Barbora OLŠANSKÁ *, Radovan KASARDA, Anna TRAKOVICKÁ,
Ondrej KADLEČÍK, Nina MORAVČÍKOVÁ [1]

IDENTIFICATION OF SELECTION SIGNALS THROUGH HAPLOTYPE STRUCTURE IN GENOME OF DUAL-PURPOSE BREEDS

SUMMARY

The aim of this study was to analyse the selection signals by the integrated haplotype homozygosity score (iHS) in dual-purpose Slovak Spotted cattle in order to clarify the effects of natural or artificial selection. Overall, 85 animals genotyped by a high-density SNP array were included in the analysis. After quality control, 35995 autosomal loci with an average adjacent SNP spacing of 69.3 kb were included in the analysis. Then, haplotypes were reconstructed for each chromosome. Next, the R package rehh was used to compute the score based on a matrix of integrated extended haplotype homozygosity statistics for both ancestral and derived alleles. The average iHS score across the genome was 0.83. The selection signals in the genome had positive iHS values. A common iHS score (higher than 2.5) was chosen to indicate genomic regions with extreme iHS frequency due to outliers (according to the boxplot distribution). In each region with selection signals, quantitative trait loci (QTLs) were identified. For BTA s 3, 6, 8 and 21, signals were identified in regions of milk production. Marbling scores, QTLs and signals within BTA 12 were placed around QTLs for calving ease. Based on the results, we can conclude that the identified regions in the genome that are affected by positive selection correspond to the breeding goals for Slovak Spotted cattle.

Keywords: artificial selection, cattle, genomic region, homozygosity, iHS score.

INTRODUCTION

The bovine genome was one of the first among livestock sequenced. The reason could be the importance of cattle in human nutrition and their evolutionary position as representative of ruminants (Tellam et al., 2009). It is assumed that since cattle domestication have been selected by the human 800 breeds, due to various economic, social and religious reasons. Those breeds represent a significant part of the world heritage (Mason et al., 1998).

[1]Barbora Olšanská (corresponding author: baska.olsanska@outlook.com), Radovan Kasarda, Anna Trakovická, Ondrej Kadlečík, Nina Moravčíková, Department of Animal Genetics and Breeding Biology, Slovak University of Agriculture in Nitra, SLOVAKIA.

        Natural and artificial selection has led to the formation of cattle breeds which are specialized in particular directions. Several specific regions in the genome were under pressure due to selection and it is generally accepted that these genomic sequences or selection signatures, which control key phenotypes are involved in specific traits. Positive selection pressure can reduce or eliminate the negative allele frequency in the genome of offsprings (Biswas and Akey, 2006). Therefore, selection signals in populations could provide genomic information to simplify selection and provide information about history of selection (Akey et al., 2002; Willing et al., 2010).

        Determination of selection signals can be accomplished using several methods. For this study, extended haplotype homozygosity (EHH) test (Sabeti et al., 2002) and integrated haplotype homozygosity score (iHS) (Voight, 2006) were chosen. EHH is used for identifying long haplotypes that carry a so-called core allele with high frequency within the population and detection of genomic regions that are candidates for having undergone recent selection (Sabeti et al., 2002). The iHS method compares EHH between ancestral and derived alleles within the population. The iHS test can avoid the impact of heterogeneous recombination rates across whole genome of cattle. This iHS method is best performed when the selected allele segregates at intermediate frequencies in the population (Zhang et al., 2017).

        Recently, selection signals have been found in several local cattle breeds in Slovakia (Kasarda et al., 2015). Based on this, the aim of our study was to find selection signals in the genome of Slovak Spotted cattle by calculation of integrated haplotype homozygosity score (iHS). For each region with selection signal was identified quantitative traits loci and explore their potential functional genes in Slovak Spotted cattle using high-density single nucleotide polymorphism (SNP) genotyping data.

MATERIAL AND METHODS

        For this study dataset of 85 Slovak spotted cattle were used to detect signatures of selection in the genome. These animals were genotyped by two platforms: Illumina BovineSNP50v2 BeadChip (sire 37) and ICBF International Dairy and Beef v3 (dams 48), respectively.

