

Evaluating signatures of selection through variation in linkage disequilibrium among different cattle breeds

Veronika Kukučková*, Nina Moravčíková, Radovan Kasarda

Slovak University of Agriculture in Nitra, Faculty of Agrobiolgy and Food Resources, Department of Animal Genetics and Breeding Biology, Nitra, Slovak Republic.

Article Details: Received: 2016-06-13 | Accepted: 2016-07-26 | Available online: 2016-09-01

<http://dx.doi.org/10.15414/afz.2016.19.si.105-110>

The aim of presented investigation was to assess the overlap between selection signature discovered through analysis of variation in linkage disequilibrium and reported genomic regions associated with economic and traits of biological importance in cattle populations. The differences across Slovak Pinzgau, Austrian Pinzgau, Simmental and Holstein cattle and thus genome signatures of production and adaptation were found. The highest peak (top 0.01 percentile) was observed between Slovak and Austrian Pinzgau on chromosome 23, between Slovak Pinzgau and Simmental on chromosome 4 and between Slovak Pinzgau and Holstein on chromosomes 1, 7 and 20. Many candidate genes found have a known role in milk production (casein genes CSN1S1, CSN2, CSN1S2, CSN3; ABCG2, HBEGF, CAPN3, DGAT1, TG, GHR), reproduction (MGAT1, FGF1), feed efficiency (R3HDM1, ZRANB3), fertility (SPOCK1) and immune response (HSPA9, CD14, ARAP3, PCDH). Results of this study could be the basis for implementation of genomic selection programs in the Slovak Pinzgau cattle.

Keywords: linkage disequilibrium, genotyping array, Pinzgau cattle, selection signatures

1 Introduction

Since domestication, significant genetic improvement has been achieved for many traits of commercial importance in cattle, including adaptation, appearance and production. In response to such intense selection pressures, the bovine genome has undergone changes at the underlying regions of functional genetic variants, which are termed “selection signatures” (Randhawa et al., 2016).

Domestication and selection are processes that alter the pattern of within- and between-population genetic variability. Recently, sequence polymorphisms at the genome-wide level have been investigated in a wide range of animals. A common approach to detect selection signatures is to compare breeds that have been selected for different breeding goals (dairy and beef cattle). However, genetic variations in different breeds with similar production aptitudes and similar phenotypes can be related to differences in their selection history (Sorbolini et al., 2015). As the technology improves and the cost of low-density genotyping platforms decreases, mating designs that utilize genomic information could assist producers in managing their herd at the genomic level (Howard et al., 2015).

A traditional method to examine the degree of differentiation is to compute Wright’s F_{ST} statistic across two populations. The use of this measure is advantageous when large differences in allele frequencies occur, such as across cattle breeds. Within a breed, small differences in allele frequencies are expected across populations and particularly when there is some degree of genetic exchange. Due to this the usefulness of F_{ST} to determine regions that are different within a breed is reduced, therefore alternative methods have been used (Howard et al., 2015). One such alternative method to characterize the genomic differences across populations is to compute the average or a specific region’s ROH (runs of homozygosity) frequency (Ferenčaković et al., 2013; Kim et al., 2013).

Signatures of selection are regions in the genome that have been preferentially increased in frequency and fixed in a population because of their functional importance in specific processes. These regions can be detected because of their lower genetic variability and specific regional linkage disequilibrium (LD) patterns. The VarLD method is a powerful tool to identify differences in LD between cattle

* **Corresponding Author:** Veronika Kukučková, Slovak University of Agriculture in Nitra, Faculty of Agrobiolgy and Food Resources, Department of Animal Genetics and Breeding Biology, Tr. A. Hlinku 2, 949 76 Nitra, Slovak Republic. E-mail: veron.kukuckova@gmail.com

populations and putative signatures of selection with potential adaptive and productive importance (Perez O'Brian et al., 2014). This method was successfully used in detection of genes associated with civilization diseases (Teo et al., 2009; Ong et al., 2010) as well as production and reproduction traits in cattle (Perez O'Brian et al., 2014; Kadri et al., 2015; Sorbolini et al., 2015; Randhawa et al. 2016).

