

## Exploring the genetic origin of Brazilian locally adapted breeds: Admixture, population history and relationship with Portuguese and indicine cattle

Lucas Lima Verardo<sup>a</sup>, Pamela Itajara Otto<sup>b</sup>, Marco Antonio Machado<sup>c</sup>, João Cláudio do Carmo Panetto<sup>c</sup>, Daniele Ribeiro de Lima Reis Faza<sup>c</sup>, Andréa Alves do Egito<sup>d</sup>, Andreia Sofia Martins Vitorino<sup>e,f</sup>, Maria Inês Carvalho Martins Carolino<sup>e,g</sup>, Nuno Pimentel Carolino<sup>e,f,g</sup>, Marcos Vinicius Gualberto Barbosa da Silva<sup>c,\*</sup>

<sup>a</sup> Universidade Federal dos Vales do Jequitinhonha e Mucuri - Campus (II) JK - Rodovia BR 367, Km 583, nº 5000, Alto da Jacuba, Diamantina - MG, 39100-000, Brazil

<sup>b</sup> Universidade Federal Santa Maria - Av. Roraima nº 1000 - Cidade Universitária, Camobi, Santa Maria - RS, 97105-900, Brazil

<sup>c</sup> Embrapa Gado de Leite - Rua Eugênio do Nascimento, 610 - Dom Bosco, Juiz de Fora - MG, 36038-330, Brazil

<sup>d</sup> Embrapa Gado de Corte - Av. Rádio Maia, 830 - Vila Popular, Campo Grande - MS, 79106-550

<sup>e</sup> Instituto de Nacional de Investigação Agrária e Veterinária - Estação Zootécnica Nacional - Polo de Investigação da Fonte Boa, 2005-048, Vale de Santarém, Santarém, Portugal

<sup>f</sup> CIISA - Centre for Interdisciplinary Research in Animal Health, Faculty of Veterinary Medicine, University of Lisbon, Av. Universidade Técnica, 1300-477 Lisboa, Portugal

<sup>g</sup> Escola Universitária Vasco da Gama, Av. José R. Sousa Fernandes 197 Lordemão, 3020-210 Coimbra, Portugal

### HIGHLIGHTS

- Portuguese Iberian and Brazilian Iberian-derived locally adapted breeds studied.
- Genetic architecture of Portuguese iberian cattle were observed.
- Portuguese breeds appear to have influenced Brazilian locally adapted population.
- Indicine breeds also seem to have influenced Brazilian locally adapted population.
- Brazilian locally adapted and Portuguese breeds as an enormous genetic reservoir.

### ARTICLE INFO

#### Keywords:

Admixture population  
SNP  
Candidate genes  
Genomic

### ABSTRACT

The domestication of bovines has given rise to distinct phenotypes resulting in diverse breeds worldwide. Subsequent artificial selection has further enriched the complexity and diversity to cattle sub-species adapting them to their various purposes and evolving environment. Among these cattle sub-species, native Portuguese cattle breeds are considered a repository of biodiversity shaped by genetic and environmental effects accumulated over the years. In Brazil, taurine animals were introduced between the 16th and 17th centuries by the European conquerors to be used for food, leather, and animal traction. Through generations of casual crossings in diverse ecosystems, these animals became adapted to a wide range of environments and displaying varying levels of phenotypic variability and improved fitness to local conditions, hereafter recognized as Iberian-derived Locally Adapted breeds. Thus, we focused in assess 13 cattle breeds representing two geographic origins: Portuguese Iberian cattle and Brazilian Iberian-derived Locally Adapted breeds. The aim is to understand the genetic architecture of Brazilian Iberian-derived Locally Adapted breeds and their potential relationship with Portuguese Iberian cattle and Brazilian indicine breeds. In our study, we observed varying extents of linkage disequilibrium across the evaluated breeds, along with differences in effective population size. Runs of homozygosity analyses revealed different genes associated with common traits, even within a cluster group. For instance, genes related to immune response, such as *MAVS* (Mertolenga), *DTX* (Barrosã), *ZBTB16* (Brava de Lide), *DUSP22* (Pantaneiro), and *IL7R* (Caracu Dairy and Caracu Beef). Overall, our results highlight that both populations studied, with their specificities, are crucial sources of animal genetic resource of animal genetic resources for food and agriculture.

\* Correspondent author.

E-mail address: [marcos.vb.silva@embrapa.br](mailto:marcos.vb.silva@embrapa.br) (M.V.G.B. da Silva).

<https://doi.org/10.1016/j.livsci.2024.105455>

Received 22 September 2023; Received in revised form 14 February 2024; Accepted 15 March 2024

Available online 16 March 2024

1871-1413/© 2024 Elsevier B.V. All rights reserved.

While consanguineous matings are somehow controlled in both Portuguese Iberian and Brazilian Iberian-derived Locally Adapted breeds, our analysis of genetic variability suggests that ongoing precautions are necessary in the management of these populations.

## 1. Introduction

The process of bovine domestication has given rise to distinct phenotypes resulting in various breeds worldwide. According to mtDNA data, there is evidence suggesting that differentiation between taurine and indicine cattle sub-species occurred approximately 10,000 years ago, predating the formal domestication process (Loftus et al., 1994). Taurine animals are believed to have originated in the Fertile Crescent encompassing modern-day Iraq, Syria, Lebanon, Israel, Palestine, Jordan, the Northeast Nile Valley regions of Egypt, the Southeast region of Turkey, and the Western fringes of Iran. In contrast, indicine cattle are thought to have emerged in the Indus Valley, covering Northeast Afghanistan, much of Pakistan, and Western and North-western India (Ajmone-Marsan et al., 2010). Distinct traits that defined domesticated cattle species became evident as domestication progressed, with artificial selection introducing even more complexity and diversity to cattle sub-species based on their intended uses and evolving environments.

Considering this, Portuguese native cattle breeds undoubtedly represent a repository of biodiversity shaped by a myriad of factors accumulated over the years. These factors include geographical, edaphoclimatic, land-related, historical, cultural, and political-economic influences, among others. Iberian cattle breeds exhibit remarkable phenotypic diversity, encompassing variations in shape, profile, body proportions, coat, and function, as extensively described by Felius et al. (2014). This diversity extends to coat colors and horn morphology, with these breeds primarily utilized for beef, work, and even fighting purposes. The origin and establishment of these cattle breeds have been investigated using microsatellite markers (Cañón et al. 2001; Beja-Pereira et al. 2003; Mateus et al. 2004; Cymbron et al. 2005; Ginja et al., 2009), mitochondrial DNA (mtDNA) (Cymbron et al. 1999; Ginja et al., 2019), chromosome sequence variation (Ginja et al. 2009) and, more recently, SNP arrays (Upadhyay et al., 2017; da Fonseca et al., 2019). Historical records and the evolution of Portuguese cattle farming suggest the influence of African and Central European cattle. In the mid-20th century, with the implementation of the first Herd Books and shift towards meat production, the current Portuguese native bovine populations became properly organized. Since the last decade of the 20th century, these breeds have been included into conservation or genetic improvement programs based on herd size and priorities defined by breeders for each breed.

In Brazil, taurine animals were introduced between the 16th and 17th centuries by the European conquerors. Their introduction aimed at various purposes, such as food supply, leather for clothes, and animal traction. Over many generations of casual crossings in diverse and variable ecosystems throughout the country, these cattle adapted to a wide range of environments. As a result, they exhibited diverse levels of phenotypic variability and improved fitness to local conditions, subsequently recognized as Iberian-derived Locally Adapted breeds (Egito et al., 2007; McManus et al., 2009). Distinct breeds, including Caracu (comprising both Caracu Caldeano lineage for dairy purposes and Caracu for beef), Crioulo Lageano (and a distinct lineage raised in Rio Grande do Sul State known as Franqueiro), Curraleiro Pé-Duro, Pantaneiro, and Mocho Nacional, were established. These breeds also appear to have influence of indicine cattle (Verardo et al., 2021). The introduction of more productive breeds from the early 20th century posed a threat to these native breeds, pushing them towards extinction. To prevent the loss of this valuable genetic material, the Brazilian Agricultural Research Corporation (Embrapa) included these populations in its genetic material conservation programs since 1983 (Mariante and Egito, 2004). The history of a population and its

demography, breeding system, and patterns of geographic subdivision can be reflected in both runs of homozygosity (ROH) and linkage disequilibrium (LD) patterns across the genome. ROH are defined as continuous homozygous segments of DNA (Gibson et al., 2006), while LD is a non-random association between alleles at different loci (Ardlie et al., 2002). The calculation of these parameters is essential for performing genomic selection (GS) and genome-wide association studies (GWAS). Studies have demonstrated that LD patterns among populations affect genome association studies and estimated breeding values (EBV) (Meuwissen et al., 2001; Marigorta et al., 2013; Hayes et al., 2009; Veroneze et al., 2014). Additionally, Luan et al. (2014) suggested that predicting genomic EBV based on ROH provided better genomic EBV prediction accuracy for simulated data compared to predictions based on linkage analysis relationships. It is also possible to identify islands of homozygosity shared among animals, which may result from selection proceedings using ROH (Zhang et al., 2015; Verardo et al., 2021). Therefore, studying LD and ROH patterns in bovine breeds is crucial for designing efficient genomic selection or conservation programs. Recently, from genome wide analyses of 19 breeds raised in Brazil, a possible genetic influence of indicine cattle (ranging from 7 to 23 %) in the Brazilian Iberian-derived Locally Adapted breeds was observed (Verardo et al., 2021). However, more than 77 % of the genetic architecture remains to be explored and understood in terms of its origin.

In this study, we assessed 13 cattle breeds representing two geographic origins, Portuguese Iberian cattle and Brazilian Iberian-derived Locally Adapted breeds. The primary aim was to understand the genetic architecture of Brazilian Iberian-derived Locally Adapted breeds and their potential relationship with Portuguese Iberian cattle, in addition to the observed relationship with indicine cattle. The specific objectives of this study were as follows: (i) to compare the overall LD decay patterns among these different populations; (ii) to estimate the effective population size ( $N_e$ ) of each population across generations; (iii) to evaluate ROH across populations; (iv) to calculate the inbreeding coefficient from ROH; and (v) to assess the breed composition across all populations.

## 2. Materials and methods

### 2.1. Data

The Iberian cattle breeds of Portugal included in this study (Cachena - CAC; Mertolenga - MER; Barrosã - BAR; Marinhova - MAR; Alentejana - ALE and Brava de Lide - BRA) are representative of the three strains of cattle formed in the Westernmost region of the Iberian Peninsula, as classified by Miranda do Vale (1949). This classification is based on ethnic traits and geographical distribution, distinguishing Mauritanian or *Bos taurus mauritanicus* (Barrosã and Cachena), Iberian or *Bos taurus ibericus* (Brava de Lide and Marinhova) and Aquitanian or *Bos taurus aquitanicus* (Alentejana and Mertolenga). The samples of Brazilian Locally Adapted breeds (Crioulo Lageano - CRI, Curraleiro Pé-Duro - CUR, Mocho Nacional - MOC, Pantaneiro - PAN, Franqueiro - FRA, Caracu Caldeano selected for milk - CCD and Caracu selected for beef - CCB) used in this study were obtained from Embrapa's DNA banks, originated from private breeders and conservation nuclei. Although the recommendation by the Commission on Genetic Resources for Food and Agriculture (FAO, 2023) suggests analyzing at least 100 animals per breed for genetic variation, it is important to note that all animals in this study were selected based on their representativeness within each breed. This approach was adopted to prevent bias resulting from the use of a

limited number of families (lineages) in the population, as also highlighted by Verardo et al. (2021). Portuguese samples were genotyped using the Affymetrix commercial genotyping array (60k), while Brazilian Locally Adapted breeds were genotyped using the Illumina Bovine HD Chip (777k SNPs). To maintain a uniform set of markers across the studied breeds, a standardized set of Single Nucleotide Polymorphism (SNPs) based on HD was generated, given that Portuguese breeds were not genotyped with the Bovine HD Chip (777k SNPs), which contains a higher number of markers. SNPs from sexual and unknown chromosomes were excluded from the analysis. Additionally, genotypes were excluded when the call rate was  $< 0.95$  and minor allele frequency  $< 0.01$ . In addition, samples showing overall call rates  $< 0.90$  were excluded from further analysis. After quality control, a total of 933 animals from Portuguese breeds and 258 from Brazilian Locally Adapted breeds were retained for the analyses (Table 1).

