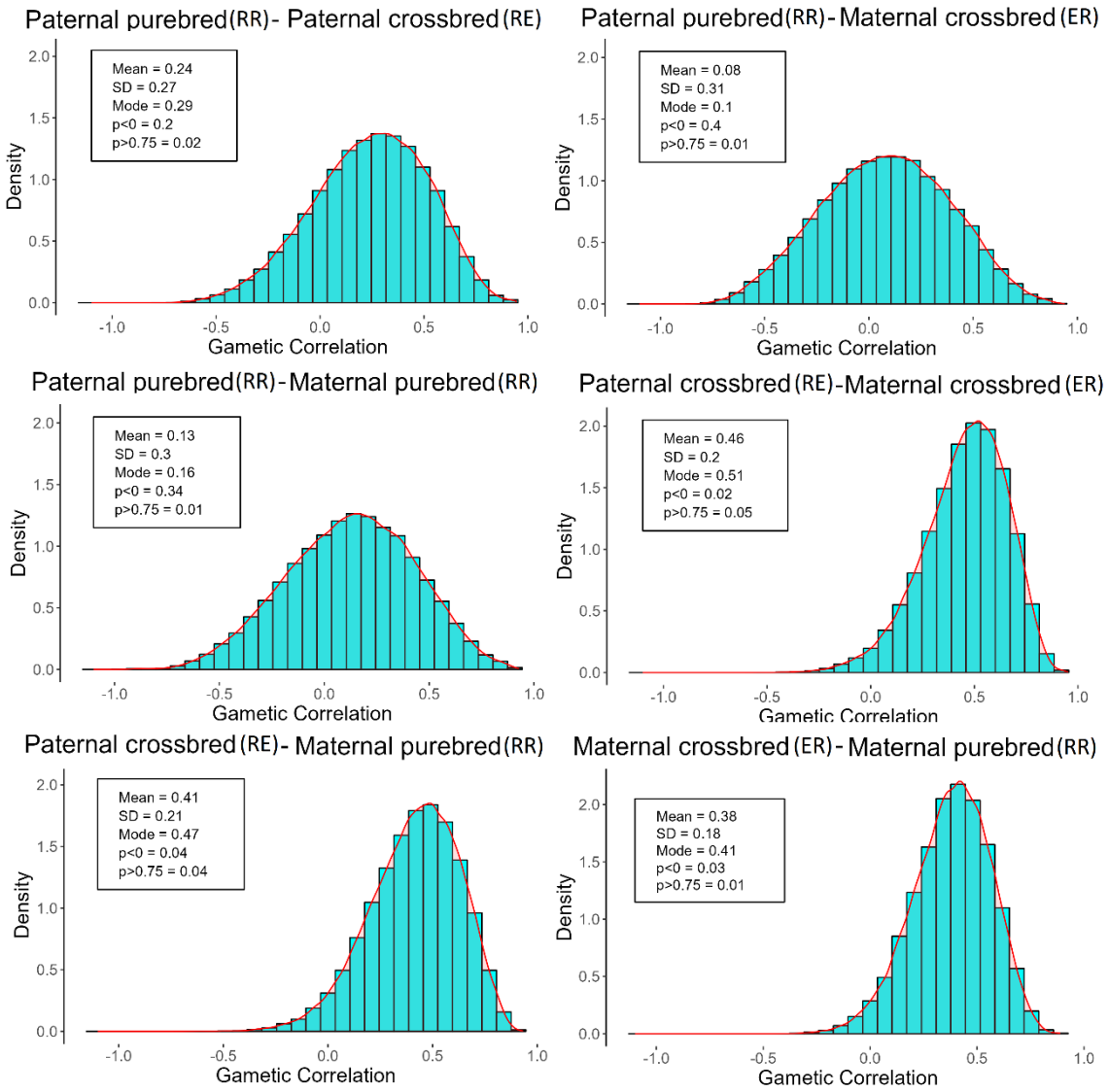


Figure S7: Posterior distribution of the correlations between the four gametic from each parental population (paternal for purebred –RR-, maternal for purebred –RR-, paternal for crossbred –RE-, and maternal for crossbred –ER-) in the Retinto population for Number Born Alive.