        Standard quality control of genotyping data was performed to exclude individuals with >10% missing genotype across the autosomal loci, markers missing in >10% of individuals, and loci with minor allele frequency <0.01 using the PLINK software (Purcell et al., 2007). After quality control remained in dataset information about 35995 autosomal loci. Analysis of selection signals using iHS statistic was based on haplotype frequencies as specified Voight et al. (2006). The haplotypes were reconstructed for each autosome by using software fastPHASE (Sheet and Stephen, 2006). The iHS statistic reflects the structure of haplotype and indicates abnormal long haplotypes carrying the ancestral and derived allele (Qanbari et al., 2011). Standardized iHS was calculated as follows:

$$iHS = ln\frac{iHH_A}{iHH_D}$$

Where iHHA represents the integrated EHH score for ancestral allele and iHHD represents integrated EHH score for derived allele. The positive iHS values indicate higher homozygosity outlying the derived allele and negative values indicate higher homozygosity outlying the ancestral allele. This analysis was performed in the R software environment using package rehh (Gautier and Vitalis, 2012).

The genome-wide significance threshold of selected SNPs with positive value of iHS was determined by the outliers according the boxplot distribution and then all SNPs above the average of significance were assigned to the genomic QTL´s location according to the Bovine genome database (animalgenome.org). For identification of these genes the Genome data viewer was used (https://www.ncbi.nlm.nih.gov/genome/gdv/?org=bos-taurus).

RESULTS AND DISCUSSION

The selection signals in Slovak spotted cattle were detected based on the set of 35995 SNPs markers, with the average distance of 69.3 kb between adjacent SNPs. The average value of iHS was 0.83 ± 0.68. The highest iHS value was observed for marker located on chromosome 15 (8.03) in position 33.6 Mb. Throughout the entire genome, the lowest iHS (0.0001) was found for SNP on chromosome 5 in genomic position 115.5 Mb. The iHS method identified 15 regions in the genome of Slovak spotted cattle that were significantly affected by positive selection (Table 1). Genome-wide distribution of the iHS score was visualized to obtain the chromosomal distribution of selection signals (Figure 1).

A)   B)
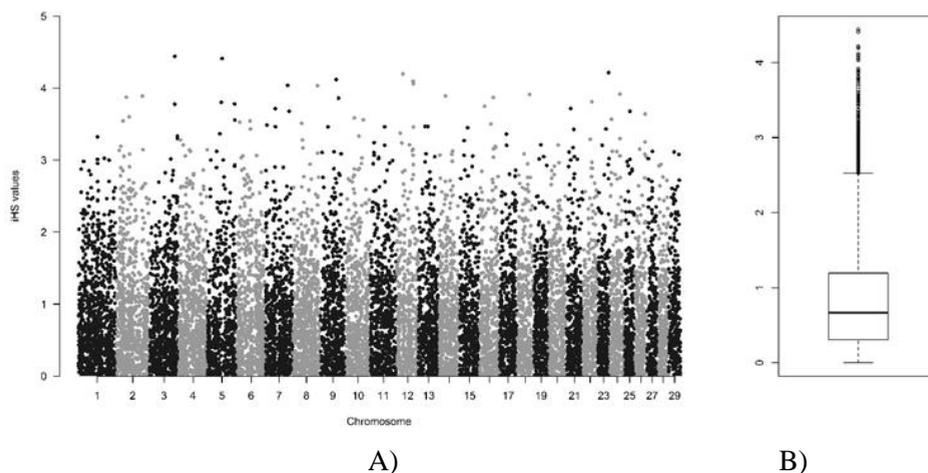
Figure 1. Genome-wide plot of the *iHS* score, A) the distribution of *iHS* values within the autosomes, B) the genome-wide significance threshold.

The genome-wide significance threshold of iHS score was set to 2.5, according to boxplot distribution, where outliers were identified as values over the maximum (Fig. 1B). Annotation of genomic regions with identified selection signals revealed several candidate genes, and those significant iHS genomic regions in Slovak Spotted cattle comprised 28 genes. BMPR1B genes were identified in the region on BTA6, Yao et al. (2018) demonstrated that by targeting BMPR1B the miR-125B regulates GC apoptosis in the yak ovary. In the region BTA21 CSK gene, affecting the immune system, was observed (Zimin et al., 2009). In the region on BTA3 the SCMH1 gene, upregulating the expression of several bone biomarkers, was found (Pei et al., 2019). The LPAR1 gene able to the response to tissue damage and infections agens was identified in the region on BTA6 (Zimin et al., 2009). In the region BTA8 PTGR1 gene which affects the prostaglandin reduce activity, was observed (Zimin et al., 2009). In the region BTA12 SOX21 gene affecting the hair follicle development, was obtained (Kiso et al., 2009). In the region BTA21 RHCG gene, able to regulate of pH, was found (Zimin et al., 2009). Using this analysis and by the describes QTL´s, we found that Slovak Spotted cattle were previously selected based on important traits for milk and beef production, muscle development, stature and reproduction (Table 1).

