Signatures of Selection for Age at First Lambing in Brazilian Local Adapted Sheep

A.M.B.O. Lôbo[1], S.R. Paiva[2], R.N.B. Lôbo[1,3].
[1]Embrapa Goats and Sheep, Sobral, Ceará, Brazil, [2] Embrapa Secretariat of International Affairs, Brasilia, DF, Brazil, [3]Federal University of Ceará, Fortaleza, Ceará, Brazil; National Council for Scientific and Technological Development (CNPq).

ABSTRACT: A total of 84 animals were genotyped on the Illumina OvineSNP50K BeadChip for identifying the genome-wide footprints of positive selection for the age at first lambing (AFL) of Santa Inês ewes. 90 SNP markers were significant by Fisher's exact test. From these SNPs, eight satisfied the imposed constraint of 100kb distance between two consecutive markers. It was identified at least two genes on chromosome 25 as probable targets of selection. The *NRG*3 comprised between the markers OAR25_38884374.1 and s19656.1, and the *PTPN*20B comprised between s69834.1 and s26762.1. The results suggest that NRG3 plays an important role in AFL of Santa Inês. This gene can have potential for genomic selection and this is particularly useful for sheep industry. The genotyping more animals of the population and the estimative of additive effect of NRG3 gene for the trait AFL may validate this data.
Keywords:
Ovis aries
Locally adapted sheep
OvineSNP50K Beadchip

Introduction

Selection for phenotypic traits leaves to footprints in several regions of the genome. These footprints are called signatures of selection and the identification of these regions can be performed by analyzing the allelic frequency of divergent animals using highly dense panels of single nucleotide polymorphisms (SNP). According Moradi et al. (2012) this issue is one of the most challenging areas of research in animal genetics. Genome-wide analyses have been performed to identify signature of selection in sheep (Fariello et al. (2013); Moioli et al. (2013); Moradi et al. (2012)). The identification of genomic regions associated with fertility is very important in sheep, particularly in Brazilian meat schemes, because this trait is one of the most economically important (Lôbo et al. (2011)). Therefore, it is expected that genomic regions (footprints) controlling this trait might be present as those have been subject of selection throughout sheep history.

The aim of this study was identify signatures of selection of the ovine genome that potentially affect the age at first lambing of the ewes, as an indicator of fertility.

Materials and Methods

Animal samples. This study used Santa Inês breed, a locally adapted breed sheep from Northeastern region of Brazil. This region is mainly characterized by long dry periods and pasture constraints. Nowadays, this breed is spread out in all regions of the country and has the biggest population size from all Brazilian breeds. The animals of this study came from a commercial flock (16°19'36.83"S / 49°31'2.34"O), located in Goiás state and they are assisted by a Breeding Program (GENECOC) developed and under supervision of Embrapa Goats and Sheep, Ceará State. The animals were submitted to semi-intensive system and the breeding season was year round with ewes grouped by lots.

Genotyping. A total of 84 animals were genotyped using the OvineSNP50 BeadChip manufactured by Illumina (San Diego, CA). Genotyping was performed by Geneseek - Neogen Corporation (Lincoln, USA). Data were analyzed by PLINK (Purcell et al. (2007)) applying quality threshold of .80, maximum per-SNP missing and maximum per-person missing of 0.1 and minor allele frequency of 0.05.

Analysis of reproductive data and animal ranking. Records of age at first lambing (AFL) of 3,248 Santa Inês ewes were used to estimate the breeding values of the animals by an animal model with the fixed effect of contemporary group. The relationship matrix contained 13,431 animals (1,691 sires).

Two tails of the distribution of the animal breeding values were used to compare allele frequency at each marker between top and worse animals for age at first lambing. The tails were divided in groups corresponding to one sixth (14 animals) and one twelfth (7 animals).

Putative signatures of selection were assessed by Fisher`s exact test of significance of differences of allele frequency between the two tails for each marker. The markers of interest were considered if satisfy the statistical significance of $P<.05$ and $P<0.001$, respectively for the 1/6 and 1/12 tails of the distribution. It was intended emphasize that the differences in allele frequencies should become higher as far as the groups become more divergent. The other condition for the interesting markers was that the distance between two consecutive markers could not be higher than 100 k bp.

Anotation of the candidate regions. The OAR v3.1 Ovine (Texel) Genome Assembly in www.livestockgenomics.csiro.au was used to obtain the coordinates of SNPs. The coordinates of SNPs were used in the NCBI Map Viewer (http://www.ncbi.nlm.nih.gov/mapview/) to discover if any

of the significant regions from the ovine whole genome analysis contained genes of interest. In the cases where was not possible perform the annotation with the ovine chromosome, it was extracted the *fasta* sequence of the region and performed a standard nucleotide blast against the database of the *Bos Taurus*.