## 2.2. Linkage disequilibrium and its decay over time

To assess the extent of linkage disequilibrium (LD), we employed the  $r^2$  statistics, a robust not sensitive to gene frequency and effective population size (Terwilliger et al., 2002; Zhao et al., 2007). Calculations were performed using PLINK v1.07 (Purcell et al., 2007) and pair-wise linkage disequilibrium between SNPs was computed with a genomic distance of 1 Mb utilizing the following command: `./plink -bfile filename -cow -ld-window 5 -ld-window-kb 1000 -ld-window-r2 0 -out outname -r2`.

Parametric nonlinear regression models have been employed to investigate the physical distance between markers, commonly referred to as LD decay. This methodology has been reported in studies by Heifetz et al. (2005), Amaral et al. (2008), Abasht et al. (2009) and Veroneze et al. (2014). In this context, the pairwise  $r_{ij}^2$  from each population were regressed against the distance between the marker pairs using a nonlinear model described by Sved (1971):

$$LD_{ijk} = \frac{1}{(1 + 4\beta_k d_{ij})e_{ijk} + 1}, \quad (1)$$

where  $LD_{ijk}$  was the observed  $r_{ij}^2$  between SNPs  $i$  and  $j$  in breed  $k$ ;  $d_{ij}$  was the distance in Kb (kilo-base pair) between SNPs  $i$  and  $j$ ;  $\beta_k$  was the coefficient that describes the LD decay with distance for breed  $k$ , and  $e_{ijk}$  was the random error. In summary, according to Sved (1971), smaller  $\beta_k$  estimates indicate a higher extent of LD.

An identity test for the nonlinear regression model (Regazzi and

**Table 1**

Number of animals by breed, number of SNPs after quality control, and number of animals by Portuguese Iberian and Brazilian Iberian-derived Locally Adapted breed origin.

Breed origin	Breed <sup>1</sup>	Animals	Markers	Sample size
Portuguese Iberian	ALE	191	34,431	933
	BAR	96	34,997	
	BRA	94	33,482	
	CAC	96	35,688	
	MAR	82	34,701	
	MER	374	35,734	
Brazilian Iberian-derived Locally Adapted	CCD	55	33,023	258
	CCB	24	35,612	
	CRI	61	37,036	
	CUR	17	31,120	
	FRA	55	35,146	
	MOC	28	36,527	
	PAN	18	32,778	
			Total	

<sup>1</sup> ALE: Alentejana; BAR: Barrosã; BRA: Brava de Lide; CAC: Cachena; MAR: Marinhoa; MER: Mertolenga; CCD: Caracu Dairy; CCB: Caracu Beef; CRI: Crioulo Lageano; CUR: Curraleiro Pé-Duro; FRA: Franqueiro; MOC: Mocho Nacional; PAN: Pantaneiro.

Silva, 2010) was employed to compare the nonlinear curves among breeds. This test allows for a statistical comparison of the LD decay parameter ( $\hat{\beta}$ ) and facilitates the visualization of clusters. Clustering was conducted using on *NbClust* package v. 3.0 (Charrad et al., 2014) in the software R (R Core Team). Groups were identified based on Euclidian distance using the Ward aggregation method, selecting the best grouping according to the maximum and minimum differences between and within group indices. Based on clustering values, a dendrogram was generated using the *hclust* function on R software v. 4.0.0 (R Core Team) to highlight similarities among the contrasting groups. The nonlinear models were fitted using the *nls* function in R software (R Core Team).

## 2.3. Effective population size

The effective population size ( $N_e$ ) in the corresponding generation (T) was estimated based on the associated estimated average of  $r^2$ . The estimate of  $N_e$  across generations followed the formula:

$$N_e = \left(\frac{1}{4c}\right) \left(\frac{1}{r^2} - 1\right), \quad (2)$$

where  $c$  is the marker distance in Morgans, assuming 1 Mb = 1 cM, and  $T = 1/2c$  (Hayes et al., 2003; Meuwissen and Goddard, 2007). From the estimated  $r^2$ , the average values were grouped into four ranges of distances ( $< 0.01$ Mb, 0.011–0.03Mb, 0.031–0.05Mb and  $> 0.05$ Mb) and used to determine the  $N_e$  in 50, 16.6, 10 and 5 generations, respectively.

## 2.4. Runs of homozygosity (ROH)

We employed the Runs of Homozygosity (ROH) approach, as outlined by Peripolli et al. (2017), utilizing the software PLINK v1.07 (Purcell et al., 2007). The parameters for each population were set as follows: minimum window length of 120 SNPs, maximum gap size between two SNPs of 1,000 kb, minimum ROH length of 1,000 kb, minimum number of potential marked SNPs of 50, one heterozygote allowed per window, maximum of five missing calls per window, sliding window length of 50 SNPs and proportion of overlapping windows that must be homozygous  $> 0.05$ .

Additionally, to identify ROH islands throughout the genome, the most observed homozygous segments shared by individuals in each breed were considered. The “-homozyg-group” function implemented in PLINK was utilized to assess ROH islands shared among individuals. GenBank annotation based on the ARS-UCD1.2 assembly of the bovine genome was then used to identify genes in the ROH regions. To further understand the biological processes associated with the identified gene sets for each breed group (Brazilian Iberian-derived Locally Adapted and Portuguese Iberian), gene networks were generated using the ClueGO plugin for Cytoscape (Bindea et al., 2009). ClueGO employs a two-sided hypergeometric test and Bonferroni correction to associate one or more sets of genes with gene ontology (GO) terms or pathways, establishing edges between each gene and the chosen GO term. This process allowed the creation of gene networks highlighting biological roles, enabling the comparison of gene clusters by visualizing their functional differences or similarities.

## 2.5. Inbreeding coefficients from runs of homozygosity

The inbreeding coefficient based on the proportion of autosomes covered in runs of homozygosity ( $F_{ROH}$ ) per individual was determined according to McQuillan et al. (2008) using the following formula:

$$F_{ROH} = \frac{\sum_{j=1}^n L_{ROH_j}}{L_{total}}, \quad (3)$$

where  $L_{ROH_j}$  is the length of  $ROH_j$ , and  $L_{total}$  is the total size of the au-

tosomes covered by markers.  $L_{total}$  was taken to be 2550,380,751 bp, based on the consensus map. The detectRUNS package of R software was used to calculate and plot the distribution of inbreeding coefficients per group based on PLINK output .hom files.

## 2.6. Breed composition

The ADMIXTURE package (Alexander et al., 2009), which categorizes individuals into K-predefined clusters based on genotype data, was employed to assess hierarchical clustering across cattle breeds. In this phase, aiming to better understand the structure of Brazilian Iberian-derived Locally Adapted and Portuguese Iberian breeds, we also included African (Muturu - MUT; Ndama - NDA; Keteku - KTE and Yakanaji - YAK), Indicine (Gir - GIR and Nelore - NEL), and Holstein - HOL breeds. Consequently, we have two sets of populations: Population 1: Portuguese Iberian breeds (CAC, MER, BAR, MAR, ALE and BRA); Population 2: Portuguese Iberian, African, Holstein, Brazilian Iberian-derived Locally Adapted and Indicine breeds (CAC, MER, BAR, MAR, ALE, BRA, MUT, NDA, KTE, YAK, HOL, CCD, CCB, CRI, CUR, FRA, MOC, PAN, GIR and NEL).

The ADMIXTURE analysis was conducted with 10 random seeds, and the most representative run for each value of K was selected. Ancestry was estimated using a reduced panel (18,375; 18,682 and 4198 SNPs) pruned by LD between subsequent markers. Randomly selected animals comprised reduced groups of up to 82 and 35 animals per breed for Population 1 and 2, respectively. It is important to note that genotype quality control was also performed for African, Indicine and Holstein breeds. The two populations were assessed by computing the number of clusters, ranging K values from 2 to 5. To evaluate the phylogenetic relationships among cattle breeds within each population, the Euclidean distance between populations was evaluated using dartR package (Gruber et al., 2018) in R. Subsequently, a dendrogram tree was constructed using a neighbor-joining approach with APE program (Paradis et al., 2004). Principal component analysis (PCA) with the first and second principal components (PC), as well as the first, second and third PC was also conducted to investigate each population using dartR package (Paradis et al., 2004) in R.

## 3. Results

### 3.1. Linkage disequilibrium

In this study, the LD decay parameter ( $\hat{\beta}$ ) was estimated from the curves of  $r^2$  values over distances (bp) for each population using a nonlinear regression model (Table 2). Additionally, statistical tests for the equality of LD decay curves suggested that the overall pattern of the inflection point differed among breeds, particularly among Brazilian Iberian-derived Locally Adapted (Fig. S1). Since smaller values of  $\hat{\beta}$  indicate a higher extent of LD, they were utilized to cluster the different breeds based on these estimates (Fig. S2). Consequently, we were able to observe six defined clusters. CUR (Cluster 1) and CRI (Cluster 3) exhibited the highest and the smallest extent of LD, respectively. Among Portuguese breeds, CAC and MER (Cluster 4) demonstrated the smallest extent of LD values, while BRA (Cluster 2) showed the highest extent of LD values.

### 3.2. Effective population size

The effective population size ( $N_e$ ) was estimated across four generations classes (5, 10, 16.6 and 50) based on marker distances in the genome. It should be noted that the shorter the marker distance, the higher the estimated generation (Fig. S3). To assess the behavior of  $N_e$  across generations in each population, the difference between generations 50 and 5 was calculated (Table S1). Among Portuguese breeds, BAR displayed the most significant difference in  $N_e$  values, ranging from

**Table 2**

LD decay parameter estimates ( $\hat{\beta}$ ), standard-errors (SE) and p-values for the significance of  $\beta$  parameter of Sved model fitted to Brazilian Iberian-derived Locally Adapted and Portuguese Iberian cattle breeds.

Origin	Breed	$\hat{\beta}$	SE	p-value*
Portuguese Iberian	Cachena	$3.517 \times 10^{-5}$	$2.6136 \times 10^{-2}$	$<10^{-6}$
	Mertolenga	$3.493 \times 10^{-5}$	$2.6411 \times 10^{-2}$	$<10^{-6}$
	Barrosã	$3.071 \times 10^{-5}$	$3.0113 \times 10^{-2}$	$<10^{-6}$
	Marinhova	$2.926 \times 10^{-5}$	$3.1482 \times 10^{-2}$	$<10^{-6}$
	Alentejana	$2.629 \times 10^{-5}$	$3.5625 \times 10^{-2}$	$<10^{-6}$
Brazilian Iberian-derived Locally Adapted	Crioulo	$4.112 \times 10^{-5}$	$2.5178 \times 10^{-2}$	$<10^{-6}$
	Lageano			
	Mocho	$3.251 \times 10^{-5}$	$3.2063 \times 10^{-2}$	$<10^{-6}$
	Nacional			
	Caracu Beef	$2.779 \times 10^{-5}$	$3.5431 \times 10^{-2}$	$<10^{-6}$
	Franqueiro	$2.650 \times 10^{-5}$	$4.0041 \times 10^{-2}$	$<10^{-6}$
	Brava de Lide	$2.141 \times 10^{-5}$	$4.3693 \times 10^{-2}$	$<10^{-6}$
	Caracu Dairy	$2.035 \times 10^{-5}$	$4.7352 \times 10^{-2}$	$<10^{-6}$
	Pantaneiro	$1.886 \times 10^{-5}$	$5.1329 \times 10^{-2}$	$<10^{-6}$
	Curraleiro Pé Duro	$1.285 \times 10^{-5}$	$6.7890 \times 10^{-2}$	$<10^{-6}$

\*Significance associated to the significance of  $\hat{\beta}$  parameter from Sved model.