Schawanrzenbacher et al. (2012) detected the significant selection signals on BTA17 and 28 for Brown Swiss and on BTA14 observed SNP under strong selection in close neighbourhood to the DGAT1 K232A polymorphism, with strong effects on milk production traits. Iso-Touru et al. (2016) identified selection signals for Finnish Ayrshire and Eastern Finncattle on BTA 20, both breeds being used mainly to dairy production and the region on BTA21 harbored 44 different genes within Yakutian and Ukrainian Gray. Maiorano et al. (2018) observed the most significant signatures of selection for Gir cattle within dairy population on chromosome 16 (23 candidate genes) and within beef population on chromosome 6 (43 candidate genes).

The SLC24A4 gene located on BTA21 was common in the dairy and beef populations (Maiorano et al., 2018). Simianer et al. (2010) observed for Holstein, that candidate regions identified through the iHS test comprised genes involved in the biological processes such as anatomical structure development, muscle development, spermatogenesis and fertilisation and these findings are consistent also with Flori et al. (2009). Zhao et al. (2015) identified 80 significant regions with selection signals by iHS test for several commercial dairy and beef breeds, including Angus, Belgian Blue, Charolais, Hereford, Simmental cattle, Limousin and Holstein-Friesian. Kasarda et al. (2015) observed only seven regions with several genes (HGD, LOC101904412, DCAF15, EYA4, NUMA1, FTO, CA10) affected by selection in Slovak Pinzgau cattle. Zang et al. (2017) identified for Chinese Jinnan cattle QTL regions associated with meat and carcass traits on chromosome 8. These findings confirm that during the evolution of each breed has been performed selection for improvement of important phenotypic traits and to achieve breeding objectives.

Table 1. Genomic regions with significant iHS score in Slovak spotted cattle.

BTA	Position	iHS value	QTL´s
3	104999360-105064402	4.1	
6	31016787-31126062	2.81	carcass weight, body weight, calving ease
6	54456990-54554982	2.61	milk yield, milk fat yield, marbling score, calving ease
8	37192348-37451820	3.1	stillbirth, clinical mastitis, dystocia, calving ease, body weight (birth)
8	102530898-102664941	3.56	
8	102664941-103083474	2.8	udder swelling score, shear force
12	35324036-35454037	2.72	shear force, milk fat yield, body weight, structural soundness, height (mature)
12	68453584-68628611	3.11	milk fat yield, milk fat content, muscle iron content
12	68628611-68784880	3.6	muscle iron content
12	69874053-69925933	3.2	muscle iron content
12	69925933-69969069	2.87	muscle iron content
12	69969069-70000094	3.45	calving ease, daughter pregnancy rate
21	21090081-21442215	3.31	marbling score, milking speed, milk fat yield, body weight
21	34165847-34339514	3.1	shear force, somatic cell score, calving ease, body weight
21	36136065-36341943	2.93	somatic cell score, body weight

CONCLUSIONS

Studies of mapping selection signals in cattle provides an ideal opportunity to investigate how the artificial selection has influenced the variability and architecture of the bovine genome. In this study, several regions under strong selection pressure have been found by a genome-wide scan of iHS score in Slovak Spotted cattle.

This test showed that the breeding history of Slovak Spotted cattle has been subject to positive selection similarly as commercial dairy and beef breeds. Our analysis discovers genes under positive selection, which are related to beef production, immunity and reproduction and revealed quantitative traits loci under positive selection, which are reflecting the selection of the required traits according to the breeding objectives. These results confirm the importance of Slovak Spotted cattle as important dual-purpose genetic resource in Slovakia.

ACKNOWLEDGEMENTS

REFERENCES

Akey, J.M., Zhang, G., Zhang, K., Jin, L., Shriver, M.D. (2002). Interrogating a high-density SNP map for signatures of natural selection, Genome Res.,12, 1805-1814.

Biswas, S. & J. M. Akey (2006). Genomic insights into positive selection, Trends Genet., 2, 437-446.

Flori, L., Fritz, S., Jaffrézic, F, Boussaha, M., Gut, I., Heath, S., Foulley, J.L.. et al. (2009). The Genome Response to Artificial Selection: A Case Study in Dairy Cattle, PLoS One, 4(8).

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

Iso-Touru, T., Tapio, M., Vilkki, J., Kiseleva, T., Ammosov, I., Ivanova, Z., Popov, R., Ozerov, M., Kantanen, J.. et al. (2016). Genetic diversity and genomic signatures of selection among cattle breeds from Siberia, eastern and northern Europe, Animal Genetics, 47, 647-657.

