

*Nina MORAVČÍKOVÁ\**, *Ondrej KADLEČÍK*,  
*Anna TRAKOVICKÁ*, *Radovan KASARDA*<sup>1</sup>

## **AUTOZYGOSITY ISLAND RESULTING FROM ARTIFICIAL SELECTION IN SLOVAK SPOTTED CATTLE**

### **SUMMARY**

The aim of this study was to identify the runs of homozygosity (ROH) segments within genome and to identify regions significantly affected by artificial selection in Slovak Spotted cattle. In total of 85 animals representing the nucleus of Slovak Spotted cattle were included in present study. The sampled population consisted of 37 AI sires and 48 dams of sires that were genotyped by using two platforms, Illumina BovineSNP50v2 BeadChip and ICBF International Dairy and Beef v3. The consensus map file, constructed to identify common SNPs within these platforms, consisted of overall 40,033 markers. Subsequently, all of loci with call rate lower than 90% and minor allele frequency lower than 0.01% were removed. The final database included genotyping information for 37,833 SNPs. The ROH segments were defined as genomic regions with 15 or more consecutive homozygous calls at density of one SNPs on every 100 kb and maximum gap between consecutive SNPs of 1 Mb. One heterozygous call was allowed for length >16 Mb. In addition, one missing call was allowed for length >4 Mb, 2 for >8 Mb and 4 for >16 Mb. The subsequent analyses of genome-wide selection signatures were based on the assumption that the most frequent homozygous regions in population (minimum ROHs with length 4 Mb) reflected the recent selection characteristic for Slovak Spotted breed. As expected due to the dual-purpose character of this breed the strongest signals of artificial selection was found direct or very close to genes associated with milk production and beef quality.

**Keywords:** Autozygosity, cattle, genotyping data, selection signatures.

### **INTRODUCTION**

Domestication followed by breed formation and selection schemes has allowed the formation of very diverse cattle breeds that are adapted to a wide variety of environments and specialized for either milk or meat production or raised as dual-purpose breeds in different regions of the globe. Natural and artificial selection strategies is likely to impose pressure on specific genome regions that control these traits (milk and meat production) as well as other

---

<sup>1</sup>Nina Moravčíková\*(corresponding author: nina.moravcikova1@gmail.com), Ondrej Kadlečík, Anna Trakovická, Radovan Kasarda, Department of Animal Genetics and Breeding Biology, Slovak University of Agriculture in Nitra, SLOVAKIA

Paper presented at the 9<sup>th</sup> International Scientific Agricultural Symposium "AGROSYM 2018".

Notes: The authors declare that they have no conflicts of interest. Authorship Form signed online.

important characteristics such as adaptation to different environments, reproduction, body conformation, behavior and resistance to diseases (de Simoni Gouveia *et al.*, 2014; Brito *et al.*, 2017). Thus, intensive selection pressure then resulted in specific changes in the patterns of variation across loci in particular genomic regions controlling traits of interest as well as in neutral loci linked to them (Zhao *et al.*, 2015). Such unique genetic patterns left behind in the genome of individuals under natural and/or artificial selection are defined as signatures of selection (Qanbari and Simianer, 2014).

The detection of selection signatures is a relevant topic since it has the potential to elucidate the identities of genes and mutations associated with phenotypic traits even if they are no longer segregating within any of the populations of interest and does not necessarily require phenotypes measures. Furthermore, this knowledge is important in order to better understand the evolution process and the mechanisms that underlie traits that have been exposed to intensive natural and artificial selection (Brito *et al.*, 2017).