Change of breeding goal and use of Total merit index to preserve dual-purpose character of Slovak Pinzgau is proposed for long time and the positive impact on population structure is expected. The use of most recent results of quantitative (Krupova et al., 2016) and molecular genetics should help to sustainably manage population. The routine collection of genomic information would be an invaluable resource for effective management of breeding programs in small, endangered populations (Mészáros et al., 2015; Šidlová et al., 2015). From the point of the molecular genetics, identification of the most valuable lines, families and individuals with the impact on population diversity and to achieve competitive genetic progress and production of progeny with higher added genetic value. In addition, the knowledge on the selection signatures of Slovak Pinzgau, the breed relationship with other European cattle populations including Austrian Pinzgau and identification of the most characteristic regions in the genome will further contribute to the development of breeding programs.

The ultimate objectives of genomic scans for evolving patterns of genetic diversity are to detect causative variants and its functional relevance to particular traits. Use of the high throughput assays for generating dense genotypes and genomic sequences will aid the fine mapping of candidate variants. Finally, functional analysis of the detected variants in the regions and genes under hotspots of positive selection will be an active area of future research to understand the biological significance of molecular variations in adaptation, appearance and production in cattle (Randhawa et al., 2016).

The regional linkage disequilibrium variation in the autosomal genome between breeds, including Slovak Pinzgau, Austrian Pinzgau, Simmental and Holstein cattle, genotyped with the Illumina Bovine50K BeadChip was measured. By comparing the differences across dairy and dual-purpose cattle types, we aimed to find genome signatures of production and adaptation.

2 Material and methods

All analyses performed in this study were based on a dataset consisting of 308 cattle individuals (4 breeds) genotyped by high-throughput technology. Information from different sources were used; all Slovak Pinzgau (SKP, 19) active bulls genotyped by the Illumina BovineSNP50 v2 BeadChip (Illumina Inc., San Diego, CA), bulls from Austrian Pinzgau (ATP, 105) that were previously used in Ferencakovic et al. (2013) as well as genotypes of Holstein (HOL, 100) and Simmental (SIM, 84) obtained from DRYAD (McTavish et al., 2013b). For the dataset retrieved from the public digital repository quality control has already been performed according to McTavish et al. (2013a). The consensus map to obtain the same number of SNPs for all evaluated breeds was created. Subsequently, quality control according McTavish et al. (2013) also for populations from other sources was performed. Afterward, we removed individuals with >10 % missing data across the markers on the 29 autosomes from our analyses. Similarly markers that were missing in >10% of individuals were discarded using PLINK v1.9 (Purcell et al. 2007). A total of 287 individuals and 41 135 SNPs were included for further analyses.

Data were processed using the varLD method that compares the regional variation of LD between populations giving a score to each window, and then compares the score across 2 breeds highlighting the regions that show high differences. The regions containing the top 5, 1, 0.1 and 0.01 percentile of signals were detected and compared with previously reported bovine genomic regions associated with production and reproduction traits to find the overlap. Variation in linkage disequilibrium detects candidate regions under positive selection by comparing genome-wide LD variation between populations (Teo et al., 2009) and it is implemented in the varLD program (Ong et al., 2010).

3 Results and Discussion

The top 0.01 percentile of signals between SKP and HOL peaked at 40-58 Mb on chromosome 7 (Table 1). This region contains many genes associated with milk production and disease resistance (SAR1B, HBEGF), reproduction traits (MGAT1, FGF1), fertility (PCSK4, SPOCK1) and immune response (HSPA9, CD14, ARAP3 and multiple members of PCDH group). Gene MGAT1 was observed also between SKP and ATP as well as SKP and SIM while gene FGF1 between SKP and SIM, however on the lower percentile. The low gene diversity genomic regions harbouring genes for feed efficiency (R3HDM1, ZRANB3) were found between SKP and HOL on chromosome 2, while casein genes (CSN1S1, CSN2, CSN1S2, CSN3) were observed between SKP-ATP and SKP-SIM on

chromosome 6. Thyroglobulin (TG, chromosome 14) responsible for meat tenderness or intramuscular fat distribution is also associated with milk yield and composition, whereas in our study was found only in SKP-HOL. Candidate genes on chromosome 2, 6, 7, 14 and more are described in the study of Randhawa et al. (2016). Intense selection to increase milk yield has had negative consequences for mastitis incidence in dairy cattle. Association signal for clinical mastitis and milk yield peaked in the 26–40 Mb region on chromosome 20 in HOL according Kadri et al. (2015) confirming also by results of our study (SKP-HOL).

