



Candidate genes for milk, growth and immune system traits in Brazilian Iberian-derived Locally Adapted cattle breeds

L.L. Verardo^{1*}, M.A. Machado², J.C.C. Panetto² and M.V.G.B. Silva²

¹ Universidade Federal dos Vales do Jequitinhonha e Mucuri, Campus JK, 39100-000, Diamantina, ² Brazilian Agricultural Research Corporation (EMBRAPA), Rua Eugênio do Nascimento, 610, 36038-330, Juiz de Fora – MG, Brazil;

* lucas.verardo@ufvjm.edu.br

Abstract

Along decades of breeding in a variable ecosystem throughout the country, the first taurine animals arrived in Brazil became adapted to a wide range of environments with different levels of improved fitness. We analyzed three cattle breeds representative of Brazilian Iberian-derived Locally Adapted cattle (Curraleiro Pé-Duro CUR with 17 animals, and both Caracu lineages, Caracu Caldeano selected for milk – CCD with 55 animals, and Caracu selected for beef – CCB with 24) aimed to evaluate runs of homozygosity (ROH), identify ROH islands and functionally analyse the identified genes. We observed the occurrence of short ROH islands in all breeds, suggesting a successful mating scheme. Genes located in ROH islands were evaluated and explored throughout their biological processes (e.g. *PRLR* related with milk and growth traits in CCD and CCB; and *CAMKK2* related with immune system in CUR) providing information about the genetic architecture of the breeds.

Introduction

The first taurine animals were introduced in Brazil between sixteenth and seventeenth centuries. Along decades of breeding, these animals adapted to environments with different levels of improved fitness. They become recognized as Iberian-derived locally adapted breeds (Egito et al., 2007; McManus et al., 2009), such as the Caracu lineages (Caracu Caldeano, selected for dairy purpose, and Caracu, selected for beef) and Curraleiro Pé-Duro.

Runs of homozygosity (ROH) patterns may be helpful in understand the history of the population, breeding system, and pattern of geographic subdivision. ROH are continuous homozygous segments of DNA (Gibson et al., 2006), and calculation of that is essential to identify islands and genes with a role in the population. The ROH islands shared among individuals, may be used to identify homozygous regions that result from selection events (Zhang et al., 2015). Moreover, genes identified in those regions can be analysed through gene networks (Verardo et al., 2021) aiming to a better understanding of their biological role in each population.

In this study, we assessed three cattle breeds representative of Brazilian Iberian-derived Locally Adapted cattle (Curraleiro Pé-Duro - CUR and both Caracu lineages, Caracu Caldeano selected for milk – CCD, and Caracu selected for beef - CCB) aimed to evaluate ROH across populations, identify ROH islands and functionally analyse the identified genes.

Materials & Methods

Biological samples were obtained from Brazilian commercial herds, and DNA was extracted at

the Embrapa Dairy Cattle Research Center. All animals were chosen according to their population representativeness in each breed, to avoid bias due to the use of a limited number of families in the population (lineages). Animals were genotyped using the Illumina Bovine HD Chip (777k SNPs). SNP genotypes from sexual chromosomes were excluded. Also, genotypes were excluded when the call rate < 0.95 or minor allele frequency < 0.01. In addition, samples showing overall call rates < 0.90 were excluded from further analysis. A total of 55, 24 and 17 animals genotyped with 624609, 674260 and 612345 markers from CCD, CCB and CUR, respectively, were analysed after quality control.

The ROH were calculated using the software PLINK v1.07 (Purcell et al., 2007) for each population with the following parameters: 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, 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. Principal component analysis (PCA) was used to identify structure in the distribution of genetic variation across breeds. The GenBank annotation based on the ARS-UCD1.2 assembly of the bovine genome was used to identify genes. Gene networks highlighting biological processes among the gene sets of 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

PCA analyses discriminated the three populations (Figure 1). The pattern of ROH differed across breeds (Figure 2A). CUR breed showed higher percentage of long-range ROH (> 31Mb; ~5%), while CCB showed the higher percentage of short range ROH (<5 Mb; ~90%). Based on ROH islands, regions were shared by more than 75% of the individuals in each breed (Table 1). Gene networks highlighting biological associations were generated based on genes found on ROH islands (Figure 2B). Several biological processes were highlighted.

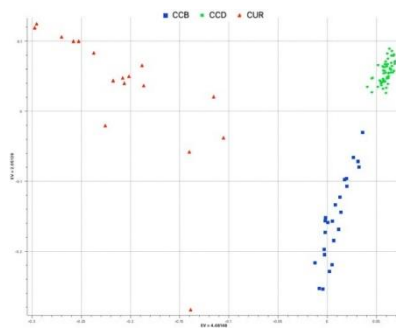


Figure 1. PCA analyses of three Brazilian Iberian-derived naturalized cattle breeds. CCA: Caracu Calderano (green); CCB: Caracu (blue); CUR: Carraleiro Pé-Duro (red).

