



# Identification of breed-specific genomic variants in Colombian Creole pig breeds by whole-genome sequencing

Rafael Suárez-Mesa<sup>1,2</sup> · Roger Ros-Freixedes<sup>1</sup> · Houda Laghouaouta<sup>1</sup> · Ramona N. Pena<sup>1</sup> · Byron Hernández-Ortiz<sup>3</sup> · Ilang Rondón-Barragán<sup>2</sup> · Joan Estany<sup>1</sup>

Received: 2 December 2022 / Accepted: 29 March 2023 / Published online: 11 April 2023  
© The Author(s) 2023

## Abstract

Dissecting genetic variation of local breeds is important for the success of conservation. In this research, we investigated the genomic variation of Colombian Creole (CR) pigs, with a focus on the breed-specific variants in the exonic region of 34 genes with reported effects on adaptive and economic traits. Seven individuals of each of the three CR breeds (CM, Casco de Mula; SP, San Pedroño; and ZU, Zungo) were whole-genome sequenced along with 7 Iberian (IB) pigs and 7 pigs of each of the four most used cosmopolitan (CP) breeds (Duroc, Landrace × Large White, and Pietrain). Molecular variability in CR (6,451,218 variants; from 3,919,242, in SP, to 4,648,069, in CM) was comparable to that in CP, but higher than in IB. For the investigated genes, SP pigs displayed less exonic variants (178) than ZU (254), CM (263), IB (200), and the individual CP genetic types (201 to 335). Sequence variation in these genes confirmed the resemblance of CR to IB and indicates that CR pigs, particularly ZU and CM, are not exempt from selective introgression of other breeds. A total of 50 exonic variants were identified as being potentially specific to CR, including a high-impact deletion in the intron between exons 15 and 16 of the leptin receptor gene, which was only found in CM and ZU. The identification of breed-specific variants in genes related to adaptive and economical traits can bolster the understanding of the role of gene-environment interactions on local adaptation and points the way for effective breeding and conservation of CR pigs.

**Keywords** Animal genetic resources · Conservation · Creole pigs · Genetic variability · *LEPR* · Whole-genome sequencing

## Introduction

Current pig production is based on highly cosmopolitan selected lines that are managed under intensive production systems. Still, there is a growing interest in the conservation of local breeds in favour of biodiversity (Ciobanu et al.,

2001) and, in particular, as a source of adaptive variation against climate change disturbances. The three officially recognized Colombian Creole (CR) pig breeds (ZU, Zungo; CM, Casco de Mula; and SP, San Pedroño) are a good example of environmental adaptation, with pigs living in areas from around sea level to 3000 m of altitude (Ocampo-Gallego, 2019; Suárez-Mesa et al., 2021). Similar to other local breeds (Kušec et al., 2015), the census of CR pigs has been continuously declining in recent decades as intensive farming has replaced traditional production systems. Latest reports indicate that the three CR pig breeds are at high risk of extinction, with only 138 CM, 99 SP, and 128 ZU censused individuals (FAO, 2018). Currently, most of these individuals are maintained in three independent nucleus farms, one per breed, which are managed by the Colombian Agricultural Research Corporation (AGROSAVIA). In each nucleus farm, pigs are distributed in family groups and subjected to a circular mating system for maintaining genetic variability (Ocampo-Gallego, 2019). These pigs are amongst the few available individuals that can be used to investigate

✉ Rafael Suárez-Mesa  
rafael.suarez@udl.cat

✉ Joan Estany  
joan.estany@udl.cat

<sup>1</sup> Department of Animal Science, University of Lleida  
Agrotecnio-CERCA Center, 191 Rovira Roure,  
25198 Lleida, Catalonia, Spain

<sup>2</sup> Research Group in Immunobiology and Pathogenesis,  
Faculty of Veterinary Medicine and Zootechnics, University  
of Tolima, Barrio Santa Helena Parte Alta, Ibagué, Colombia

<sup>3</sup> Research and Innovation Group in Animal Health  
and Welfare Germplasm Animal Bank, Agrosavia,  
Bogotá 250047, Colombia

whether CR pigs harbour specific genetic variants that merit conservation.

Conservation of local breeds depends on their utility and prospects as a research, social, or economic resource (Barker, 1999). The Iberian pig, the most likely ancestor of CR pigs (Burgos et al., 2013), can be referred to as a model of how a local breed, previously at risk, is now abundantly used for producing premium pork products (García-Gudiño et al., 2021) and in genetic research (Crespo-Piazuelo et al., 2020). The identification of breed-specific polymorphisms in genes related to adaptive and performance traits (Bovo et al., 2020) can be a useful first approach to enhance the genetic value of local breeds (Herrero-Medrano et al., 2013; Muñoz et al., 2018). However, this has not yet been done in CR pigs. As in many other local breeds, the first attempts to genetically characterize CR local breeds were based on a small set of neutral markers such as microsatellites (Oslinger et al., 2006; Gélvez et al., 2015), but their relationship with relevant traits is not straightforward (Kirk & Freeland, 2011).

Next-generation DNA sequencing empowers geneticists to identify genetic variants at higher resolution than previously. Whole-genome sequencing has already been used in some European endangered local breeds (D'alessandro et al., 2019; Herrero-Medrano et al., 2014), but not in CR, where only a few variants associated with meat quality and fertility have been studied (Hernández et al., 2008; Pardo, 2016). In order to ensure the continuity of CR breeds, a more profound assessment of their genetic diversity is needed. In particular, CR breeds may carry breed-specific variants of genes related to adaptive and economic traits, such as those reported to affect relevant morphological, reproductive, growth, disease resilience, or meat quality traits. Thus, this research aims at identifying and characterizing the genetic variation in genes with potential effects on adaptive and economic traits in CR breeds through whole-genome sequencing.

## Material and methods

### Animals

Seven representative individuals from each of the three CR breeds (ZU, CM, and SP) were randomly sampled across available families (one per family to ensure representativeness) from the AGROSAVIA germplasm breeding nucleus of La Libertad (Department of Meta), for CM, El Nus (Department of Antioquia), for SP, and Turipaná (Department of Córdoba), for ZU, from April to July 2019 (Suárez-Mesa et al., 2021). In addition, 7 Iberian (IB) and 21 cosmopolitan (CP) pigs (7 Duroc, 7 Pietrain and 7 Landrace × Large White) from the UdLGIM (University of Lleida) biobank (Estany et al., 2014) were also randomly sampled. The IB and CP pigs were chosen for being, respectively, the most likely ancestors

of CR pigs and the current most representative transboundary genetic types. Finally, to better estimate the allele distribution across genetic types, the genotypes of 101 additional pigs from the Iberian trunk (IT; 53 IB and 48 Alentejano) and 194 CP (2 Pietrain, 151 Duroc, and 41 Landrace × Large White) for the 44 preselected markers (see below) were retrieved from either public data (Muñoz et al., 2018) or the UdLGIM biobank.

### Isolation of genomic DNA and whole-genome sequencing

Genomic DNA isolation was performed from blood samples. Briefly, blood samples were washed with TE buffer, then lysed in the presence of proteinase K, and DNA was purified by phenol:chloroform extraction, followed by ethanol precipitation. Finally, the DNA was resuspended and stored in TE buffer. The quantification and estimation of the quality and purity of the genomic DNA were done by spectrophotometer (NanoDrop N-1000, Thermo Fisher Scientific, Wilmington, USA). The integrity of the DNA was tested by electrophoresis on a 0.8% agarose gel and visualized by staining with ethidium bromide under UV illumination. Following the requirements of the National Center for Genomic Analysis (CNAG-CRG, Barcelona, Spain), all samples had a minimum concentration of 50 ng/μl. The concentration was estimated in a fluorometer (Qubit 4, Thermo Fisher Scientific).

The short-insert paired-end libraries for the whole-genome sequencing were prepared with a PCR-free protocol using the KAPA HyperPrep kit (Roche, Basel, Switzerland), with some modifications. In short, depending on the starting DNA available, 0.4 to 1.0 μg of genomic DNA was sheared on a Covaris™ LE220-Plus (Covaris, Brighton, UK) in order to reach the fragment size of ~400 bp. The fragmented DNA was size-selected for the fragment range of 220–550 bp with AMPure XP beads (Agencourt, Beckman Coulter, Nyon, Switzerland). The size-selected genomic DNA fragments were end-repaired and adenylated, and adaptors with unique dual indexes and unique molecular identifiers compatible with the Illumina platform (Integrated DNA Technologies, Leuven, Belgium) were ligated. The libraries were quality-controlled on an Agilent 2100 Bioanalyzer with the DNA 7500 assay (Agilent, Madrid, Spain) for size and quantified by KAPA Library Quantification Kit for Illumina platforms (Roche). The libraries were sequenced on a NovaSeq6000 (Illumina, San Diego, CA, USA) platform in paired-end mode with a read length of 2 × 151 + 17 + 8 bp following the manufacturer's protocol for dual indexing. Image analysis, base calling, and quality scoring of the run were processed using the manufacturer's software Real Time Analysis (RTA 3.4.4, Illumina) and followed by generation of FASTQ sequence files. A minimum of 20 Gb of sequencing data was generated per sample.