Candidate genes ABCG2 (SKP-SIM), CAPN3 (SKP-SIM, SKP-HOL) and DGAT1 (SKP-HOL) on chromosomes 6, 10, 14, respectively, were found corresponding with Sorbolini et al. (2015). In addition, they observed MSTN and FTO gene which are connected with beef production and thus were not preserved in our study with dairy and dual-purpose cattle. Kasarda et al. (2015) found only 2 SNPs close to gene FTO using the integrated Haplotype Score (iHS) in Slovak and Austrian cattle. The highest peak (0.01%) between Slovak and Austrian on chromosome 23, between Slovak Pinzgau and Simmental on chromosome 4 and between Slovak Pinzgau and Holstein on chromosomes 1, 7 and 20 was observed. Similarly, Perez O'Brian et al. (2014) noticed that the largest varLD scores are found when comparing different production types within a subspecies. According Table 1 is obvious that only variance in LD between Slovak and Austrian Pinzgau showed signals in all autosomes (at least on the top 5%) most likely because SKP has been selected mainly for milk production and ATP mainly for beef production.

Table 1 Regions with different signal between all evaluated pairs of breed containing the top 5, 1, 0.1 and 0.01 percentile of signals

Chromosome	Slovak Pinzgau – Austrian Pinzgau	Slovak Pinzgau – Simmental	Slovak Pinzgau – Holstein
1	5, 5, 5, 5, 1	5, 0.1, 5, 5, 5, 5, 5, 5	5, 5, 0.01
2	1, 5	5, 5, 5	5, 5, 5
3	5	5	5, 5
4	1, 5	5, 5, 0.01	5
5	5, 5, 1	1, 1	5
6	5, 5	5, 5, 5, 5	-
7	5	1, 5	5, 5, 5, 0.01, 5
8	5	5, 5	5
9	1, 1	1, 5, 5	-
10	0.1, 5, 5	5, 5	5, 5, 5
11	1, 5	5, 1, 5, 1, 5	5
12	5	5	5
13	5	1, 5	5, 1, 1, 5
14	5	-	5, 5, 5, 5, 5, 1
15	5, 5, 1	5	-
16	5, 5, 5	1, 0.1, 5	5, 5
17	1, 5	1	5, 5, 1
18	1, 5	5, 1	-
19	1, 0.1, 5	5, 5, 0.1, 0.1	5
20	5	5, 5	0.01
21	5	5	-
22	5, 0.1	-	5, 5
23	5, 5, 5, 1, 0.01	5, 5, 1, 5, 5, 5	-
24	5	5	5, 5
25	1, 5, 0.1	5, 5, 1	-
26	5, 5	5	5
27	5	1, 5	-
28	5, 5	5, 5, 1	-
29	5	-	-

The differences among pairs of breeds were obtained in expression of a specific region of the genome. Between Slovak and Austrian Pinzgau there was only one region assumed as the most significant signal on chromosome 23 (Figure 1, 2). Šidlová et al. (2015) found KHDRBS2 gene on

chromosome 23, while results of this study showed signals on the same chromosome (top 0.01 % between SKP-ATP, top 1 % between SKP-SIM) in different region. Richardson et al. (2016) identified candidate region and imputation of this region to sequence data pinpointed the association signal in the introns of the FKBP5 gene, which is involved in immune response. The variance of LD between SKP-ATP (top 1%) and SKP-SIM (top 5 %) peaked also in this region (Figure 1, 2A, 2B, 2C).

