Livestock Science 197 (2017) 36–45

Contents lists available at ScienceDirect

Livestock Science

journal homepage: www.elsevier.com/locate/livsci

Genome-wide search for signatures of selection in three major Brazilian locally adapted sheep breeds

João José de Simoni Gouveia[a,b,*], Samuel Rezende Paiva[c], Concepta M. McManus[d], Alexandre Rodrigues Caetano[e], James W. Kijas[f], Olivardo Facó[g], Hymerson Costa Azevedo[h], Adriana Mello de Araujo[i], Carlos José Hoff de Souza[j], Michel Eduardo B. Yamagishi[k], Paulo Luiz Souza Carneiro[l], Raimundo Nonato Braga Lôbo[g], Sônia Maria Pinheiro de Oliveira[m], Marcos Vinicius G.B. da Silva[n]

[a] Colegiado Acadêmico de Zootecnia, Universidade Federal do Vale do São Francisco, Petrolina/PE 56300-000, Brazil
[b] Programa de Doutorado Integrado em Zootecnia (PDIZ-UFC/UFPB/UFRPE), Universidade Federal do Ceará, Fortaleza/CE 60020-181, Brazil
[c] Secretaria de Relações Internacionais, Embrapa Sede, Brasília/DF 70770-901, Brazil
[d] INCT Pecuária, Faculdade de Agronomia e Medicina Veterinária, Universidade de Brasília, Brasília/DF 70910-900, Brazil
[e] Embrapa Recursos Genéticos e Biotecnologia, Brasília/DF 70770-917, Brazil
[f] CSIRO Animal, Food and Health Sciences, St Lucia, Brisbane, Queensland 4067, Australia
[g] Embrapa Caprinos e Ovinos, Sobral/CE 62011-970, Brazil
[h] Embrapa Tabuleiros Costeiros, Aracaju/SE 49025-040, Brazil
[i] Embrapa Meio-Norte, Teresina/PI 64006-220, Brazil
[j] Embrapa Pecuária Sul, Bagé/RS 96401-970, Brazil
[k] Embrapa Informática Agropecuária, Campinas/SP 13083-970, Brazil
[l] Departamento de Ciências Biológicas, Universidade Estadual do Sudoeste da Bahia, Jequié/BA 45200-000, Brazil
[m] Departamento de Zootecnia, Universidade Federal do Ceará, Fortaleza/CE 60020-181, Brazil
[n] Embrapa Gado de Leite, Juiz de Fora/MG 36038-330, Brazil

ARTICLE INFO

Keywords:
Ovis aries
Selective sweep
Climate change
Conservation animal genetic resources
SNP markers

ABSTRACT

The study of locally adapted breeds has the potential to underpin the discovery of genes involved in economically and ecologically important traits. Brazilian locally adapted sheep breeds have distinctive characteristics that could be of value for specialized production systems. Therefore, the main objective of the present study was to identify genomic regions that may have been under selection and therefore may explain ecological and production differences observed among three important Brazilian locally adapted sheep breeds. Animals from the Brazilian Creole, Morada Nova and Santa Ines breeds were genotyped using the Illumina Ovine SNP50 BeadChip. The identification of selection signatures was based on two groups of methodologies: differentiation among populations (F_{ST}) and linkage disequilibrium (iHS and RsB). Taken together, these analyses allowed for the identification of 86 candidate genes. Functional analysis revealed genes related to immunity, nervous system development, reproduction and sensory perception. A number of genes are of particular interest including: *RXFP2*, which has recently been associated with the presence/absence and morphology of horns in sheep; the *TRPM8* gene, involved in regulation of body temperature at low temperatures; *DIS3L2*, *PLAG1* and *NIPBL*, associated with height variation; and finally, *SPEF2* and *SPAG6*, important for spermatogenesis. Selective sweeps were identified using multiple methods, and in a number of cases sweep regions contained genes with a demonstrated role in phenotypic variation. The genomic distribution of the sweep regions differed between populations, suggesting that breed specific signatures were successfully identified that may reflect the consequence of local adaptation.

1. Introduction

Identification of genomic regions subject to selection in livestock may assist in both understanding the processes involved in genome evolution as well as in the discovery and validation of genomic regions involved in the manifestation of traits of economic and ecological

* Corresponding author at: Colegiado Acadêmico de Zootecnia, Universidade Federal do Vale do São Francisco, Petrolina/PE 56300-000, Brazil.
E-mail address: joao.gouveia@univasf.edu.br (J.J. de Simoni Gouveia).

http://dx.doi.org/10.1016/j.livsci.2017.01.006
Received 3 March 2016; Received in revised form 5 January 2017; Accepted 6 January 2017
1871-1413/

interest (Hayes et al., 2008; Nielsen, 2001, 2005; Schlötterer, 2003).

The domestication process, breed development and the more recent progress in animal breeding methods such as genetic improvement programs, have led to species modifications which are of interest to animal science from morphological, behavioral, productive, adaptive and, consequently, genetic points of view. A wide variety of breeds and genetic groups adapted to diverse environmental conditions can be observed for farm animal species (Diamond, 2002; Toro and Mäki-Tanila, 2007; Flori et al., 2009).

Brazilian sheep are descendants of animals imported during the colonization period and, since then, have been subject to systematic and non-systematic evolutionary processes. This has led to the development of local breeds adapted to the diverse Brazilian environmental conditions (Mariante et al., 2009). The Brazilian hair sheep breeds, for example, have better quality pelts (Jacinto et al., 2004, 2011) and lower levels of infection from gastrointestinal nematodes when compared to exotic wool breeds (Amarante et al., 2009; McManus et al., 2008).

The recent publication of a draft sheep genome sequence (Jiang et al., 2014), coupled with the availability of SNP genotyping platforms, has meant studies that aim to identify genomic regions under selection are now possible for sheep. These tools were used in a study of 74 sheep breeds that tested nearly 50,000 SNP loci to identify regions likely to have undergone positive selection across sheep breeds (Kijas et al., 2012). The study, however, did not search for evidence of selection that may be breed specific to one or a few locally adapted breeds. Therefore, the main objective of this study was to identify genomic regions that may have been under selection in three major Brazilian locally adapted sheep breeds. Candidate genes located in identified regions were listed, based on publically available annotation information.

2. Material and methods

2.1. Animals and SNPs

Ninety-two animals from Brazilian Creole (BCS), Morada Nova (BMN) and Santa Ines (BSI), the three main Brazilian locally adapted sheep breeds, were genotyped as part of the International Sheep Genomics Consortium (Kijas et al., 2012), using the Illumina Ovine SNP50 BeadChip.. Of these, a total of 87 unrelated individuals (22 Brazilian Creole, 20 Morada Nova and 45 Santa Ines) were used to carry out this study. Information related to individual sampling locations is shown in Table 1.

The Illumina Ovine SNP50 BeadChip. contains 54,241 SNPs from the sheep genome. Pre-processing and filtering of the raw data resulted in a total of 49,034 high quality useful SNPs (Kijas et al., 2012). Filtered data were then subjected to additional processing to exclude non-autosomal markers, according to version 3.1 of the sheep genome (ISGC et al., 2010), and markers with MAF < 0.05 (considering the 87 individuals used in this study). A total of 44,850 SNPs remained at the end of processing.

Table 1
Information about the individuals used in the study.

Breed	N	Herd	Location	Conservation Program*
Brazilian Creole	5	BCS_CAS	Caçapava do Sul, RS	Yes
	6	BCS_EMB	Bagé, RS	Yes
	5	BCS_LAG	Lages, SC	No
	6	BCS_POA	Ponte Alta, SC	No
Morada Nova	13	BMN_EBD	Itapetinga, BA	Yes
	7	BMN_EMB	Sobral, CE	Yes
Santa Ines	6	BSI_EBD	Itapetinga, BA	Yes
	5	BSI_ECO	Sobral, CE	Yes
	13	BSI_EMN	Teresina, PI	Yes
	10	BSI_ETC	Aracaju, SE	Yes
	11	BSI_GAA	Inhumas, GO	No

* Herds belonging to Embrapa conservation nuclei.

2.2. Population structure

Principal component analysis (PCA) was performed using EIGENSOFT software v5.0.2 (Patterson et al., 2006) based on a subset containing 5620 unlinked SNPs obtained using the command –indep-pairwise 20 5 0.05 of PLINK (Purcell et al., 2007).