Results and Discussion

The age at first lambing of Santa Inês sheep presents great variation as expected and it was used divergent animals. The average for this sample was 961.03±396.40 days. The average breeding values for the two tails of AFL trait was reported in Table 1.

After data pruning 44,263 SNPs genotyped in 84 animals passed on quality control and were used in the subsequent analysis. 90 markers where the significance were $P_{1/6}<0.05$ and $P_{1/12}<0.001$ were selected by Fisher's exact test (Table 2). Eight markers satisfied the condition to have 100 Kb distance between two consecutive markers (Table 3). Two genes on chromosome 25 were annotated within the regions identified as probable targets of selection. The Neuregulin 3 (*NRG*3) in the region comprised between the markers OAR25_38884374.1 and s19656.1, and the Protein Tyrosine Phosphatase, Non-Receptor Type 20B (*PTPN*20B) in the region comprised between the markers s69834.1 and s26762.1.

The NRG3 gene promotes, among other aspects, the mammary differentiation during embryogenesis. This gene is a member of the epidermal growth factor receptor family related to the proliferation of granulosa cells, steroidogenesis and oocyte maturation (Conti et al., (2006)). Kezele et al. (2005) suggested that NRG is a candidate for the regulation early follicle development. Yang et al. (1995) demonstrated that NRG is controlled by pregnancy hormones in the local mesenchyme and it induces lobulo-alveolar development and the milk production of the mammary gland epithelium. According Gratão (2007), the initiation of mammalian puberty requires the activations of luteinizing hormone-releasing hormone (LHRH) and the neuregulin is one the factors produced in hypothalamic astrocytes that stimulate LHRH release. According Ma et al. (1990) the erbB-2-erbB-4 (direct ligand to NRG3) receptor appears signaling the process by which astrocytes facilitate the acquisition of female reproductive capacity in mammals. Age at puberty is directly related to age at first lambing and have moderately heritable and would respond to selection (Schoeman & Albertyn, 1991). This suggests that NRG3 plays an important role in age at first lambing of Santa Inês. This gene can have potential for genomic selection and this is particularly useful for sheep industry but it is necessary assigning the functional effects for the identified regions. The genotyping more animals of the population and the estimative of additive effect of NRG3 gene for the trait age at first lambing may validate this data.

The PTPN20B is a member of phosphatase family. The protein tyrosine phosphatases play key roles in a diverse range of cellular processes such as differentiation, cell proliferation, apoptosis, immunological signaling, and cytoskeletal function (Seo et al. (2013)). This suggests that selection might also have occurred at genes related to other cellular events.

In the regions comprised between the markers OAR14_42449381.1 and OAR14_43565704.1, and between OAR2_258621731.1 and OAR2_258668680.1 we did not find annotations in ovine genome.

Conclusion

In this study, four genomic regions contributing to within-breed allelic differences were detected. A putative promising candidate gene for fertility in Santa Inês sheep or a putative signature of selection for age at first lambing of the ewes was identified. The results suggest that selection might also have occurred at genes related to other cellular events. However, this is a preliminary study and the validation of these results is necessary.

Literature Cited

Conti, M., Hsieh, M., Park, J-Y., Su, Y-Q. (2006). Molec. Endoc. 20:715-723.

Fariello, M. I., Servin, B., Tosser-Klopp, G., et al. (2013). bioRxiv. doi: 10.1101/001453

Gratão, A. A. (2007). Thesis. Faculty of Veterinary Medicine of the Ludwing of the Ludwig-Maximilians University, Munich. http://edoc.ub.uni-muenchen.de/6575/2/Gratao_Ana.pdf

Kezele, P. R., Ague, J. M., Nilsson, E., Skinner, M. K. (2005). Biol. Reprod. 72:241-255.

Lôbo, R. N. B., Pereira, I. D. C., Facó, O., et al. (2011). Small Rum. Res. 96: 93-100.

Ma, Y. J., Hill, D. F., Creswick, K. E., et al. (1999). J. Neurosci. 19:9913-9927.

Moioli, B., Scatà, M. C., Steri, R., et al. (2013). BMC Genetics. 14:1-7.

Moradi, M. H., Nejati-Javaremi, A., Moradi-Shahrbabak, M., et al. (2012). BMC Genetics. 13:1-15;

Purcell, S., Neale, B., Todd-Brown, K., et al. (2007). Am. J. of Hum. Genet. 81:559-575.

Seo, H., Lee I-S., Park, J. E., et al. (2013). PLoS ONE. 8:1-7.