In recent years, the availability of high-density genomic data and parallel progress in statistical techniques have allowed the identification of genomic regions that have been subjected to natural and artificial selection in cattle (Gutiérrez-Gil *et al.*, 2015). One of the methods that can be used to determine the impact of artificial selection on the genome architecture is the analysis of ROH segments distribution in the genome. Runs of homozygosity (ROH), defined as contiguous homozygous regions of the genome where the two haplotypes inherited from the parents are identical, reflect essentially the autozygosity of an individual (Ferenčaković *et al.*, 2013; Curik *et al.*, 2014). The length and frequency of ROH can be used for example to describe the history of the population in which an individual occurs, to describe the trend of increase in inbreeding within population or to reveal recent population bottleneck (Szmatola *et al.*, 2016). Moreover, it has been showed that the distribution of ROH islands in the genome isn't random and due to the various selection events the ROH are distributed and shared among related individuals (Zhang *et al.*, 2015). Because of this and based on the fact that the genomic regions sharing ROH potentially contain alleles associated with genetic improvement in cattle, they can be used to identify the genomic regions under the strongest selection pressure (Marras *et al.*, 2015; Peripolli *et al.*, 2018).

The aim of this study was to determine the impact of artificial selection on the Slovak Spotted cattle genome based on the analysis of genome-wide ROH distribution and to identify genomic regions resulting from the intensive selective breeding for traits of interest during the development of this breed.

## **MATERIAL AND METHODS**

In total of 85 animals representing the nucleus of Slovak Spotted cattle were included in present study. The sampled population consisted of 37 AI sires and 48 dams of sires that were genotyped by using two platforms, Illumina

BovineSNP50v2 BeadChip (sires) and ICBF International Dairy and Beef v3 (dams), in commercial lab.

The subsequent quality control of SNP genotyping data was performed by using PLINK 1.9 (Chang *et al.*, 2015). The quality control was carried out to remove all of SNP markers assigned to unmapped regions or with unknown chromosomal position according to the latest bovine genome assembly (Btau 4.0) and SNPs positioned to sex chromosomes. In next step, because of the two different genotyping platforms used for animals' genotyping, the consensus map have to be constructed.

The final consensus map file consisted of 40,033 markers. In the subsequent SNP pruning only samples with lower than 10 % of missing genotypes, autosomal SNPs with call rate higher than 90 % and minor allele frequency higher than 1 % that adhered to mendelian inheritance patterns were retained. The final database included genotyping information for 37,833 SNPs.

The ROH segments were defined as proposed Ferenčaković *et al.* (2013) as genomic regions with 15 or more consecutive homozygous calls with maximum gap between consecutive SNPs of 1 Mb and minimum density of one SNPs on every 100 kb. Because of the theoretical relationship between the distribution of identity by descent (IBD) fragments and the number of generation since common ancestor the minimum length of ROH segments was set to 1 Mb. Due to the fact that ROH segments with specific length can reflect different generation of ancestors (Ferenčaković *et al.*, 2013; Curik *et al.*, 2014) the distribution of ROH segments in the genome were analysed separately for five length categories (>1 Mb, >2 Mb, >4 Mb, >8 Mb, and >16 Mb). Heterozygous calls were not allowed across ROH categories, except length >16 Mb with one permissible call. Missing calls per windows were not allowed for lengths >1 Mb and >2 Mb, while one missing call were accepted for length >4 Mb, 2 for >8 Mb and 4 for >16 Mb. Subsequently, the total number of ROH detected, the average length of ROH (in Mb) and the sum of all ROH segments by animals were calculated for each ROH length category. The proportion of autosomes covered by ROH was then expressed as the pools of overlapping segments within animals per each breed by using R package detectRUNS (Biscarini *et al.*, 2018).

The analysis of genome-wide selection signatures was based on the assumption that the identified autozygosity islands across the genome of Slovak Spotted cattle are a results of selective breeding for traits of interest defined in their breeding objectives. The autozygosity islands, characterized by SNPs with extreme frequency in ROH segments >4 Mb across specific genomic regions, were determined by the calculation of runs incidence per each SNP using R package detectRUNS (Biscarini *et al.*, 2018). The genome-wide occurrence of SNPs in ROH was then expressed as the frequency (%) of overlapping ROH shared among individuals. Genomic regions under selection were defined by the top 0.01 percentile of signals. For identifying genes located directly in the detected regions the Genome data viewer of the bovine genome UMD3.1.1 was used (<https://www.ncbi.nlm.nih.gov/genome/gdv/>).