2.3. Identification of signatures of selection based on differentiation among populations

Analysis for selective sweeps was performed using the BayeScan v2.1 software (Foll and Gaggiotti, 2008). To control the number of false positives, the *a priori* probability of the neutral model in relation to the model assuming selection was considered as equal to 10 and the definition of the significant SNPs was carried out considering the q-value (the minimum FDR for a *locus* to be considered as significant) as equal to 0.01. The F_{ST} analyses were performed considering the three populations together and also separately for each pairwise comparison.

Additionally, the F_{ST} analysis was performed comparing the north-eastern hair breeds (Morada Nova and Santa Ines) against the southern wool breed (Brazilian Creole).

2.4. Identification of signatures of selection based on linkage disequilibrium

The haplotypes of each individual were separately reconstructed for each chromosome within each breed, using fastPHASE software (Scheet and Stephens, 2006). Population recombination rates ($\rho=4\times N_e\times r$) were estimated separately for each chromosome within each breed, using LDhat (McVean et al., 2004). These estimates were converted into genetic distances (cM) considering the effective population size of 371 individuals for Brazilian Creole, 327 for Morada Nova and 520 for Santa Ines (Kijas et al., 2012). A genetic reference map was created from the arithmetic mean of the maps from each breed. To assign the ancestral status at each marker, SNP genotyping was performed on animals from nine out-group species (Supplementary Table 1). Genotypes with GC score < 0.8 were discarded, before alleles were compared across species. A total of 35,008 SNP had their ancestral allele assigned where segregation was allowed in no more than one out-group species, and no out group was fixed for the alternative allele.

Identification of selection signatures based on linkage disequilibrium was performed using two complementary methodologies: iHS proposed by Voight et al. (2006), and RsB proposed by Tang et al. (2007). These two methods are derived from the long-range haplotype (LRH) test, proposed by Sabeti et al. (2002), and they are based on the extended haplotype homozygosity (EHH) estimates of linkage disequilibrium.

Both the iHS and RsB analyses were performed using the rehh package (Gautier and Vitalis, 2012) in R. To identify loci subject to selection, the iHS and RsB values were transformed into piHS ($piHS = -\log[1 - 2\times(\Phi_{(iHS)}-0,5)]$) and pRsB ($pRsb = -\log[1 - 2\times(\Phi_{(Rsb)}-0,5)]$), where $\Phi_{(x)}$ represents the Gaussian cumulative distribution function (Gautier and Vitalis, 2012). Assuming that iHS and RsB values are normally distributed under neutrality, piHS and pRsB values may be interpreted as $-\log 10(P-value)$.

The iHS analysis was performed separately for each of the three breeds studied and the SNPs that exhibited piHS ≥4 ($P-value = 0.0001$) were considered significant. RsB analysis was carried out for each two-by-two comparison between the breeds studied, and the SNPs that exhibited pRsB ≥4 ($P-value = 0.0001$) were considered significant.

2.5. Annotation of significant regions

SNPs showing statistically significant F_{ST}, iHS and RsB values were used to identify genes which are candidates under selection. Identification of these genes was made based on version 3.1 of the sheep genome (Jiang et al., 2014), downloaded from http://www.livestockgenomics.csiro.au/sheep/oar3.1.php. For each analysis, the significant SNPs that were less than 50 kb apart were clustered in a single region and the SNP of greatest statistical value (F_{ST}, iHS or RsB) within each region was considered as the significant SNP for the region.

When a significant SNP was contained within a gene, the gene was considered as a candidate under selection. For the significant SNPs not located within genes, the genes nearest this SNP were considered candidates, up to a maximum distance of 25 kb.

2.6. Functional analysis of the candidate genes

The DAVID platform (Huang et al., 2009) was used to perform functional analysis of the genes identified as candidates under selection. In addition, two text mining tools were used to assist in the search for scientific literature related to the candidate genes: IHOP (Hoffmann and Valencia, 2004) and CHILIBOT (Chen and Sharp, 2004).

3. Results

3.1. Population structure

To access the genetic structure of the studied populations, the PCA was performed using 5620 SNPs selected to be unlinked. The first three components accounted for 7.93% (PC1), 4.58% (PC2) and 3.46% (PC3) of the variation. The first component clearly separates the northeastern hair breeds (Morada Nova and Santa Ines) from the southern wool breed (Brazilian Creole). The second component permits the separation within the hair breeds and the combination of the first and the third components allows the insight into a possible substructure within the Brazilian Creole breed (Fig. 1).

3.2. Selection identified using F_{ST}

SNP Genotypes were used to estimate allele differentiation (measured as F_{ST}) in pairwise comparison between breeds, and also using all breeds together. The genome wide distribution of F_{ST} is shown in Fig. 2, and clear signals can be seen that differs between comparisons suggesting the identification of breed specific selection. A total of 58 significant SNPs (q-value ≤0.01) were identified in the single analysis considering all three breeds (top panel, Fig. 2; Supplementary Table 2). Pairwise comparisons revealed 33 significant SNPs between Brazilian Creole and Morada Nova, 53 between Brazilian Creole and Santa Ines, and 23 between Morada Nova and Santa Ines (Supplementary Table 2). Significant SNPs were identified on all chromosomes except for OAR21 and OAR23, and clusters that defined clear peaks were found on OAR2, OAR3, OAR16 and OAR25 (Fig. 2).

Results from each pair-wise breed analysis were compared to identify significant SNPs in both the analysis considering all three breeds and at least one pairwise comparison (Supplementary Fig. 2). A total of 42 SNPs distributed across 40 distinct regions from 17 chromosomes met these criteria and were used for defining genomic regions and were manually annotated (Supplementary Table 3). In summary, the following genes were identified in regions under selection by the F_{ST} analysis: *TRPM8* (OAR1) observed for the comparison between BCS x BSI; *RXFP2* (OAR10) for the comparison BCS x BSI; *WDR70* (OAR16) for the comparison between BCS x BSI and *TARBP1* (OAR25) for the comparison between BCS x BMN (Table 2).

The analysis comparing the northeastern hair breeds (Morada Nova and Santa Ines) against the southern wool breed (Brazilian Creole) revealed 46 significant SNPs (q-value ≤0.01) (Supplementary Fig. 7 and Supplementary Table 9) distributed across 44 distinct regions (Supplementary Table 10). Inside these regions, 21 genes were identified as candidates to be under selection (Supplementary Table 11).

3.3. Selection identified using iHS

A separate analysis, based on phased haplotypes, was performed using iHS. Analysis of the Brazilian Creole identified 13 significant SNPs (piHS ≥4) distributed across 8 chromosomes (Supplementary Table 4). In the Morada Nova, eight significant SNPs were identified on five chromossomes, and in the Santa Ines, 70 significant SNPs were identified on seven chromosomes (Supplementary Table 4).

Clusters of significant SNPs were observed on chromosomes OAR2, OAR3 and OAR16 (Fig. 3), in a total of 13 regions in the Brazilian Creole, eight regions in the Morada Nova, and 59 regions in the Santa Ines (Supplementary Table 5). The main regions identified in iHS analysis were located on chromosomes OAR3 and OAR16.

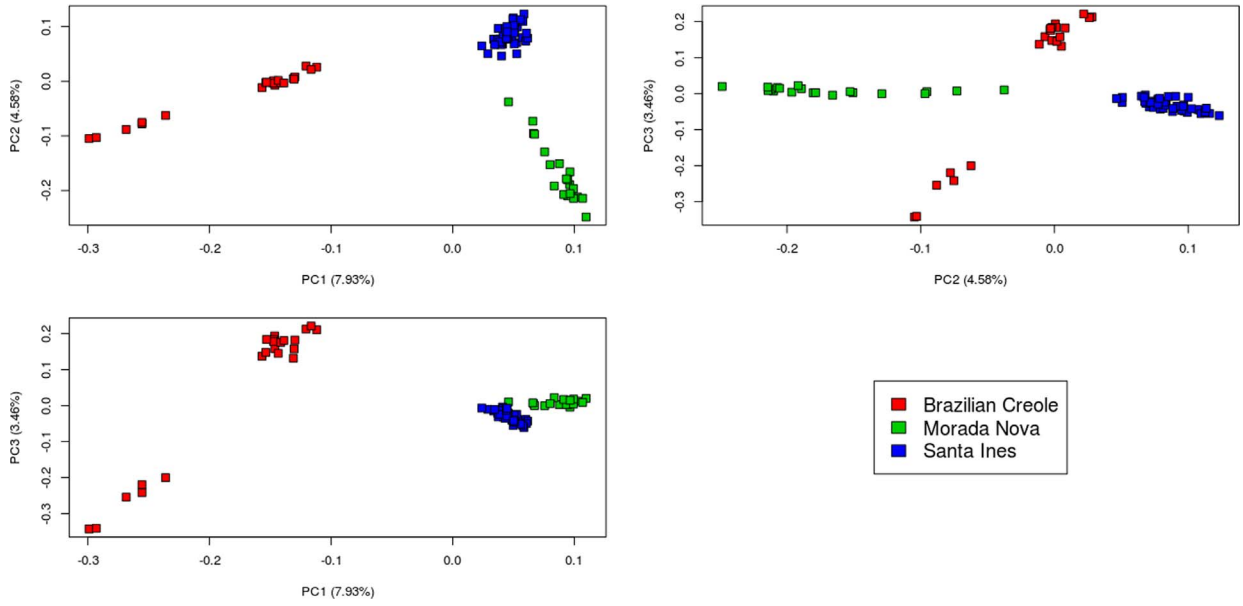

Fig. 1. Principal component analysis results based on a subset containing 5620 SNPs. Each individual is plotted according to its coordinates on the biplot.