## Sequence data processing and variant discovery

The sequence reads were pre-processed using Trimmomatic (Bolger et al., 2014) to remove the adapters from the sequences DNA. The reads were aligned to the reference genome Sscrofa11.1 (GenBank accession: GCA\_000003025.6) using the BWA-MEM algorithm (Li, 2013). Duplicates were marked for exclusion with Picard (<http://broadinstitute.github.io/picard/>). Single nucleotide polymorphisms (SNPs) and short insertions and deletions (indels) were identified with the variant caller GATK HaplotypeCaller (GATK 3.8.0) (DePristo et al., 2011; Poplin et al., 2017) using default settings. The average realized sequencing coverage was  $7.9 \times$  (SD  $2.4 \times$ ). Variant discovery with GATK HaplotypeCaller was performed separately for each individual and then the individuals in each population were jointly genotyped by extracting the variant positions from all the individuals. We retained all biallelic variants for further analyses with VCFtools (Danecek et al., 2011). Variants with minor allele frequency below 0.01 (jointly considering the sequenced individuals from all genetic types) or with a genotyping rate below 90% were removed using PLINK 1.9 software (Chang et al., 2015). This software was also used

to perform a principal component analysis of genomic data to investigate population structure.

## Genetic variants in candidate genes

We preselected 44 SNPs in 34 genes with reported effects on morphological, reproductive, and response to disease-related traits (Table 1) and on growth, fatness, and meat and fat quality traits (Table 2) in pigs. The alleles for all variants are described based on the positive strand of the Sscrofa11.1 pig genome assembly. The genotypes of the sequenced pigs for these 44 SNPs were retrieved from the whole-genome sequence data, and the frequency of each allele in each breed was calculated.

In addition, we also called all the variants along the exonic regions of the genes in Tables 1 and 2 that are annotated in the Sscrofa11.1 assembly of the pig genome (27 of 34 genes). These variants were grouped according to whether they were common (i.e. called in all genetic types) or breed-specific (i.e. called in only one breed or group of breeds, in particular to CR) and to their predicted impact using the Ensembl Variant Effect Predictor (VEP) tool (McLaren et al., 2016).

**Table 1** Investigated genetic variants in genes with reported effects on morphological, reproductive and adaptive (i.e. response to disease) traits

Variant	Gene	SSC	Position (bp)	Reference allele	Alternative allele	Trait	Reference
ACTN1_1	<i>ACTN1</i>	7	92,555,961	T	C	Fertility, piglets born alive	Wimmers et al. (2005)
ADIPOQ_1	<i>ADIPOQ</i>	13	124,643,017	G	A	Morphology, fatness	Zhang et al. (2014)
AHR_1	<i>AHR</i>	9	86,550,830	G	T	Litter size	Bosse et al. (2014)
AHR_2	<i>AHR</i>	9	86,549,936	A	C	Age at puberty	Zhang et al. (2020)
AHR_3	<i>AHR</i>	9	86,551,088	T	C	Female age at puberty	Zhu et al. (2017)
FUT1_1	<i>FUT1</i>	6	54,079,560	T	C	Diseases resistance	Wang et al. (2012)
GBP5_1	<i>GBP1</i>	4	127,301,202	G	T	Early host response to PRRS virus	Kommadath et al. (2017); Jeon et al. (2021)
HSP70_1	<i>HSP70</i>	7	23,925,510	C	A	Sperm concentration, sperm motility	Huang et al. (2002)
HSP70_2	<i>HSP70</i>	7	23,914,842	C	A	Sperm concentration, sperm motility	Huang et al. (2002)
HSP70_3	<i>HSP70</i>	7	23,914,955	T	C	Sperm concentration, sperm motility	Huang et al. (2002)
HSP70_4	<i>HSP70</i>	7	23,925,859	T	C	Sperm concentration, sperm motility	Huang et al. (2002)
KIT_1	<i>KIT</i>	8	41,488,472	C	T	Coat colour	Johansson et al. (2005); Fontanesi et al. (2010)
MC1R_1	<i>MC1R</i>	6	181,461	T	C	Coat colour	Kijas et al. (1998); Kijas et al. (2001)
MC1R_2	<i>MC1R</i>	6	181,697	A	G	Coat colour	Kijas et al. (1998); Kijas et al. (2001)
MC1R_3	<i>MC1R</i>	6	181,818	C	T	Coat colour	Kijas et al. (1998); Kijas et al. (2001)
MC1R_4	<i>MC1R</i>	6	181,825	A	G	Coat colour	Kijas et al. (1998); Kijas et al. (2001)
MC1R_5	<i>MC1R</i>	6	181,905	C	T	Coat colour	Kijas et al. (1998); Kijas et al. (2001)
MUC4_1	<i>MUC4</i>	13	134,226,654	C	G	Resistance to colibacteriosis	Schroyen et al. (2012)
NR6A1_1	<i>NR6A1</i>	1	265,347,265	A	G	Number of vertebrae	Fontanesi et al. (2014)
PPARD_1	<i>PPARD</i>	7	31,281,804	G	A	Ear size	Ren et al. (2011)
TAS2R39_1	<i>TAS2R39</i>	18	7,068,883	T	G	Growth	Ribani et al. (2017)

**Table 2** Investigated genetic variants in genes with reported effects on growth, fatness, and meat and fat quality traits

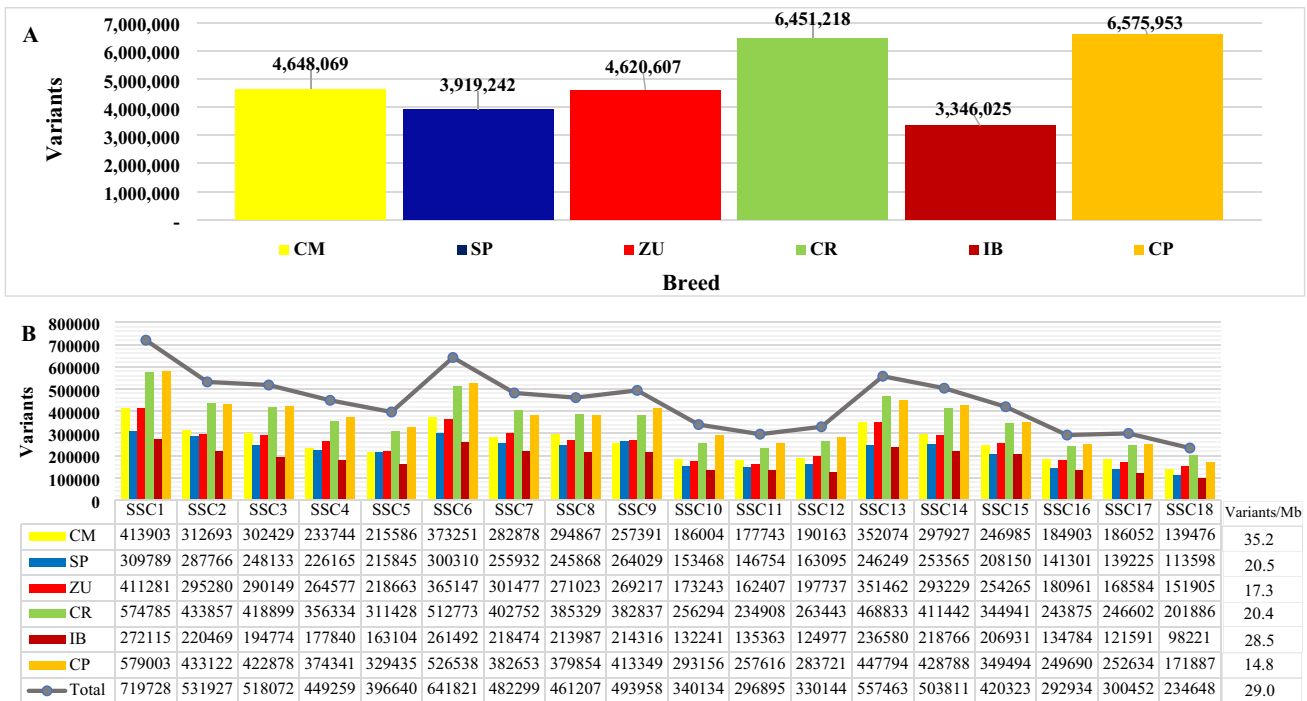
Variant	Gene	SSC	Position (bp)	Reference allele	Alternative allele	Traits affected	Reference
ACACA_1	<i>ACACA</i>	12	3,862,4687	G	A	Carcass fatness, meat quality, fat composition	Muñoz et al. (2013)
CAPNS1_1	<i>CAPNS1</i>	6	45,514,212	A	C	Meat quality	Gandolfi et al. (2011a)
CAST_1	<i>CAST</i>	2	103,299,934	G	A	Meat quality (tenderness)	Meyers & Beever (2008)
CAST_2	<i>CAST</i>	2	103,327,456	G	A	Meat quality	Gandolfi et al. (2011b)
CTSK_1	<i>CTSK</i>	4	98,393,909	G	A	Carcass fatness, meat quality (IMF)	Fontanesi et al. (2010)
CYB5A_1	<i>CYB5A</i>	1	149,737,752	G	T	Meat quality (boar taint)	Peacock et al. (2008); Lin et al. (2005)
CYP2E1_1	<i>CYP2E1</i>	14	141,702,809	A	G	Meat quality (boar taint)	Lin et al. (2006)
FADS2_1	<i>FADS2</i>	2	9,667,336	C	T	Fat composition	Gol et al. (2018)
FASN_1	<i>FASN</i>	12	926,299	G	A	Carcass fatness, meat quality, fat composition	Muñoz et al. (2007)
FTO_1	<i>FTO</i>	6	31,460,242	A	T	Growth, carcass fatness	Dvořáková et al. (2012)
LEP_1	<i>LEP</i>	18	20,111,759	A	G	Growth, carcass fatness, feed intake	Kennes et al. (2001)
LEPR_1	<i>LEPR</i>	6	146,829,589	G	A	Growth, fatness, meat quality (IMF), feed intake	Óvilo et al. (2010); Óvilo et al. (2005)
MC4R_1	<i>MC4R</i>	1	160,773,437	G	A	Feed intake, growth, carcass fatness	Kim et al. (2000)
MSTN_1	<i>MSTN</i>	15	94,629,248	C	T	Growth, carcass fatness	Tu et al. (2014)
MTTP_1	<i>MTTP</i>	8	120,821,998	G	A	Meat quality, fat composition	Estellé et al. (2009)
PCK1_1	<i>PCK1</i>	17	57,932,233	A	C	Growth, carcass fatness, meat quality (IMF)	Latorre et al. (2016)
PHKG1_1	<i>PHKG1</i>	3	16,830,320	C	A	Carcass fatness, meat quality (pH)	Ma et al. (2014)
PLIN1_1	<i>PLIN1</i>	7	55,250,707	C	T	Growth	Gol et al. (2016)
PLIN2_1	<i>PLIN2</i>	1	203,694,497	A	G	Growth, carcass fatness	Gol et al. (2016)
PPARGC1A_1	<i>PPARGC1A</i>	8	17,867,068	A	T	Meat quality	Gandolfi et al. (2011a)
PRKAG3_1	<i>PRKAG3</i>	15	120,863,537	C	T	Meat quality (pH)	Ciobanu et al. (2001); Milan et al. (2000)
RYR1_1	<i>RYR1</i>	6	47,357,966	T	C	Growth, carcass fatness, meat quality (pH)	Roberts et al. (2001); Fujii et al. (1991)
SCD_1	<i>SCD</i>	14	111,461,751	C	T	Fat composition	Estany et al. (2014)