## RESULTS AND DISCUSSION

The final database of SNP genotyping data consisted of 37,833 markers covering overall length 2,496,829 kb of the genome. The average obtained distance between adjacent SNPs ( $66.05 \pm 70.27$  kb) was in accordance with previous studies that utilized 50K SNP panel for cattle genotyping (Flury *et al.*, 2010; Beghain *et al.*, 2012). The minimum distance between adjacent SNP markers was 0.001 kb and the maximum was 4428.95 kb. All of animals under consideration showed call rate greater than 90 %. The average call rate at level 99.16 % was comparable with results published for dairy as well as beef cattle (Cooper *et al.*, 2013; Mullen *et al.*, 2013).

As expected based on the previous studies the ROH segments were detected in the genome of all analysed animals. The average sum of ROHs calculated per animal was  $70.65 \pm 14.95$  and the average length of ROHs was  $153.24 \pm 51.52$  Mb. Overall, the ROH segments represents 6.14 % of the Slovak Spotted genome covered by SNP markers. However, the distribution of ROH segments across the autosomal genome wasn't uniform. The major fraction of autosome residing in ROH was found on BTA6 (13.5 %), whereas the lowest proportion of ROHs showed BTA24. The obtained proportion of overall autozygosity per animals was comparable with other cattle breeds (Kim *et al.*, 2015; Mastrangelo *et al.*, 2016; Szmatoła *et al.*, 2016; Kukučková *et al.*, 2017).

The analysis of ROHs distribution showed that in the genome of Slovak Spotted cattle across all of length ROH categories in total of 8,632 segments were located. As showed many studies the different length categories of ROH as well as the frequency of ROH occurrence in the genome give insight into the history of target population (Szmatoła *et al.*, 2016) and can be used to estimate the age of inbreeding (Curik *et al.*, 2014). It is generally accepted that very long autozygous segments (>16 Mb) are originated from recent common ancestors, whereas most of the short segments are likely derived from more remote ancestors (Curik *et al.*, 2017). In cattle, Ferenčaković *et al.* (2013) showed that the ROH >1 Mb date back ~50 generations, >2 Mb ~25 generations, >4 Mb ~12.5 generations, >8 Mb ~ 6 generations, and >16 Mb ~3 generations. In Slovak Spotted cattle the total length of ROHs was composed mostly from short segments >1 Mb that accounted for 69.57 %. The ROH segments of 2 – 4 Mb long, representing the 25 – 12.5 generations from common ancestor, accounted for 28.30 %. The lowest proportion within the total length of ROH was found for the longest segments (>16 Mb) that accounted for 0.43 % of all segments identified in the genome. From the point of view of genetic diversity it can be conclude that overall 1.47 % (genome coverage of ROH >16 Mb) of the Slovak Spotted genome can be affected by the increase in inbreeding in recent generation.

The analysis of selection signatures was based on the assumption that the genomic regions with extreme ROH frequencies are most likely consequences of selective breeding for traits of interest during the formation of target breed. Only ROH segments greater than 4 Mb were taken into account mainly due to the fact

that this length category date back ~12.5 generations of common ancestors, which correspond to the breeding history of Slovak Spotted cattle in Slovakia. Genomic regions under selection were defined by the top 0.01 percentile of signals (Figure 1).

