Genome-wide scan for loci under selection in local Slovak and Austrian populations of Pinzgau cattle

R. Kasarda¹, N. Moravčíková¹, G. Mészáros², J. Sölkner², V. Kukučková¹, A. Trakovická¹ & O. Kadlečík¹

¹Slovak University of Agriculture in Nitra, Department of Animal Genetics and Breeding Biology, Tr. A. Hlinku 2, 94976 Nitra, Slovakia radovan.kasarda@uniag.sk (Corresponding Author)
²University of Natural Resources and Life Sciences, Division of Livestock Sciences, Gregor Mendel Str. 33, 1180 Vienna, Austria

Summary

The aim of this study was to identify loci under selection pressure within two local populations of Pinzgau cattle and to differentiate between potential outliers under directional and balancing selection. The genome-wide scan was carried out using high-density genotyping data of 257 animals. In total 105 Austrian sires and 152 Slovak purebred animals registered by their breeder associations were genotyped using Illumina BovineHD and Illumina BovineSNP50 v2 BeadChip. After quality control and SNP pruning, the final dataset composed of a total of 37,775 autosomal loci. The $F_{ST}$ outlier approach was used to identify loci under directional (positive) or balancing selection. From a general point of view, low levels of genetic differentiation between analysed populations were found ($F_{ST} = 0.13$) that confirmed the previously stated assumption of a common gene pool between Slovak and Austrian populations. Despite the weak signals of selection signatures across the genome, the applied approach was sensitive to detect genomic regions under strong selection pressure. From a total of 1,602 identified outlier loci, 1,278 markers were recognized as directional outliers and the rest were assigned as signal of balancing selection. The genomic response to differences in breeding schemes for both Austrian and Pinzgau populations resulted in the changes of allele frequencies in loci located mainly in genomic regions responsible for muscle formation and beef production (CAPN, MSTN, MYO3A) as well as control of the immunity system (IGF1, BOLA-DRA). The results of this study can help to understand the biological limits and mechanisms of selection and to detect genes of interest to breeding strategy.

Keywords: high-density genotyping data, local populations, Pinzgau cattle, selection signatures

Introduction

Long-term use of artificial selection in cattle breeding led to the diversification of local breeds and populations that are specialized for either milk or meat production or rose as dual-purpose breeds. In response to such selection strategies, the particular genomic regions controlling traits of interest as well as other characteristics including disease resistance and immune response have undergone changes of functional genetic variants (Zhao et al., 2015; Randhawa et al., 2016). The unique genetic pattern left behind in the genome of individuals under selection pressures is referred to as selection signatures (Qanbari et al., 2014). The detection of such selection signatures has the potential to elucidate the identities of genes and
mutations associated with important phenotypic traits and may have practical implications for implementation of genomic selection (Brito et al., 2017).

The nucleus of Pinzgau cattle, most often bred across Europe in dual-purpose type, is based upon local populations registered in Austrian and Slovakian herd books of. The Pinzgau cattle, as an original Alpine breed, was introduced into the mountains regions of northern Slovakia approximately two centuries ago. The genetic reserve of Slovak population is bred under intensive farming conditions of submountaineous regions, whereas in Austria more than 65% of cows are in the suckling cow system (Kasarda et al., 2016). Both Austrian and Slovakian breeders are interested in the future development of this valuable breed for example through genetic evaluation of growth intensity (Bulla et al., 2013), longevity (Mészáros et al., 2010) as well as traits that are part of selection index (Kasarda et al., 2009).

The aim of this study was to identify loci reflecting genetic differentiation among local populations of Pinzgau cattle and to differentiate between potential outliers under directional and balancing selection.