## Results

### Number of the called variants

We present here the first whole-genome sequence data of CR pig breeds (Fig. 1A). The total number of genetic variants called across all sequenced types was 7,971,714. Of them, 6,451,218 were called in CR, which was in line with the number of variants found in CP (6,575,953). The number of variants in each CR breed ranged from 3,919,242, in SP, to 4,648,069, in CM, which was higher than in IB (3,346,025) and similar to CP types (from 3,723,941, in Duroc, to 5,068,938, in Landrace × Large

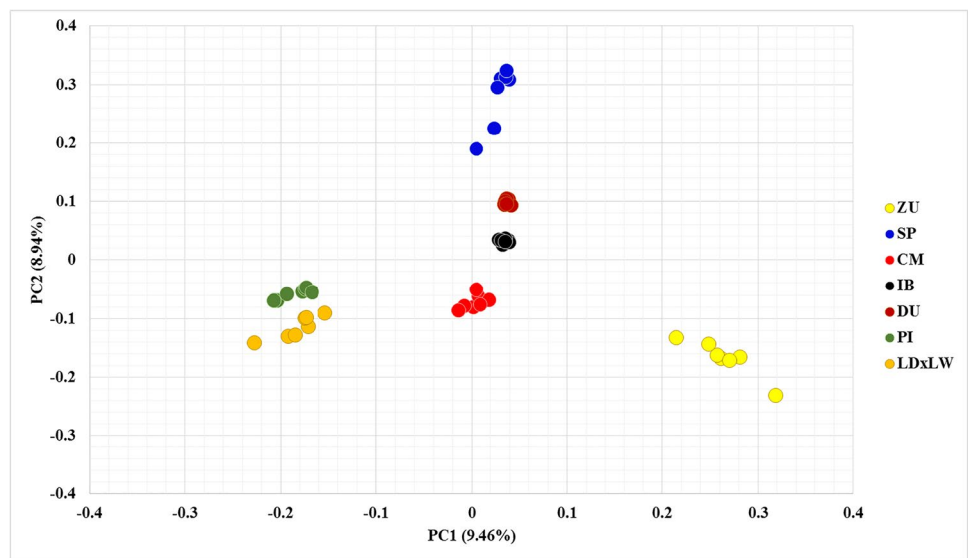
White). As expected, the number of variants per chromosome increased with chromosome size. The CR pigs presented a higher average variant density (20.4 variants/Mb) than the CP pigs (14.8 variants/Mb) (Fig. 1B). Despite the limited number of pigs per breed, sequence data revealed a sensible population structure. As can be seen from the outcome of the principal component analysis of the so-called variants, pigs of the same breed get clustered together (Fig. 2). The first principal component clearly differentiated white types (Pietrain and Landrace × Large White) from ZU, while the second principal component distinguished the other four breeds, particularly CM from SP.



**Fig. 1** Total number of the so-called variants by breed (A) and chromosome (B) in Colombian Creole (CR, *n* = 21, with CM, Casco de Mula; ZU, Zungo; SP, San Pedroño), Iberian (IB, *n* = 7), and cosmo-

politan (CP, *n* = 21) pigs. A number of the so-called variants in the individual CP genetic types were 3,723,941 in Duroc, 5,068,938 in Landrace × Large White, and 4,391,019 in Pietrain

**Fig. 2** Scatter plot of the two first principal components (PC1 and PC2, in parenthesis the variance explained) for genome sequence in Colombian Creole (CM, Casco de Mula; SP, San Pedroño; and ZU, Zungo), Iberian (IB) and Cosmopolitan pig breeds (PI: Pietrain; DU: Duroc; and LDxLW: Landrace × Large White). Each circle represents a subject



**Allele frequency of preselected variants from candidate genes**

The frequency of the alternative allele (as annotated in the reference genome Sscrofa11.1) for each of the 44 SNPs in Tables 1 and 2 is given in Tables 3 and 4, respectively. The CR pigs showed relatively high frequencies for

some of the alternative alleles. As a result, the number of variants that were fixed (i.e. only one of the two alleles of the variant was present in the sample of pigs used for this study) was lower in CR (6) than in IT (13), although higher than in CP (4). There was also variability in the alternative allele frequency of the preselected variants across CR breeds. Twenty-eight of them were fixed in at

**Table 3** Frequency of the alternative allele in the investigated variants for morphological, reproductive and adaptive (i.e. response to disease) traits in Colombian Creole, Iberian trunk and cosmopolitan pig breeds

Variant <sup>2</sup>	Alternative allele	Breed <sup>1</sup>					
		Colombian Creole					
		CM <i>n</i> = 7	SP <i>n</i> = 7	ZU <i>n</i> = 7	CR <i>n</i> = 21	IT <i>n</i> = 108	CP <i>n</i> = 215
ACTN1_1	T	0.29	0.79	0.33	0.47	0.58	0.41
ADIPOQ_1	A	0.21	0.57	0.21	0.33	0.03	0.13
AHR_1	T	0.75	1.00	0.79	0.85	0.94	0.33
AHR_2	C	0.71	1.00	0.86	0.86	1.00	0.36
AHR_3	C	0.79	1.00	0.86	0.88	1.00	0.32
FUT1_1	C	0.50	0.36	0.93	0.60	0.94	0.75
GBP5_1	T	0.07	0.36	0.64	0.36	0.00	0.19
HSP70_1	A	0.43	0.21	0.00	0.21	0.60	0.38
HSP70_2	C	0.00	0.50	0.50	0.33	0.75	0.63
HSP70_3	A	0.00	0.00	0.00	0.00	0.17	0.01
HSP70_4	C	1.00	1.00	0.50	0.83	0.88	0.76
KIT_1	T	0.07	0.00	0.00	0.02	0.00	0.00
MC1R_1	T	0.00	0.14	0.29	0.14	0.01	0.00
MC1R_2	C	1.00	1.00	1.00	1.00	1.00	0.67
MC1R_3	G	1.00	1.00	1.00	1.00	1.00	0.67
MC1R_4	T	1.00	0.86	0.79	0.88	0.76	0.67
MC1R_5	G	0.00	0.14	0.21	0.12	0.01	0.00
MUC4_1	G	0.00	0.00	0.07	0.02	0.03	0.29
NR6A1_1	G	0.07	0.00	0.00	0.02	0.14	0.00
PPARD_1	A	0.00	0.00	0.50	0.17	0.02	0.01
TAS2R39_1	G	1.00	1.00	1.00	1.00	0.96	0.79

Bold type indicates allele fixation. <sup>1</sup>Colombian Creole breeds (*CM*, Casco de Mula; *SP*, San Pedroño; *ZU*, Zungo; *CR*, all three Colombian Creole breeds); *IT*, Iberian trunk pigs (60 Iberian and 48 Alentejano); and *CP*, cosmopolitan breeds (9 Pietrain; 158 Duroc; and 48 Landrace × Large White). <sup>2</sup>See Table 1 for variant description

least one CR breed, with SP having the largest number of fixed variants (21), ZU the fewest (10), while CM had an intermediate number (16). Ten of the 13 variants (from 8 of 34 genes) that were fixed in IT were also fixed in SP, including *AHR\_2*, *AHR\_3*, *LEPR\_1*, *MC4R\_1*, and *SCD\_1*, but only 6 (from 5 genes) and 3 (from 3 genes) in ZU and CM, respectively. Interestingly, the maximum differences in allele frequency across genetic types occurred for *LEPR\_1*, where the frequency of the alternative allele ranged from values lower than 0.3, in CP, CM, and ZU, to 1.00, in IT and SP, and for *AHR* polymorphisms, where the frequency of the alternative allele was much higher in CR and IT pigs (0.71–1.00) than in CP pigs (0.32–0.36). The other three fixed variants in IT (*CAST\_2*, *PLIN2\_1*, and *GBP5\_1*) segregated in all three CR breeds, except for *PLIN2\_1* in ZU, for which the same allele as in IT was fixed. In general, the alleles fixed in CR segregated at a very high frequency in IT.