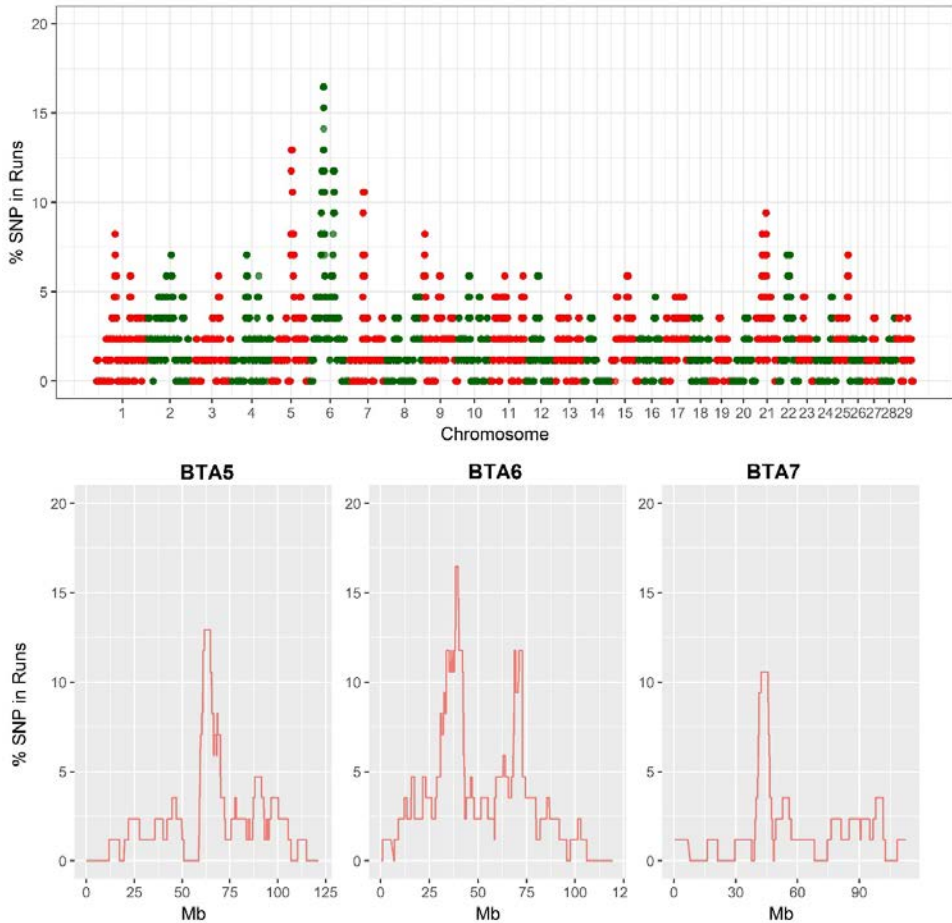


Figure 1. Genome-wide occurrence of SNPs in ROH (minimum ROH length set to 4 Mb) across autosomal genome of Slovak Spotted cattle.

According to this criteria overall four genomic regions located within three autosomes were found: BTA5 (59,665,562-68,880,383 bp), BTA6 (32,489,075-42,866,573 bp and 68,546,212-72,969,608 bp), and BTA7 (41,321,459-46,354,401 bp). Inside identified regions a number of genes were detected, including those involved in genetic control of milk production (ABCG2, SPP1, FAM13A), feed efficiency and growth (FAM184B, PKD2), muscle formation and body composition (PACRGL), reproduction (GDF9), and coat colour pattern (KIT, KDR).

The observed intensive selection pressure particularly in the genomic areas associated with economically important production traits was confirmed also in previous studies. The extensive study of The Bovine HapMap consortium (2009) showed that within the genomes of 19 geographically and biologically distinct cattle breeds the most genomic regions displaying selection signatures were related especially to milk yield, meat quality and feed efficiency.

Another studies identified in several dairy and beef breeds the extreme signals of selection mainly in the regions involved in genetic control of coat colour (MCR1, KIT) and body size (McClure *et al.*, 2010; Lee *et al.*, 2014; O'Brien *et al.*, 2014; Porto-Neto *et al.*, 2014; Kim *et al.*, 2017). Similarly, Rothammer *et al.* (2013) showed that the selection signatures are mostly limited to sequences of genes associated with QTL regions controlling economically important traits of cattle (TG, ABCG2, DGAT1, GH1, GHR, casein family genes).

### CONCLUSIONS

Our study showed that the autozygosity islands covered in average 6.14 % of the Slovak Spotted genome expressed by the SNP markers under consideration. The distribution of ROH segments across the autosomal genome wasn't uniform and the major fraction of autosome residing in ROH was found on BTA6. The intensive pressure of artificial selection on genomic regions on BTA6 was confirmed also by the subsequent analysis of selection signatures.