Schoeman, S. J., Albertyn, J. R. (1991). S. Afr. J. Anim. Sci. 21:169-172.

Yang, Y., Spitzer, E., Meyer, D., et al. (1995). J. Cell Biol. 13:215-226.

Table 1. Average of animal breeding values for divergent groups for age at first lambing

Category	N	Mean (day)	Std
Top sixth	14	-134.66	25.16
Worse sixth	14	55.06	25.76
Top twelfth	7	-154.32	17.91
Worse twelfth	7	77.20	8.94

Table 2. Markers presenting different allele frequencies between divergent animals for age at first lambing

Chromosome	Marker	Coordinate on OAR v3.1
1	s47644.1	2657803
1	s04088.1	3336925
1	OAR1_34876908.1	34118937
1	OAR1_36621076.1	35661180
1	OAR1_174976013.1	162220155
1	s14177.1	162526531
1	OAR1_187611331.1	174117310
1	s57353.1	174989613
1	OAR1_188962060.1	175347678
1	OAR1_234437646.1	217364552
1	OAR1_235455300.1	218233462
1	OAR1_253155103.1	234490177
2	s63576.1	14664664
2	OAR2_28037853.1	27144318
2	s56017.1	28252533
2	s70928.1	28360954
2	OAR2_34693156.1	33332770
2	OAR2_35283978.1	33913809
2	OAR2_37856600.1	36454190
2	OAR2_41188882.1	39534629
2	OAR2_100622769.1	93520861
2	OAR2_137914666.1	129518436
2	s40725.1	175346551
2	s55941.1	217421910
2	s57699.1	240393589
2	OAR2_258621731.1	244684889
2	OAR2_258668680.1	244731375
2	s20190.1	244962782
2	s73346.1	245349075
3	s73962.1	93903353
3	s62934.1	133025037
3	OAR3_179813294.1	167375732
4	OAR4_15813361.1	15562966
4	OAR4_18647481.1	18399255
4	OAR4_26459299.1	25176950
5	s16712.1	827819
5	OAR5_1151915_X.1	933361
5	s06747.1	77635110
7	OAR7_33119139.1	29387533
7	s54436.1	69672264
7	OAR7_76941227.1	70209501
7	OAR7_96394687.1	88666908
7	OAR7_97925176.1	90050938
8	s18363.1	22752132
8	OAR8_90705948.1	84048731
9	OAR9_63148438.1	60167949
9	OAR9_65483377.1	62234049
9	OAR9_65626804.1	62369406
9	OAR9_69552784.1	65707443
9	OAR9_72877955.1	68828865
9	OAR9_73284295.1	69233179
9	OAR9_91661512.1	86455870
9	OAR9_95791776.1	90206411
10	OAR10_48898826.1	48135136
10	OAR10_56091191.1	55038366
10	OAR10_69870667.1	67725438

Table 2. Markers presenting different allele frequencies between divergent animals for age at first lambing *(Continued)*

12	OAR12_21638552.1	18758822
14	OAR14_22212176.1	21618920
14	OAR14_33000004.1	31698036
14	s23054.1	35673098
14	OAR14_43449381.1	41692276
14	OAR14_43565704.1	41696728
15	OAR15_3963856.1	4594655
15	OAR15_6384151.1	6847178
15	OAR15_33766559.1	32201996
15	s32555.1	78871439
16	s02237.1	742915
16	OAR16_58984443.1	54110779
18	OAR18_60724030.1	56966995
18	s75539.1	62770038
18	s38764.1	65829542
19	OAR19_6316471.1	6043844
20	OAR20_40967781.1	37515740
20	OAR20_41289146.1	37842627
20	OAR20_41756644.1	38294283
20	OAR20_46680243.1	43001737
22	OAR22_30915840.1	26597689
22	OAR22_37614558.1	32820513
23	OARX_34048710.1	26356619
24	OAR24_26676939.1	24190286
24	OAR24_38653760.1	35459203
24	s23797.1	36250095
24	OAR24_41337655.1	38162634
25	OAR25_38884374.1	37084341
25	s19656.1	37160326
25	OAR25_39517297.1	37695570
25	s12325.1	40690849
25	s69834.1	42098554
25	s26762.1	42153979
25	DU510943_400.1	43238638

Table 3. Consecutive markers no more than 100 Kbp distant (dist.) from each other

Chrom	Marker 1	Dist. (bp)	Marker 2
2	OAR2_258621731.1	46486	OAR2_258668680.1
14	OAR14_43449381.1	4452	OAR14_43565704.1
25	OAR25_38884374.1	75985	s19656.1
25	s69834.1	55425	s26762.1