### Additional variants in preselected candidate genes

In annotated genes (Table 5), exonic sequence variation in CR (334 variants; Supporting Table S1) was greater than in IB (200 variants) but lower than in CP (369 variants). The SP pigs had less exonic variants (178) than the two other CR (254, for ZU, and 263, for CM), IB (200), and CP breeds (201, for Pietrain, to 335, for Landrace × Large White). A total of 106 of these variants were common to all breeds, while 50 of them were specific to CR (Fig. 3). As compared to CR, the number of specific variants was similar in IB (53 variants) and higher in CP (83 variants). However, CR breeds shared less variants with IB that were not called in CP (4) than variants with CP that were not called in IB (143). Variants were categorized according to their predicted impact over mRNA transcription and protein translation and functionality (Table 6). In CR, ZU had the highest number of breed-specific variants (16), of low

**Table 4** Frequency of the alternative allele in the investigated variants for growth, fatness and meat and fat quality traits in Colombian Creole, Iberian trunk and cosmopolitan pig breeds

Variant <sup>2</sup>	Alternative allele	Breed <sup>1</sup>						
		Colombian Creole						
		CM <i>n</i> =7	SP <i>n</i> =7	ZU <i>n</i> =7	CR <i>n</i> =21	IT <i>n</i> =108	CP <i>n</i> =215	
ACACA_1	A	0.57	0.50	0.79	0.62	0.57	0.36	
CAPNS1_1	C	0.71	0.50	0.93	0.71	0.48	0.52	
CAST_1	A	0.21	0.50	0.43	0.38	0.83	0.35	
CAST_2	A	0.21	0.14	0.43	0.26	0.00	0.35	
CTSK_1	A	0.00	0.00	0.00	0.00	0.00	0.08	
CYB5A_1	T	0.00	0.00	0.43	0.14	0.48	0.07	
CYP2E1_1	G	0.57	0.07	0.79	0.48	0.79	0.40	
FADS2_1	T	1.00	0.36	0.93	0.76	0.92	0.55	
FASN_1	A	0.50	1.00	0.29	0.60	0.75	0.73	
FTO_1	T	0.50	0.36	0.21	0.36	0.67	0.37	
LEP_1	G	0.93	0.64	0.29	0.62	0.20	0.89	
LEPR_1	A	0.21	1.00	0.29	0.50	1.00	0.23	
MC4R_1	A	0.79	0.00	0.07	0.29	0.00	0.26	
MSTN_1	T	0.50	0.50	0.71	0.57	0.88	0.54	
MTTP_1	A	0.36	0.64	0.93	0.64	0.33	0.43	
PCK1_1	C	0.36	0.00	0.43	0.26	0.03	0.55	
PHKG1_1	A	0.00	0.00	0.14	0.05	0.02	0.10	
PLIN1_1	T	0.93	0.64	0.79	0.79	0.59	0.44	
PLIN2_1	G	0.79	0.21	1.00	0.67	1.00	0.81	
PPARGC1A_1	T	0.29	0.64	0.79	0.57	0.14	0.64	
PRKAG3_1	T	0.07	0.29	0.33	0.23	0.67	0.25	
RYR1_1	C	1.00	1.00	1.00	1.00	1.00	0.83	
SCD_1	T	0.57	1.00	0.93	0.83	1.00	0.80	

Bold type indicates allele fixation. <sup>1</sup>Colombian Creole breeds (*CM*, Casco de Mula; *SP*, San Pedroño; *ZU*, Zungo; *CR*, all three Colombian Creole breeds); *IT*, Iberian trunk pigs (60 Iberian and 48 Alentejano); and *CP*, cosmopolitan breeds (9 Pietrain; 158 Duroc; and 48 Landrace × Large White). <sup>2</sup>See Table 2 for variant description

(11) or moderate (5) predicted impact. These variants were located in the *ACTN1*, *FADS2*, *FTO*, *FUT1*, *PHKG1*, *PLIN2*, *PRKAG3*, and *TAS2R39* genes. The breed-specific variants found in *CM* (13) were in another set of genes (*ADIPOQ*, *CAPNS1*, *CAST*, *FTO*, *KIT*, *LEPR*, and *PPARGC1A*) and were of low (9) or moderate (4) predicted impact. The *SP* breed presented only 4 breed-specific variants, 3 in the *ACACA* gene and 1 in the *FADS2* gene, all of them of low predicted impact.

In total, there were 18 variants with high predicted impact in *CR* (Supporting Table S1). Nine of them were shared amongst *CR* breeds and all but one (in *MC1R* and segregating in *CM*) were in *ZU*. Six genes (*ACTN1*, *ADIPOQ*, *LEPR*, *PCK1*, *PRKAG3*, and *TAS2R39*) harboured the other 8 high-impact variants. The three high-impact variants in *ACTN1* were not observed in *SP* and only one of the two that were identified in *PCK1*, *PRKAG3* and *ADIPOQ* were observed in *CM*. The high-impact variants in *LEPR* and *TAS2R39* were found in *CM* and *ZU* but not in *SP*. Of all high-impact

variants, only the splice-donor polymorphism located in the *LEPR* gene was specific to *CR*. High-impact variants were mostly frameshift indels (Table 6).

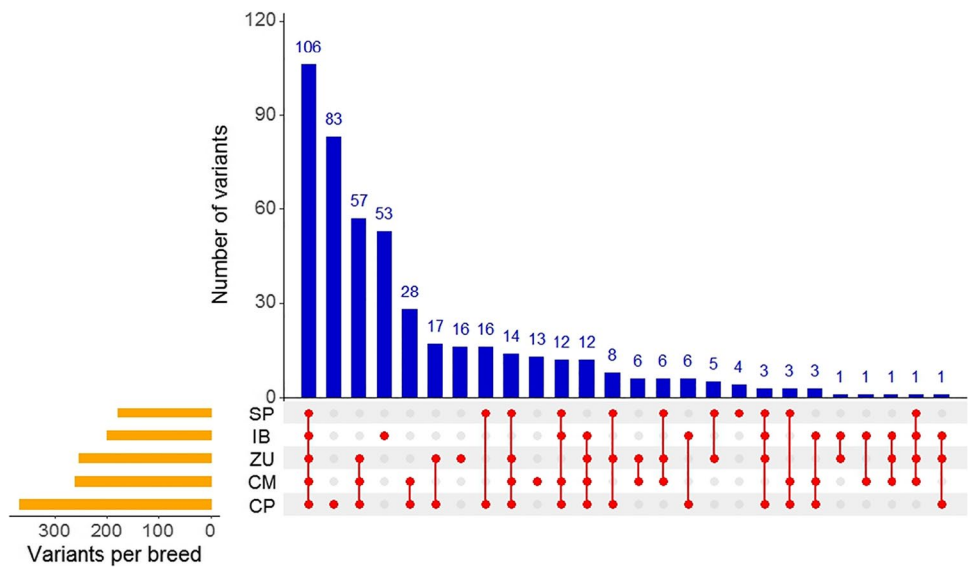
## Discussion

The few analyses of genetic variation carried out so far in *CR* pigs were limited to a small set of neutral markers (Vargas et al., 2016). Whole-genome sequencing provides a more comprehensive resolution of the genetic variation within and across populations across all genomic regions (Ros-Freixedes et al., 2022). Here, we focused on a set of 34 candidate genes with reported effects on adaptive or economic traits (Tables 1 and 2). Six of the preselected variants in these genes were not observed in *CR* (*CTSK\_1*, *HSP70\_3*, *MC1R\_2*, *MC1R\_3*, *RYR1\_1*, and *TAS2R39\_1*), and seven were only seen in one of the *CR* breeds (*CYB5A\_1*, *HSP70\_4*, *KIT\_1*, *MUC4\_1*, *NR6A1\_1*, *PHKG1\_1*, and





**Fig. 3** UpSet plot of the variants found in the investigated genes (Table 5). Set size is the total number of variants in each breed (CM, Casco de Mula; ZU, Zungo; SP, San Pedroño; IB, Iberian; and CP, Cosmopolitan breeds). Interaction size represents the number of variants in the intersections of the breeds as indicated by the black dots below the bars



**Table 6** Number of exonic variants by breed and predicted impact over mRNA transcription and translation

Predicted impact <sup>2</sup>	Breed <sup>1</sup>					
	Colombian Creole					
	CM <i>n</i> =7	SP <i>n</i> =7	ZU <i>n</i> =7	CR <i>n</i> =21	IB <i>n</i> =7	CP <i>n</i> =21
High	15	12	17	18	14	20
Frameshift indel	11	10	12	13	12	14
Splice acceptor	1	1	1	1	1	2
Splice donor	2	-	2	2	1	2
Start lost	1	-	1	1	-	1
Stop lost	-	1	1	1	-	1
Moderate	83	57	82	106	55	124
In-frame deletion	3	2	3	3	2	4
Missense variant	80	55	79	103	53	120
Low	165	109	155	210	131	225
Splice region	37	21	33	45	18	47
Synonymous	128	88	122	165	113	178
Total	263	178	254	334	200	369 <sup>3</sup>

Bold type indicates totals. <sup>1</sup>Colombian Creole breeds (CM, Casco de Mula; SP, San Pedroño; ZU, Zungo; CR, all three Colombian Creole breeds); IB, Iberian; and CP, cosmopolitan breeds (7 Pietrain; 7 Duroc; and 7 Landrace×Large White). <sup>2</sup>Predicted impact using the Ensembl Variant Effect Predictor (VEP). <sup>3</sup>Duroc: 223 exonic variants in total; Landrace×Large White: 335; Pietrain: 201

PPARD\_1). The variants in the *MC1R*, *NR6A1*, *PPARD*, and *TAS2R39* genes are missense mutations that might have been selected for environmental adaptation.