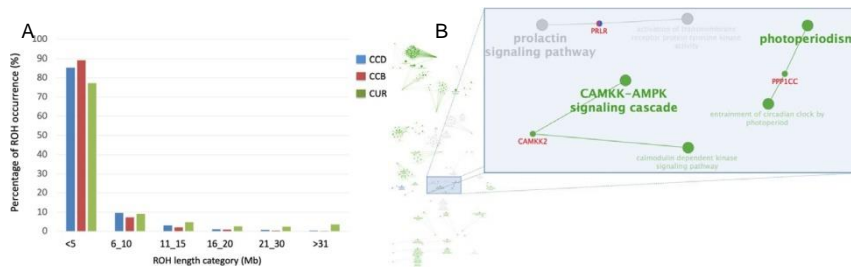


Figure 2. A) Percentage of runs of homozygosity (ROH) per length categories in three Brazilian Iberian-derived naturalized cattle breeds. The sum of ROH was calculated per animal within each ROH length category. B) Gene networks highlighting biological associations of genes found on ROH islands in three Brazilian Iberian-derived naturalized cattle breeds. Blue and red nodes are genes identified in CCA and CCB, respectively. Green nodes are biological processes related with genes identified in CUR breed. Grey nodes are biological processes shared between two breeds genes. Enriched terms are in bold.

Table 1. 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 three Brazilian Iberian-derived naturalized cattle breeds.

Breed	Chr.	Begin	End	Length (bp)	N _o SNPs	ROH freq. (%)	Number of genes	Annotations	
CCD	2	0	4	1	3	75.55	3	<i>PRLR, RAI14, RAD1, CIQTNF3, LOC112443006,</i>	
	3	0	3	1	5		5	<i>SLC45A2, ADAMTS12, BRX1, TTC23L, AMACR, RXFP3, AGXT2, DNAJC21, LOC107131570,</i>	
	8	4	5	1	1		1	<i>LOC112443059</i>	
	9	0	0	0	0		0	0	
	0	8	4	1	4		1	<i>LOC112443042</i>	
	5	9	9	0	0		0	0	
	8	7	7	0	0		0	0	
	4	0	0	1	8		1	1	<i>NPR3, TARS1,</i>
	0	0	1	3	6		6	6	<i>ADAMTS12, LOC107131571</i>
	4	0	9	0	0		0	0	
2	1	1	9	9	9	9			
4	0	0	0	0	0	0			
8	8	8	8	8	8	8			
8	8	8	8	8	8	8			
9	9	9	9	9	9	9			
CCB	2	0	4	1	4	75.00	4	<i>PRLR, RAI14, RAD1, CIQTNF3, LOC112443006,</i>	
	3	0	6	2	7		7	<i>SLC45A2, ADAMTS12, BRX1, TARS1, TTC23L,</i>	
	8	2	3	9	9		9	9	<i>AMACR, RXFP3, AGXT2, DNAJC21,</i>
	8	3	0	8	3		3	3	
	9	0	0	3	3		3	3	
	3	9	1	8	8		8	8	
	2	1	1	1	1		1	1	
	5	3	3	3	3		3	3	
	3	3	3	3	3		3	3	
	3	3	3	3	3		3	3	
CUR	17	53,133,247	54,716,438	44,354	10	75.00	10	<i>LOC107131571, LOC107131570, TRNAT-UGU, LOC112443059, LOC112443042</i>	
	1,583,191	181				82.35	82.35	<i>SKP2, NADK2</i>	
CCB	17	53,133,247	54,716,438	44,354	10	75.00	10	<i>KDM2B, RHOF, CAMKK2, MORN3, DIABLO, MLXIP, HVCN1, ANAPC5, IFTS1, TCTN1, RAD9B, ATP2A2, TMEM120B, RNF34, GPN3, PPP1CC, ORAI1, P2RX7, HPD, P2RX4, FAM216A, VPS29, MYL2, BCL7A, SETD1B,</i>	
	1,583,191	181				82.35	82.35	<i>VPS33A, ARPC3, PSMD9, LRRC43, CFAP251, ANAPC7, B3GNT4, PPTC7, CCDC63, LOC100300938, LOC104974641, LOC104968522, LOC101905029, LOC101904098, LOC101903644, LOC100335936, LOC783255, TRNAE-UUC, LOC112442115, LOC112442104, LOC112442039, LOC112442038, LOC112442004</i>	

Discussion

We evaluated the run of homozygosity in three Brazilian Iberian-derived Locally Adapted cattle. Since 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 the three breeds showed a similar occurrence of short ROH. This indicates a proper matting scheme to prevent higher inbreeding.