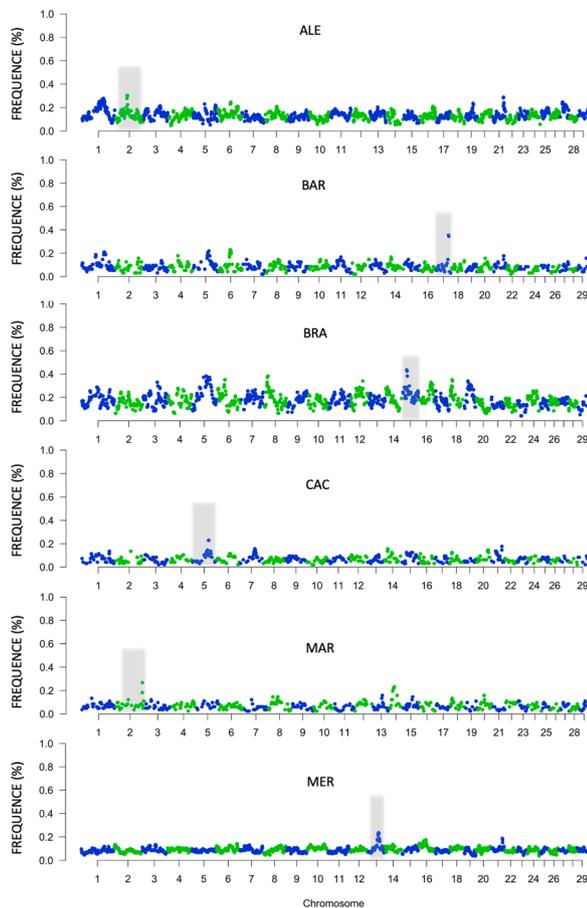
41.86 at generation 50 to 16.48 at generation 5, while BRA exhibited the smallest values, ranging from 29.19 to 12.62 at generation 50 and 5, respectively. In addition, Brazilian Iberian-derived Locally Adapted breeds showed the largest difference and lowest value of  $N_e$ , as observed in PAN and CUR, respectively.

### 3.3. Runs of homozygosity

The pattern of runs of homozygosity (ROH) varied across breeds (Fig. S4). Among Portuguese breeds, MER exhibited the highest percentage of long-range ROH ( $>31\text{Mb}$ ;  $>6\%$ ), while MAR, BAR and CAC showed higher percentages of short-range ROH ( $<5\text{Mb}$ ;  $>50\%$ ). In Brazilian Iberian-derived Locally Adapted breeds, CUR displayed the highest percentage of long-range ROH ( $>31\text{Mb}$ ;  $>10\%$ ), while PAN, CCB, and CRI showed higher percentages of short-range ROH ( $<5\text{Mb}$ ;  $>50\%$ ). Furthermore, we conducted for the most shared regions in homozygosity within each breed (Fig. 1 and Fig. 2), followed by gene identification (Table S2). Among the identified genes, the biological process network (Fig. S5) highlighted MAVS and CDC25B genes, both identified in the MER population, which are related to the immune system via regulation of IP-10 production and germ cell development, respectively. Similarly, in germ cell development, we observed the ZBTB16 gene from BRA, which shares immune-related biological processes with the DTX1 gene identified in the BAR breed. The BAR breed also showed the SDS gene related to amino acids metabolic process. In CAC breed, the FAR2 gene was found to be related with wax and fatty acids metabolic processes. Concerning the Brazilian Iberian-derived Locally Adapted breeds, we observed genes related also with the immune system (e.g. IRF4, EXOC2, IL7R, DUSP22, and SERPINB9), prolactin (PRLR) and genetic imprinting (KDM1B), as observed in Fig. S6 and S7.

### 3.4. Inbreeding coefficients from runs of homozygosity

Descriptive statistics of the inbreeding coefficient in the breeds are summarized in Table 3. Among the Brazilian Iberian-derived Locally Adapted breeds, the average  $F_{ROH}$  of CUR was the highest ( $F_{ROH} = 0.1492$ ), while the average  $F_{ROH}$  of CRI (0.0351) was the lowest. Considering the Portuguese Iberian breeds, CAC had the lowest average  $F_{ROH}$  (0.0476), while BRA was the highest, at 0.1540. At the individual level, individuals with the highest  $F_{ROH}$  appeared in the CUR breed (0.4539), which also exhibited more individuals with extreme values



**Fig. 1.** Manhattan plot of the distribution of runs of homozygosity (ROH) hotspots in the Portuguese Iberian cattle breeds genome. The x-axis represents the SNP genomic coordinate, and the y-axis shows the frequency (%) of overlapping ROH shared among individuals. Grey boxes highlight the most shared ROH in each population. CAC: Cachena; MER: Mertolenga; BAR: Barroã; MAR: Marinhoa; ALE: Alentejana; BRA: Brava de Lide.

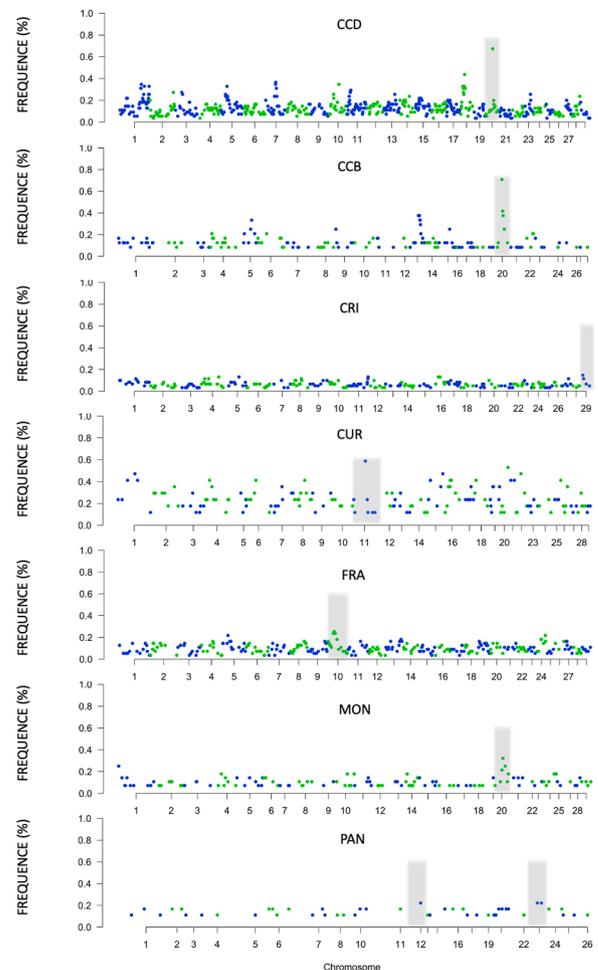
compared with other populations (Fig. S8).

### 3.5. Breed composition

The breed composition for each Portuguese breed was determined using the admixture analysis (Fig. S9 and Fig. S10). With  $K = 2$ , it was observed that BRA exhibited distinct genotypes (red; >90 %). At  $K = 3$  another genotype was discriminated (light green) in ALE and MER, while at  $K = 4$ , a third ancestry was observed in MAR (purple). Finally, at  $K = 5$ , the MER population displayed a distinct ancestry (yellow).

Furthermore, the breed composition for each population was evaluated when African (MUT, NDA, KTE and YAK), indicine (GIR and NEL), and Holstein breeds were included in the admixture analysis (Fig. S11 and Fig. 3). At  $K = 2$ , it was observed that GIR and NEL showed distinct genotypes (red; >95 %). At  $K = 3$  other genotypes were discriminated (light green) in African breeds, while at  $K = 4$ , a third ancestry was observed in Brazilian Iberian-derived Locally Adapted breeds (purple). Finally, at  $K = 5$ , the HOL population displayed a distinct ancestry (orange).

The neighbor-joining analysis of the Portuguese breeds indicated three distinct and representative groups: *Bos taurus mauritanicus* (CAC and BAR), *Bos taurus ibericus* (MAR and BRA), and *Bos taurus aquitanicus* (ALE and MER) (Fig. 4a). Comparing the breeds within the groups, it is possible to observe that the CAC and BAR breeds showed a higher relationship, as expected. When evaluating Portuguese Iberian, Brazilian Iberian-derived Locally Adapted, Indicine and Holstein breeds



**Fig. 2.** Manhattan plot of the distribution of runs of homozygosity (ROH) hotspots in the Brazilian Iberian-derived Locally Adapted cattle breeds genome. The x-axis represents the SNP genomic coordinate, and the y-axis shows the frequency (%) of overlapping ROH shared among individuals. Grey boxes represent the most shared ROH in each population. CCD: Caracu Dairy; CCB: Caracu Beef; CRI: Crioulo Lageano; CUR: Curraleiro Pé Duro; FRA: Franqueiro; MON: Mocho Nacional; PAN: Pantaneiro.

together, it was observed that the Portuguese breeds were more distantly related to Africans and Indicine breeds than the Brazilian Iberian-derived Locally Adapted, while the African and Indicine were more closely grouped (Fig. 4b).

In the PCA plot with PC1 and PC2 (Fig. 5a), the animals were classified into three groups (Group1: ALE, Group2: MER, and Group3: MAR, CAC, BAR and BRA), where ALE and MER were the breeds more distantly related. The inclusion of the PC3 allowed a better distinction of MAR, CAC, BAR and BRA breeds (Fig. 5b), showing that BRA animals were more distantly related to MAR, CAC and BAR. CAC and BAR breeds were more distantly related to MAR, CAC and BAR. CAC and BAR breeds were more closely grouped. Thus, Portuguese breeds were grouped in four different groups: Group1- ALE; Group2 - MER; Group3 - MAR, CAC and BAR; and Group4 - BRA.

The cattle groups were also clearly discriminated according to the history of development when the Portuguese Iberian, Brazilian Iberian-derived Locally Adapted, Indicine, and Holstein were evaluated in the PCA analyses. The GIR and NEL were clustered together, as well as Portuguese Iberian breeds. The Brazilian Iberian-derived Locally Adapted breeds were also well-grouped, while the African breeds were clustered into two distinct groups, one including YAK and KTE, and the second including MUT and NDA (Fig. 5c and 5d).

**Table 3**  
Descriptive statistics of inbreeding coefficients ( $F_{ROH}$ ) in Portuguese Iberian and Brazilian Iberian-derived Locally Adapted cattle breeds.

Group	Breed*	$F_{ROH}$			
		Median	Mean	SE	Range
Portuguese Iberian	ALE	0.1045	0.1166	$4.378 \times 10^{-3}$	0.0244 - 0.3283
	BAR	0.0432	0.0626	$6.610 \times 10^{-3}$	0.0126 - 0.4023
	BRA	0.1499	0.1540	$9.306 \times 10^{-3}$	0.0009 - 0.3902
	CAC	0.0273	0.0476	$5.887 \times 10^{-3}$	0.0021 - 0.3296
	MAR	0.0414	0.0519	$5.153 \times 10^{-3}$	0.0014 - 0.3163
	MER	0.0604	0.0808	$3.679 \times 10^{-3}$	0.0006 - 0.4279
Brazilian Iberian-derived Locally Adapted	CCD	0.0954	0.0914	$2.852 \times 10^{-2}$	0.0410 - 0.1324
	CCB	0.0437	0.0482	$5.124 \times 10^{-3}$	0.0077 - 0.1025
	CRI	0.0202	0.0351	$5.624 \times 10^{-3}$	0.0007 - 0.2800
	CUR	0.1486	0.1492	$3.445 \times 10^{-2}$	0.0023 - 0.4539
	FRA	0.048	0.0691	$8.044 \times 10^{-3}$	0.0013 - 0.2536
	MON	0.0379	0.0413	$6.492 \times 10^{-3}$	0.0009 - 0.1326
PAN	0.0181	0.0392	$7.095 \times 10^{-3}$	0.0043 - 0.1087	

\* CAC: Cachena; MER: Mertolenga; BAR: Barrosã; MAR: Marinhoa; ALE: Alentejana; BRA: Brava de Lide; CCD: Caracu Dairy; CCB: Caracu Beef; CRI: Crioulo Lageano; CUR: Curraleiro Pé Duro; FRA: Franqueiro; MON: Mocho Nacional; PAN: Pantaneiro.

## 4. Discussion

### 4.1. Linkage disequilibrium

The LD decay differed among breeds from each group of origin (Portuguese Iberian and Brazilian Iberian-derived Locally Adapted) based on LD curves parameter estimates ( $\hat{\beta}$ ) (Table 1). Among breeds, the pattern of LD diverged, and the resulting decay curves were displayed according to the group of origin (Fig. S1a and S1b). All Portuguese Iberian breeds followed a similar pattern of decreasing LD as the physical distance increased in the genome. Portuguese cattle breeds, in general, have a history of inbreeding, which can increase the levels of linkage disequilibrium in the population. This can make it more challenging to separate the effects of individual genes and make traditional selective breeding techniques less effective.