The *MC1R* gene has a great impact in the determination of coat colour due to its key role regulating the synthesis of eumelanin (black/brown) and pheomelanin (yellow/red) in the melanocytes (Barsh, 1996; Fang et al., 2009). At least six haplotypes, tagged by 5 SNPs (MC1R\_1 to MC1R\_5, Table 1) and one deletion (g.182126CC > \*), have been described in this gene (Muñoz et al., 2018). In CR, we only

found three of these six haplotypes (Supporting Table S2), the so-called MC1R\*2 (GCGCA\*\*), MC1R\*3 (GCATG\*\*), and MC1R\*6 (GCATGCC), all of which are associated with black coat or spotting. The predominance of MC1R\*3 in SP (frequency of 85.7%) and in ZU (frequency of 71.4%) is consistent with IB (likely, Lampiño) origin (Ocampo-Gallego and Abuabara-Pérez, 2021), since this haplotype is fixed in old black-coated and hairless IB strains such as Lampiño (Alves et al., 2007; Fernández et al., 2004). However, MC1R\*3 was residual in CM, where MC1R\*6 was the

predominant haplotype (frequency of 85.7%), as happens in current commercial IB strains (Muñoz et al., 2018). The presence of the MC1R\*2 in SP (frequency of 14.3%) and ZU (frequency 21.4%), which has been previously detected in Large Black, provides evidence of introgression of black alleles from Asian origin into these two breeds. Likewise, CM does not seem to be completely free of introgression from transboundary breeds, as indicates the presence of the T allele in the KIT\_1 variant (belted phenotype), which is absent in IT breeds (Muñoz et al., 2018). The absence of the MC1R\*4 (ATGTG\*\*) haplotype in CR pigs indicates that they have not been crossbred with Duroc.

The *TAS2R39* gene is a member of the bitter-taste receptor family that has been related to fatness (Ribani et al., 2017). In agreement with findings in European local breeds (Muñoz et al., 2018), the G allele at *TAS2R39\_1* is fixed in the three CR breeds, thereby suggesting a selective pressure towards defensive bitter taste. The A allele at *NR6A1\_1* was fixed in SP and ZU but not in CM. This allele increases the vertebrae number in pigs, resulting in longer carcasses (Mikawa et al., 2007). This could imply that CM pigs could have been less intensively selected for body size than SP and ZU. On the other hand, *PPARD\_1*, as well as *CBY5A\_1* and *HSP70\_4*, only segregated in ZU and at intermediate frequencies. Since the ZU pigs are found in the Atlantic coastal area, where the weather is especially hot, it is worth exploring whether these three variants might be related to heat resistance, as it has been described before. For instance, the missense mutation *PPARD\_1* (A allele; Table 1) increases ear size in pigs (Ren et al., 2011), with implications on skin homeostasis and fat deposition. The A allele is found in Asian but not in European breeds. The fact that the A allele segregated in ZU at a frequency of 50% adds evidence of Asian introgression into this breed, which, on the other hand, is characterized by having large and droopy ears (FAO, 1992). The T allele at *CYB5A\_1* has been associated with low fat and androstenone levels (Peacock et al., 2008; Lin et al., 2005). While this may be desirable to reduce the risk of boar taint in carcasses from entire ZU males, it may jeopardize reproduction success. The *HSP70* variants have also been related to male reproduction. In particular, the T allele at *HSP70\_4* has been associated to larger ejaculates and semen quality. In a previous research we showed that CR boars produced less normal and motile sperm per ejaculate than CP boars (Suárez-Mesa et al., 2021). Allele frequency patterns of *HSP70* variants across breeds do not provide further evidence for an association of these markers with male fertility.

The rest of variants that appeared as fixed in a single CR breed were mostly found in SP. These variants were also either fixed or at very high frequency in IT (*AHR\_1* to *AHR\_3*, *LEPR\_1*, *MC4R\_1*, *PCK1\_1*, and *SCD\_1*). The fatty nature of these breeds is consistent with the presence of the A allele at *LEPR\_1* (Table 2), which has been

documented to increase feed intake and fatness and to impair reproductive and maternal abilities (Ros-Freixedes et al., 2016; Solé et al., 2021). This allele co-segregated with the T allele at *AHR\_1*, which has a negative impact on prolificacy (Bosse et al., 2014). The joint presence of these two fixed alleles in SP can compromise the reproductive outcome of this breed. Given the sample size per breed, no clear-cut pattern can be inferred from the allele frequency of *MC4R\_1*, *PCK1\_1*, and *SCD\_1* across breeds, except that, amongst CR, SP was the closest to IT and CM the most differentiated. The CM pigs showed higher frequencies of the alleles associated with increased fatness (A allele at *MC4R\_1*; Kim et al., 2004; and C allele at *PCK1\_1*; Latorre et al., 2016) and saturated fatty acid abundance (C allele at *SCD\_1*; Estany et al., 2014).

A total of 27 of the studied genes are annotated in the Sscrofa11.1 assembly of the pig genome. Therefore, in a second step, we went further to search for new variants into the coding region of these genes using whole-genome sequence data. In CR, we found 18 variants of high impact on mRNA sequence and protein translation. Only one of them was specific to CR. This is a splice-donor variant in *LEPR* that consists of a 7-bp deletion extending upstream on intron between exons 15 and 16 (*LEPR\_2*: SSC6:146,829,573–146,829,580 bp) that affects the three transcripts of the gene. This deletion was only observed in CM and ZU (frequency of 7.1% and 28.6%, respectively) and was fully linked to the G (non-fatty) allele in *LEPR\_1*, but not vice versa. Since these two variants are separated by only 9 bp, we can hypothesize that the 7-bp deletion appeared later from a haplotype with the G allele at *LEPR\_1*. No homozygous individuals for the deletion allele were found, even though the probability of sampling at least one in ZU was around 45%. Apart from *LEPR\_1*, other seven variants of moderate impact were detected in CR for *LEPR* (Table 5; Supporting Table S3). In line with results for *LEPR\_2*, SP pigs did not show the alternative allele in these variants (*LEPR\_3*: SSC6:146,831,558 bp, frequency of 57.1%, only in CM; *LEPR\_4*: SSC6:146,838,276 bp, frequency of 7.1%, in CM, and 50.0%, in ZU; *LEPR\_5*: SSC6:146,838,380 bp, frequency of 21.4.0%, only in ZU; *LEPR\_6*: SSC6:146,847,237 bp, frequency of 21.4%, only in CM; *LEPR\_7*: SSC6:146,861,093 bp, frequency of 42.9%, only in CM; *LEPR\_8*: SSC6:146,861,094 bp, frequency of 42.9%, only in CM; and *LEPR\_9*: SSC6:146,861,105 bp, frequency of 64.3%, in CM, and 14.3%, in ZU). Interestingly, for *LEPR\_1* to *LEPR\_9*, the same allele was fixed in SP and IB. However, considering all *LEPR* variants in Table 5, we can infer that, in SP, all of these variants reside in a single haplotype of 35,019 bp (from SSC6:146,826,086 bp to SSC6:146,861,105 bp), while, in IB, they are inherited in two haplotypes due to a specific in-between missense variant at SSC6:146,830,356 bp (frequency of 54.2%). More

detailed studies are needed to decipher the connection between *LEPR* variants and their effects on phenotypes. Nevertheless, findings so far provide further evidence on the IB origin of SP and give clues about the role that *LEPR*, as a key element of the endocrine control of energy balance (Friedman, 2019), may have played in the adaptation of CR breeds to different geographical locations and dietary regimes.

Besides those in *LEPR*, we identified 12 more missense mutations. Three of them were present in more than one CR breed and affected genes involved in coat colour (*KIT*, in the three CR breeds, and *MC1R*, in SP and ZU). The remaining 9 were observed in CM and ZU (*FASN*) or only in CM (*ADIPOQ*, *CAST*, and *PPARGCIA*) or ZU (*FADS2*, *FUT1*, *ACTN1*, and *PRKAG3*). No breed-specific missense mutations were found in SP, a result that would confirm that molecular variability is lower in this breed as compared to CM and ZU. As a whole, our findings support that genome-wide characterization is a useful tool to identify patterns of genetic variation between and within CR breeds.

## Conclusions

This is the first study that characterizes genetic variation at the whole-genome sequence level in CR pigs. The molecular variability of the three CR breeds is comparable to CP breeds, although higher in ZU and CM than in SP. Despite the limited sample size per breed, the sequence variation of the 34 investigated genes would confirm the relationship between CR and IB pigs, but also that they are not exempt from selective introgression of transboundary breeds, particularly ZU and CM. Differential allele distribution across breeds provides evidence to understand the genetic makeup of the CR breeds for body size, fatness, skin colour, ear size, and boar taint. The identification of 50 sequence variants that are potentially specific to CR points the way forward for further research and adds new data to inform breed development and conservation decisions. The discovery of a novel variant of *LEPR* in CM and ZU can give new clues on the role of *LEPR*-environment interactions on local adaptation. Our findings reinforce the need for ad hoc phenotyping schemes in order to experimentally validate in silico predictions of the impact of such variants on adaptive and economical traits and to develop effective breeding and conservation programmes for CR breeds.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11250-023-03557-9>.

**Acknowledgements** We acknowledge AGROSAVIA, Colombian Corporation for Agricultural Research, for enabling us to use the Creole pigs from the Colombian porcine germplasm programme and the personnel there for their cooperation and technical assistance.