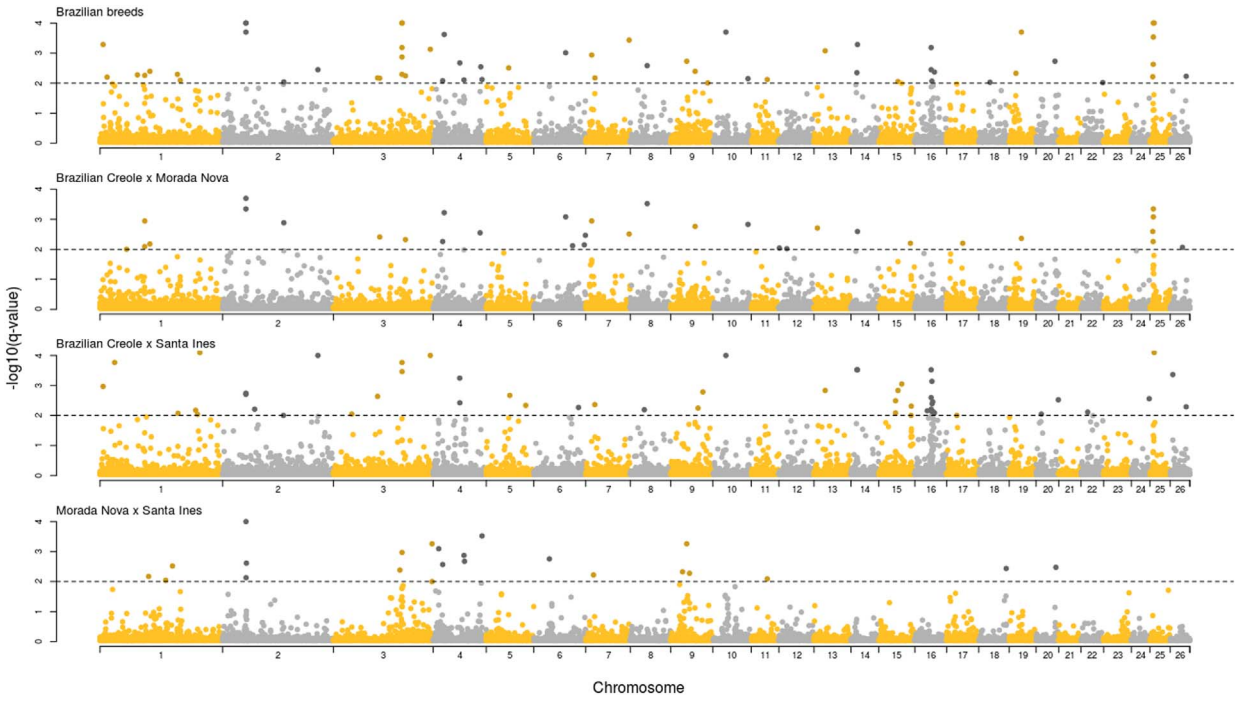

Fig. 2. Manhattan plot for FST analysis. Odd-numbered chromosomes are represented in yellow while even-numbered chromosomes are represented in gray. The darker points above the dotted line indicate the significant SNPs. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

On chromosome OAR3, a cluster of three significant adjacent SNPs was identified in the Santa Ines and a cluster of two significant SNPs was identified in the Morada Nova (Supplementary Table 5). Inside these regions, the gene *GNS* (Santa Ines) and *MAPKAP1* (Morada Nova) were identified as candidates to be under selection (Table 2). On chromosome OAR16, six regions containing clusters of significant SNPs were identified also in Santa Ines (Supplementary Table 5) and the genes *WDR70, MGC134093* and *SPEF2*) were identified as candidates.

3.4. Selection Identified using RsB

RsB analysis revealed 21 significant SNPs (pRsB ≥4) distributed across eight chromosomes in the Brazilian Creole x Morada Nova comparison. In the Brazilian Creole x Santa Ines comparison, 204 significant SNPs, distributed across six chromosomes were identified. For the Morada Nova x Santa Ines comparison, 26 significant SNPs distributed across six chromosomes were identified (Supplementary Table 4). Clusters of significant SNPs were observed on chromosomes OAR2, OAR3, OAR8, OAR10, OAR12, OAR13 and OAR16 (Fig. 4). Significant SNPs were clustered in 19, 124 and 22 regions, when considering Brazilian Creole x Morada Nova, Brazilian Creole x Santa Ines and Morada Nova x Santa Ines comparisons, respectively (Supplementary Table 7). The main regions identified through the RsB analyses contain the genes: *DIS3L2* and *FGR* (OAR2), identified in the BCS x BSI comparison; *MANEA* (OAR8), identified in the BMN x BSI comparison; *LIFR*, *WDR70* and *MGC134093* (OAR16) identified for the BCS x BSI comparison and *SPEF2* (OAR16), identified both for BCS x BMN and BCS x BSI comparisons.

3.5. Overlap between selection metrics

In total, the different analyses revealed signatures of selection in 246 genomic regions across 24 of the 26 autosomal chromosomes of the sheep genome. Coincident signatures were identified by more than one methodology on chromosomes OAR3, OAR8, OAR9, OAR10, OAR13 and OAR16. Of these, 5 were identified by three methodologies, 37 by two and 204 by only one of the methodologies (Supplementary Table 8).

Mannual annotation of these regions revealed a total of 86 candidate genes, 61 of which contained SNP identified in the original analysis (Table 2). Three genes (*CDH6*, *SPEF2* and *WDR70*) were identified by all three methodologies (F_{ST}, iHS and RsB), while seven genes (*GNS*, *KRT8*, *LOC515823*, *C9*, *FYB*,*MGC134093* and *SPAG6*) were identified with at least two methodologies.

3.6. Functional analysis of the candidate genes

Functional analysis revealed that the identified candidate genes have been associated with diverse biological functions, such as immunity (*ADAMTS12, AMHR2, AQP3, ARHGAP24, C6, C9, COL1A1, COPS7B, DAB2, DROSHA, FGR, FYB, GDNF, GOLPH3, GPR158, GPR65, IL1RL1, KRT8, LANCL2, MACROD2, MAPKAP1, MSRB3, NIPBL, PIK3CB, PLCB1, SKAP2, SMAD6, SNX27, SPEF2* and *TRPM8*), nervous system development (*CNTNAP2, FUT9, GDNF, ISPD, LIFR, MACROD2, MAPKAP1, NIPBL, NTNG1, PLAG1, PLCB1* and *SPEF2*), sensory perception (*COL1A1, NIPBL, PDE6D* and *TRPM8*) and reproduction (*AMHR2, KRT8, NIPBL, PLAG1, PLCB1, RXFP2, SP1, SPAG6* and *SPEF2*).

4. Discussion

A large number of selection signatures (246) were identified in the present study in comparison with previous reports performed with multiple sheep breeds (Kijas et al., 2012; Moradi et al., 2012; Fariello et al., 2013). This may reflect application of multiple analytical approaches which differ in their sensitivity. The great majority (204) of the detected signatures were identified using only a single methodology, and these are likely to be enriched for false positives when compared to regions identified by multiple approaches. Thirty seven coincident signatures were detected by two methodologies and five coincident signatures were detected using all three approaches. Discrepancies found among results may arise from the fact that each methodology used is based on different parameters, therefore capturing distinct traces left in the genome by selection pressures over time (Sabeti et al., 2002; Beaumont, 2005; Storz, 2005; Voight et al., 2006; Tang et al., 2007).