The Brazilian Iberian-derived Locally Adapted breeds showed higher variation among them compared to the Portuguese Iberian breeds. At short marker distances, CRI exhibited lower LD for autosomes compared to either Portuguese or Brazilian breeds, while CUR showed a higher LD. When the distance between markers was 25 kb, the average LD ( $r^2$ ) observed for CRI and CUR was 0.20 and 0.45, respectively, which is lower to the LD reported for indicine cattle (0.25) but equal to the value previously reported for taurine breeds in Australia (0.45) (Porto-Neto et al., 2014). The difference among Portuguese Iberian breeds is not as clear for markers separated by a larger physical distance (> 70 kb) where LD quickly approached background levels, and  $r^2$  values were ~ 0.10 in both studies. Nevertheless, we found that LD decay for the Brazilian breeds was, in general, slower than the Portuguese.

Population mating patterns and admixture can strongly influence LD. In general, LD decays more rapidly in outcrossed populations, and that might explain the result for CRI. In fact, Ribeiro (1993) reported experimental crossings with the Nellore and Charolais breeds in projects approved in 1984. In addition, Spritze et al. (2003), when evaluating

different generations of the CRI breed, mentioned that zootechnical data from the beginning of the 90's indicate that crossbred animals may have been kept in the herds. Studies with other molecular markers, in addition to historical reports, also demonstrate the possible introgression of indicine genes (Nellore) in the past, where these animals were sampled (Spritze et al. 2003; Egito et al. 2007). Also, Verardo et al. (2021), performing similar studies using 19 breeds raised in Brazil, observed that CRI breed was the Brazilian Iberian-derived Locally Adapted breed with a more variable genetic composition and LD decay partner similar as Canchim and Girolando breeds (*Bos taurus* x *Bos indicus* crossbreed).

Otherwise, Egito et al. (2007) reported that among the five Creole breeds investigated, the highest inbreeding was detected in CUR, and the reason was the very limited number of bulls available for this breed. Additionally, historical data showed the difficulty of exchanging animals between different herds, as well as the fact that this breed was only officially recognized by the Brazilian government in 2012, which prevented the commercialization of frozen semen. Moreover, the indicine introgression in CUR may be less pronounced. The LD results found in our research corroborate the fact that the observed LD in this breed is higher in relation to the other breeds. This might be due to less effective recombination in inbred breeds, where individuals are more likely to be homozygous, than in outbred populations.

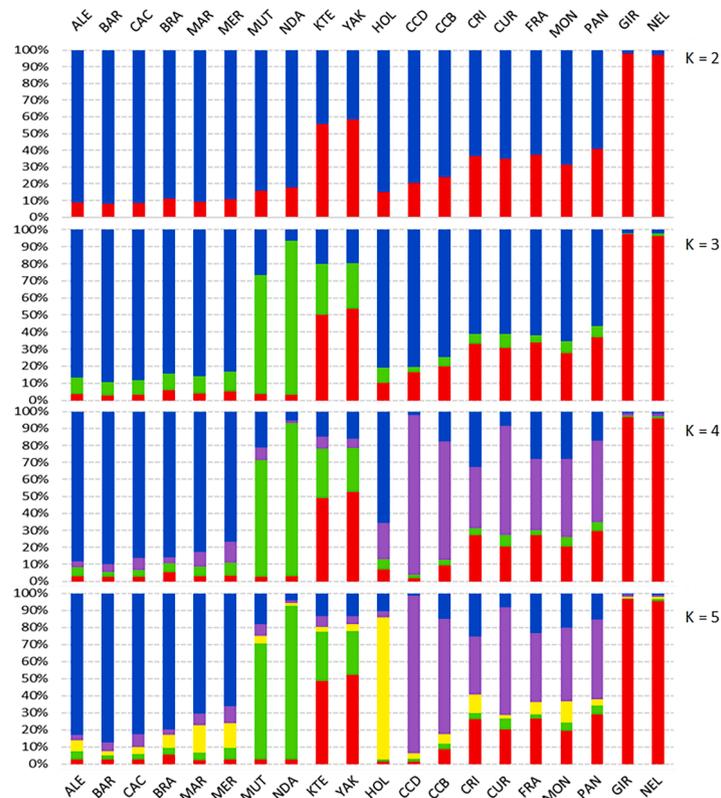
Many different authors working with distinct taurine and indicine breeds using the average  $r^2$  showed the same pattern of LD decrease (McKay et al., 2007; Lu et al., 2012; Zhu et al., 2013; O'Brien et al., 2014; Mokry et al., 2014; Toro Ospina et al., 2019; Xu et al., 2019), although they did not use a linear model to analyze the equality of the curves as performed in our study. As presented by Verardo et al. (2021), the use of a linear model allowed not only to verify the existence of breed differences but also to make a pairwise comparison across all breeds.

According to the comparison of LD decay among breeds, six clusters were identified (Fig. S2). In this current study, clustered breeds showed a similar LD decay pattern, suggesting that a similar marker density could be used in genomic studies for each one of these breeds. The LD extent provides insight into the number of SNPs required for genomic selection (GS) and genome-wide association studies (GWAS). However, this does not signify that the same markers are suitable for a particular breed cluster group since different set of markers may be segregating in each breed. For example, in cluster 2, we observed the presence of Brazilian and Portuguese breeds (CCD and BRA, respectively) highlighting the similarity of LD extent, although they may have different genetic backgrounds. Thus, the set of SNPs that could provide the best prediction accuracy at a given marker density should be different between CCD and BRA. We also observed different extents of LD between clusters 1 and 4, with higher LD observed in cluster 4 and lower in cluster 1. This information implies that different marker densities may be used for GS and GWAS for these breeds, which could also influence the accuracy of GS. However, in practical terms, it is important to state that a minimal marker density of 60 kb would be useful for all breeds, in addition to removing monomorphic SNPs.

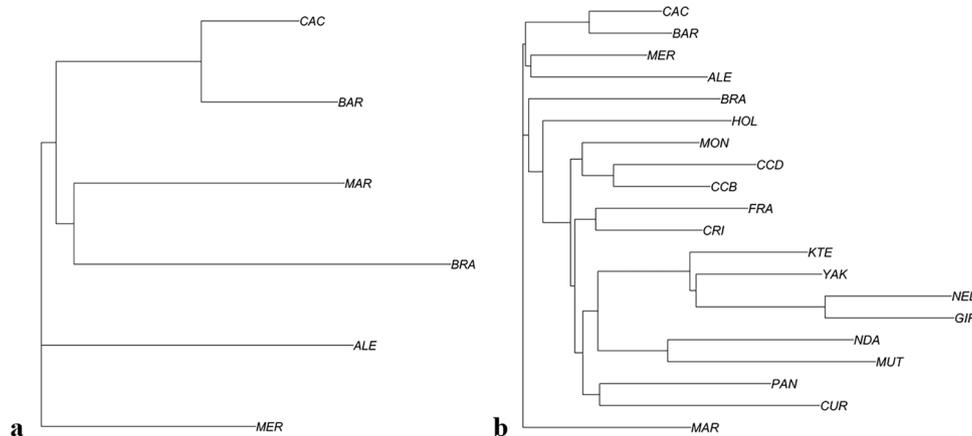
### 4.2. Effective population size

The effective population size in cattle refers to the number of individuals in a population that contribute to its genetic diversity and future generations. This number is often smaller than the actual population size due to factors such as inbreeding, unequal mating and reduced fertility. Effective population size is a crucial metric for the conservation and management of cattle populations, as it helps determine the rate of genetic drift and loss of diversity over time. According to the Commission on Genetic Resources for Food and Agriculture (FAO, 2022), the ideal population size for cattle is generally considered to be over 6000 individuals to avoid being considered at risk. However, an effective population size of about 50 animals per generation is accepted, leading to a rate of inbreeding of 1 % per generation (FAO, 1998).

Based on our results, we observe that in all Portuguese Iberian



**Fig. 3. Admixture analysis reflecting genetic clustering of twenty breeds grouped by Portuguese Iberian, African, Holstein, Brazilian Iberian-derived Locally Adapted and Indicine animals.** Individual vertical bars along the x-axis represent individual cattle which are grouped by breed. ADMIXTURE analysis using the first 5 K solutions for twenty populations. ALE: Alentejana; BAR: Barrosã; CAC: Cachena; BRA: Brava de Lide; MAR: Marinhoa; MER: Mertolenga; MUT: Muturu; NDA: Ndama; KTE: Keteku; YAK: Yakanaji; HOL: Holstein; CCD: Caracu Dairy; CCB: Caracu Beef; CRI: Crioulo Lageano; CUR: Curraleiro Pé-Duro; FRA: Franqueiro; MOC: Mocho Nacional; PAN: Pantaneiro; GIR: Gir; NEL: Nelore.

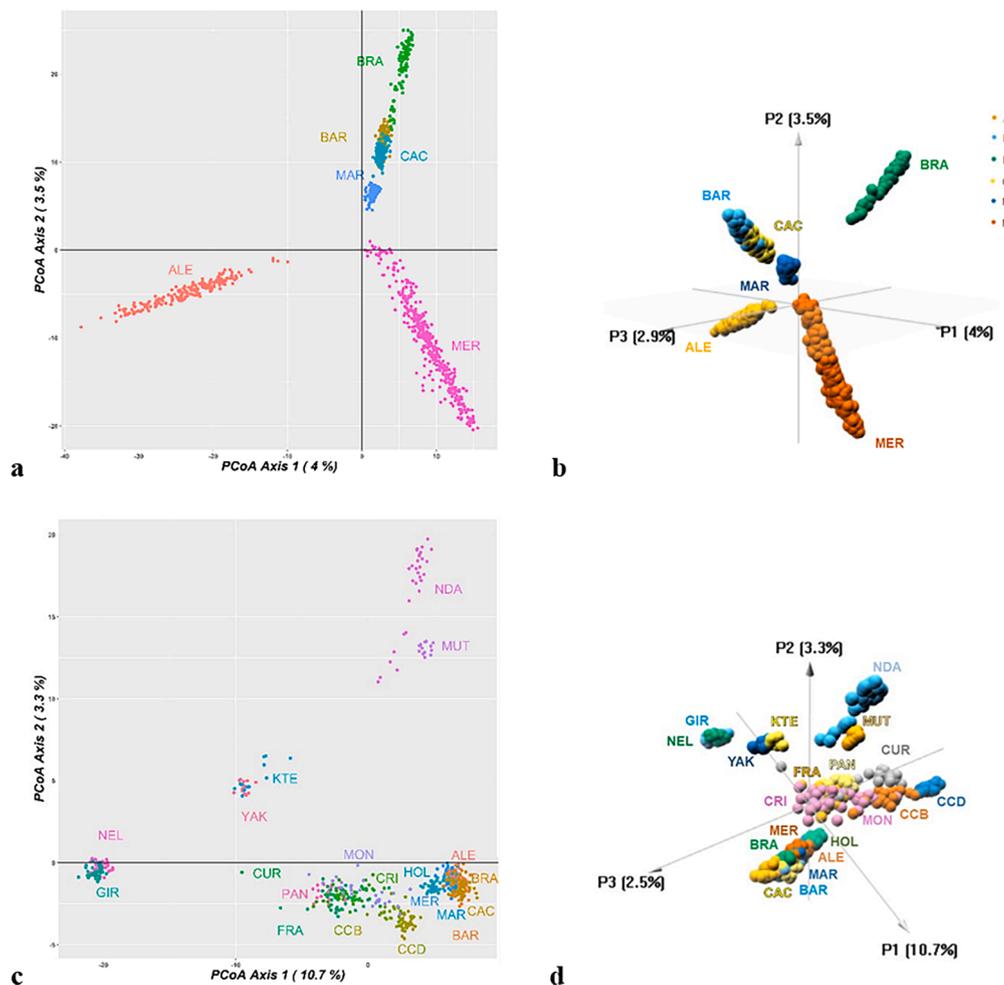


**Fig. 4. Dendrogram showing genetic relationships among twenty cattle breeds.** Dendrogram constructed using a neighbour-joining approach using Euclidian distance. a) Clustering including six Portuguese Iberian breeds; b) Clustering including Portuguese Iberian, African, Holstein, and Brazilian Iberian-derived Locally Adapted and Indicine animals. CAC: Cachena; MER: Mertolenga; BAR: Barrosã; MAR: Marinhoa; ALE: Alentejana; BRA: Brava de Lide; MUT: Muturu; NDA: Ndama; KTE: Keteku; YAK: Yakanaji; HOL: Holstein; CCD: Caracu Dairy; CCB: Caracu Beef; CRI: Crioulo Lageano; CUR: Curraleiro Pé-Duro; FRA: Franqueiro; MOC: Mocho Nacional; PAN: Pantaneiro; GIR: Gir; NEL: Nelore.

breeds, the  $N_e$  across generations was lower than 50. Using pedigree information, Maia (2020) and Carolino et al. (2020a, b) found similar results for  $N_e$  in the Barrosã, Cachena, and Mertolenga breeds, respectively. In the Brazilian Iberian-derived Locally Adapted breeds,  $N_e$  for CUR and FRA was below 50 in all generations. For CCD, CCB, CRI, MON, and PAN the  $N_e$  values were above this threshold only for generation 50. Dias et al. (2011), Pezzini et al. (2018), and Ribeiro et al. (2021) obtained analogous results, respectively, for CCB, CRI, and CUR breeds

using pedigree information. Interestingly, it should be recalled that when a pedigree is uninformative, as in a small set database, in some cases,  $N_e$  is underestimated, and thus, the estimations using pedigree and genotypes may vary.