RHO islands were identified in the three breeds from which genes were retrieved and their biological processes analysed. In Caracu lineages (CCD and CCB), we identified islands in chromosome 20 with a gene associated, for example, with Prolactin signalling pathway (*PRLR*). *PRLR* encodes for 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). This gene was also cited to present ‘slick coat’ mutations that confer termotolerance to tropically adapted *Bos taurus* cattle (Littlejohn et al. 2014; Florez Murillo et al. 2021). Since CCD and CCB has been selected for dairy and beef, respectively, it makes sense to be fixing this gene, with a possible role in milking, growth and heat resilience. ROH island was observed on chromosome 17 of CUR breed. The biological process network a gene (*CAMKK2*) related with CAMKK-AMPK signalling pathway. This gene encodes for the calcium/calmodulin dependent protein kinase kinase 2 protein and have been cited to be involved in regulating important physiological and pathophysiological processes, including energy balance, adiposity, glucose homeostasis and inflammation (Raciopi and Means, 2012). Thus, it is possible that this gene may play a role in some of these processes in CUR breed.

We observed the occurrence of short ROH islands in all breeds, suggesting a successful matting scheme, also highlighted by a lower occurrence of long ROH. Genes located in ROH islands were explored via biological processes, provid genetic insights about the breeds.

Acknowledgments

The authors would like to thank 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). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*, 25(8), 1091-1093. <https://doi.org/10.1093/bioinformatics/btp101>
- Egito A.A., Paiva S.R., Maria do Socorro M., Mariante A.S., Almeida L.D., *et al.*, (2007). Microsatellite based genetic diversity and relationships among ten Creole and commercial cattle breeds raised in Brazil. *BMC genetics*, 8(1), 1-14. <https://doi.org/10.1186/1471-2156-8-83>
- Gibson J., Newton E.M. and Collins A. (2006). Extended tracts of homozygosity in outbred human populations. *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). Novel multilocus measure of linkage disequilibrium to estimate past effective population size. *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). Genomic runs of homozygosity record population history and consanguinity. *PLoS One* 5:e13996. <https://doi.org/10.1371/journal.pone.0013996>
- Littlejohn M.D., Henty K.M., Tiplady K., Johnson T., Harland C. *et al.*, (2014). Nature communications, 5(1), 1-8. <https://doi.org/10.1038/ncomms6861>
- Lü A., Hu X., Chen H., Dong Y., and Pang, Y. (2011). Single nucleotide polymorphisms of

the prolactin receptor (PRLR) gene and its association with growth traits in Chinese cattle. *Molecular biology reports*, 38(1), 261-266. <https://doi.org/10.1007/s11033-010-0103-5> McManus C., Prescott E., Paludo G.R., Bianchini E., Louvandini, H., et al., (2009). Heat tolerance in naturalized Brazilian cattle breeds. *Livestock Science*, 120(3), 256-264.

<https://doi.org/10.1016/j.livsci.2008.07.01>

Murillo J.M.F., Landaeta-Hernández A.J., Kim E.S., Bostrom J.R., Larson S.A., et al., (2021). *Animal genetics*, 52(1), 132-134. <https://doi.org/10.1111/age.13027>

Peripolli E., Munari D.P., Silva M.V.G.B., Lima A.L.F., Irgang R., et al. (2017). Runs of homozygosity: current knowledge and applications in livestock. *Animal genetics*, 48(3), 255-271. <https://doi.org/10.1111/age.12526>

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics*, 81(3), 559-575. <https://doi.org/10.1086/519795>

Racioppi, L., and Means, A. R. (2012). Calcium/calmodulin-dependent protein kinase kinase 2: roles in signaling and pathophysiology. *Journal of Biological Chemistry*, 287(38), 31658- 31665. <https://doi.org/10.1074/jbc.R112.356485>

Verardo L.L., e Silva F.F., Machado M.A., do Carmo Panetto J.C., Faza D.R.D.L.R., et al. (2021). Genome-Wide Analyses Reveal the Genetic Architecture and Candidate Genes of Indicine, Taurine, Synthetic Crossbreds, and Locally Adapted Cattle in Brazil. *Frontiers in genetics*, 12: 702822. <https://doi.org/10.3389/fgene.2021.702822>

Zhang J., Zan L., Fang P., Zhang F., Shen G., et al., (2008). Genetic variation of PRLR gene and association with milk performance traits in dairy cattle. *Canadian journal of animal science*, 88(1), 33-39. <https://doi.org/10.4141/CJAS07052>

Zhang Q., Guldbbrandtsen B., Bosse M., Lund M.S., and Sahana, G. (2015). Runs of homozygosity and distribution of functional variants in the cattle genome. *BMC genomics*, 16(1), 1-16. <https://doi.org/10.1186/s12864-015-1715-x>