**Author contribution** RSM, IRB, and JE conceived and designed the experiment. RSM, RRF, HL, and RNP performed the analyses. RSM, RRF, and JE wrote the original draft. All authors read and approved the final manuscript.

**Funding** Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. The authors acknowledge the financial support from the Spanish Ministry of Science, Innovation & Universities and the EU Regional Development Funds (grants RTI2018-101346-B-I00 and PID2021-125689OB-I00, Spain). RSM acknowledges the financial support of the Tolima Government and Administrative Department of Science, Technology and Innovation (COLCIENCIAS, call 755 for the formation of high-level human capital for Tolima; Colombia) and of the Animal Breeding and Genetics Research Group, University of Lleida, for complementary funding.

**Data availability** Please contact author for data requests.

## Declarations

**Ethical approval** The experimental protocol was approved by the Ethical Committee on Animal Experimentation of the University of Tolima.

**Consent to participate** Not applicable.

**Consent for publication** All the authors read and agree to the content of this paper and its publication.

**Conflict of interest** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Alves, E., Fernández, A., Barragán, C., Rodríguez, C., & Silió, L., 2007. Bayesian analysis of genetic differentiation between populations of the Duroc and Iberian pig breeds. *International Symposium on the Mediterranean Pig CIHEAM* 76, 23–28.
- Barker, B.S., 1999. Conservation of livestock breed diversity. *Anim Genet Resour* 25, 33–43. [http://www.cattlenetwork.net/docs/agri/agri25\\_2.pdf](http://www.cattlenetwork.net/docs/agri/agri25_2.pdf)
- Barsh, G.S., 1996. The genetics of pigmentation: From fancy genes to complex traits. *Trends in Genetics* 12(8), 299–305. [https://doi.org/10.1016/0168-9525\(96\)10031-7](https://doi.org/10.1016/0168-9525(96)10031-7)
- Bolger, A.M., Lohse, M., Usadel, B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Bosse, M., Megens, H.J., Frantz, L.A.F., Madsen, O., Larson, G., Paudel, Y., Duijvesteijn, N., Harlizius, B., Hagemeyer, Y., Crooijmans, R., & Groenen, M.A.M., 2014. Genomic analysis reveals selection for Asian genes in European pigs following

- human-mediated introgression. *Nature Communications* 5, 1–8. <https://doi.org/10.1038/ncomms5392>
- Bovo, S., Ribani, A., Muñoz, M., Alves, E., Araujo, J.P., Bozzi, R., Čandek-Potokar, M., Charneca, R., Di Palma, F., Etherington, G., Fernandez, A. I., García, F., García-Casco, J., Karolyi, D., Gallo, M., Margeta, V., Martins, J. M., Mercat, M. J., Moscatelli, G., ... Fontanesi, L., 2020. Whole-genome sequencing of European autochthonous and commercial pig breeds allows the detection of signatures of selection for adaptation of genetic resources to different breeding and production systems. *Genetics Selection Evolution* 52(1), 1–19. <https://doi.org/10.1186/s12711-020-00553-7>
- Burgos, W., Souza, C., Megens, H., Ramayo, Y., Melo, M., Lemús, C., Caal, E., Soto, H., Martínez, R., Álvarez, L., Aguirre, L., Iñiguez, V., Revidatti, M., Martínez, O., Llambi, S., Esteve, A., Rodríguez, M., Crooijmans, R., Paiva, S., ... Pérez, M., 2013. Porcine colonization of the Americas : a 60k SNP story. *Heredity* 110, 321–330. <https://doi.org/10.1038/hdy.2012.109>
- Ciobanu, D.C., Day, A.E., Nagy, A., Wales, R., Rothschild, M.F., Plastow, G.S., 2001. Genetic variation in two conserved local Romanian pig breeds using type 1 DNA markers. *Genetics Selection Evolution* 33(4), 417–432. <https://doi.org/10.1051/gse:2001125>
- Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M. Lee, J.J., 2015. Second generation PLINK: rising to the challenge of larger and richer datasets, *Gigascience* 4, s13742–015–0047–8. <https://doi.org/10.1186/s13742-015-0047-8>
- Crespo-Piazuelo, D., Criado-Mesas, L., Revilla, M., Castelló, A., Noguera, J.L., Fernández, A.I., Ballester, M., Folch, J.M., 2020. Identification of strong candidate genes for backfat and intramuscular fatty acid composition in three crosses based on the Iberian pig. *Scientific Reports* 10(1), 1–17. <https://doi.org/10.1038/s41598-020-70894-2>
- Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., Durbin, R., 2011. The variant call format and VCFtools. *Bioinformatics* 27(15), 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>
- D'Alessandro, E., Giosa, D., Sapienza, I., Giuffrè, L., Cigliano, R.A., Romeo, O., Zumbo, A., 2019. Whole genome snps discovery in nero siciliano pig. *Genetics and Molecular Biology* 42(3), 594–602. <https://doi.org/10.1590/1678-4685-gmb-2018-0169>
- Dvořáková, V., Bartenschlager, H., Stratil, A., Horák, P., Stupka, R., Cítek, J., Sprysl, M., Hrdlicová, A., Geldermann, H., 2012. Association between polymorphism in the FTO gene and growth and carcass traits in pig crosses. *Genetics, Selection, Evolution : GSE* 44, 13. <https://doi.org/10.1186/1297-9686-44-13>
- DePristo, M., Banks, E., Poplin, R., Garimella, K., Maguire, J., Hartl, C., Philippakis, A., Del-Angel, G., Rivas, M., Hanna, M., McKenna, A., Fennell, T., Kernysky, A., Sivachenko, A., Cibulskis, K., Gabriel, S., Altshul, D., Daly, M., 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43(5), 491–498. <https://doi.org/10.1038/ng.806.A>
- Estany, J., Ros-Freixedes, R., Tor, M., Pena, R.N., 2014. A functional variant in the stearoyl-CoA desaturase gene promoter enhances fatty acid desaturation in pork. *PLoS ONE* 9(1), 1–11. <https://doi.org/10.1371/journal.pone.0086177>
- Estellé, J., Fernández, A.I., Pérez-Enciso, M., Fernández, A., Rodríguez, C., Sánchez, A., Noguera, J.L., Folch, J.M., 2009. A non-synonymous mutation in a conserved site of the MTTP gene is strongly associated with protein activity and fatty acid profile in pigs. *Animal Genetics* 40(6), 813–820. <https://doi.org/10.1111/j.1365-2052.2009.01922.x>
- Fang, M., Larson, G., Ribeiro, H.S., Li, N., Andersson, L., 2009. Contrasting mode of evolution at a coat color locus in wild and domestic pigs. *PLoS Genetics* 5(1). <https://doi.org/10.1371/journal.pgen.1000341>
- FAO., 1992. *Animal genetic resources information* 9. AGRI, AGAP, FAO. <https://www.fao.org/3/a1250s/a1250s.pdf>
- FAO., 2018. *Sistema de Información sobre la diversidad de los animales domésticos*. <http://www.fao.org/dad-is/browse-by-count-ry-and-species/es/>
- Fernández, A., Fabuel, E., Alves, E., Rodríguez, C., Silió, L., Óvilo, C., 2004. DNA tests based on coat colour genes for authentication of the raw material of meat products from Iberian pigs. *Journal of the Science of Food and Agriculture* 84(14), 1855–1860. <https://doi.org/10.1002/jsfa.1829>
- Fontanesi, L., D'Alessandro, E., Scotti, E., Liotta, L., Crovetto, A., Chiofalo, V., Russo, V., 2010. Genetic heterogeneity and selection signature at the KIT gene in pigs showing different coat colours and patterns. *Animal Genetics* 41(5), 478–492. <https://doi.org/10.1111/j.1365-2052.2010.02054.x>
- Fontanesi, L., Ribani, A., Scotti, E., Utzeri, V.J., Veličković, N., Dall'Olio, S., 2014. Differentiation of meat from European wild boars and domestic pigs using polymorphisms in the MC1R and NR6A1 genes. *Meat Science* 98(4), 781–784. <https://doi.org/10.1016/j.meatsci.2014.07.026>
- Friedman, J.M., 2019. Leptin and the endocrine control of energy balance. *Nature Metabolism* 1(8), 754–764. <https://doi.org/10.1038/s42255-019-0095-y>
- Fujii, J., Otsu, K., Zorzato, F., De Leon, S., Khanna, V.K., Weiler, J.E., O'Brien, P.J., MacLennan, D.H., 1991. Identification of a mutation in porcine ryanodine receptor associated with malignant hyperthermia. *Science* 253(5018), 448–451. <https://doi.org/10.1126/science.1862346>
- Gandolfi, G., Cinar, M.U., Ponsuksili, S., Wimmers, K., Tesfaye, D., Looft, C., Jünger, H., Tholen, E., Phatsara, C., Schellander, K., Davoli, R., 2011. Association of PPARGC1A and CAPN1 gene polymorphisms and expression with meat quality traits in pigs. *Meat Science* 89(4), 478–485. <https://doi.org/10.1016/j.meatsci.2011.05.015>
- Gandolfi, G., Pomponio, L., Ertbjerg, P., Karlsson, A.H., Nanni Costa, L., Lametsch, R., Russo, V., Davoli, R., 2011. Investigation on CAST, CAPN1 and CAPN3 porcine gene polymorphisms and expression in relation to post-mortem calpain activity in muscle and meat quality. *Meat Science* 88(4), 694–700. <https://doi.org/10.1016/j.meatsci.2011.02.031>
- García-Gudiño, J., Blanco-Penedo, I., Font-i-Furnols, M., Angón, E., Perea, J.M., 2021. Analysis of the sustainability of fattening systems for Iberian traditional pig production through a technical and environmental approach. *Animals* 11(2), 1–13. <https://doi.org/10.3390/ani11020411>
- Gélvez, I.M., Pérez, E.P., Martínez, T.C., 2015. Genetic variation in domestic pig *Sus scrofa domestica* populations in Córdoba-Colombia based on microsatellite markers. *Revista Mexicana De Ciencias Pecuarias* 6(4), 443–452. <https://doi.org/10.22319/rmcp.v6i4.4104>
- Gol, S., Pena, R.N., Rothschild, M.F., Tor, M., Estany, J., 2018. A polymorphism in the fatty acid desaturase-2 gene is associated with the arachidonic acid metabolism in pigs. *Scientific Reports* 8(1), 1–9. <https://doi.org/10.1038/s41598-018-32710-w>
- Gol, S., Ros-Freixedes, R., Zambonelli, P., Tor, M., Pena, R.N., Braglia, S., Zappaterra, M., Estany, J., Davoli, R., 2016. Relationship between perilipin genes polymorphisms and growth, carcass and meat quality traits in pigs. *Journal of Animal Breeding and Genetics* 133(1), 24–30. <https://doi.org/10.1111/jbg.12159>
- Hernández, D., Posso, Á.M.T., Muñoz, J.E.F., 2008. Detección de una mutación puntual en el gen receptor Ryanodina (Ryr 1) en cerdos criollos colombianos. *Acta Agronómica* 57(4), 275–278. [http://www.revistas.unal.edu.co/index.php/acta\\_agronomica/article/view/9266/9927](http://www.revistas.unal.edu.co/index.php/acta_agronomica/article/view/9266/9927)
- Herrero-Medrano, J.M., Megens, H.J., Groenen, M.A.M., Bosse, M., Pérez-Enciso, M., Crooijmans, R., 2014. Whole-genome sequence