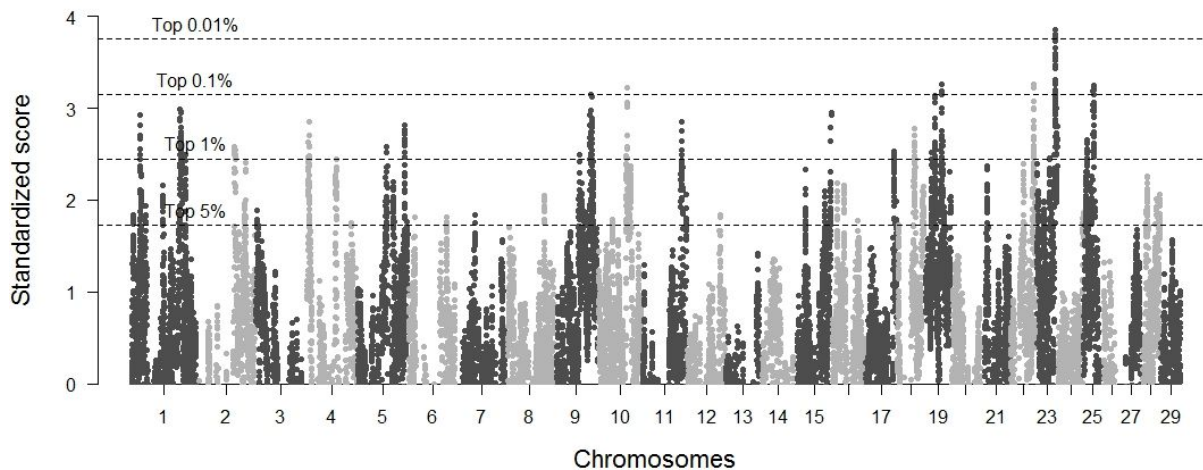


Figure 1 Comparison between Slovak and Austrian Pinzgau based on variation in linkage disequilibrium analysis. Manhattan plot demonstrates the presence of significant signals in the regions on several autosomes. Upper abline represents signal with whole-genome significance threshold

Signatures of selection for adaptation are mainly attributed to tolerance in new climates, feed resources and resistance to different disease agents in various cattle breeds. The changes in genetic aspects of behavioural control for new adaptations, from survive to thrive, and using available resources have also been detected under positive selection in several populations (Randhawa et al., 2016). Kim et al. (2013) found that several of the regions that had differing levels of ROH across populations were associated with economically important traits including milk, fat and protein yield. The same approach was utilized in this study to detect signatures of selection in common and different across populations. Kasarda et al. (2015) identified the evidence of recent selection based on estimation of the iHS , F_{ST} and characterized affected regions near QTL associated with traits under strong selection in Pinzgau cattle. Recent investigations of the structural variation in cattle genome suggest that selective forces, in addition to the genotypic and haplotypic patterns, operate on the copy number variation (CNV) in candidate genes and can be helpful to characterize the effects of domestication, breed formation and artificial selection (Perez O'Brian et al., 2014).

4 Conclusions

The highest peak between Slovak and Austrian on chromosome 23, between Slovak Pinzgau and Simmental on chromosome 4 and between Slovak Pinzgau and Holstein on chromosomes 1, 7 and 20 was observed. Many candidate genes associated with milk production, reproduction and health were estimated, whereas genes associated with meat production were not found. Results of this study could be the basis for implementation of genomic selection programs in the Slovak Pinzgau cattle.

Acknowledgments

This study was supported by the Slovak Research and Development Agency under the Contract No. APVV-14-0054.