Table 2
Putative genes under selection.

Chromosome	SNP	SNP in gene	Method	Comparison	RefSeq	Gene symbol	Gene name
1	OAR1_121001304	No	F_{ST}	bcsxbmn	NM_001083790	C3H1orf226	chromosome 3 open reading frame, human C1orf226
1	s16767	No	RsB	bmnxbsi	XM_002686328; XM_616138	CACHD1	cache domain containing 1
1	OAR1_241765164	Yes	F_{ST}	FST(Hair X Wool)	NM_001075420	NMD3	NMD3 ribosome export adaptor
1	OAR1_89287467	No	F_{ST}	FST(Hair X Wool)	NM_001080369	NTNG1	netrin G1
1	OAR1_268303279_X	Yes	F_{ST}	FST(Hair X Wool)	NM_001206047	PIK3CB	phosphatidylinositol–4,5-bisphosphate 3-kinase catalytic subunit beta
1	OAR1_127962239	Yes	RsB	bcsxbmn	NM_001192122	SFT2D2	SFT2 domain containing 2
1	OAR1_107797839	Yes	F_{ST}	bcsxbmn	NM_001098956	SNX27	sorting nexin family member 27
1	s11273	No	F_{ST}	FST(Hair X Wool)	NM_001205892	TMPRSS7	transmembrane protease, serine 7
1	s49628	Yes	F_{ST}	bcsxbsi;FST(Hair X Wool)	NM_001206066	TRPM8	transient receptor potential cation channel, subfamily M, member 8
2	s23721	Yes	RsB	bcsxbsi	NM_001206409	AHDC1	AT hook, DNA binding motif, containing 1
2	s36668	No	iHS	bsi	NM_001079794	AQP3	aquaporin 3 (Gill blood group)
2	OAR2_200920481	Yes	iHS	bcs	XM_003581848	CNTNAP5	contactin associated protein-like 5
2	OAR2_245680195	No	RsB	bcsxbsi	NM_001046612	COPS7B	COP9 signalosome subunit 7B
2	s67306	Yes	RsB	bcsxbsi	NM_001035346	DHDDS	dehydrodolichyl diphosphate synthase
2	OAR2_246045571	Yes	RsB	bcsxbsi	NM_001192506	DIS3L2	DIS3 like 3'-5' exoribonuclease 2
2	s20468	No	F_{ST}	bmnxbsi	XM_002689701; XM_001250503	FAM221B	family with sequence similarity 221, member B
2	OAR2_251955154	Yes	RsB	bcsxbsi	NM_001098991	FGR	FGR proto-oncogene, Src family tyrosine kinase
2	OAR2_226474680	Yes	F_{ST}	bcsxbsi;FST(Hair X Wool)	XM_002685559; XM_003581860	IKZF2	IKAROS family zinc finger 2 (Helios)
2	OAR2_245597717	No	RsB	bcsxbsi	NM_174420	PDE6D	phosphodiesterase 6D, cGMP-specific, rod, delta
2	s35831	No	iHS	bsi	NM_001034408	POLR1E	polymerase (RNA) I polypeptide E, 53 kDa
3	OAR3_142056197	Yes	iHS	bsi	NM_001205328	AMHR2	anti-Mullerian hormone receptor, type II
3	s22657; s68180	No	RsB	bmnxbsi	NM_001046435	CALCOCO1	calcium binding and coiled-coil domain 1
3	OAR3_168639353	Yes	iHS	bsi	XM_002684155; XM_002687401	FAM19A2	family with sequence similarity 19 (chemokine (C-C motif)-like), member A2
3	OAR3_165741097; OAR3_165766186; OAR3_165801399	Yes	iHS; RsB	iHS (bsi); RsB (bcsxbsi)	NM_001075562	GNS	glucosamine (N-acetyl)-6-sulfatase
3	OAR3_105506271	Yes	F_{ST}	bcsxbsi	NM_001206302	IL1RL1	interleukin 1 receptor-like 1
3	s20793	Yes	iHS; RsB	iHS (bsi); RsB (bmnxbsi)	NM_001033610	KRT8	keratin 8
3	s20793	Yes	iHS; RsB	iHS (bsi); RsB (bmnxbsi)	XR_083785; XR_082839	LOC515823	keratin, type II cytoskeletal 8 pseudogene
3	s42902; OAR3_11033570	Yes	iHS	bmn	NM_001081603	MAPKAP1	mitogen-activated protein kinase associated protein 1
3	OAR3_173071993	Yes	F_{ST}	bcsxbmn	NM_001038091	MARS	methionyl-tRNA synthetase
3	OAR3_138290871	Yes	iHS	bsi	NM_001037622	MRPL42	mitochondrial ribosomal protein L42
3	OAR3_165050963	Yes	F_{ST}	bcsxbsi; bmnxbsi	XM_002687541; XM_001250112; XM_002687544; XM_001250168; XM_002687543; XM_001250065	MSRB3	methionine sulfoxide reductase B3
3	s62626	Yes	F_{ST}	bcsxbsi;FST(Hair X Wool)	NM_001046468	PACSIN2	protein kinase C and casein kinase substrate in neurons 2
3	OAR3_142190582; s62934	No	RsB	bmnxbsi	NM_001078027	SP1	Sp1 transcription factor
3	s09462	Yes	F_{ST}	bcsxbmn;FST(Hair X Wool)	NM_001102086	STARD7	StAR-related lipid transfer (START) domain containing 7
3	OAR3_114955879	No	RsB	bcsxbsi	XM_003582172	TPH2	tryptophan hydroxylase 2
3	OAR3_168114192	Yes	iHS	bsi	NM_001046430; XM_003586089	USP15	ubiquitin specific peptidase 15
4	OAR4_117150483	Yes	F_{ST}	bmnxbsi	XM_002687137	CNTNAP2	contactin associated protein-like 2
4	OAR4_25990541	Yes	F_{ST}	bcsxbmn	XM_002686679	ISPD	isoprenoid synthase domain containing
4	OAR4_114392312	Yes	F_{ST}	FST(Hair X Wool)	XM_001788707	LOC519916	olfactory receptor 2A1/2A42
4	OAR4_73170316	Yes	F_{ST}	bmnxbsi	NM_001038214	SKAP2	src kinase associated phosphoprotein 2
6	OAR6_110447914	Yes	F_{ST}	FST(Hair X Wool)	NM_001102234	ARHGAP24	Rho GTPase activating protein 24

(continued on next page)

