Runs of homozygosity patterns in beef cattle

Barbora Olšanská*, Kristína Lehocká, Radovan Kasarda, Nina Moravčíková
Slovak University of Agriculture in Nitra, Faculty of Agrobiology and Food Resources, Slovakia

The study presents an approach to detect selection signals by the scan of runs of homozygosity (*ROH*) in two beef cattle populations bred Slovakia. The frequency and distribution of runs of homozygosity in the genome are affected by natural and artificial selection, recombination rate and structure of the population. After quality control, the final data set included 43,427 single nucleotide polymorphisms with overall length 2,504 Mb. Across the genome, sixteen regions under strong selection pressure with a total length of 73.94 Mb were identified. The functional analysis of selection signals revealed several quantitative trait loci for body structure, fitness and milk production. In the region with a high frequency of *ROH* reflecting the intense artificial selection genes related mainly to muscle development (MSNT, ROCK1, LAMTOR5) were observed. Besides, genes related to the control of the immune system (PTX3, FGL2) and reproduction (ADCYAP1R1) were localized within selection signals. The results confirm the intention to improve the production and reproduction traits of Charolais and Limousine cattle according to established breeding objectives for each breed.

Keywords: beef production, genomic region, homozygosity, quantitative traits, signatures of selection

1   Introduction

Since domestication, significant genetic improvement has been achieved for many commercial important cattle traits including adaptation, conformation, milk and beef production. In response to strong selection pressure, the bovine genome has undergone changes mainly in regions that control the preferred phenotype. Access to the bovine genome sequence and high-density genotypic panels provides remarkable resources to study the effects of domestication and selection on the structure of the cattle genome (Randhawa et al., 2016; Moravčíková et al., 2018). Selection of the best animals reduced the diversity of haplotypes and increased homozygosity around the target loci resulting in the formation of runs of homozygosity (*ROH*) segments across the genome. The runs of homozygosity are defined as regions containing consecutive homozygous genotypes. The identification and description of homozygous segments can provide insight into the evolution of population history, structure and demography. Such population phenomena can affect patterns of homozygosity in the genome and can be detected by the identification of runs of homozygosity (Zavarez et al., 2015; Williams et al., 2015; Zhang et al., 2015).

The Charolais and Limousine are characterized by excellent growth potential, feeding efficiency and good carcass quality. The breeding objectives of both breeds were focused to increase body weight while maintaining of calving ease (Sifuentes et al., 2015; Clarke et al., 2009). The aim of this study was evaluated the effect of artificial selection on the genome structure of Charolais and Limousin cattle based on the *ROH* patterns and performed functional analysis of regions under selection pressure by identifying QTL traits and genes located in these regions.

2   Material and methods

The study was performed on 85 animals (Charolais 68, Limousine 17). Genomic DNA was extracted from the hair roots samples of analysed animals. The initial SNP pruning performed in the PLINK v1.9 environment (Chang et al., 2015).

Corresponding Author: Barbora Olšanská, Slovak University of Agriculture in Nitra, Faculty of Agrobiology and Food Resources, Department of Animal Genetics and Breeding Biology, Slovakia; e-mail: baska.olsanska@outlook.com

From the database, all markers on gonosomes and markers with unknown position were extracted. All individuals and markers with missing more than 10% of genotypes were removed. Analysis of *ROH* segments distribution was performed according to Ferenčaković et al. (2013). *ROH* segments were defined as 15 consecutive homozygous SNP markers with a maximum distance between markers 1 Mb, a minimum density 1 SNP marker per each 100 kb and minimum *ROH* length set to 4Mb. The occurrence of a heterozygous genotype was not allowed. Only one missing genotyped within homozygous segments was permissible. The selection signals were defined based on the SNP markers with extreme frequency in specific *ROH* segments. For the analysis of *ROH* segments was used the PLINK v 1.9 (Chang et al., 2015) and for visualisation was used R package qqman (Turner, 2017). The cut-off value defining the selection signal across the genome was set by upper quartile of a boxplot. The outliers above the upper quartile reflected the regions significantly affected by intensive selection. Subsequently, all SNPs above the cut-off value was assigned to the genomic QTL location according to the QTLcattledb (animalgenome.org) and genes according to the genome data viewer (https://www.ncbi.nlm.nih.gov/genome/gdv/).

3   Results and discussion

After applying the quality control the final dataset consisted of 43,427 SNP markers that covered overall 2,503,601 kb of the autosomal genome with average spacing 57.69 kb. The cut-off value for selection signals was set to 5. The analysis revealed 16 genomic regions under strong selection pressure (Figure 1). The average number of SNPs markers residing in *ROH* was 1.38 ±1.69 and overall length of homozygous regions was 73.94 Mb. In regions under selection pressure QTLs for body structure (body weight, stature, body depth, foot angle, rump width, chest depth, feet and length conformation), reproduction (calving ease, sperm motility, daughter pregnancy rate, stillbirth, calving interval, first service conception) and milk production (milk fat yield, milk yield, somatic cell score, milk protein yield) were identified. Szmatoła et al. (2016) identified 15 regions with increased *ROH* frequency for Limousine cattle. Purfield et al. (2012) reported that genomic regions under selection pressure in the beef genome are located mainly on BTA 7, 14, 16 and 18 and include important QTL and genes related with the immune system, musculature and calving ease. Purfield et al. (2012) find out that the total length of *ROH* varied significantly across dairy and beef breeds (e.g. Holstein 178.5 Mb, for Limousine 71.3 Mb and for Simmental 78.1 Mb).