- analysis reveals differences in population management and selection of European low-input pig breeds. *BMC Genomics* 15(1), 1–12. <https://doi.org/10.1186/1471-2164-15-601>
- Herrero-Medrano, J. M., Megens, H. J., Groenen, M. A. M., Ramis, G., Bosse, M., Pérez-Enciso, M., Crooijmans, R., 2013. Conservation genomic analysis of domestic and wild pig populations from the Iberian Peninsula. *BMC Genomics* 14, 1–13. <https://doi.org/10.1186/1471-2156-14-106>
- Huang, S.Y., Chen, M.Y., Lin, E.C., Tsou, H.L., Kuo, Y.H., Ju, C.C., Lee, W.C., 2002. Effects of single nucleotide polymorphisms in the 5'-flanking region of heat shock protein 70.2 gene on semen quality in boars. *Animal Reproduction Science* 70(1–2), 99–109. [https://doi.org/10.1016/S0378-4320\(01\)00202-0](https://doi.org/10.1016/S0378-4320(01)00202-0)
- Jeon, R. L., Cheng, J., Putz, A.M., Dong, Q., Harding, J.C.S., Dyck, M.K., Plastow, G.S., Fortin, F., Lunney, J., Rowland, R., Canada, P.G., Dekkers, J.C.M., 2021. Effect of a genetic marker for the GBP5 gene on resilience to a polymicrobial natural disease challenge in pigs. *Livestock Science* 244(June 2020), 104399. <https://doi.org/10.1016/j.livsci.2021.104399>
- Johansson, A., Pielberg, G., Andersson, L., Edfors-Lilja, I., 2005. Polymorphism at the porcine Dominant white/KIT locus influence coat colour and peripheral blood cell measures. *Animal Genetics* 36(4), 288–296. <https://doi.org/10.1111/j.1365-2052.2005.01320.x>
- Kennes, Y.M., Murphy, B.D., Pothier, F., Palin, M.F., 2001. Characterization of swine leptin (LEP) polymorphisms and their association with production traits. *Animal Genetics* 32(4), 215–218. <https://doi.org/10.1046/j.1365-2052.2001.00768.x>
- Kijas, J.M.H., Moller, M., Plastow, G., & Andersson, L. 2001. A frameshift mutation in MC1R and a high frequency of somatic reversions cause black spotting in pigs. *Genetics* 158(2), 779–785. <https://doi.org/10.1093/genetics/158.2.779>
- Kijas, J.M.H., Wales, R., Törnsten, A., Chardon, P., Moller, M., Andersson, L., 1998. Melanocortin receptor 1 (MC1R) mutations and coat color in pigs. *Genetics* 150(3), 1177–1185.
- Kim, K.S., Larsen, N., Short, T., Plastow, G., Rothschild, M.F., 2000. A missense variant of the porcine melanocortin-4 receptor (MC4R) gene is associated with fatness, growth, and feed intake traits. *Mammalian Genome* 11(2), 131–135. <https://doi.org/10.1007/s003350010025>
- Kim, K.S., Reecy, J.M., Hsu, W.H., Anderson, L.L., Rothschild, M.F., 2004. Functional and phylogenetic analyses of a melanocortin-4 receptor mutation in domestic pigs. *Domestic Animal Endocrinology* 26(1), 75–86. <https://doi.org/10.1016/j.domaniend.2003.12.001>
- Kirk, H., Freeland, J.R., 2011. Applications and implications of neutral versus non-neutral markers in molecular ecology. *International Journal of Molecular Sciences* 12(6), 3966–3988. <https://doi.org/10.3390/ijms12063966>
- Kommadath, A., Bao, H., Choi, I., Reecy, J.M., Koltes, J.E., Fritz-Waters, E., Easley, C.J., Grant, J.R., Rowland, R.R.R., Tuggle, C.K., Dekkers, J.C.M., Lunney, J.K., Guan, L.L., Stothard, P., Plastow, G.S., 2017. Genetic architecture of gene expression underlying variation in host response to porcine reproductive and respiratory syndrome virus infection. *Scientific Reports* 7(April), 1–11. <https://doi.org/10.1038/srep46203>
- Kušec, G., Dovč, P., Karolyi, D., Čandek Potokar, M., 2015. Local pig breeds and pork products in Croatia and Slovenia – Unexploited treasure. *Poljoprivreda* 21(1), 16–21. <https://doi.org/10.18047/poljo.21.1.sup.3>
- Latorre, P., Burgos, C., Hidalgo, J., Varona, L., Carrodeguas, J.A., López-Buesa, P., 2016. C.A2456C-substitution in Pck1 changes the enzyme kinetic and functional properties modifying fat distribution in pigs. *Scientific Reports* (6), 19617. <https://doi.org/10.1038/srep19617>
- Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv, 1303.3997v1 [q – bio.GN]. <http://arxiv.org/abs/1303.3997>
- Lin, Z., Lou, Y., Peacock, J., Squires, E.J., 2005. A novel polymorphism in the 5' untranslated region of the porcine cytochrome b5 (CYB5) gene is associated with decreased fat androstenone level. *Mammalian Genome* 16(5), 367–373. <https://doi.org/10.1007/s00335-004-2439-4>
- Lin, Z., Lou, Y., Squires, E.J., 2006. Functional polymorphism in porcine CYP2E1 gene: Its association with skatole levels. *Journal of Steroid Biochemistry and Molecular Biology* 99(4–5), 231–237. <https://doi.org/10.1016/j.jsbmb.2005.07.001>
- Ma, J., Yang, J., Zhou, L., Ren, J., Liu, X., Zhang, H., Yang, B., Zhang, Z., Ma, H., Xie, X., Xing, Y., Guo, Y., Huang, L., 2014. A Splice Mutation in the PHKG1 Gene Causes High Glycogen Content and Low Meat Quality in Pig Skeletal Muscle. *PLoS Genetics* 10(10). <https://doi.org/10.1371/journal.pgen.1004710>
- McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., Flicek, P., Cunningham, F., 2016 et al. The Ensembl Variant Effect Predictor. *Genome Biol* 17, 122. <https://doi.org/10.1186/s13059-016-0974-4>
- Meyers, S.N., & Beever, J.E., 2008. Investigating the genetic basis of pork tenderness: Genomic analysis of porcine CAST. *Animal Genetics* 39(5), 531–543. <https://doi.org/10.1111/j.1365-2052.2008.01765.x>
- Mikawa, S., Morozumi, T., Shimanuki, S.I., Hayashi, T., Uenishi, H., Domukai, M., Okumura, N., Awata, T., 2007. Fine mapping of a swine quantitative trait locus for number of vertebrae and analysis of an orphan nuclear receptor, germ cell nuclear factor (NR6A1). *Genome Research* 17(5), 586–593. <https://doi.org/10.1101/gr.6085507>
- Milan, D., Jeon, J.T., Looft, C., Amarger, V., Robic, A., Thelander, M., Rogel-Gaillard, C., Paul, S., Iannucci, N., Rask, L., Ronne, H., Lundström, K., Reinsch, N., Gellin, J., Kalm, E., Le Roy, P., Chardon, P., Andersson, L. 2000. A mutation in PRKAG3 associated with excess glycogen content in pig skeletal muscle. *Science* 288(5469), 1248–1251. <https://doi.org/10.1126/science.288.5469.1248>
- Muñoz, G., Alves, E., Fernández, A., Óvilo, C., Barragán, C., Estellé, J., Quintanilla, R., Folch, J. M., Silió, L., Rodríguez, M. C., Fernández, A.I., 2007. QTL detection on porcine chromosome 12 for fatty-acid composition and association analyses of the fatty acid synthase, gastric inhibitory polypeptide and acetyl-coenzyme A carboxylase alpha genes. *Animal Genetics* 38(6), 639–646. <https://doi.org/10.1111/j.1365-2052.2007.01668.x>
- Muñoz, M., Bozzi, R., García, F., Núñez, Y., Geraci, C., Crovetto, A., García-Casco, J., Alves, E., Škrlep, M., Charneca, R., Martins, J.M., Quintanilla, R., Tibau, J., Kušec, G., Djurkin-Kušec, I., Mercat, M.J., Riquet, J., Estellé, J., Zimmer, C., ... Óvilo, C., 2018. Diversity across major and candidate genes in European local pig breeds. *PLoS ONE* 13(11), 1–30. <https://doi.org/10.1371/journal.pone.0207475>
- Muñoz, M., Fernández, A.I., Benítez, R., Pena, R.N., Folch, J.M., Rodríguez, M. del C., Silió, L., Alves, E., 2013. Disentangling Two QTL on Porcine Chromosome 12 for Backfat Fatty Acid Composition. *Animal Biotechnology* 24(3), 168–186. <https://doi.org/10.1080/10495398.2012.763130>
- Ocampo-Gallego, R.J., 2019. Análisis de diversidad genética en cerdo criollo san pedroño utilizando datos de pedigrí. *Ecosistemas y Recursos Agropecuarios* 6(17), 333. <https://doi.org/10.19136/era.a6n17.2049>
- Ocampo-Gallego, R.J., Abuabara-Pérez, J.Y. 2021. Evaluation of the genetic diversity of the Zungo creole pig by pedigree analysis [Evaluación de la diversidad genética del cerdo criollo Zungo por análisis de pedigrí]. *Livestock Research for Rural Development* 33(5). <http://www.lrrd.org/lrrd33/5/337Irocam.html>
- Oslinger, A., Muñoz, J.E., Álvarez, L.Á., Ariza, F., Moreno, F., Posso, A., 2006. Characterization of Colombian creole pigs by RAMS Caracterización de cerdos criollos colombianos mediante la