Moreover, the effective population size in cattle per generation can vary greatly depending on various factors such as mating system, population size, genetic structure, and the system of selection or conservation. The values found for all Portuguese and Brazilian breeds for



**Fig. 5.** PCA analyses showing the relationships between twenty breeds grouped by Portuguese Iberian, African, Holstein, Brazilian Iberian-derived Locally Adapted and Indicine animals. a) PCA analyses with PC1 and PC2 (PC) using six Portuguese Iberian breeds; b) PCA analyses with first, second, and third PC. c) PCA analyses with first and second PC using twenty breeds grouped by Portuguese Iberian, African, Holstein, Brazilian Iberian-derived Locally Adapted and Indicine animals; d) PCA analyses with first, second and third PC using twenty breeds grouped by Portuguese Iberian, African, Holstein, Brazilian Iberian-derived Locally Adapted and Indicine animals. CAC: Cachena; MER\_VER: Red Mertolenga; MER\_ROS: Roan Mertolenga; MER\_MAL: Red Spotted Mertolenga; MER\_MAL: Mertolenga Malhado; BAR: Barrosã; MAR: Marinhoa; ALE: Alentejana; BRA: Brava de Lide; MUT: Muturu; NDA: Ndama; KTE: Keteku; YAK: Yakanaji; HOL: Holstein; CCD: Caracu Dairy; CCB: Caracu Beef; CRI: Crioulo Lageano; CUR: Curraleiro Pé-Duro; FRA: Franqueiro; MOC: Mocho Nacional; PAN: Pantaneiro; GIR: Gir; NEL: Nelore.

different types of generations were lower than the minimum value ( $N_e$  minimum = 50) proposed by FAO (FAO, 1998).  $N_e$  affects the dynamics of the population, which in turn can influence the rate and size of the allele frequency of a given genotype. Nevertheless, it is important to note that the  $N_e$  can change over time and may be influenced by decisions of breeders, demographic, genetic, and environmental factors. Regular monitoring of the effective population size and adjusting the breeding programs as necessary can help maintaining genetic diversity and ensure the long-term viability of cattle populations used in this study. This fact is extremely important for Brazilian Locally Adapted breeds since, with the exception of the Caracu and Mocho Nacional (included as hornless variation of Caracu breed), the other have breeder associations formed less than 25 years ago and a recent history of introduction into breeding programs.

#### 4.3. Runs of homozygosity (ROH) and inbreeding coefficients from roh

In the present study, we initially identified runs of homozygosity (ROH) in six Portuguese Iberian and seven Brazilian Iberian-derived Locally Adapted breeds. The characterization of ROHs, particularly their size and frequency, provides information about relatedness within

a population. In this step, it is important to note that a higher density SNP chip, instead of a medium, as used here to maintain a uniform set of markers among the studied breeds, would be more informative. However, this choice may be justified in terms of the unique set of animals available for this study, serving as a valuable genetic resource from Brazil and Portugal. Our findings indicated that most individual samples (>75 %) showed ROHs <10 Mb. Since long ROH fragments are indicators of recent inbreeding, we can hypothesize that consanguineous matings are somehow controlled in the Portuguese Iberian and Brazilian Iberian-derived Locally Adapted farming systems. In Brazilian breeds, this hypothesis corroborates the work conducted at the in situ Conservation Nuclei that are part of the Animal Genetics Resource Conservation Program developed by Embrapa in partnership with universities and private breeders. In the Pantaneiro breed, for example, efforts since the implementation of the in situ conservation nucleus, in 1984, have been focused on increasing the representativeness of the animals, as well as maintaining the genetic variation and inbreeding rate at acceptable levels through reproductive management (Mazza et al., 1994). However, CUR showed the highest frequency (>10 %) of long ROH (>31 Mb), probably for the reasons mentioned above, also associated with the breeding system of this breed due to the cultural and family nature of the

activity (Fioravanti et al., 2021). It is hoped that with the unification of breeders' associations and the official recognition of the breed, this situation can be reversed. Mainly because this is particularly problematic as long autozygous ROHs are generally associated with an increased incidence of strongly deleterious mutations as well as mild deleterious mutations. Since ROH represents the probability that all consecutive markers on a pair of homologous chromosome segments, in the same or different individual(s), display identical alleles (Hayes et al., 2003), the extent and frequency of ROH may provide information about the ancestry of an animal and its population. Thus, we used the same parameters and thresholds for all populations, and we were able to observe significant different patterns of ROH among them. Furthermore, inbreeding may be suggested from the presence of long ROH, with longer segments indicating recent inbreeding within a population (Kirin et al., 2010), suggesting a small effective population size. Therefore, we also evaluated the  $N_e$  to better understand the relationship between ROH and  $N_e$ .

Among Portuguese Iberian breeds, although BRA showed the lowest  $N_e$  in the estimated generation 5, we observed a similar occurrence of the longest ROH in all Portuguese breeds. Additionally, when compared with the Brazilian breeds, they showed larger ranges of  $F_{ROH}$ , even though the median was lower, except in BRA breed. In general, this indicates a proper mating scheme in Portuguese breeds to prevent higher inbreeding, but special attention to the BRA breed is needed to prevent higher inbreeding rates (e.g., the design of mating to minimize the kinship between parents). These results observed in the BRA breed may be related to the excessive use of some breeding males. Moreover, breeders may have preferences regarding the lineages to be used and, consequently, use breeding males in very different ways.

As also observed by Verardo et al. (2021) in a comparison of 19 breeds raised in Brazil, the Brazilian Iberian-derived Locally Adapted CUR breed in this study presented the lowest  $N_e$  value in the generation 5. It suggests a recent inbreeding event, supported by the occurrence of the longest ROH and higher inbreeding coefficients ( $F_{ROH}$ ) found in this study. Using microsatellite markers, Egito et al. (2007, 2015) also detected inbreeding in this breed, and according to Carvalho (2015), this breed comprises approximately five thousand individuals only. Egito et al. (2015) suggested that in herds where genealogical record is difficult to implement, genetic management could be carried out by exchanging animals between more genetically distant properties after their molecular evaluation. Based on our results, special attention should be given to mating schemes in this breed to preserve the germplasm or other conservation strategies based on low or non-inbred mating. There is a need to plan and implement, as soon as possible, management strategies for these genetic resources in order to minimize a possible allelic loss.

All breeds in this study exhibited shared runs of homozygosity (ROH) islands among individuals. From these islands, genes were retrieved, and their biological processes were investigated to understand their roles in each breed. In the Portuguese Iberian breeds, the biological process network highlighted the *MAVS* and *CDC25B* genes, both identified in the MER population, which are related to the immune system via regulation of IP-10 production and germ cell development, respectively. *MAVS* encodes mitochondrial antiviral signaling protein, known to be required for defense against viruses (Kumar et al., 2006; Seth et al., 2005). *CDC25B* encodes a cell division cycle 25B protein, which was indicated to be important for normal mouse oocyte maturation, fertilization, and early embryonic development, together with other proteins like *YWHA* (Eisa et al., 2019). Moreover, *CDC25B* gene shares its biological process with the *ZBTB16* gene from the BRA breed, which was also associated with the regulation of T-cell differentiation. This gene encodes a zinc finger and BTB domain containing 16 proteins and has been related to innate and innate-like lymphoid lineage development in mice (Mao et al., 2017) and defense response to gram-positive bacteria in goats (Rout and Verma, 2021). Also, its overexpression is suggested to promote white adipogenesis and induce brown-like adipocyte formation in

bovine white intramuscular preadipocytes (Wei et al., 2018).

Sharing the regulation of the T-cell differentiation process with *ZBTB16*, we observed the *DTX1* gene, which encodes a deletion in E3 ubiquitin ligase 1 protein and is reported to play a regulatory role in immune responses in teleost fish (Zheng et al., 2021). This gene was highlighted in the BAR breed, in which we also observed the *SDS* gene which is associated with amino acid metabolic processes. *SDS* encodes a serine dehydratase protein and has been found to be involved in gluconeogenesis in rats and potential association with weight gain (Imam et al., 2003). Moreover, in the CAC breed, we observed the *FAR2* gene, which encodes fatty acyl-CoA reductase 2 and has been associated with feed conversion in pigs (Xiang et al., 2020).

Concerning the Brazilian Iberian-derived Locally Adapted breeds, we observed genes related to the immune system (e.g. *IRF4*, *EXOC2*, *IL7R*, *DUSP22*, and *SERPINB9*), prolactin (*PRLR*) and genetic imprinting (*KDM1B*) in the biological process network. The *DUSP22* gene, identified in the PAN breed, shared the negative regulation of the lymphocyte-mediated immunity process with the *IL7R* gene, identified in the CCD and CCB breeds. *DUSP22* encodes the dual specificity phosphatase 22 protein and has been related to immune responses in humans and mice (Chuang et al., 2016; Melard et al., 2016). In humans, variations in the *IL7R* (interleukin 7 receptor) gene have been linked to autoimmune risk (Al-Mossawi et al., 2019), while in livestock animals, it has been described to be associated with growth traits (Jia et al., 2019; Zhu et al., 2019). Given the known adaptability and rusticity of PAN, CCD, and CCB breeds and the immune system-related genes identified in this study, we could suggest that these breeds have been successfully selected for rusticity. In Caracu lineages (CCD and CCB), we identified islands in chromosome 20 with a gene associated, for example, with the Prolactin signaling pathway (*PRLR*). This gene encodes the prolactin receptor protein and was previously reported to be a useful genetic marker for milk production (Zhang et al., 2008) and a promising candidate gene for growth traits in cattle (Lü et al., 2011). Also, there is evidence that this gene contains the 'slick hair' mutations that confer thermotolerance to tropically adapted *Bos taurus* cattle (Littlejohn et al., 2014; Flórez Murillo, 2021), and Faza et al. (2021) have shown that Caracu populations contain the SNP 39,136,666 mutation. Since CCD and CCB have been selected for dairy and beef, respectively, in Brazil, it makes sense to suggest that this gene could have been fixed and could be linked to milk production, growth and heat stress tolerance.

Some studies conducted with Brazilian Iberian-derived Locally Adapted breeds revealed that they are efficient in building innate immunity response (Juliano et al., 2016); had a greater responsiveness to vaccines, indicating a profile of resistance in combating intracellular infectious agents (Maggioli et al., 2013), and better adaptability to heat stress (Bianchini et al., 2006; Cardoso et al., 2016). Even though the robustness genes observed in the Brazilian Iberian-derived Locally Adapted breeds could originate from indicine breeds, Verardo et al. (2021) evaluating the RHO islands of 19 breeds raised in Brazil, including indicine and Brazilian Iberian-derived Locally Adapted breeds, did not observe shared genes or ROH island between the two groups.