The genomic scan for homozygous regions across the 29 autosomes of analysed breeds pointed on several genes with different biological and molecular function (Table 1). On BTA 1, gene PTX3 which is involved in the innate immune response and fertility was identified (Gaudet et al., 2011). Gene MSTN, which play a role as a regulator of muscle
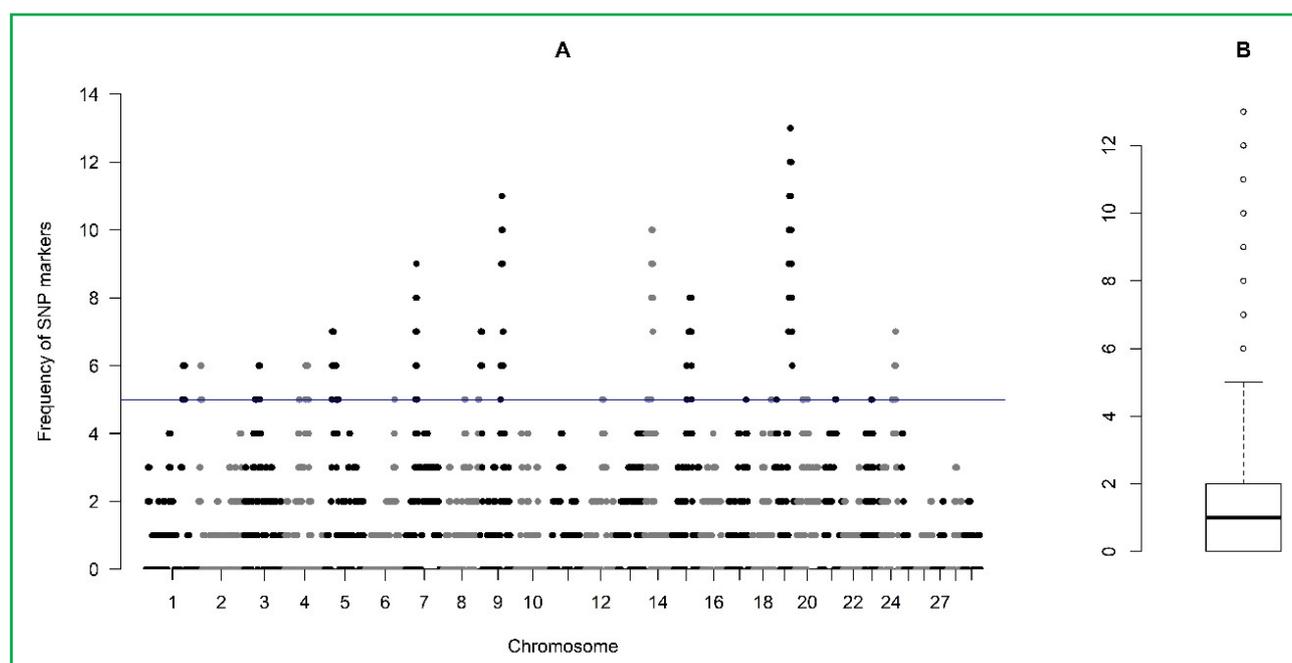

Figure 1   A – the genome-wide occurrence of SNP markers with high frequency in the beef genome; B – boxplot showing the genome-wide threshold for selection signals

Table 1   Description of 16 overlapping genomic regions under selection pressure for beef cattle

CHR	Position (Mb)	Genes
1	109.03–116.93	PTX3, KCNAB1, SLC33A1, MME, RAP2B, P2RY1, MBNL1, LEKR1
2	5.54–10.62	INPP1, MSTN, ASNSD1, WDR75, COL3A1, GULP1, TFPI, CALCRL, ITGAV
3	32.94–37.39	KCNA2, LAMTOR5, CYM, PROK1, SLC16A4, CSF1, GNAT2, GNAI3, SARS
3	44.23–48.31	SNX7, DPYD
4	42.69–45.02	MAGI2, CCDC146, FGL2, LRRC17, ARMC10, NAPEPLD, PMPCB, PSMC2, DNAJC2
4	61.02–71.60	ADCYAP1R1, GHRHR, AQP1, CRHR2
5	20.74 -27.09	KERA, LUM, DCN, GLYCAM1, PDE1B, ITGA5, HNRNPA1, SMUG1, ATP5G2, SP1
5	30.28–34.52	TMBIM6, TUBA1C, TUBA1A, PRKAG1, WNT1, CCNT1, LALBA, COL2A1, VDR, RAPGEF3
5	36.39–40.87	PUS7L, CNTN1, MUC19
7	33.91–35.43	
9	60.53–62.75	MAP3K7, BACH2, ANKRD6, UBE2J1, RNGTT, GJA10, MDN1, MIR7862, GABRR1, GABRR2, MIR2903, RRAGD
14	10.31–10.72	
14	15.54–16.92	NSMCE2, TRIB1, KIAA0196
15	43.75–45.48	SWAP70, TMEM9B, AKIP1, RPL27A, TRIM66, STK33, LMO1, TUB
19	71.64–77.31	ANKFN1, NOG
24	34.38–45.90	ROCK1, ADCYAP1, MYL12A, LAMA1, NDUFV2, CIDEA, MC2R