Kasarda, R. et al. Moravčíková, N., Trakovická, A., Mészáros, G., Kadlečík, O. (2015). Genome-wide selection signatures in Pinzgau cattle, Potravinárstvo. 9(1), 268-274.

Kiso, M. Tanaka,S., Saba, R., Matsuda, S., Shimizu, A., Ohyama, M., Okano, H.J., Shiroishi, T., Okano, H., Saga, Yet al. (2009). The disruption of Sox21-mediated hair shaft cuticle differentiation causes cyclic alopecia in mice, Proc Natl Acad Sci U S A, 106(23), 9292-9297.

Maiorano, A. M., Lourenco, D.L., Tsuruta, S., Ospina, A.M.T., Stafuzza, N.B., Masuda, Y., Filho, A.E.V., Cyrillo, J.N.D.S.G., Curi, R.A., Silva, J.A.I.V. et al. (2018). Assessing genetic architecture and signatures of selection of dual purpose Gir cattle populations using genomic information, PLoS ONE, 13(8).

Mason, I. L. (1998). World Dictionary of Livestock Breeds. Wallingford, UK: CAB International.

Pei, Y. F., Liu, L., Liu, T.L., Yang, X.L., Zhang, H., Wei, X.T., Feng, G.J., Hai, R., Ran, S., Zhang, L.e t al. (2019). Joint Association Analysis Identified 18 New Loci for Bone Mineral Density, Journal of bone and mineral research.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., Sham, P.C. et al. (2007). PLINK: a toolset for whole-genome association and population-based linkage analysis, Am J Hum Genet, 81, 559-575.

Qanbari, S.,Gianola, D., Hayes, B., Schenkel, F., Miller, S., Moore, S., Thaller, G., Simianer, H. et al. (2011). Application of site and haplotype-frequency based approaches for detecting selection signatures in cattle, BMC Genomics, 12, 318.

Sabeti, P. C., Reich, D.E., Higgins, J.M., Levine, H.Z., Richter, D.J., Schaffner, S.F., Gabriel, S.B., Platko, J.V., Patterson, N.J., McDonald, G.J., Ackerman, H.C., Campbell, S.J., Altshuler, D., Cooper, R., Kwiatkowski, D., Ward, R., Lander, E.S. et al. (2002). Detecting recent positive selection in the human genome from haplotype structure, Nature, 419, 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, American Journal of Human Genetics, 78, 629-644.

Schwarzenbacher, H., Dolezal, M., Flisikowski, K., Seefried, F., Wurmser, C., Schlötterer, C., Fries, R. et al. (2012). Combining evidence of selection with association analysis increases power to detect regions influencing complex traits in dairy cattle, BMC Genomic 13(48).

Simianer, H.,Qanbari, S., Gianola, Det al. (2010). Detection of selection signatures within and between cattle populations, Proceedings of 9th World Congress on Genetics Applied to Livestock Production.

Tellam R.L. (2009). Unlocking the bovine genome, BMC Genomics,10.

Voight, B. F. (2006). A map of recent positive selection in the human genome, PLoS Biol, 4(72)

Willing, E. M., Bentzen, P., van Oosterhout, C., Hoffmann, M., Cable, J., Breden, F., Weigel, D., Dreyer, C.. et al. (2010). Genome-wide single nucleotide polymorphisms reveal population history and adaptive divergence in wild guppies, Mol. Ecol, 19, 968-984

Yao, Y. et al., Niu, J., Sizhu, S., Li, B., Chen, Y., Li, R., Yangzong, Q., Li, Q., Xu, Y. (2018). microRNA-125b Regulates Apoptosis by Targeting Bone Morphogenetic Protein Receptor 1B in Yak Granulosa Cells, DNA Cell Biol, 37(11), 878-887.

Zhang, Y., He, D., Cheng, J., Xu, F., Li, B., Jin, G., Zhang, X. et al. (2017). Detection of selection signatures based on the integrated haplotype score in Chinese Jinnan cattle, Emirates Journal of Food and Agriculture. 29(7), 562-566.

Zhao, F., et al McParland, S., Kearney, F., Du, L., Berry, D.P. (2015). Detection of selection signatures in dairy and beef cattle using high-density genomic information, Genet. Sel. Evol. 47(49).

Zimin, A.V. et al., Delcher, A.L., Florea, L., Kelley, D.R., Schatz, M.C., Puiu, D., Hanrahan, F., Pertea, G., Van Tassell, C.P., Sonstegard, T.S., Marçais, G., Roberts, M., Subramanian, P., Yorke, J.A., Salzberg, S.L. (2009). A whole-genome assembly of the domestic cow, Bos taurus, Genome Biology, 10(r42).