In general, our results identified candidate genes for important economical traits in both Portuguese Iberian and Brazilian Iberian-derived Locally Adapted breeds, such as those related to immune systems, fertility, milk, heat stress, and meat production. Some gene functions corroborate with the breed's known specificities, such as those related to the immune systems, and thus, the rusticity observed in Brazilian breeds and some of the Portuguese breeds (e.g.: MER), as cited by Monteiro et al. (1981). However, it is important to note that different genes were found to be related to common traits even within a cluster group, for example, those related to immune response such as *MAVS* (MER), *DTX* (BAR), *ZBTB16* (BRA), *DUSP22* (PAN) and *IL7R* (CCD and CCB). This could indicate that each breed traced its genetic path in defining its main attributes probably due to selection pressure in totally different environments and, even though they have some common

attributes, they also have their specific genetic architecture.

#### 4.4. Breed composition

The admixture model, neighbor-joining, and PCA analyses allowed us to evaluate the breed composition of Portuguese Iberian, Brazilian Iberian-derived Locally Adapted, Holstein, African, and Indicine cattle. Firstly, it is important to note that the Portuguese and Brazilian Iberian-derived Locally Adapted breeds have a taurine origin, similar to the Holstein breed. On the other hand, African breeds Muturu and N'dama are of taurine origin, while the Keteku breed is a crossbred between Muturu and Zebu animals, and the Yakanaji breed, also known as White Fulani, has Zebuine origin (Joshi et al. 1957; Adebambo, 2001; Gwaza and Momoh, 2016). Miscegenation between different subspecies occurs at distinct levels in African breeds due to the extensive breeding system (Joshi et al. 1957; Adebambo, 2001; Gwaza and Momoh, 2016).

Initially, the study focused on the six Portuguese Iberian breeds were studied, followed by a comparison with others breeds. Based on admixture analysis ( $K = 4$ ), CAC and BAR animals formed one cluster, while the other breeds formed independent groups. Previous studies using the admixture model, neighbor-joining, and PCA analyses also described CAC and BAR as closely related breeds (Ginja et al., 2009; Martín-Burriel et al., 2011), which it is consistent with their evolution history. According to Leite and Dantas (2000), historical evidence supports a close genetic relationship between these breeds, where they were considered a single population until 1998 when BAR was recognized as an independent breed. Martín-Burriel et al. (2011) described BRA as one of the most divergent breeds within the 40 populations evaluated in their study, supporting our results found with ancestral population identified from  $K = 2$ .

The inclusion of African and indicine breeds in the analyses allowed the identification of a Zebuine-like genotype (red) in Portuguese Iberian animals, likely originating from African breeds. This result is consistent with previous studies that evaluated mtDNA sequences and Y haplotypes, detecting African alleles in Portuguese Iberian breeds, suggesting that Zebu gene flow in Portugal with African origins (Beja-Pereira et al., 2006; Ginja et al., 2009). At  $K = 3$ , a light green genotype was found, discriminating mainly African taurine breeds, identified to a lesser extent in Portuguese Iberian and Brazilian Iberian-derived Locally Adapted breeds. This explains the greater genetic distance of African MUT and NDA breeds, and KTE and YAK closer to the breeds of Zebuine origin when evaluating PCA results. This result highlights the influence of African cattle in Iberian breeds, consistent with their demographic history resulting from the Arab occupation of the Iberian Peninsula for more than seven centuries since AD 711 or due to the later importation of African animals carried out by Portuguese colonizers in Africa (Cymbron et al., 1999). Similarly, at  $K = 5$ , a *Bos taurus* fraction detected in the HOL breed (yellow) was found in low density in Portuguese cattle. The Brazilian Iberian-derived Locally Adapted breeds CCD, CCB, CRI, CUR, FRA, MOC, and PAN were grouped in neighbor-joining and PCA analyses. The admixture results ( $K = 2$ ) displayed an Indicine genome, as also observed by Verardo et al. (2021) and others authors using STR markers (Egito et al., 2007; Ginja et al., 2010). This result was expected, as historical data show that animals of Indian origin were introduced in South America from the end of the 19th century, with the predominant importation of males of Indian origin. Although Zebu breeds are predominant in Brazil, it is known that only around 7000 animals were imported, which were used for absorptive crossing, replacing the Brazilian Iberian-derived Locally Adapted breeds. Thus, it is expected that even after these years, there will be alleles of Zebu origin in the Locally Adapted base. Cytogenetic studies have shown the presence of the Indicine (acrocentric) Y chromosome in bulls of Locally Adapted breeds (Issa et al., 2009), and it was also possible to verify this introgression through historical reports (Mazza et al., 1994) and other specific Y markers (Ginja et al., 2019). Several genetic studies support these facts, including this study, where one could observe in  $K = 2$  that NEL and GIR

breeds still had a slight taurine influence (Egito et al., 2007a; Ginja et al., 2019). Moreover, the majority of Brazilian Indicine breeds have mitochondrial haplotypes of taurine origin (Meirelles et al., 1999; Egito, 2007; Ginja et al., 2019).

At  $K = 3$ , the PCA and the dendrogram generated also reveal the influence of African breeds on the Brazilian ones. Evaluating mitochondrial DNA studies (Miretti et al., 2002; 2004; Egito, 2007; Ginja et al. 2019), it appears that most of the existing haplotypes are taurine of African origin, including one that occurs at high frequency only in Brazil, including in Indicine breeds. It is suggested that there may be three possible hypotheses, which are not mutually exclusive. First, this genome was introduced from the Iberian breeds, originating from North Africa, probably as the result of the Arab occupation prior to European migration to the New World, and due to the bottleneck effect and genetic drift; it remained in the Brazilian breeds (Lenstra et al., 2014). It is also possible to suggest the possibility that these animals were introduced through undocumented imports on ships that brought slaves to Brazil or more recently, due to the official importation of African breeds by private breeders. At  $K = 4$ , a purple background taurine ancestry was highlighted, discriminating the Brazilian Iberian-derived Locally Adapted from the other evaluated breeds. On the other hand, Portuguese Iberian and Indicine breeds seem to have a genetic influence on Brazilian Iberian-derived Locally Adapted animals, which could be explained by the Portuguese origin of these breeds and the intense use of adapted animals in crossbreeding cattle at the beginning of livestock farming in Brazil (Alderson, 1992; Madalena et al., 2002). Nevertheless, the remarkable purple background ancestry in Brazilian Iberian-derived Locally Adapted breeds was not found in the Portuguese Iberian, African, Holsteins, or Indicine breeds. Historical reports demonstrate that animals of Spanish origin also participated in the formation of Locally Adapted breeds, as the current delimitation of Brazil took place long after the introduction of cattle in the South American continent, which also had a strong Spanish colonization (Mazza et al., 1994). The genetic divergence observed in Brazilian Iberian-derived Locally Adapted breeds, as also observed by Verardo et al. (2021), probably occurred as a result of many years of natural selection in distinct environments during Brazil's colonization, associated with miscegenation of different breeds. This could have contributed to the formation of bovine populations with higher adaptation to the tropical environment with a unique gene pool. Thus, the Brazilian Iberian-derived Locally Adapted breeds genetic represent an important genetic resource for Brazil and other tropical countries since these breeds seem to contain a repository of alleles associated with adaptation to different climates and regions of Brazil. In this way, more attention should be invested in the conservation and breeding programs of CCD, CCB, CRI, CUR, FRA, MOC, and PAN breeds.

Moreover, a highly homogeneous indicine ancestry in the indicine GIR and NEL breeds was observed from  $K = 2$  up to 5. This reinforces that the ancestry of these breeds is very conserved, as described by Verardo et al. (2021), which found  $\geq 99\%$  of indicine background at  $K = 5$ . This may be a relevant source to support the international market between Brazil and foreign countries interested in importing these indicine-improved breeds, which are suitable for use in commercial *Bos indicus* x *Bos taurus* crossbred production systems in tropical regions of the globe.

## 5. Conclusions

This study revealed varying extents of linkage disequilibrium (LD) across the assessed breeds, along with differences in effective population size ( $N_e$ ), particularly attributable to specific genetic programs or their absence, particularly evident in some Locally Adapted Brazilian Iberian breeds. Furthermore, the analysis of runs of homozygosity (ROH) identified different genes associated with common traits, even within a cluster groups, highlighting the importance of continued caution in managing these populations. Despite controlled consanguineous mating in Portuguese Iberian and Brazilian Iberian-derived Locally Adapted

breeds farming systems, the analysis of genetic variability suggests the need for ongoing precautions in their management.

Importantly, certain ancestral components observed in Locally Adapted Brazilian Iberian breeds were absent in Portuguese Iberian and other Taurine, African, and Zebu breeds. This discrepancy may be a consequence of extensive artificial and natural selection over many years after the colonization of Brazil, contributing to the formation of cattle populations with enhanced adaptation to the tropical environment. Consequently, the genetic background of Brazilian Iberian-derived Locally Adapted breeds emerge as a crucial genetic resource for Brazil and for other tropical countries, potentially harboring alleles associated with adaptation to different climates and regions of Brazil. Therefore, conservation and breeding programs for these breeds warrant increased attention. Additionally, both sets of populations, Locally Adapted Brazilian and Portuguese breeds, represent substantial reservoirs of animal genetic resources for food and agriculture, thriving in different challenging environments in two geographically distant regions, yet sharing similarities in facing heat and water scarcity for a significant part of the year.

## Funding

MVGBS was supported by Embrapa (Brazil) SEG 02.13.05.011.00.00 and CNPq 310199/2015–8 “Detecting signatures of selection from Next Generation Sequencing Data,” MCTI/CNPq/INCT-CA and APQ-02750–23. LLV was supported by MCTI/CNPq/INCT-CA and CNPq (314532/2021–8).

## CRedit authorship contribution statement

**Lucas Lima Verardo:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Pamela Itajara Otto:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Marco Antonio Machado:** Writing – review & editing, Writing – original draft, Data curation. **João Cláudio do Carmo Panetto:** Writing – review & editing, Writing – original draft, Formal analysis. **Daniele Ribeiro de Lima Reis Faza:** Writing – review & editing, Writing – original draft, Formal analysis. **Andréa Alves do Egito:** Writing – review & editing, Data curation, Conceptualization. **Andreia Sofia Martins Vitorino:** Writing – review & editing, Writing – original draft, Formal analysis. **Maria Inês Carvalho Martins Carolino:** Writing – review & editing, Writing – original draft, Formal analysis. **Nuno Pimentel Carolino:** Writing – review & editing, Writing – original draft, Funding acquisition, Data curation, Conceptualization. **Marcos Vinicius Gualberto Barbosa da Silva:** .

## Declaration of competing interest

The authors declare no conflicts of interest.

## Acknowledgments

We wish to thank the several Portuguese Cattle Breeders Associations for their support in providing the genetic material from Portuguese cattle breed for genotyping, namely, Associação dos Criadores de Bovinos da Raça Alentejana (ACBRA), Associação dos Criadores de Bovinos de Raça Barrosã (AMIBA), Associação Portuguesa de Criadores de Toiros de Lide (APCTL), Associação dos Criadores da Raça Cachena (ACRC), Estação de Apoio à Bovinicultura Leiteira (EABL) and Associação de Criadores Bovinos Mertolengos (ACBM), respectively, of the Alentejana, Barrosã, Brava de Lide, Cachena, Marinhola and Mertolenga breeds.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.livsci.2024.105455](https://doi.org/10.1016/j.livsci.2024.105455).