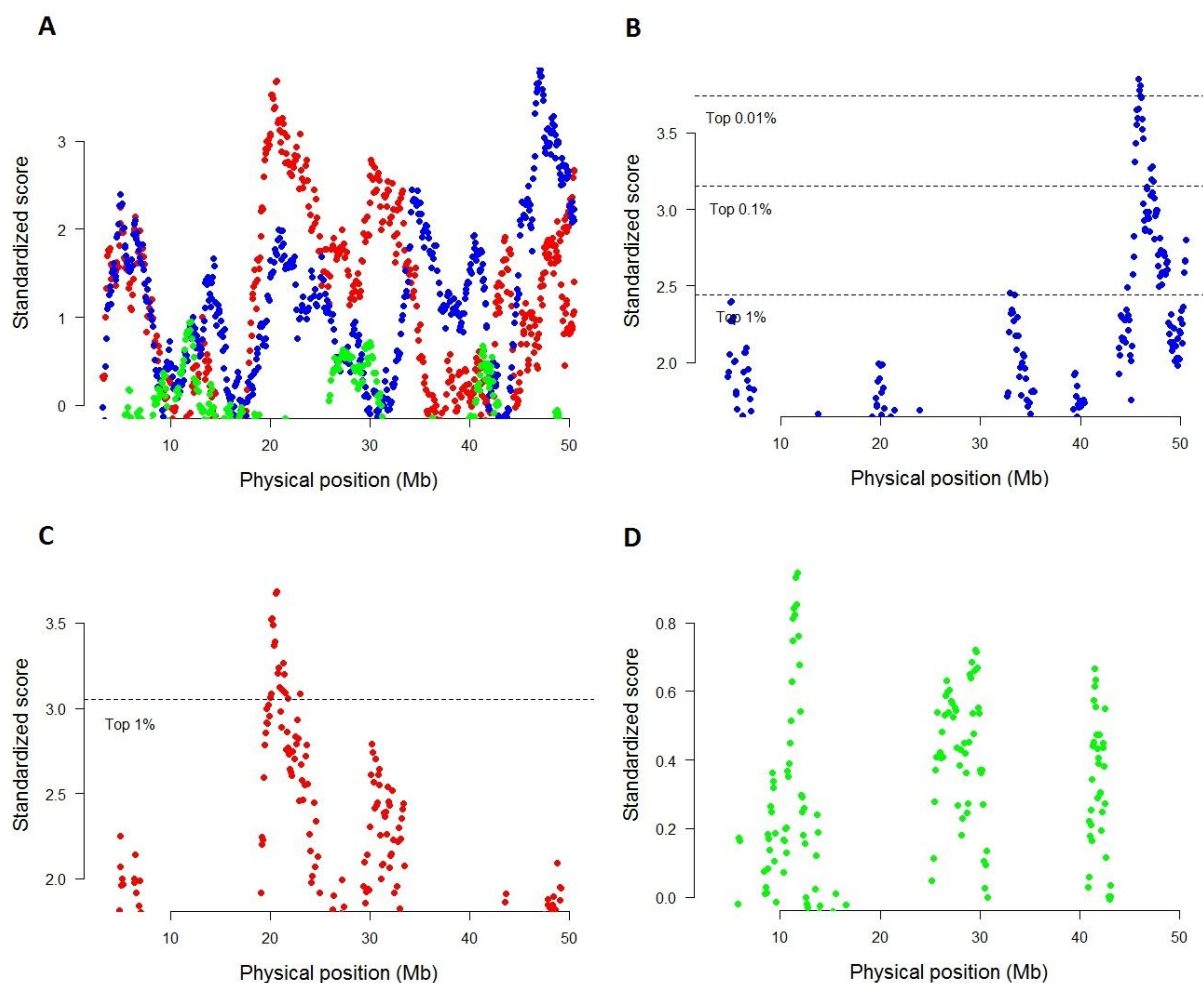


Figure 2 Differences among different breed pairs on chromosome 23: A – The comparison of all evaluated breed pairs, B – Slovak Pinzgau v. Austrian Pinzgau (blue), C – Slovak Pinzgau v. Simmental (red) and D– Slovak Pinzgau v. Holstein (green)

