Candidate genes for disease, reproduction and meat quality traits in Portuguese native breeds

M.V.G.B. Silva[1*], L.L. Verardo[2], M.A. Machado[1], J.C.C. Panetto[1], I. Carolino[3] and N. Carolino[3]

[1]Brazilian Agricultural Research Corporation (EMBRAPA), Rua Eugênio do Nascimento, 610, 36038-330, Juiz de Fora – MG, Brazil; [2] Universidade Federal dos Vales do Jequitinhonha e Mucuri, Campus JK, 39100-000, Diamantina, Brazil; [3] Instituto Nacional de Investigação Agrária e Veterinária, Fonte Boa, 2005-048, Vale de Santarém, Portugal;
[*]marcos.vb.silva@embrapa.br

Abstract

Native Portuguese cattle breeds are a "biodiversity deposit" resulting from genetic and environmental effects accumulated over the years. We assessed two cattle breeds representative of Portuguese native cattle (Barrosã - BAR and Brava de Lide – BRA) aiming the evaluation of ROH across populations, identification of ROH islands and functional analysis of the identified genes. The pattern of ROH differed across breeds mainly from short to median range. Based on ROH islands analysis, three regions were observed to be shared by more than 35% of the individuals in the two breeds. Besides, gene networks highlighted biological associations based on genes found on ROH islands with biological processes related with reproductive traits (*RBM19*) and immune systems (*DTX1*) in BAR breed, and immune systems and adipogeneses processes (*ZBTB16*) in BRA breed.

Introduction

Native Portuguese cattle breeds are a "biodiversity deposit" resulting from genetic and environmental effects accumulated over the years, namely, geographical, edapho-climatic, land, historical, cultural, political-economic, among others. In the middle of the 20th century, with the implementation of the first Herd Books and the reorientation of production, mainly for beef, the current native bovine populations became properly organized. Since the 1990s, according to the size of the herd and the priorities defined by breeders for each breed, they have been included in conservation or genetic improvement programs. Except for the Brava de Lide breed, whose selection is oriented towards behavioural traits, the other breeds have been selected essentially for maternal and reproductive traits.

History of the population, breeding system, and pattern of geographic subdivision may be reflected in runs of homozygosity (ROH) patterns throughout the genome. ROH are continuous homozygous segments of DNA (Gibson et al., 2006), and its calculation is essential to identify islands. The islands of homozygosity shared among individuals, may be helpful in identify possible results of selection events (Zhang et al., 2015). Moreover, genes identified in those regions can be functionally analysed through gene networks (Verardo et al., 2021) aiming to a better understand of their biological role in each population. Thus, we assessed two cattle breeds representative of Portuguese native cattle, with different breeding goals (Barrosã - BAR and

Brava de Lide – BRA) aiming the evaluation of ROH across populations, identification of ROH islands and functional analysis of the identified genes.

Materials & Methods

The cattle breeds of Portugal used in this study are representative of the two trunks of cattle that were formed in the westernmost region of the Iberian Peninsula and classified by Miranda do Vale (1949), according to their ethnic characteristics and geographical distribution: Mauritanian or *Bos taurus mauritanicus* (Barrosã - BAR) and Iberian or *Bos taurus ibericus* (Brava de Lide - BRA). Animals were genotyped using the commercial genotyping array of Affymetrix (60k). SNP genotypes from sexual chromosomes were excluded from the analysis. Also, genotypes were excluded when the call rate < 0.95 and minor allele frequency < 0.01. In addition, samples showing overall call rates < 0.90 were excluded from further analysis. A total of 96 and 94 animals were analysed after quality control with 34997 and 33482 SNPs in the genotypes of BAR and BRA, respectively.

The ROH approach using the software PLINK v1.07 (Purcell et al., 2007) for each population with the following parameters were used: 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 a homozygous > 0.05. To identify ROH islands throughout the genome, the most observed homozygous segments shared by individuals in each breed were considered as an indication. The "--homozyg-group" function implemented in PLINK was used to assess ROH islands shared among individuals. The GenBank annotation based on the ARS-UCD1.2 assembly of the bovine genome was used to identify genes in ROH regions.

Gene networks highlighting biological processes among the gene sets identified for each breed were generated using the ClueGO plugin for Cytoscape (Bindea et al., 2009). This procedure allowed the creation of gene networks highlighting biological roles, and the comparison of gene clusters by visualizing their functional differences or similarities.