## References

- Abasht, B., et al., 2009. Extent and consistency of linkage disequilibrium and identification of DNA markers for production and egg quality traits in commercial layer chicken populations. *BMC. Genomics*. 10 (Suppl 2), S2. <https://doi.org/10.1186/1471-2164-10-S2-S2>, 2009.
- Adebambo, 2001. The Muturu: a rare sacred breed of cattle in Nigeria. In: *Animal Genetic Resources/Resources Génétiques Animales/Recursos Genéticos Animales*, 31, pp. 27–36. <https://doi.org/10.1017/S101423390001450>.
- Alderson, L., 1992. The categorization of types and breeds of cattle in Europe. *Arch. Zootec.* 41, 325–334. <https://dialnet.unirioja.es/servlet/articulo?codigo=278708> [Accessed June 22, 2023].
- Alexander, D.H., et al., 2009. Fast Model-Based Estimation of Ancestry in Unrelated Individuals. *Genome Res.* 19, 1655–1664. <https://doi.org/10.1101/gr.094052.109>.
- Ajomone-Marsan, P., et al., 2010. On the origin of cattle: how aurochs became cattle and colonized the world. *Evol. Anthropol.* 19, 148–157. <https://doi.org/10.1002/evan.20267>.
- Al-Mossawi, H., et al., 2019. Context-specific regulation of surface and soluble IL7R expression by an autoimmune risk allele. *Nat. Commun.* 10, 4575. <https://doi.org/10.1038/s41467-019-12393-1>.
- Amaral, A.J., et al., 2008. Linkage disequilibrium decay and haplotype block structure in the pig. *Genetics* 179, 569–579. <https://doi.org/10.1534/genetics.107.084277>.
- Ardlie, K.G., et al., 2002. Patterns of linkage disequilibrium in the human genome. *Nat. Rev. Genet.* 3, 299. <https://doi.org/10.1038/nrg777>.
- Beja-Pereira, A., et al., 2003. Genetic characterization of southwestern European bovine breeds: a historical and biogeographical reassessment with a set of 16 microsatellites. *J. Hered.* 94 (3), 243–250. <https://doi.org/10.1093/jhered/esg055>.
- Beja-Pereira, A., et al., 2006. The origin of European cattle: evidence from modern and ancient DNA. *PNAS* 103, 8113–8118. <https://doi.org/10.1073/pnas.0509210101>.
- Bianchini, E., et al., 2006. Características corporais associadas com a adaptação ao calor em bovinos naturalizados brasileiros. *Pesqui. Agropecu. Bras.* 41, 1443–1448. <https://doi.org/10.1590/S0100-204x2006000900014>.
- Bindea, G., et al., 2009. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*. 25, 1091–1093. <https://doi.org/10.1093/bioinformatics/btp101>.
- Cañón, J., et al., 2001. Genetic diversity measures of local European beef cattle breeds for conservation purposes. *Genet. Select. Evol.* 33 (3), 311–332. <https://doi.org/10.1186/1297-9686-33-3-311>.
- Cardoso, C.C., et al., 2016. Heat tolerance in Curraleiro Pe-Duro, Pantaneiro and Nelore cattle using thermographic images. *Animals* 6 (2), 9. <https://doi.org/10.3390/ani6020009>.
- Carolino, N., et al., 2020a. Raça Bovina Cachena - Caracterização Genética por Análise Demográfica –2020. Instituto Nacional de Investigação Agrária e Veterinária, I.P., Polo de Investigação da Fonte, Boa, Portugal.
- Carolino, N., et al., 2020b. Genetic Diversity in the Portuguese Mertolenga Cattle Breed Assessed by Pedigree Analysis. *Animals* 10 (11), 1990. <https://doi.org/10.3390/ani10111990>.
- Carvalho, 2015. Curraleiro Pé-Duro: Germoplasma Estratégico Do Brasil. Embrapa Press, Brasília.
- Charrad, M., et al., 2014. Nbclust: an R package for determining the relevant number of clusters in a data set. *J. Stat. Softw.* 61, 1–36. <https://doi.org/10.18637/jss.v061.i06>.
- Chuang, H.C., et al., 2016. Downregulation of the phosphatase JKAP/DUSP22 in T cells as a potential new biomarker of systemic lupus erythematosus nephritis. *Oncotarget*. 7, 57593–57605. <https://doi.org/10.18632/oncotarget.11419>.
- Cymbron, T., et al., 1999. Mitochondrial sequence variation suggests an African influence in Portuguese cattle. *Proc. R. Soc. Lond. B. Biol. Sci.* 266 (1419), 597–603. <https://doi.org/10.1098/rspb.1999.0678>.
- Cymbron, T., et al., 2005. Microsatellite diversity suggests different histories for Mediterranean and Northern European cattle populations. *Proc. Royal Soc. B.* 272 (1574), 1837–1843. <https://doi.org/10.1098/rspb.2005.3138>.
- da Fonseca, R.R., et al., 2019. Consequences of breed formation on patterns of genomic diversity and differentiation: the case of highly diverse peripheral Iberian cattle. *BMC. Genomics*. 20 (1), 1–13. <https://doi.org/10.1186/s12864-019-5685-2>.
- Dias, A.C., et al., 2011. Análise da estrutura populacional raça Caracu: consanguinidade e Tamanho Efetivo. *RBZ* 360, 283. Abstract retrieved from Abstracts in 48 Reunião Anual da Sociedade Brasileira de Zootecnia.
- Egito, A.A., et al., 2007. Microsatellite-based genetic diversity and relationships among ten creole and commercial cattle breeds raised in Brazil. *BMC. Genet.* 8, 83. <https://doi.org/10.1186/1471-2156-8-83>.
- Egito, 2007. Diversidade genética, ancestralidade individual e miscigenação nas raças bovinas no Brasil com base em Microsatélites e Haplótipos de DNA Mitocôndrial: subsídios para a conservação. 246 f. Thesis (Doutorado em Biologia Molecular)- Universidade Brasília, Brasília, 2007. Available at [https://repositorio.unb.br/bitstream/10482/1136/1/Tese\\_2007\\_AndreaEgito.pdf](https://repositorio.unb.br/bitstream/10482/1136/1/Tese_2007_AndreaEgito.pdf).
- Egito, A.A., et al., 2015. Ferramentas moleculares para a gestão, uso e conservação das raças Curraleiro Pé-Duro e Pantaneira. Abstract retrieved from Simpósio Internacional de Raças Nativas: sustentabilidade e Propriedade Intelectual. Teresina