- técnica molecular RAMs. *Acta Agronómica*, 55(4), 45–52. [https://revistas.unal.edu.co/index.php/acta\\_agronomica/articulo/view/479](https://revistas.unal.edu.co/index.php/acta_agronomica/articulo/view/479)
- Óvilo, C., Fernández, A., Noguera, J.L., Barragán, C., Letón, R., Rodríguez, C., Mercadé, A., Alves, E., Folch, J.M., Varona, L., Toro, M., 2005. Fine mapping of porcine chromosome 6 QTL and LEPR effects on body composition in multiple generations of an Iberian by Landrace intercross. *Genetical Research* 85(1), 57–67. <https://doi.org/10.1017/S0016672305007330>
- Óvilo, A., Fernández, A.I., Folch, J.M., Varona, L., Benítez, R., Nuñez, Y., Rodríguez, C., Silió, L., 2010. Hypothalamic expression of porcine leptin receptor (LEPR), neuropeptide y (NPY), and cocaine- and amphetamine-regulated transcript (CART) genes is influenced by LEPR genotype. *Mammalian Genome* 21(11–12), 583–591. <https://doi.org/10.1007/s00335-010-9307-1>
- Pardo, C., 2016. Evaluación de la calidad de la carne y caracterización de genes asociados a la calidad de tres razas de cerdos criollos colombianos. Universidad Nacional de Colombia.
- Peacock, J., Lou, Y., Lundström, K., Squires, E.J., 2008. The effect of a c.-8G>T polymorphism on the expression of cytochrome b5A and boar taint in pigs. *Animal Genetics* 39(1), 15–21. <https://doi.org/10.1111/j.1365-2052.2007.01674.x>
- Poplin, R., Ruano-Rubio, V., DePristo, M.A., Fennell, T.J., Carneiro, M.O., Van der Auwera, G.A., Kling, D.E., Gauthier, L.D., Levy-Moonshine, A., Roazen, D., Shakir, K., Thibault, J., Chandran, S., Whelan, C., Lek, M., Gabriel, S., Daly, M. J., Neale, B., MacArthur, D.G., Banks, E., 2017. Scaling accurate genetic variant discovery to tens of thousands of samples. *BioRxiv* 1–22. <https://doi.org/10.1101/201178>
- Ren, J., Duan, Y., Qiao, R., Yao, F., Zhang, Z., Yang, B., Guo, Y., Xiao, S., Wei, R., Ouyang, Z., Ding, N., Ai, H., Huang, L., 2011. A missense mutation in PPARD causes a major QTL effect on ear size in pigs. *PLoS Genetics* 7(5). <https://doi.org/10.1371/journal.pgen.1002043>
- Ribani, A., Bertolini, F., Schiavo, G., Scotti, E., Utzeri, V.J., Dall'Olio, S., Trevisi, P., Bosi, P., Fontanesi, L., 2017. Next generation semiconductor based sequencing of bitter taste receptor genes in different pig populations and association analysis using a selective DNA pool-seq approach. *Animal Genetics* 48(1), 97–102. <https://doi.org/10.1111/age.12472>
- Roberts, M.C., Mickelson, J.R., Patterson, E.E., Nelson, T.E., Jane Armstrong, P., Brunson, D.B., Hogan, K., 2001. Autosomal dominant canine malignant hyperthermia is caused by a mutation in the gene encoding the skeletal muscle calcium release channel (RYR1). *Anesthesiology* 95(3), 716–725. <https://doi.org/10.1097/0000542-200109000-00026>
- Ros-Freixedes, R., Gol, S., Pena, R.N., Tor, M., Ibañez-Escriche, N., Dekkers, J.C.M., Estany, J., 2016. Genome-wide association study singles out SCD and LEPR as the two main loci influencing intramuscular fat content and fatty acid composition in Duroc pigs. *PLoS One* 11(3), 1–18. <https://doi.org/10.1371/journal.pone.0152496>
- Ros-Freixedes, R., Valente, B.D., Chen, C.Y., Herring, W.O., Gorjanc, G., Hickey, J.M., Johnsson, M., 2022. Rare and population-specific functional variation across pig lines. *Genet Sel Evol* 54(39). <https://doi.org/10.1186/s12711-022-00732-8>
- Schroyen, M., Stinckens, A., Verhelst, R., Niewold, T., Buys, N., 2012. The search for the gene mutations underlying enterotoxigenic Escherichia coli F4ab/ac susceptibility in pigs: A review. *Veterinary Research* 43(1), 1–8. <https://doi.org/10.1186/1297-9716-43-70>
- Solé, E., Ros-Freixedes, R., Tor, M., Reixach, J., Pena, R.N., Estany, J., 2021. Antagonistic maternal and direct effects of the leptin receptor gene on body weight in pigs. *PLoS ONE* 16, 1–12. <https://doi.org/10.1371/journal.pone.0246198>
- Suárez-Mesa, R., Estany, J., Rondón-Barragán, I., 2021. Semen quality of Colombian Creole as compared to commercial pig breeds. *Tropical Animal Health and Production* 53(1), 129. <https://doi.org/10.1007/s11250-021-02557-x>
- Tu, P.A., Lo, L.L., Chen, Y.C., Hsu, C.C., Shiau, J.W., Lin, E.C., Lin, R.S., Wang, P.H., 2014. Polymorphisms in the promoter region of myostatin gene are associated with carcass traits in pigs. *Journal of Animal Breeding and Genetics* 131(2), 116–122. <https://doi.org/10.1111/jbg.12053>
- Vargas, J.C., Velásquez, F.J., Chacón, E., 2016. Estructura genética y caracterización molecular del cerdo criollo (sus scrofa domestica) de Ecuador. Utilizando marcadores microsatélites. *Acta Agronómica* 65(3), 292–297. <https://doi.org/10.15446/acag.v65n3.45661>
- Wang, S.J., Liu, W.J., Yang, L.G., Sargent, C.A., Liu, H.B., Wang, C., Liu, X.D., Zhao, S.H., Affara, N.A., Liang, A.X., Zhang, S.J., 2012. Effects of FUT1 gene mutation on resistance to infectious disease. *Molecular Biology Reports* 39(3), 2805–2810. <https://doi.org/10.1007/s11033-011-1039-0>
- Wimmers, K., Lin, C.L., Tholen, E., Jennen, D.G.J., Schellander, K., Ponsuksili, S., 2005. Polymorphisms in candidate genes as markers for sperm quality and boar fertility. *Animal Genetics* 36(2), 152–155. <https://doi.org/10.1111/j.1365-2052.2005.01267.x>
- Zhang, C.Y., Wang, Z., Bruce, H. L., Janz, J., Goddard, E., Moore, S., Plastow, G.S., 2014. Associations between single nucleotide polymorphisms in 33 candidate genes and meat quality traits in commercial pigs. *Animal Genetics* 45(4), 508–516. <https://doi.org/10.1111/age.12155>
- Zhang, Q., Huang, R., Ma, X., Jiang, N., Zhou, W., Li, P., 2020. Association of Rs339939442 in the AHR Gene with Litter Size are Inconsistent among Chinese Indigenous Pigs and Western Commercial Pigs. *Animals* 10, 11. <https://doi.org/10.3390/ani10010011>
- Zhu, Y., Li, W., Yang, B., Zhang, Z., Ai, H., Ren, J., Huang, L., 2017. Signatures of selection and interspecies introgression in the genome of Chinese domestic pigs. *Genome Biology and Evolution* 9(10), 2592–2603. <https://doi.org/10.1093/gbe/evx186>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.