growth factor, was detected on BTA 2 (Trukhachev et al., 2015). On BTA 3, gene LAMTOR5, which supports cell growth in response to growth factor, was observed (Gaudet et al., 2011). Gene FGL2, which play a role in the immune system through the production of immunoglobulin, was identified on BTA 4 (Connor et al., 2008). Gene ADCYAP1R1, which may be involved in spermatogenesis and sperm motility, was found on BTA 4 (Gaudet et al., 2011). On BTA 5 was detected gene LALBA, which allows the synthesis of lactose, the main component of milk (Wyatt et al., 2013). Gene ROCK1, which is involved in the regulation of smooth muscle contraction, was identified on BTA 24 (Gaudet et al., 2011). Szmatoła et al. (2016) identified for Limousine also gene MSTN on BTA 2 and they reported that this gene can be recognized as a candidate gene for the double-muscling. The similar regions on BTA 2 identified Marras et al. (2015). Szmatoła et al. (2019) identified regions affected by artificial selection on BTA 5 and BTA 6 for Charolais cattle that contains mainly genes related to growth and coat colour. Comparison of our results with other studies showed that the distribution of *ROH* regions in the different beef breeds is similar as a consequence of common breeding objectives across breeds.

4   Conclusion

We can conclude that the influence of natural and artificial selection in order to improve production traits definitely shaped up of genome of breeds Limousine and Charolais. The result showed that the strong selection pressure was mainly targeted to regions controlling reproduction traits (ADCYAP1R1), muscle development (MSNT, ROCK1, LAMTOR5) and immune system (PTX3, FGL2). This study provides a genome-wide map of regions under selection pressure in beef cattle and the basis for further functional analysis of biological mechanisms behind complex phenotype traits of the analysed breeds.

Acknowledgement

References

CLARKE, A.M. et al. (2009). Intake, live animal scores/measurements and carcass composition and value of late-maturing beef and dairy breeds. *Livestock Science,* 126, 57–68.

CONNOR, E.E. et al. (2008). Effects of increased milking frequency on gene expression in the bovine mammary gland. *BMC Genom,* 9, 362.

FERENČAKOVIĆ, M. et al. (2013). Estimating autozygosity from high-throughput information: effects of SNP density and genotyping errors. *Genetics Selection Evolution,* 76(4), 325–329.

GAUDET, P. et al. (2011). Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. *Brief Bioinform*., 12(5), 449–462.

CHANG, C.C. et al. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience,* 25(4).

MARRAS, G. et al. (2015). Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy. *Animal Genetics,* 4,110–121.

MORAVČÍKOVÁ, N. et al. (2018). Genomic Response to Natural Selection within Alpine Cattle Breeds. *Czech J. Anim. Sci.,* 63(4), 136–143.

PURFIELD, D.C. et al. (2012). Runs of homozygosity and population history in cattle. *BMC Genet*, 13.

RANDHAWA, I.A.S. et al. (2016). A Meta-Assembly of Selection Signatures In cattle. *PLOS One,* 11(4).

SIFUENTES, A.M. et al. (2015). Loci asociados con enfermedades genéticas y calidad de carne en bovinos Charolais mexicanos. *Revista Mexicana De Ciencias Pecuarias,* 6(4), 361–375.

SZMATOŁA, T. et al. (2016). Characteristics of runs of homozygosity in selected cattle breeds maintained in Poland. *Livestock Science,* 188, 72–80.

SZMATOŁA, T. et al. (2019). A Comprehensive Analysis of Runs of Homozygosity of Eleven Cattle Breeds Representing Different Production Types. *Animals,* 9(12).

TRUKHACHEV, V. et al. (2015). Myostatin gene (MSTN) polymorphism with a negative effect on meat productivity in Dzhalginsky Merino sheep breed. *J. BioSci. Biotechnol.,* 4(2), 191–199.

TURNER, S.D. (2017). qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *bioRxiv*.

WILLIAMS, J.L. (2015). Inbreeding and purging at the genomic Level: the Chillingham cattle reveal extensive, non-random SNP heterozygosity. *Animal Genetics*, 47(1), 19–27.

WYATT, A.R. et al. (2013). Protease-activated alpha-2-macroglobulin can inhibit amyloid formation via two distinct mechanisms. *FEBS Lett*, 587(5), 398–403.

ZAVAREZ, L.B. et al. (2015). Assessment of autozygosity in Nellore cows (*Bos indicus*) through high-density SNP genotypes. *Frontiers in Genetics,* 6(5).

ZHANG, Y. et al. (2015). Runs of homozygosity and distribution of functional variants in the cattle genome. *BMC Genom*ics, 16, 542.