- Pl. 19 a 22 de agosto de 2015. Available in <https://www.alice.cnptia.embrapa.br/bitstream/doc/1025674/1/Ferramentasmolecularesparagestaouseconservacaoda racasCurraleiroPeDuroePantaneira.pdf>.
- Eisa, A.A., et al., 2019. YWHA (14-3-3) protein isoforms and their interactions with CDC25B phosphatase in mouse oogenesis and oocyte maturation. *BMC Dev. Biol.* 19 (1), 1–22. <https://doi.org/10.1186/s12861-019-0200-1>.
- FAO Commission on genetic resources for food and agriculture, 2022. Item 4 of the Provisional Agenda. In: Intergovernmental technical working group on animal genetic resources for food and agriculture. Twelfth Session, Rome, 18–20. January 2022. <https://www.fao.org/3/nl061en/nl061en.pdf> [Accessed February 14, 2023].
- Faza, D.R.L.R., et al., 2021. Development of tetra-primer ARMS-PCR protocol to genotype the prolactin receptor SNP 39136666 and assessment of this SNP in Brazilian Locally Adapted cattle breeds. *Arq. Bras. Med. Vet. Zootec.* 73, 534–538. <https://doi.org/10.1590/1678-4162-12104>.
- Fioravanti, M.C.S., et al., 2021. O uso dos bovinos locais brasileiros em sistemas de produção sustentáveis: raças Curraleiro Pé-Duro e Pantaneiro. *Orgs. In: OLIVEIRA, I. P., LEANDRO, W.M., COSTA, K.P., FURTINI NETO, A.E. (Eds.), Sistema Agropecuário De Produção Sustentável. Cegraf UFG, Goiânia, pp. 1352–1425, 2021.*
- Flórez Murillo, J.M., et al., 2021. Three novel nonsense mutations of prolactin receptor found in heat-tolerant *Bos taurus* breeds of the Caribbean Basin. *Anim. Genet.* 52 (1), 132–134. <https://doi.org/10.1111/age.13027>.
- FAO Commission on genetic resources for food and agriculture, 2023. Item 10.2 of the Provisional Agenda. METHODS FOR ESTIMATION OF WITHIN-POPULATION GENETIC VARIATION. Ninetenth Regular Session, Rome, 17-21. July, 2023. <http://www.fao.org/3/cc6007em/cc6007em.pdf> [Access Jan 21st 2024].
- Food and Agriculture Organization of the United Nations (FAO), 1998. Secondary guidelines for the development of national farm animal genetic resources management plans: management of small populations at risk. <https://www.fao.org/3/w9361e/w9361e.pdf> [Accessed February 14, 2023].
- Felius, M., et al., 2014. On the history of cattle genetic resources. *Diversity.* (Basel) 6.4, 705–750. <https://doi.org/10.3390/d6040705>.
- Gibson, J., et al., 2006. Extended tracts of homozygosity in outbred human populations. *Hum. Mol. Genet.* 15, 789–795. <https://doi.org/10.1093/hmg/ddi493>.
- Ginja, C., et al., 2009. Y chromosome haplotype analysis in Portuguese cattle breeds using SNPs and STRs. *J. Hered.* 100 (2), 148–157. <https://doi.org/10.1093/jhered/esn080>.
- Ginja, C., et al., 2010. Analysis of STR markers reveals high genetic structure in Portuguese native cattle. *J. Hered.* 101 (2), 201–210. <https://doi.org/10.1093/jhered/esp104>.
- Ginja, C., et al., 2019. The genetic ancestry of American Creole cattle is inferred from uniparental and autosomal genetic markers. *Sci. Rep.* 9, 11486. <https://doi.org/10.1038/s41598-019-47636-0>.
- Gruber, B., et al., 2018. Darrt: an R package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Mol. Ecol. Resour.* 18 (3), 691–699. <https://doi.org/10.1111/1755-0998.12745>.
- Gwaza, Momoh, 2016. Endangered Indigenous Cattle Breeds of Nigeria a case For Their Conservation and Management, 30. World Scientific News, pp. 68–88 file:///Users/lucasverardo/Downloads/Endangered\_indigenous\_cattle\_breeds.pdf.
- Hayes, B.J., et al., 2009. Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genet. Sel. Evol.* 41, 51. <https://doi.org/10.1186/1297-9686-41-51>.
- Hayes, B.J., et al., 2003. Novel multilocus measure of linkage disequilibrium to estimate past effective population size. *Genome Res.* 13, 635–643. <https://doi.org/10.1101/gr.387103>.
- Heifetz, E.M., et al., 2005. Extent and consistency across generations of linkage disequilibrium in commercial layer chicken breeding populations. *Genetics* 171, 1173–1181. <https://doi.org/10.1534/genetics.105.040782>.
- Imam, S., et al., 2003. Quantity as Well as Quality of Dietary Protein Affects Serine Dehydratase Gene Expression in Rat Liver. *J. Nutr. Sci. Vitaminol.* 49, 33–39. <https://doi.org/10.3177/jnsv.49.33>.
- Issa, E.C., et al., 2009. Cytogenetic analysis of the y chromosome of native brazilian bovine breeds: preliminary data. *Arch. de Zootec.* 58 (221), 93–101. <https://www.alice.cnptia.embrapa.br/bitstream/doc/656857/1/SP19829ID32133.pdf>.
- Jia, C., et al., 2019. Identification of genetic loci associated with growth traits at weaning in yak through a genome-wide association study. *Anim. Genet.* 51, 300–305. <https://doi.org/10.1111/age.12897>.
- Joshi, N.R., et al., 1957. Types and breeds of African cattle. *FAO Agricultural Studies No. 37.* (Rome, Italy 1957).
- Juliano, R.S., et al., 2016. Curraleiro Pé Duro cattle breed (*Bos taurus*) differ in phagocytic neutrophil function related to crossbreed cattle (*Bos taurus* x *Bos indicus*). *Actas iberoram* 7, 24–30 v. <http://www.alice.cnptia.embrapa.br/alice/handle/doc/1048281>.
- Kirin, M., et al., 2010. Genomic runs of homozygosity record population history and consanguinity. *PLoS One* 5, e13996. <https://doi.org/10.1371/journal.pone.0013996>.
- Kumar, H., et al., 2006. Essential role of IPS-1 in innate immune responses against RNA viruses. *J. Exp. Med.* 203, 1795–1803. <https://doi.org/10.1084/jem.20060792>.
- Leite, Dantas, 2000. Cachena. editor. In: Garcia, EY (Ed.), *Catálogo De Raças Autóctones De Castela e Leão (Espanha)—Região Norte de Portugal: Espécies bovina e equina. Salamanca (Spain). Fundação Rei Afonso Henriques, pp. 77–90.*
- Lenstra, J.A., et al., 2014. Meta-analysis of mitochondrial DNA reveals several population bottlenecks during worldwide migrations of cattle. *Diversity.* (Basel) 6 (1), 178–187. <https://doi.org/10.3390/d6010178>.
- Littlejohn, M.D., et al., 2014. Functionally reciprocal mutations of the prolactin signalling pathway define hairy and slick cattle. *Nat. Commun.* 5 (1), 5861. <https://doi.org/10.1038/ncomms5861>.
- Loftus, R.T., et al., 1994. Evidence for two independent domestications of cattle. *Proc. Natl. Acad. Sci.* 91 (7), 2757–2761. <https://doi.org/10.1073/pnas.91.7.2757>.
- Luan, T., et al., 2014. Genomic prediction based on runs of homozygosity. *Genet. Sel. Evol.* 46 (64) <https://doi.org/10.1186/s12711-014-0064-6>.
- Lü, A., et al., 2011. Novel SNPs of the bovine PRLR gene associated with milk production traits. *Biochem. Genet.* 49, 177–189. <https://doi.org/10.1007/s10528-010-9397-1>.
- Lu, D., et al., 2012. Linkage disequilibrium in Angus, Charolais, and crossbred beef cattle. *Front. Genet.* 3 (152) <https://doi.org/10.3389/fgene.2012.00152>.
- Madalena, F.E., et al., 2002. Genetic improvement in medium-to low-input systems of animal production: experiences to date. In: *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production.*
- Maggioli, M.F., et al., 2013. Cellular immune response of Curraleiro Pé-duro and Nellore calves following *Mycobacterium bovis*-BCG vaccination. *Pesqui. Vet. Bras.* 33 (12), 1403–1408. <https://doi.org/10.1590/S0100-736x2013001200002>.
- Maia, 2020. Data from: caracterização genética por análise demográfica da raça Barrosã. <https://www.amiba.pt/2020%20Caracterizacao%20Genetica%20Analise%20Demografica.pdf>.
- Mao, A.P., et al., 2017. A shared Runx1-bound Zbtb16 enhancer directs innate and innate-like lymphoid lineage development. *Nat. Commun.* 8 (1), 1–14. <https://doi.org/10.1038/s41467-017-00882-0>.
- Marigorta, Navarro, 2013. High trans-ethnic replicability of GWAS results implies common causal variants. *PLoS Genet.* 9, e1003566 <https://doi.org/10.1371/journal.pgen.1003566>.
- Martín-Burriel, I., et al., 2011. Genetic diversity, structure, and breed relationships in Iberian cattle. *J. Anim. Sci.* 89 (4), 893–906. <https://doi.org/10.2527/jas.2010-3338>.
- Mateus, J.C., et al., 2004. Contributions of Portuguese cattle breeds to genetic diversity using marker-estimated kinships. *Anim. Genet.* 35 (4), 305–313. <https://doi.org/10.1111/j.1365-2052.2004.01168.x>.
- Mazza, M.C.M., et al., 1994. *Etnobiologia e conservação do bovino Pantaneiro. Embrapa-SPI. Corumbá, 1ª ed., p. 61*
- McManus, C., et al., 2009. Heat tolerance in naturalized Brazilian cattle breeds. *Livest. Sci.* 120, 256–264. <https://doi.org/10.1016/j.livsci.2008.07.014>.
- McQuillan, R., et al., 2008. Runs of homozygosity in European populations. *Am. J. Hum. Genet.* 83, 359–372. <https://doi.org/10.1016/j.ajhg.2008.08.007>.
- McKay, S.D., et al., 2007. Whole genome linkage disequilibrium maps in cattle. *BMC Genet.* 8 (74) <https://doi.org/10.1186/1471-2156-8-74>.
- Melard, P., et al., 2016. Molecular alterations and tumor suppressive function of the DUSP22 (Dual Specificity Phosphatase 22) gene in peripheral T-cell lymphoma subtypes. *Oncotarget.* 7, 68734–68748. <https://doi.org/10.18632/oncotarget.11930>.
- Meirelles, F.V., et al., 1999. Is the American zebu really *Bos indicus*? *Genet. Mol. Biol.* 22, 543–546. <https://doi.org/10.1590/S1415-47571999000400013>.
- Meuwissen, Goddard, 2007. Multipoint identity-by-descent prediction using dense markers to map quantitative trait loci and estimate effective population size. *Genetics* 176, 2551–2560. <https://doi.org/10.1534/genetics.107.070953>.
- Meuwissen, T.H., et al., 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829. <https://doi.org/10.1093/genetics/157.4.1819>.
- Miranda do, Vale, 1949. *Gado Bissulco. A Terra e o Homem. Coleção de Livros Agrícolas. Livraria Sá da Costa, Lisboa.*
- Miretti, M.M., et al., 2004. Predominant African-derived mtDNA in Caribbean and Brazilian Creole cattle is also found in Spanish cattle (*Bos taurus*). *J. Hered.* 95 (5), 450–453. <https://doi.org/10.1093/jhered/esh070>.
- Miretti, M., et al., 2002. African derived mitochondria in South American native cattle breeds (*Bos taurus*): evidence of a new taurine mitochondrial lineage. *J. Hered.* 93, 323–330. <https://doi.org/10.1093/jhered/93.5.323>.
- Mokry, F.B., et al., 2014. Linkage disequilibrium and haplotype block structure in a composite beef cattle breed. *BMC Genomics.* 15, S6. <https://doi.org/10.1186/1471-2164-15-S7-S6>.
- Bovinos em Portugal Monteiro, A.P., et al., 1981. Raça bovina Mertolenga. In: Rodrigues, A. (Ed.), Direção Geral dos Serviços Veterinários—Direção De Serviços de Fomento e Melhoramento Animal: Lisboa, Portugal, pp. 197–216.**
- O'Brien, A.M.P., et al., 2014. Linkage disequilibrium levels in *Bos indicus* and *Bos taurus* cattle using medium and high-density SNP chip data and different minor allele frequency distributions. *Livest. Sci.* 166, 121–132. <https://doi.org/10.1016/j.livsci.2014.05.007>.
- Paradis, E., et al., 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics.* 20 (2), 289–290. <https://doi.org/10.1093/bioinformatics/btg412>.
- Peripolli, E., et al., 2017. Runs of homozygosity: current knowledge and applications in livestock. *Anim. Genet.* 48, 255–271. <https://doi.org/10.1111/age.12526>.
- Pezzini, T., et al., 2018. Population structure of the Brazilian Crioula Lageana cattle (*Bos taurus*) breed. *Rev. Colomb. Cienc. Pec.* 31 (2), 93–102. <https://doi.org/10.17533/udea.rccp.v31n2a02>.
- Porto-Neto, L.R., et al., 2014. The extent of linkage disequilibrium in beef cattle breeds using high-density SNP genotypes. *Genet. Sel.* 46, 1–5. <https://doi.org/10.1186/1297-9686-46-22>.
- Purcell, S., et al., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. <https://doi.org/10.1086/519795>.
- Regazzi, Silva, 2010. Tests for model identity and parameter equality with nonlinear regression models in data from randomized complete block design. *Rev. Ceres* 57 (3), 315–320. <https://doi.org/10.1590/S0034-737x2010000300005>.
- Ribeiro, 1993. *Gado Crioulo Lageano, uma alternativa sustentada para as pastagens naturais do Planalto Catarinense. Simpósio Da 30ª Reunião Anual da SBZ. Anais... 245-261, p. 1993. Rio de Janeiro.*

- Ribeiro, N.L., et al., 2021. Analysis of the population structure of a cattle conservation nucleus Curraleiro Pé Duro. *Arq. Bras. Med. Vet. Zootec.* 73 (1), 231–238. <https://doi.org/10.1590/1678-4162-12042>.
- Rout, Verma, 2021. Post-translational modifications of milk proteins in geographically diverse goat breeds. *Sci. Rep.* 11 (1), 1–16. <https://doi.org/10.1038/s41598-021-85094-9>.
- Seth, R.B., et al., 2005. Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF- $\kappa$ B and IRF3. *Cell* 122, 669–682. <https://doi.org/10.1016/j.cell.2005.08.012>.
- Spritzke, A., et al., 2003. Genetic characterization of Criollo Lageano cattle by RAPD markers. *Pesqui. Agropecu. Bras.* 38, 1157–1164. <https://doi.org/10.1590/S0100-204x2003001000004>.
- Sved, 1971. Linkage disequilibrium and homozygosity of chromosome segments in finite populations. *Theor. Popul. Biol.* 2, 125–141. [https://doi.org/10.1016/0040-5809\(71\)90011-6](https://doi.org/10.1016/0040-5809(71)90011-6).
- Terwilliger, J.D., et al., 2002. A bias-ed assessment of the use of SNPs in human complex traits. *Curr. Opin. Genet. Dev.* 12 (6), 726–734. [https://doi.org/10.1016/S0959-437X\(02\)00357-X](https://doi.org/10.1016/S0959-437X(02)00357-X).
- Toro Ospina, A.M., et al., 2019. Linkage disequilibrium and effective population size in GIR cattle selected for yearling weight. *Reprod. Domest. Anim.* 54, 1524–1531. <https://doi.org/10.1111/rda.13559>.
- Upadhyay, M.R., et al., 2017. Genetic origin, admixture, and population history of aurochs (*Bos primigenius*) and primitive European cattle. *Heredity (Edinb)* 118, 169–176. <https://doi.org/10.1038/hdy.2016.79>.
- Verardo, L.L., et al., 2021. Genome-Wide Analyses Reveal the Genetic Architecture and Candidate Genes of Indicine, Taurine, Synthetic Crossbreds, and Locally Adapted Cattle in Brazil. *Front. Genet.* 12, 702822 <https://doi.org/10.3389/fgene.2021.702822>.
- Veroneze, R., et al., 2014. Linkage disequilibrium patterns and persistence of phase in purebred and crossbred pig (*Sus scrofa*) populations. *BMC. Genet.* 15, 126. <https://doi.org/10.1186/s12863-014-0126-3>.
- Wei, S., et al., 2018. ZBTB16 Overexpression Enhances White Adipogenesis and Induces Brown-Like Adipocyte Formation of Bovine White Intramuscular Preadipocytes. *Cell. Physiol. Biochem.* 48 (6), 2528–2538. <https://doi.org/10.1159/000492697>.
- Xiang, 2020. Integrating a Genome-Wide Association Study With Transcriptome Analyses to Identify Candidate Genes and Pathways for Feed Conversion Ratio in Yorkshire Pigs. *Res. Sq.* <https://doi.org/10.21203/rs.3.rs-104076/v1>.
- Xu, L., et al., 2019. Evaluation of linkage disequilibrium, effective population size, and haplotype block structure in Chinese cattle. *Animals* 9, 83. <https://doi.org/10.3390/ani9030083>.
- Zhang, J., et al., 2008. Genetic variation of PRLR gene and association with milk performance traits in dairy cattle. *Can. J. Anim. Sci.* 88, 33–39. <https://doi.org/10.4141/CJAS07052>.
- Zhang, Q., et al., 2015. Runs of homozygosity and distribution of functional variants in the cattle genome. *BMC. Genomics.* 16, 542. <https://doi.org/10.1186/s12864-015-1715-x>.
- Zhao, H., et al., 2007. Evaluation of linkage disequilibrium measures between multi-allelic markers as predictors of linkage disequilibrium between single nucleotide polymorphisms. *Genet. Res.* 89 (1), 1–6. <https://doi.org/10.1017/S0016672307008634>.
- Zheng, W., et al., 2021. Circular RNA circDtx1 regulates IRF3-mediated antiviral immune responses through suppression of miR-15a-5p-dependent TRIF downregulation in teleost fish. *PLoS. Pathog.* 17 (3), e1009438 <https://doi.org/10.1371/journal.ppat.1009438>.
- Zhu, F., et al., 2019. Genome-wide association study of growth and feeding traits in Pekin ducks. *Front. Genet.* 10, 702. <https://doi.org/10.3389/fgene.2019.00702>.
- Zhu, M., et al., 2013. Linkage disequilibrium estimation of Chinese beef Simmental cattle using high-density SNP panels. *Asian Australas. J. Anim.* 26, 772–779. <https://doi.org/10.5713/ajas.2012.12721>.