References

- FERENČAKOVIČ, M., SOLKNER, J. and CURIK, I. (2013) Estimating autozygosity from high-throughput information: effects of SNP density and genotyping errors. *Genetics Selection Evolution*, vol. 45, no. 1, pp. 42. doi: <http://dx.doi.org/10.1186/1297-9686-45-42>
- HOWARD, J.T. et al. (2015) Characterizing homozygosity across United States, New Zealand and Australian Jersey cow and bull populations. *BMC Genomics*, vol. 16, p. 187. doi: <http://dx.doi.org/10.1186/s12864-015-1352-4>
- KADRI, N.K. et al. (2015) Genetic dissection of milk yield traits and mastitis resistance quantitative trait loci on chromosome 20 in dairy cattle. *Journal of Dairy Science*, vol. 98, pp. 9015–9025. doi: <http://dx.doi.org/10.3168/jds.2015-9599>
- KASARDA, R. et al. (2015) Genome-wide selection signatures in Pinzgau cattle. *Potravinárstvo*, vol. 9, no. 1, pp. 268–274. doi: <http://dx.doi.org/10.5219/478>
- KIM, E. et al. (2013) Effect of artificial selection on runs of homozygosity in U.S. Holstein cattle. *PLoS ONE*, vol. 8, no. 11, e80813. doi: <http://dx.doi.org/10.1371/journal.pone.0080813>
- Krupová, Z. et al. (2016) Economic values for health and feed efficiency traits of dual-purpose cattle in marginal areas. *Journal of Dairy Science*, vol. 99, pp. 644–656. doi: <http://dx.doi.org/10.3168/jds.2015-9951>
- McTavish, E.J. et al. (2013a) New World cattle show ancestry from multiple independent domestication events. *PNAS*, vol. 110, no. 15, E1398–E1406. doi: <http://dx.doi.org/10.1073/pnas.1303367110>

McTavish, E.J. et al. (2013b) New World cattle show ancestry from multiple independent domestication events. Dryad Digital Repository. doi: <http://dx.doi.org/10.5061/dryad.42tr0.2>

Mészáros, G. et al. (2015) Genomic analysis for managing small and endangered populations: A case study in Tyrol Grey cattle. *Frontiers in Genetics*, vol. 6, pp. 173. doi: <http://dx.doi.org/10.3389/fgene.2015.00173>

ONG, R.T.-H. and TEO, Y.-Y. (2010) varLD: a program for quantifying variation in linkage disequilibrium patterns between populations. *Bioinformatics application notes*, vol. 26, no. 9, pp. 1269–1270. doi: <http://dx.doi.org/10.1093/bioinformatics/btq125>

PÉREZ O'BRIAN, A.M. et al. (2014) Assessing signatures of selection through variation in linkage disequilibrium between taurine and indicine cattle. *Genetics Selection Evolution*, vol. 46, pp. 19. doi: <http://dx.doi.org/10.1186/1297-9686-46-19>

R DEVELOPMENT CORE TEAM. (2011) *R: A Language and Environment for Statistical Computing (v3.1.2)*. [Software]. Vienna: R Foundation for Statistical Computing. Retrieved January 10, 2016 from <http://www.r-project.org>

RANDHAWA, I.A.S. et al. (2016) A Meta-Assembly of Selection Signatures in Cattle. *PLOS One*, vol. 11, no. 4, e0153013. doi: <http://dx.doi.org/10.1371/journal.pone.0153013>

SAS INSTITUTE Inc. (2011) *SAS/STAT Software : Version 9.3*. Cary NC: SAS Institute Inc.

RICHARDSON, I.W. et al. (2016) A genome-wide association study for genetic susceptibility to *Mycobacterium bovis* infection in dairy cattle identifies a susceptibility QTL on chromosome 23. *Genetics Selection Evolution*, vol. 48, pp. 19. doi: <http://dx.doi.org/10.1186/s12711-016-0197-x>

SORBOLINI, S. et al. (2015) Detection of selection signatures in Piemontese and Marchigiana cattle, two breeds with similar production aptitudes but different selection histories. *Genetics Selection Evolution*, vol. 47, no. 1, pp. 52. doi: <http://dx.doi.org/10.1186/s12711-015-0128-2>

ŠIDLOVÁ, V. et al. (2015) Production type of Slovak Pinzgau cattle in respect of related breeds. *Acta Fytotechnica et Zootechnica*, vol. 18, no. 2, pp. 25-29. doi: <http://dx.doi.org/10.15414/afz.2015.18.02.25-29>

TEO, Y.Y. et al. (2009) Genome-wide comparisons of variation in linkage disequilibrium. *Genome Research*, vol. 19, no. 10, pp. 1849–60. doi: <http://dx.doi.org/10.1101/gr.092189.109>