As expected due to the dual-purpose character of Slovak Spotted cattle the majority part of selection signatures was detected directly or very close to genes associated with milk production, feed efficiency and growth, muscle formation, body composition, reproduction, and coat colour pattern. The detailed analysis of SNP markers located in those regions can be valuable for genetic improvement of Slovak Spotted cattle in the future.

### ACKNOWLEDGEMENT

This study was supported by the Slovak Research and Development Agency (APVV-14-0054 and APVV-17-0060) and VEGA (1/0742/17).

### REFERENCES

- Beghain J., Boitard S., Weiss B., Boussaha M., Gut I., Rocha D. (2012). Genome wide linkage disequilibrium in the Blonde d'Aquitaine cattle breed, *Journal of Animal Breeding and Genetics*, Vol. 130, pp. 294-302.
- Biscarini F., Cozzi P., Gaspa G., Marras G. 2018. detectRUNS: Detect Runs of Homozygosity and Runs of Heterozygosity in Diploid Genomes.
- Brito L.F., Kijas J.W., Ventura R.V., Sargolzaei M., Porto-Neto L.R., Cánovas A., Feng Z., Jafarikia M., Schenkel F.S. (2017). Genetic diversity and signatures of selection in various goat breeds revealed by genome-wide SNP markers, *BMC Genomics*, Vol. 18, pp. 229.
- Chang C.C., Chow C.C., Tellier L.C.A.M., Vattikuti S., Purcell S.M., Lee J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets, *GigaScience*, Vol. 4, pp. 7.

- Cooper T.A., Wiggans G.R., VanRaden P.M. (2013). Short communication: relationship of call rate and accuracy of single nucleotide polymorphism genotypes in dairy cattle, *Journal of Dairy Science*, Vol. 96, pp. 3336-3339.
- Curik I., Ferenčaković M., Sölkner J. (2014). Inbreeding and runs of homozygosity: A possible solution to an old problem, *Livestock Science*, Vol. 166, pp. 26-34.
- Curik, I., Ferenčaković, M., Sölkner, J. (2017). Genomic dissection of inbreeding depression: a gate to new opportunities, *Revista Brasileira de Zootecnia*, Vol. 46 (9).
- de Simoni Gouveia J.J., da Silva M.V.G.B., Paiva S.R., de Oliveira S.M.P. (2014). Identification of selection signatures in livestock species, *Genet Mol Biol*, Vol. 37 (2), pp. 330-342.
- Ferenčaković M., Hamzić E., Gredler B., Solberg T.R., Klemetsdal G., Curik I., Sölkner J. (2013). Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations, *Journal of Animal Breeding and Genetics*, Vol. 130, pp. 286-293.
- Flury C., Tapio M., Sonstegard T., Drogemuller C., Leeb T., Simianer H., Hanotte O., Rieder S. (2010). Effective population size of an indigenous Swiss cattle breed estimated from linkage disequilibrium, *J Anim Breed Genet*, Vol. 127, pp. 339-347.
- Gutiérrez-Gil B., Arranz J.J., Wiener P. (2015). An interpretive review of selective sweep studies in *Bos taurus* cattle populations: identification of unique and shared selection signals across breeds, *Front Genet*, Vol. 13 (6), pp.167.
- Kim J., Cho S., Caetano-Anolles K., Kim H., Ryu Y.C. (2015). Genome-wide detection and characterization of positive selection in Korean Native Black Pig from Jeju Island, *BMC Genetics*, Vol. 16 (3), pp. 1-9.
- Kim S.J., Ka S., Ha J.W., Kim J., Yoo D., Kim K., Lee H.K., Lim D., Cho S., Hanotte O., Mwai O.A., Dessie T., Kemp S., Oh S.J., Kim H. (2017). Cattle genome-wide analysis reveals genetic signatures in trypanotolerant N'Dama, *BMC Genomics*, Vol. 18, pp. 1-18.
- Kukučková V., Moravčíková N., Ferenčaković M., Simčič M., Mészáros G., Sölkner J., Trakovická A., Kadlečík O., Curik I., Kasarda R. (2017). Genomic characterization of Pinzgau cattle: genetic conservation and breeding perspectives, *Conservation Genetics*, Vol. 18 (4), pp. 893-910.
- Lee S.H., Park B.H., Sharma A., Dang C.G., Lee S.S., Choi T.J., Choy Y.H., Kim H.C., Jeon K.J., Kim S.D., Yeon S.H., Park S.B., Kang H.S. (2014). Hanwoo cattle: origin, domestication, breeding strategies and genomic selection, *Journal of Animal Science and Technology*, Vol. 56 (2), pp. 1-8.
- Marras G., Gaspa G., Sorbolini S., DimAURO C., Ajmone-Marsan P., Valentini A., Williams J.L., Macciotta N.P. (2015). Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy, *Animal Genetics*, Vol. 46 (2), pp. 110-121.
- Mastrangelo S., Tolone M., Di Gerlando R., Fontanesi L., Sardina M.T., Portolano B. (2016). Genomic inbreeding estimation in small populations: evaluation of runs of homozygosity in three local dairy cattle breeds, *Animal*, Vol. 10, pp. 746-754.
- McClure M.C., Morsci N.S., Schnabel R.D., Kim J.W., Yao P., Rolf M.M., McKay S.D., Gregg S.J., Chapple R.H., Northcutt S.L., Taylor J.F. (2010). A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle, *Animal Genetics*, Vol. 41 (6), pp. 597-607.
- Mullen M.P., McClure M.C., Kearney J.F., Waters S.M., Weld R., Flynn P., Creevey C.J., Cromie A.R., Berry D.P. (2013). Development of a Custom SNP Chip for