Table 2 *(continued)*

Chromosome	SNP	SNP in gene	Method	Comparison	RefSeq	Gene symbol	Gene name
6	OAR6_40311379	Yes	iHS	bsi	NM_001083663	HERC3	HECT and RLD domain containing E3 ubiquitin protein ligase 3
7	OAR7_106207879	Yes	F_{ST}	bcsxbmn	NM_001101303	GPR65	G protein-coupled receptor 65
7	OAR7_13638892	No	F_{ST}	bcsxbmn	NM_001206145	SMAD6	SMAD family member 6
8	OAR8_40885595	Yes	F_{ST}	bcsxbmn	NM_001076059	FBXL4	F-box and leucine-rich repeat protein 4
8	OAR8_43900420	No	RsB	bcsxbmn	NM_174735	FUT9	fucosyltransferase 9 (alpha (1,3) fucosyltransferase)
8	OAR8_44479113; OAR8_44532355; OAR8_44575878_X	No	RsB	bmnxbsi	NM_001206725	MANEA	mannosidase, endo-alpha
9	OAR9_38670488	No	RsB	bmnxbsi	NM_001046335	IMPAD1	inositol monophosphatase domain containing 1
9	OAR9_38099807	No	F_{ST}	bmnxbsi	XM_585648; XM_002692655	PLAG1	pleiomorphic adenoma gene 1
10	OAR10_29538398	Yes	F_{ST}	bcsxbsi;FST(Hair X Wool)	XM_002684281; XM_002691757	RXFP2	relaxin/insulin-like family peptide receptor 2
11	OAR11_38437597	No	F_{ST}	bmnxbsi	NM_001034039	COL1A1	collagen, type I, alpha 1
12	s69940	Yes	iHS	bcs	XM_003583233; XM_002694045	KIF1B	kinesin family member 1B
13	OAR13_25941147	No	RsB	bcsxbmn	XM_002701036; XM_003586768	C13H10orf67	chromosome 13 open reading frame, human C10orf67
13	OAR13_49971490	Yes	iHS	bcs	NM_001078046	CDS2	CDP-diacylglycerol synthase (phosphatidate cytidylyltransferase) 2
13	OAR13_27814492	Yes	F_{ST}	bcsxbsi	NM_001206442	GPR158	G protein-coupled receptor 158
13	OAR13_9356390	Yes	RsB	bcsxbsi	NM_001105029	MACROD2	MACRO domain containing 2
13	OAR13_1628753	Yes	RsB	bcsxbsi	NM_174817	PLCB1	phospholipase C, beta 1 (phosphoinositide-specific)
13	OAR13_25099166	Yes	RsB	bcsxbsi;FST(Hair X Wool)	NM_001038185	SPAG6	sperm associated antigen 6
15	OAR15_45152619	Yes	F_{ST}	bcsxbsi;FST(Hair X Wool)	XM_002693106; XM_001252217	DENND5A	DENN/MADD domain containing 5 A
15	OAR15_78548290	Yes	F_{ST}	FST(Hair X Wool)	NM_177496	EXT2	exostosin glycosyltransferase 2
16	s06837; OAR16_43573089	Yes	RsB	bcsxbsi	NM_001192609	ADAMTS12	ADAM metallopeptidase with thrombospondin type 1 motif, 12
16	OAR16_36156919; OAR16_36192133	No	RsB	bcsxbsi	NM_001045979	C6	complement component 6
16	OAR16_38221926; s40592	Yes	iHS; RsB	iHS (bsi); RsB (bcsxbsi)	NM_001035364	C9	complement component 9
16	OAR16_45680234	No	iHS; RsB	iHS (bsi); RsB (bcsxbsi);FST (Hair X Wool)	NM_001034640	CDH6	cadherin 6, type 2, K-cadherin (fetal kidney)
16	OAR16_50704170; s58708; OAR16_50765168	Yes	RsB	bcsxbsi	NM_001105253	CDH9	cadherin 9, type 2 (T1-cadherin)
16	OAR16_38040585; OAR16_38083665	No	RsB	bcsxbsi	NM_001193246	DAB2	Dab, mitogen-responsive phosphoprotein, homolog 2 (Drosophila)
16	s27395; OAR16_45550332; OAR16_45590568; OAR16_45631306; OAR16_45680234	Yes	RsB	RsB (bcsxbmn and bcsxbsi)	XM_591998; XM_002696390	DROSHA	drosha, ribonuclease type III
16	OAR16_39143709; s49984	Yes	RsB	bcsxbsi	NM_001083478	EGFLAM	EGF-like, fibronectin type III and laminin G domains
16	OAR16_38355680; OAR16_38414913	Yes	iHS; RsB	iHS (bsi); RsB (bcsxbsi)	NM_001105414	FYB	FYN binding protein
16	OAR16_39780130	No	RsB	bcsxbmn	NM_001192849	GDNF	glial cell derived neurotrophic factor
16	OAR16_44884811; 250506CS3900211600001_1041	Yes	RsB	bcsxbsi	NM_001075948	GOLPH3	golgi phosphoprotein 3 (coat-protein)
16	OAR16_38988953; s74559; OAR16_39062204; s25960	Yes	RsB	RsB (bcsxbmn and bcsxbsi)	NM_001192263	LIFR	leukemia inhibitory factor receptor alpha
16	OAR16_41320024	Yes	RsB	bcsxbsi	NM_001076087	LOC537188	UDP-glucuronosyltransferase 3A1-like
16	s59948; OAR16_45398511; s17055	Yes	RsB	bcsxbsi	XM_001252695	LOC784368	Golgi phosphoprotein 3-like
16	OAR16_40589074; OAR16_40715706_X; OAR16_40749182; OAR16_40846568_X; s11229; s34284	Yes	iHS; RsB	iHS (bsi); RsB (bcsxbmn and bcsxbsi)	NM_001077070	MGC134093	uncharacterized protein MGC134093
16	OAR16_44759148	Yes	RsB	bcsxbsi	XM_002696402; XM_614802	MTMR12	myotubularin related protein 12
16	OAR16_40426113	Yes	RsB	bcsxbsi	NM_001206584	NIPBL	Nipped-B homolog (Drosophila)
16	OAR16_44325630; s23014	Yes	RsB	bcsxbsi	NM_174127	NPR3	natriuretic peptide receptor 3
16	s32951; OAR13_34211514; OAR16_41617317; OAR16_41681642; OAR16_41711634; OAR16_41737441; OAR16_41779659; OAR16_41804913	Yes	F_{ST}; iHS; RsB	fst; (bcsxbsi); iHS (bsi); RsB (bcsxbmn and bcsxbsi);FST (Hair X Wool)	XM_002696381	SPEF2	sperm flagellar 2
16	OAR16_43715998; OAR16_43747463; OAR16_43758108	Yes	RsB	bcsxbsi	NM_001034370	TARS	threonyl-tRNA synthetase

(continued on next page)

Table 2 (*continued*)

Chromosome	SNP	SNP in gene	Method	Comparison	RefSeq	Gene symbol	Gene name
16	OAR16_39780130; DU529574_332; OAR16_39840363; OAR16_39888776; OAR16_39987451; OAR16_40016364	Yes	F_{ST}; iHS; RsB	fst; (bcsxbsi); iHS (bsi); RsB (bcsxbsi);FST(Hair X Wool)	NM_001076238	WDR70	WD repeat domain 70
19	s31918	No	F_{ST}	FST(Hair X Wool)	NM_001102256	LANCL2	LanC like 2
25	OAR25_23974308	Yes	iHS	bcs	NM_001079638	CTNNA3	catenin (cadherin-associated protein), alpha 3
25	DU388965_586	Yes	F_{ST}	bcsxbmn	XM_002698790; XM_607409	PCNXL2	pecanex-like 2 (Drosophila)
25	s30024; s25195	Yes	F_{ST}	bcsxbmn;FST(Hair X Wool)	XM_002698809	TARBP1	TAR (HIV-1) RNA binding protein 1
26	s45791	No	F_{ST}	bcsxbsi	NM_001080280	SLC20A2	solute carrier family 20 (phosphate transporter), member 2

bcs= Brazilian Creole, bmn= Morada Nova, bsi=Santa Ines, bcsxbmn= Brazilian Creole x Morada Nova, bcsxbsi = Brazilian Creole x Santa Ines and bmnxbsi = Morada Nova x Santa Ines.

The identification of signatures of selection by more than one methodology may be seen as strong evidence of the activity of selection in a particular genomic region. However, a gene or genomic region identified as being subject to selection by one methodology and not by another does not exclude the possibility that selection has indeed occurred (Hohenlohe et al., 2010; Oleksyk et al., 2010).

The size of the genomic regions identified ranged from 50 Kb to 215.2 Kb, and highlighted the lack of precision resulting from use of a modest number of loci. Also, the strength of the selection can influence the selection signatures. The size of the regions identified and the quantity of genes contained within these regions often makes it difficult to identify the target of selection.

A total of six genes (*CNTNAP5*, *DIS3L2*, *KIF1B*, *RXFP2*, *TARBP1* and *MSRB3*) found as candidates to be under selection in this study were previously reported in other studies involving sheep and cattle (Fariello, et al., 2014; Flori et al., 2009; Gautier et al., 2009; Kijas et al., 2012; Stella et al., 2010; Zhang et al., 2013; Wei et al., 2015).

The *RXFP2* gene is in a region containing a selection signature detected through F_{ST} analysis. It was identified in the comparison jointly considering all three Brazilian breeds, in the pairwise comparison of Brazilian Creole and Santa Ines and in the comparison between the northeastern hair breeds x southern wool breed (Fig. 2 and Supplementary Fig. 7). This gene as recently associated with horn type in sheep (Johnston et al., 2011) and it was identified by Kijas et al. (2012) to be present in a region containing a strong selection signature in a study involving 74 sheep breeds from around the world (which also included the animals used for carrying out the present study). Morada Nova and Santa Ines have been under selection for this trait, as breed standards indicate that these animals should be polled, but scurs (rudimentary horns) are permitted in the Santa Ines. In the Brazilian Creole breed, both polled and horned animals are accepted for registry. Moreover, animals that have more than one pair of horns (polycerate) are also accepted.