- **74th EAAP ANNUAL MEETING -LYON , France 2023 August 26th to September 1st, 2023**

A MULTIVARIATE GAMETIC MODEL FOR THE ANALYSIS FROM RECIPROCAL CROSSES FOR TWOIBERIAN VARIETIES . H. Srihi., D. López-Carbonell., N. Ibáñez-Escriche., J. Casellas,P., Hernández, S., Negro and L. Varona.

- **XX Jornadas de Producción Anima AIDA , Zaragoza 13 - 14 june 2023**

COMPARACIÓN DE PROCEDIMIENTOS PARA LA ASIGNACIÓN DE FASESHAPLOTÍPICAS EN UN CRUCE DIALÉLICO ENTRE ESTIRPES DE CERDO IBÉRICO . Srihi*, H., López-Carbonell, D., Ibañez-Escriche, N., Casellas, J., Hernández, P.,Rosas, J. P., y Varona, L.

- **World Congress on Genetics Applied to Livestock Production (WCGALP) , Rotterdam, The Netherlands, 3 - 8 July 2022.**

BAYESIAN ANALYSIS OF PATERNAL AND MATERNAL GAMETIC EFFECTS IN A RECIPROCALCROSS BETWEEN TWO IBERIAN VARIETIES . H. Srihi*, N. Ibáñez-Escriche, J. Casellas,J. L. Noguera, P. Hernández, M. Martín de Hijas, M. Vázquez-Gómez, S. Negro,J. P. Rosas and L. Varona .

- **XX Reunión Nacional de Mejora Genética Animal, Madrid 1 - 3 June 2022**

ANÁLISIS BAYESIANO DE LOS EFECTOS GAMÉTICOS PATERNOS Y MATERNOS EN UN CRUCERECÍPROCO ENTRE DOS ESTIRPES DE CERDO IBÉRICO. H. Srihi, N. Ibáñez-Escriche, J. Casellas, J. L. Noguera, P. Hernández, M.Martín de Hijas, M. Vázquez-Gómez, S. Negro, J. P. Rosas & L. Varona.

- **72nd EAAP Annual Meeting in Davos, Switzerland from 30th of August to3rd of September 2021.**

ADDITIVE AND DOMINANT GENOMIC ANALYSIS FOR LITTER SIZE IN PURE AND CROSSBREEDIBERIAN PIGS. Srihi, H, Noguera, JL, Topayan, V, Martín-de-Hijas, M, Ibañez-Escriche, N, Casellas, J, Vázquez-Gómez4, M, Martínez-Castillero, M, Rosas, JPy Varona, L.ç

- **XIX Jornadas sobre Producción Animal AIDA , Zaragoza 1 - 2 june 2021.**

CORRELACIONES GENÉTICAS ADITIVAS Y DOMINANTES PARA TAMAÑO DECAMADA ENTRE LAS ESTIRPES ENTREPELADO Y RETINTO Y SU CRUCE Srihi, H, Noguera, JL, Topayan, V, Martín-de-Hijas, M, Ibañez-Escriche, N,Casellas, J, Vázquez-Gómez, M, Martínez-Castillero1, M, Rosas, JP y Varona, L

RESEARCH STAY AT AARHUS UNIVERSITY

In addition to the research conducted as part of this thesis, I had the privilege of undertaking a research stay at the esteemed Center for Quantitative Genetics and Genomics, Aarhus University, Denmark. As a visiting PhD student, from September 2022 to December 2022, this opportunity significantly enriched my academic journey. Under the expert guidance of Professor Guosheng Su, I delved deeply into the realm of genomic prediction for purebred and crossbred pigs. This period was marked by intensive study and research, where I focused specifically on the breed of origin in Iberian pigs. This experience not only broadened my understanding of complex genetic mechanisms but also provided invaluable insights into advanced genomic prediction techniques. The knowledge and skills acquired during this enriching sojourn at Aarhus University have been instrumental in shaping the research presented in this thesis and will undoubtedly continue to influence my future work in the field of animal genetics.