- Dairy and Beef Cattle Breeding, Parentage and Research, *Interbull Bulletin*, Vol. 47, pp. 23-25.
- O'Brian A.M., Mészáros G., Utsunoomiya Y.T., Sonstegard T.S., Garcia J.F., Van Tassell C.P., Carvalheiro R., da Silva M.V.B., Solkner J. (2014). Linkage disequilibrium levels in *Bos indicus* and *Bos taurus* cattle using medium and high density SNP chip data and different minor allele frequency distributions, *Livestock Science*, Vol. 166, pp. 121-132.
- Peripolli E., Stafuzza N.B., Munari D.P., Lima A.L.F., Irgang R., Machado M.A., Panetto J.C.D.C., Ventura R.V., Baldi F., da Silva M.V.G.B. (2018). Assessment of runs of homozygosity islands and estimates of genomic inbreeding in Gyr (*Bos indicus*) dairy cattle, *BMC Genomics*, Vol. 19 (1), pp. 34.
- Porto-Neto L.R., Lee S.H., Sonstegard T.S., Van Tassell C.P., Lee H.K., Gibson J.P., Gondro C. (2014). Genome-wide detection of signatures of selection in Korean Hanwoo cattle, *Animal Genetics*, Vol. 45 (2), pp. 180-190.
- Qanbari S., Simianer H. (2014). Mapping signatures of positive selection in the genome of livestock, *Livestock Science*, Vol. 166, pp. 133-143.
- Rothhammer S., Seichter D., Förster M., Medugorac I. (2013). A genome-wide scan for signatures of differential artificial selection in ten cattle breeds, *BMC genomics*, Vol. 14, pp. 908.
- Szmatoła T., Gurgul A., Ropka-Molik K., Jasielczuk I., Ząbek T., Bugno-Poniewierska M. (2016). Characteristics of runs of homozygosity in selected cattle breeds maintained in Poland, *Livestock Science*, Vol. 188, pp. 72-80.
- The Bovine HapMap consortium (2009). Genome-survey of SNP variation uncovers the genetic structure of cattle breeds, *Science*, Vol. 324, pp. 528-532.
- Zhang Q., Guldbandsen B., Bosse M., Lund M.S., Sahana G. (2015). Runs of homozygosity and distribution of functional variants in the cattle genome, *BMC Genomics*, Vol. 16, pp. 542.
- Zhao F., 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*, Vol. 47 (1), pp. 49.