The *TRPM8* gene was found to be present in a region containing a selection signature by F_{ST} analysis (Table 2 and Supplementary Table 10). This gene was previously detected as a candidate to be under selection in worldwide sheep populations (Fariello et al., 2014) and is involved in regulation of body temperature at low temperatures both through autonomic mechanisms (induction of the thermogenic activity of the brown adipose tissue and vasoconstriction in the skin) and behavioral alterations (Tajino et al., 2011; Almeida et al., 2012; Gavva et al., 2012). This signature of selection was identified both in the comparison considering all three breeds, in the comparison of Brazilian Creole and Santa Ines and in the comparison between the northeastern hair breeds x southern wool breed (Fig. 2 and Supplementary Fig. 7). Brazilian Creole sheep have been developed in the Southern region of Brazil, where subtropical conditions are found and sub-zero temperatures can be frequently observed during the winter. Conversely, Morada Nova and Santa Ines sheep have been bred in Northeastern Brazil, in a tropical semi-arid climate where temperatures are high throughout the year.

Among the breeds analyzed, there is considerable variation in relation to the adult size of the animals. While the Santa Ines is considered a large animal, the Brazilian Creole and Morada Nova breeds are considered small. Some genes (*DIS3L2*, *PLAG1*, *NIPBL*) associated with height variation were identified in this study. *NIPBL* was identified by RsB analysis in the Brazilian Creole x Santa Ines comparison (Fig. 4). Mutations in this gene are associated with *Cornelia de Lange* syndrome, a development disorder that may be characterized by facial deformations and deformations in the upper extremities, hirsutism, heart defects, low stature and mental retardation (Gillis et al., 2004; Kemper et al., 2012). A region identified by F_{ST} analysis in both all three breeds and in the Morada Nova x Santa Ines comparison (Fig. 2) included the *PLAG1* gene. Recent studies indicate the association of this gene with stature and birth weight variation in cattle (Karim et al., 2011; Littlejohn et al., 2011; Utsunomiya et al.,

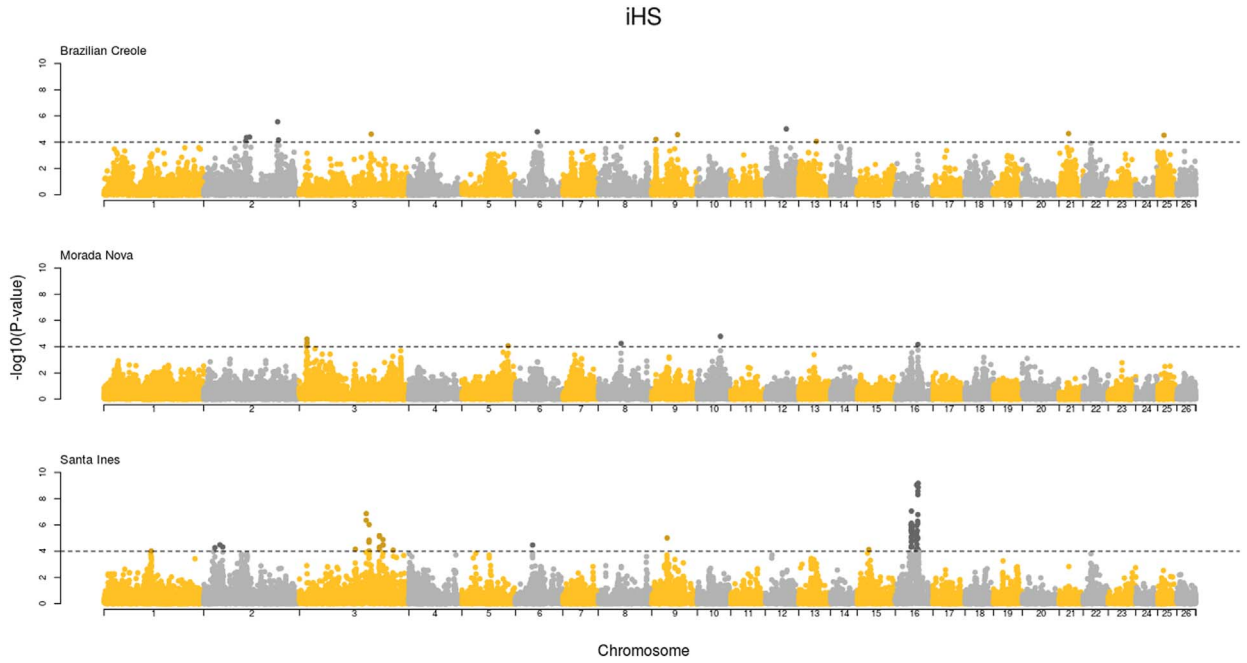

Fig. 3. Manhattan plot for iHS analysis Odd-numbered chromosomes are represented in yellow while even-numbered chromosomes are represented in gray. The darker points above the dotted line indicate the significant SNPs. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2013). The gene *DIS3L2* was identified by RsB analysis in the Santa Ines x Brazilian Creole comparison. This gene is also a candidate under selection in the cattle (Gautier et al., 2009), and in humans it is associated with the Perlman syndrome, a disorder characterized by overgrowth and susceptibility to development of tumors (Astuti et al., 2012).

Adult size is a trait directly related to production efficiency in farm animals, as larger animals generally have higher maintenance requirements (Gomes et al., 2013). The identification of candidate genes for adult size in comparisons with the Santa Ines breed may be a consequence of the recent breed history, as major increases in the mean stature of the breed have been reported in recent years (McManus et al., 2010).

Candidate genes involved with reproductive processes were found in regions containing selection signatures. *SPEF2* and *SPAG6* are important for spermatogenesis (Escalier, 2006; Sironen et al., 2010, 2012), and mutations in these genes have been reported to be associated with infertility in males of various species of mammals (Escalier, 2006; Sironen et al., 2006, 2011). *SPAG6* has also been associated with male infertility in humans (Williamson et al., 2007; Barreiro et al., 2008).

The sheep breeds analyzed in the present study diverge in coat

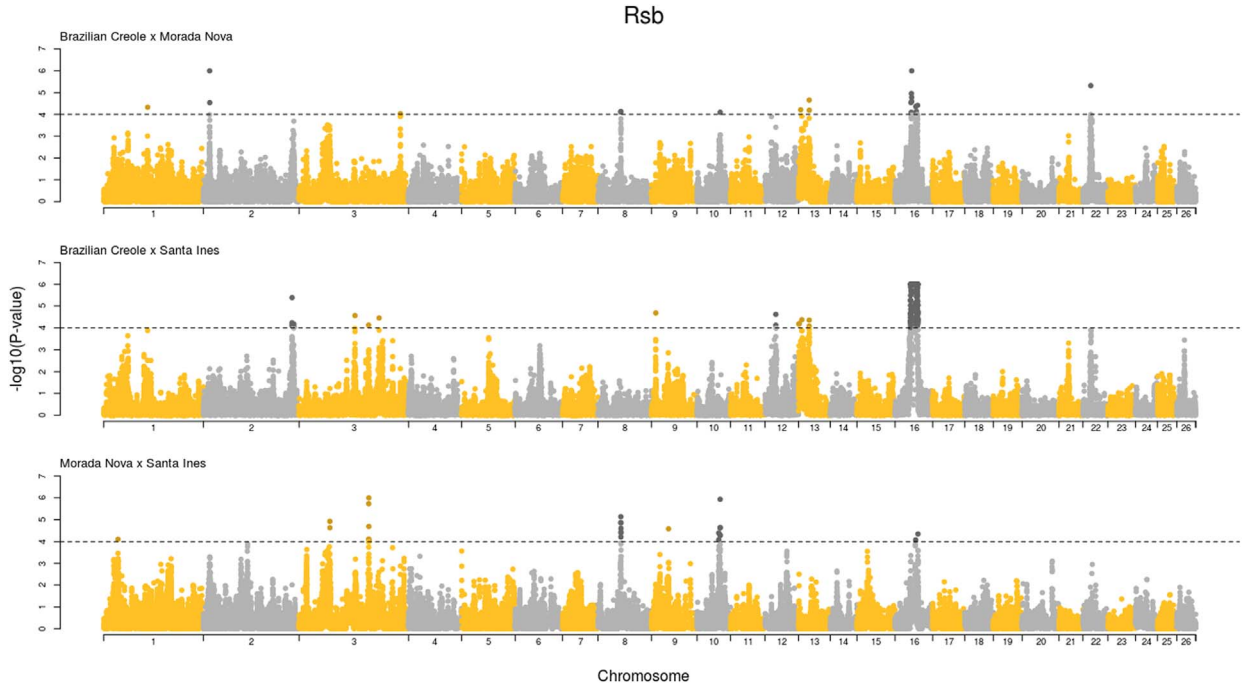

Fig. 4. Manhattan plot for RsB analysis Odd-numbered chromosomes are represented in yellow while even-numbered chromosomes are represented in gray. The darker points above the dotted line indicate the significant SNPs. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

characteristics. Brazilian Creole sheep is a wool breed developed in the Southern region of Brazil. Conversely, Morada Nova and Santa Ines sheep have been bred in Northeastern Brazil and are classified as hair sheep breeds. To pinpoint putative genes involved in these traits, we performed a F_{ST} analysis comparing the northeastern hair breeds against the southern wool breed (Supplementary Table 10). No gene found in this study was identified as a candidate to be associated with wool production characteristics in a GWAS study (Wang et al., 2014).