Results

The pattern of ROH differed across breeds mainly from short to median range (Figure 1 - A). BAR and BRA breeds showed similar percentage of long-range ROH (> 31Mb; <5%), while BAR showed higher percentages of short range ROH (<5 Mb; >50%). In addition, based on ROH islands analysis, three regions were observed to be shared by more than 35% of the individuals in the two breeds (Table 1). Gene networks highlighting biological associations were generated based on genes found on ROH islands (Figure 1 - B). Several biological processes were highlighted, including Regulation of T cell differentiation and positive regulation of embryonic development.

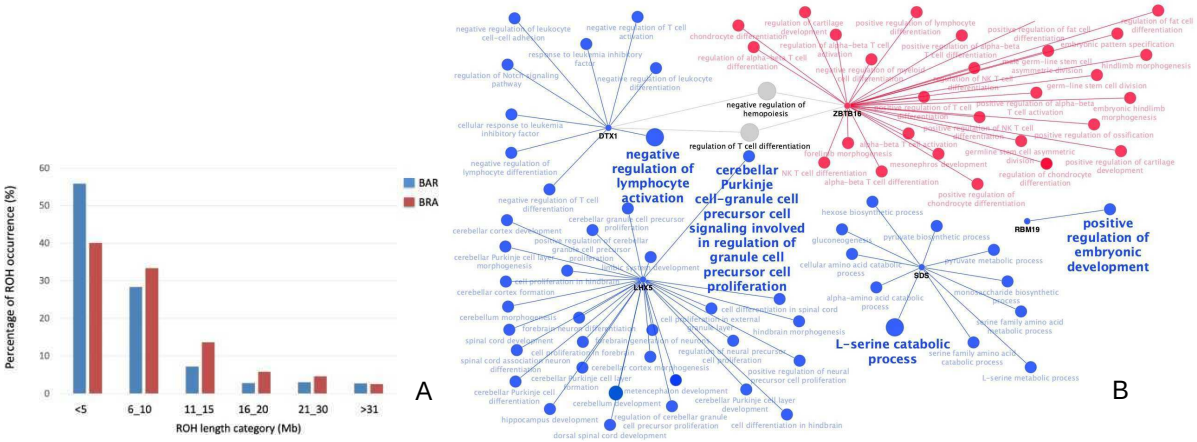

Figure 1. A) Percentage of runs of homozygosity (ROH) per length categories in two Portuguese Iberian cattle breeds. The sum of ROH was calculated per animal within each ROH length category. BAR: Barrosã; BRA: Brava de Lide. B) Gene networks highlighting biological associations of genes found on ROH islands in two Portuguese Iberian cattle breeds. Blue and red nodes are biological processes related with genes

identified in BAR and BRA breeds, respectively. Grey nodes are biological processes shared between both breeds genes.

Table 2. Chromosomal position, length in base pairs (bp), number of markers and gene content of runs of consensus homozygosity (ROH) islands most shared (ROH freq.) identified in two Portuguese Iberian breeds.

Breed[1]	Chr.	Begin	End	Length (bp)	Nº SNP	ROH freq. (%)	Annotated genes
BAR	17	60,517,694	61,094,324	576,630	11	35.42	*PLBD2, SDS, SDSL, DTX1, LHX5, LOC104974667, LOC101907408, TRNAG-CCC, RBM19, LOC100847943*
BRA	15	24,406,959	24,891,500	484,541	7	43.62	*ZBTB16, HTR3A, HTR3B, LOC511161, LOC101907641, LOC782610, TRNAG-CCC, LOC112441707*
	15	24,166,187	24,231,362	65,175	3	43.62	*TMPRSS5*

[1]BAR: Barrosã; BRA: Brava de Lide

Discussion

We first aimed to evaluate the run of homozygosity in two Portuguese breeds, Barrosã and Brava de Lide. A run of homozygosity is 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 individual and its population. Moreover, inbreeding may be inferred from the presence of long ROH, with longer segments indicating recent inbreeding within a population (Kirin et al., 2010). We observed that both breeds showed small and similar occurrence of longest ROH found in this study (<5%). This indicates a proper matting scheme in those breeds to prevent higher inbreeding. Besides, BAR showed higher prevalence of small ROH indicating more genetic diversity.