Despite of this, the gene *SPEF2* was identified as a candidate in F_{ST} (Brazilian Creole x Santa Ines and Hair x Wool comparisons), iHS (Santa Ines) and RsB (Brazilian Creole x Morada Nova and Brazilian Creole x Santa Ines comparisons) analyses. This gene was previously related to late feathering phenotype in male chickens (Elferink et al., 2008) and also identified as candidate to be related to the slick-hair coat in tropically adapted cattle breeds (Huson et al., 2014).

Although many genes have been suggested as targets of selection and some show evidence of participation in important physiological mechanisms, it is important to emphasize that the aim of the present study was not to define a cause-effect relationship between genotype and phenotype. In addition, there is the possibility that some of the signatures identified in this study are not due to selection, but rather to other evolutionary processes (e.g., population expansion, subdivision and bottlenecking) (Randhawa et al., 2016).

In spite of this, the present study opens the possibility for a more detailed investigation, in the sense of validating genotype-phenotype associations that may explain ecological and productive differences/particularities among Brazilian locally adapted sheep breeds. The results of this study might even be important as additional tool to select donors for the National germplasm bank maintained by Embrapa.

5. Conclusions

The results presented here provide evidence of selection that may be specific to one or a few locally adapted breeds using a combination of two groups of methodologies (population differentiation and linkage disequilibrium). The candidate regions identified contain genes related to important adaptative traits that can be considered in further studies aiming to confirm the selection signals and to identify the causal mutations that confer a selective advantage in these populations.

Conflicts of interest statement

The authors declare that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in the manuscript intitled "*Genome-wide search for signatures of selection in three major Brazilian locally adapted sheep breeds*".

Acknowledgements

We would like to acknowledge all Public Institutions and farmers that donated the samples for this work. This study was performed with resources from Brazilian Agricultural Research Corporation (Embrapa) and National Council for Scientific and Technological Development (CNPq).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.livsci.2017.01.006.

References

Almeida, M.C., Hew-Butler, T., Soriano, R.N., Rao, S., Wang, W., Wang, J., Tamayo, N., Oliveira, D.L., et al., 2012. Pharmacological blockade of the cold Receptor TRPM8 attenuates autonomic and behavioral cold defenses and decreases deep body temperature. J. Neurosci. 32 (6), 2086–2099.

Amarante, A.F.T., Bricarello, P.A., Rocha, R.A., Gennari, S.M., 2009. Resistance of Santa Ines Suffolk and Ile de France sheep to naturally acquired gastrointestinal nematode infections. Vet. Parasitol. 120 (1–2), 91–106.

Astuti, D., Morris, M.R., Cooper, W.N., Staals, R.H.J., Wake, N.C., Fews, G.A., Gill, H., Gentle, D., Shuib, S., Ricketts, C.J., et al., 2012. Germline mutations in DIS3L2 cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility. Nat. Genet. 44, 277–284.

Barreiro, L.B., Laval, G., Quach, H., Patin, E., Quintana-Murci, L., 2008. Natural selection has driven population differentiation in modern humans. Nat. Genet. 40 (3), 340–345.

Beaumont, M.A., 2005. Adaptation and speciation: what can Fst tell us? Trends Ecol. Evol. 20 (8), 435–440.

Chen, H., Sharp, B.M., 2004. Content-rich biological network constructed by mining PubMed abstracts. BMC Bioinform. 8 (5).

Diamond, J., 2002. Evolution consequences and future of plant and animal domestication. Nature 418, 700–707.

Elferink, M.G., Vallée, A.A.A., Jungerius, A.P., Crooijmans, R.P.M.A., Groenen, M.A.M., 2008. Partial duplication of the PRLR and SPEF2 genes at the late feathering locus in chicken. BMC Genom. 9 (1).

Escalier, D., 2006. Knockout mouse models of sperm flagellum anomalies. Hum. Reprod. Update 12 (4), 449–461.

Fariello, M.I., Boitard, S., Naya, H., Sancristobal, M., Servin, B., 2013. Detecting signatures of selection through haplotype differentiation among hierarchically structured populations. Genetics 193 (3), 929–941.

Fariello, M.-I., Servin, B., Tosser-Klopp, G., Rupp, R., Moreno, C., et al., 2014. Selection signatures in worldwide sheep populations. PLoS One 9 (8), e103813.

Flori, L., Fritz, S., Jaffrézic, F., Boussaha, M., Gut, I., Heath, S., Foulley, J.L., Gautier, M., 2009. The genome response to artificial selection: a case study in dairy cattle. PLoS One 4 (8), e6595, (2009).

Foll, M., Gaggiotti, O., 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics 180, 977–993.

Gautier, M., Vitalis, R., 2012. Rehh: an R package to detect footprints of selection in genome-wide SNP data from haplotype structure. Bioinformatics 28 (8), 1176–1177.

Gautier, M., Flori, L., Riebler, A., Jaffrézic, F., Laloé, D., Gut, I., Moazami-Goudarzi, K., Foulley, J.L., 2009. A whole genome Bayesian scan for adaptive genetic divergence in West African cattle. BMC Genom. 10, 550.

Gavva, N.R., Davis, C., Lehto, S.G., Rao, S., Wang, W., Zhu, D.X.D., 2012. Transient receptor potential melastatin 8 (TRPM8) channels are involved in body temperature regulation. Mol. Pain 8, 36.

Gillis, L.A., Mccallum, J., Kaur, M., Descipio, C., Yaeger, D., Mariani, A., Kline, A.D., Li, H.-H., Devoto, M., Jackson, L.G., Krantz, I.D., 2004. NIPBL Mutational analysis in 120 individuals with Cornelia de Lange syndrome and evaluation of genotype-phenotype correlations. Am. J. Hum. Genet. 75 (4), 610–623.

Gomes, E.F., Louvandini, H., Dallago, B.S.L., Canozzi, M.E.A., Bernal, F.E.M., McManus, C.M., 2013. Productivity in ewes of different genetic groups and body sizes. J. Anim. Sci. Adv. 3 (5), 243–255.

Hayes, B.J., Lien, S., Nilsen, H., Olsen, H.G., Berg, P., Maceachern, S., Potter, S., Meuwissen, T.H.E., 2008. The origin of selection signatures on bovine chromosome 6. Anim. Genet. 39, 105–111.

Hoffmann, R., Valencia, A.A., 2004. Gene network for navigating the literature. Nat. Genet. 36 (664).

Hohenlohe, P.A., Phillips, P.C., Cresko, W.A., 2010. Using population genomics to detect selection in natural populations: Key concepts and methodological considerations. Int J. Plant Sci. 171 (9), 1059–1071.

Huang, D.W., Sherman, B.T., Lempicki, R.A., 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4 (1), 44–57.

Huson, H.J., Kim, E.-S., Godfrey, R.W., Olson, T.A., McClure, M.C., Chase, C.C., Rizzi, R., O´Brien, A.M.P., Van Tassel, C.P., Garcia, J.F., Sonstegard, T.S., 2014. Genome-wide association study and ancestral origins of the slick-hair coat in tropically adapted cattle. Front. Genet. 5 (101).

Jacinto, M.A.C., Silva Sobrinho, A.G., Costa, R.G., 2004. Características anátomo-estruturais da pele de ovinos (Ovis aries) lanados e deslanados, relacionadas com o aspecto físico-mecânico do couro. Rev. Bras. Zoot. 33 (4), 1001–1008.

Jacinto, M.A.C., Vargas Junior, F.M., Martins, C.F., Pinto, G.S., Reis, F.A., Oliveira, A.R., 2011. Influence of genotype on the quality of sheep leather. Rev. Bras. Zoot. 40 (8), 1830–1836.

Jiang, Y., Xie, M., Chen, W., Talbot, R., Maddox, J.F., Faraut, T., Wu, C., Muzny, D.M., Li, Y., Zhang, W., et al., 2014. The sheep genome illuminates biology of the rumen and lipid metabolism. Science 344, 1168–1173.