Moreover, both breeds showed ROH islands. It is important to note that with few sampled genotypes, it is quite likely that some ROH are artifacts. This is particularly important when pointing to specific candidate regions/genes as these might be misleading. In this study, genes and their biological processes were analysed aiming to understand their roles in each breed. In BAR, we identified an island in chromosome 17 with genes associated, for example, with positive regulation of embryonic development (*RBM19*) and negative regulation of lymphocyte activation (*DTX1*) in the biological process network. *RBM19* encoding for RNA binding motif protein 19, may be involved in regulating ribosome biogenesis and was previously reported to be essential for preimplantation during mouse embryonic development (Zhang et al., 2008). *DTX1* encode for deltex E3 ubiquitin ligase 1 protein being reported to play a regulatory role in immune responses in teleost fish (Zheng et al., 2021). Thus, these two genes may play a role on reproductive traits (*RBM19*) and immune systems (*DTX1*) in BAR breed.

Also, ROH islands were observed on chromosome 15 of BRA breed. In those islands, based on biological process network, we observed a gene (*ZBTB16*) related with regulation of T cell differentiation which is also shared with the previous discussed gene of BAR breed (*DTX1*). The *ZBTB16* gene encodes for the zinc finger and BTB domain containing 16 protein and have been related with innate and innate-like lymphoid lineage development in mouse (Mao et al., 2017) and defense response to gram-positive bacteria in goat (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). It suggests a role of this gene in immune systems and adipogeneses processes of BRA breed.

In this study, short ROH were observed in both Portuguese Iberian breeds BAR and BRA, suggesting a successful matting scheme, also highlighted by a lower occurrence of long ROH. Genes located in ROH islands were evaluated and explored throughout their biological processes, which provided genetic insights about the breeds. Genes were associated with

important traits (e.g., *RBM19* with reproductive traits and *DTX1* with immune systems in BAR breed; and *ZBTB16* with immune systems and adipogeneses processes in BRA breed).

Acknowledgments
The authors would like to thank the *Associação dos Criadores de Bovinos de Raça Barrosã* and *Associação Portuguesa de Criadores de Toiros de Lide*, and the *Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)/Instituto Nacional de Ciência e Tecnologia—Ciência Animal (INCT-CA)* for their financial support.

References
Bindea G., Mlecnik B., Hackl H., Charoentong P., Tosolini M., *et al.* (2009). Bioinformatics, 25(8), 1091-1093. https://doi.org/10.1093/bioinformatics/btp101
Do Vale M. (1949) Gado Bissulco — Suínos, Bovinos, Arietinos, Caprinos. Livraria Sá da Costa, Lisboa, PT.
Gibson J., Newton E.M. and Collins A. (2006). Human Molecular Genetics, 15, 789–95. https://doi.org/10.1093/hmg/ddi493
Hayes B.J., Visscher P.M., McPartlan H.C. and Goddard M.E. (2003). Genome Res. 13, 635-643. http://www.genome.org/cgi/doi/10.1101/gr.387103
Kirin M., McQuillan R., Franklin C. S., Campbell H., McKeigue P. M., *et al.* (2010). PLoS One 5:e13996. https://doi.org/10.1371/journal.pone.0013996
Mao A.P., Ishizuka I.E., Kasal D.N., Mandal M., and Bendelac A. (2017). Nature communications, 8(1), 1-14. https://doi.org/10.1038/s41467-017-00882-0
Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A., *et al.* (2007). The American journal of human genetics, 81(3), 559-575. https://doi.org/10.1086/519795
Rout P.K., and Verma M. (2021). Scientific reports, 11(1), 1-16. https://doi.org/10.1038/s41598-021-85094-9
Verardo L.L., e Silva F.F., Machado M.A., do Carmo Panetto J.C., Faza D.R.D.L.R., *et al.* (2021). Frontiers in genetics, 12: 702822. https://doi.org/10.3389/fgene.2021.702822
Wei S., Zhang M., Zheng Y., and Yan, P. (2018). Cellular Physiology and Biochemistry, 48(6), 2528-2538. https://doi.org/10.1159/000492697
Zhang J., Tomasini A.J., and Mayer, A.N. (2008). BMC developmental biology, 8(1), 1-14. https://doi.org/10.1186/1471-213X-8-115
Zhang Q., Guldbrandtsen B., Bosse M., Lund M.S., and Sahana, G. (2015). BMC genomics, 16(1), 1-16. https://doi.org/10.1186/s12864-015-1715-x
Zheng W., Chu Q., Yang L., Sun L., and Xu, T. (2021). PLoS pathogens, 17(3), e1009438. https://doi.org/10.1371/journal.ppat.1009438