Johnston, S.E., Mcewan, J.C., Pickering, N.K., Kijas, J.W., Beraldi, D., Pilkington, J.G., Pemberton, J.M., Slate, J., 2011. Genome-wide association mapping identifies the genetic basis of discrete and quantitative variation in sexual weaponry in a wild sheep population. Mol. Ecol. 20, 2555–2566.

Karim, L., Takeda, H., Lin, L., Druet, T., Arias, J.A.C., Baurain, D., Cambisano, N., Davis, S.R., Farnir, F., Grisart, B., et al., 2011. Variants modulating the expression of a chromosome domain encompassing PLAG1 influence bovine stature. Nat. Genet. 43, 405–413.

Kemper, K.E., Visscher, P.M., Goddard, M.E., 2012. Genetic architecture of body size in

mammals. Genome Biol. 13, 244.
Kijas, J.W., Lenstra, J.A., Hayes, B., Boitard, S., Porto Neto, L.R., Cristobal, M.S., Servin, B., Mcculloch, R., Whan, V., Gietzen, K., et al., 2012. Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. PLoS Biol. 10 (2), e1001258.
Littlejohn, M., Grala, T., Sanders, K., Walker, C., Waghorn, G., Macdonald, K., Coppieters, W., Georges, M., Spelman, R., Hillerton, E., et al., 2011. Genetic variation in PLAG1 associates with early life body weight and peripubertal weight and growth in Bos taurus. Anim. Genet. 43, 591–594.
Mariante, A.S., Albuquerque, M.S.M., Egito, A.A., McManus, C., Lopes, M.A., Paiva, S.R., 2009. Present status of the conservation of livestock genetic resources in Brazil. Livest. Sci. 120 (3), 204–212.
McManus, C., Paiva, S.R., Araújo, R.O., 2010. Genetics and breeding of sheep in Brazil. Rev. Bras. Zoot. 39, 236–246.
McManus, C., Louvandini, H., Paiva, S.R., Oliveira, A.A., Azevedo, H.C., Melo, C.B., 2008. Genetic factors of sheep affecting gastrointestinal infections in the Distrito Federal, Brazil. Vet. Parasitol. 166, 308–313.
McVean, G.A.T., Myers, S., Hunt, S., Deloukas, P., Bentley, D., Donnelly, P., 2004. The fine-scale structure of recombination rate variation in the human genome. Science 304, 581–584.
Moradi, M.H., Nejati-Javaremi, A., Moradi-Shahrbabak, M., Dodds, K.G., McEwan, J.C., 2012. Genomic scan of selective sweeps in thin and fat tail sheep breeds for identifying of candidate regions associated with fat deposition. BMC Genet. 13, 10.
Nielsen, R., 2001. Statistical tests of selective neutrality in the age of genomics. Heredity 86, 641–647.
Nielsen, R., 2005. Molecular signatures of natural selection. Annu Rev. Genet. 39, 197–218.
Oleksyk, T.K., Smith, M.W., O'Brien, J., 2010. Genome-wide scans for footprints of natural selection. Philos Trans. R. Soc. Lond. B Biol. Sci. 365, 185–205.
Patterson, N., Price, A.L., Reich, D., 2006. Population structure and eigenanalysis. PLoS Genet. 2 (12), e190.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., Sham, P.C., 2007. PLINK: a toolset for whole-genome association and population-based linkage analysis. Am. J. Hum. Genet., 81.
Randhawa, I.A.S., Khatkar, M.S., Thomson, P.C., Raadsma, H.W., 2016. A meta-assembly of selection signatures in cattle. PLoS ONE 11 (4), e0153013.
Sabeti, P.C., Reich, D.E., Higgins, J.M., Levine, H.Z.P., Richter, D.J., Schaffner, S.F., Gabriel, S.B., Platko, J.V., Patterson, N.J., Mcdonald, G.J., et al., 2002. Detecting recent positive selection in the human genome from haplotype structure. Nature 419 (24), 832–837.
Scheet, P., Stephens, M., 2006. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. Am. J. Hum. Genet. 78, 629–644.
Schlötterer, C., 2003. Hitchhiking mapping: functional genomics from the population genetics perspective. Trends Genet. 19 (1), 32–38.
Sironen, A., Uimari, P., Iso-Touru, T., Vilkki, J., 2012. L1 insertion within SPEF2 gene is associated with increased litter size in the Finnish Yorkshire population. J. Anim. Breed. Genet. 129, 92–97.
Sironen, A., Thomsen, B., Andersson, M., Hola, V., Vilkki, J., 2006. An intronic insertion in KPL2 results in aberrant splicing and causes the immotile short-tail sperm defect in the pig. Proc. Natl. Acad.. Sci. U.S.A. 103 (13), 5006–5011.
Sironen, A., Hansen, J., Thomsen, B., Andersson, M., Volkki, J., Toppari, J., Kotaja, N., 2010. Expression of SPEF2 during mouse spermatogenesis and identification of IFT20 as an interacting protein. Biol. Reprod. 82, 580–590.
Sironen, A., Kotaja, N., Mulhern, H., Wyatt, T.A., Sisson, J.H., Pavlik, J.A., Miiluniemi, M., Fleming, M.D., Lee, L., 2011. Loss of SPEF2 function in mice results in spermatogenesis defects and primary ciliary dyskinesia. Biol. Reprod. 85, 690–701.
Stella, A., Ajmone-Marsan, P., Lazzari, B., Boettcher, P., 2010. Identification of selection signatures in cattle breeds selected for dairy production. Genetics 185, 1451–1461.
Storz, J.F., 2005. Using genome scans of DNA polymorphism to infer adaptive population divergence. Mol. Ecol. 14, 671, (668).
Tajino, K., Hosokawa, H., Maegawa, S., Matsumura, K., Dhaka, A., Kobayashi, S., 2011. Cooling-sensitive TRPM8 is thermostat of skin temperature against temperature regulation. Mol. Pain 8 (36).
Tang, K., Thornton, K.R., Stoneking, M., 2007. A new approach for using genome scans to detect recent positive selection in the human genome. PLoS Biol. 5 (7), e171.
The International Sheep Genomics Consortium (ISGC), Archibald, A.L., Cockett, N.E., Dalrymple, B.P., Faraut, T., Kijas, J.W., Maddox, J.F., Mcewan, J.C., Hutton Oddy, V., Raadsma, H.W., Wade, C., Wang, J., Wang, W., Xun, X., 2010. The sheep genome reference sequence: a work in progress. Anim. Genet. 41, 449–453.
Toro, M., Mäki-Tanila, A., 2007. Genomics reveals domestication history and facilitates breed development. In: Wageningen, Oldenbroek K. (Ed.), Utilisation and Conservation ofFarm Animal Genetic Resources. Academic Publishers, Wageningen, 75–102.
Utsunomiya, Y.T., Carmo, A.S., Carvalheiro, R., Neves, H.H.R., Matos, M.C., Zavarez, L.B., O'Brien, A.M.P., Sölkner, J., McEwan, J.C., Cole, J.B., et al., 2013. Genome-wide association study for birth weight in Nellore cattle points to previously described orthologous genes affecting human and bovine height. BMC Genet. 14, 52.
Voight, B.F., Kudaravalli, S., Wen, X., Pritchard, J.K., 2006. A map of recent positive selection in the human genome. PLoS Biol. 4 (3), 446–458.
Wang, Z., Zhang, H., Yang, H., Wang, S., Rong, E., Pei, W., Wang, N., 2014. Genome-wide association study for wool production traits in a chinese merino sheep population. PLoS ONE 9 (9), e107101.
Wei, C., Wang, H., Liu, G., Wu, M., Cao, J., Liu, Z., Liu, R., Zhao, F., 2015. Genome-wide analysis reveals population structure and selection in Chinese indigenous sheep breeds. BMC Genom. 16, 194.
Williamson, S.H., Hubisz, M.J., Clark, A.G., Payseur, B.A., Bustamante, C.D., Nielsen, R., 2007. Localizing recent adaptive evolution in the human genome. PLoS Genet. 3 (6), e90.
Zhang, L., Mousel, M.R., Wu, X., Michall, J.J., Zhou, X., Ding, B., Dodson, M.V., El-Halawany, N.K., Lewis, G.S., Jiang, Z., 2013. Genome-wide genetic diversity and differentially selected regions among suffolk, rambouillet, columbia, olypay, and targhee sheep. PLoS One 8 (6), e65942.