Material and methods

The genome-wide scan for outlier loci under selection was performed using genotyping data of 257 animals of two genetically interconnected local populations of Pinzgau cattle. In total 105 Austrian sires and 152 Slovak purebred animals registered by their breeding association were genotyped using Illumina BovineHD and Illumina BovineSNP50 v2 BeadChip, respectively. The sampled population of Slovak Pinzgau cattle covered living animals: 37 sires representing active breeding bulls (19 animals) and AI doses deposited in gene bank (18 animals), 35 dams (nucleus cows) and 80 cows. To merge datasets consistently the consensus map with common SNPs across populations was prepared and overall 42,444 SNPs were retained in the reduced panel of markers. From these, all markers assigned to unmapped regions or loci with unknown chromosomal position according to the latest bovine genome assembly (Btau 4.6) and SNPs positioned to sex chromosomes were removed. The subsequent quality control of genotyping data were performed according to Purcell et al., (2007) to exclude samples with more than 10% missing genotypes (2), autosomal loci with call rate lower than 90% (1523), and minor allele frequency lower than 0.01 (3262). After applying quality control, the final dataset was composed of a total of 37,775 autosomal loci.

The $F_{ST}$ outlier approach implemented in LOSITAN (Antao et al., 2008) was used to identify loci under directional (positive) or balancing selection. The value of $F_{ST}$ was used as an indicator of genetic differentiation among breeds and consequential locus-specific $F_{ST}$ values significantly differing from the average as a signal of genomic regions under selection pressure. LOSITAN analysis was run using 100,000 iterations with a 99.5% confidence interval, false discovery rate (FDR) of 0.05, attempted $F_{ST}$ of 0.132, subsample size of 50 and simulated $F_{ST}$ 0.139. Both recommended options i.e. “Neutral” mean $F_{ST}$ and “Force” mean $F_{ST}$ were applied to increase the reliability of the estimated initial average $F_{ST}$ and to approximate a desired average $F_{ST}$ for the dataset based on bisection algorithm over repeated simulations (Antao et al., 2008). The loci under selection were recognized if the genome-wide distribution of $F_{ST}$ values was unusually high (directional selection) or low (balancing selection) than expected (Beaumont & Balding, 2004). SNPs with appropriate levels of genome-wide significance were assigned to the genomic location according to the NCBI database (www.ncbi.nlm.nih.gov).
Results and discussion

Overall, 1,602 outlier loci were detected in the analysed dataset based on the $F_{ST}$ approach and LOSITAN analysis. Most of them were recognized as directional outliers (1,278), whereas 324 markers were identified as balancing outliers. Figure 1 shows the distribution of loci identified to be under selection. The highest proportion of SNPs subject to the positive and/or balancing selection were found on chromosomes 1 (7.43%) and 4 (6.31%).

![Figure 1. Proportion of outlier loci under selection across bovine autosomes.](image)

The lowest total number of potential outliers subject to the positive and/or balancing selection was identified on chromosome 25 (1.37%). Several of the identified outlier loci were found to be located directly in the genomic region or close to genes involved in control of production traits and immunity. Overall 17 loci were located within the sequence of annotated genes (Table 1).

Table 1. Summary of outlier loci under direction selection located in the sequence of genes controlling important phenotypic traits in cattle.

<table>
<thead>
<tr>
<th>BT A</th>
<th>No. of SNPs</th>
<th>Position (Mb)</th>
<th>Annotated genes in regions according to NCBI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>153.75 - 154.30</td>
<td>calpain 7 (<em>CAPN</em>)</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>6.7</td>
<td>myostatin (<em>MSTN</em>)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>105.35 - 107.46</td>
<td>insulin like growth factor binding protein (<em>IGFBP</em>)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>186.46</td>
<td>cytochrome P450, family 27 (<em>CYP27A1</em>)</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>65.48</td>
<td>growth hormone releasing hormone receptor (<em>GHRHR</em>)</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>67.78</td>
<td>insulin like growth factor 1 (<em>IGF1</em>)</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>87.2</td>
<td>casein beta (<em>CSN2</em>)</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>59.23</td>
<td>cytochrome P450, family 19 (<em>CYP19A1</em>)</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>6.65 - 7.21</td>
<td>interleukin family</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>88.13</td>
<td>collagen type IV alpha 1 chain (<em>COL4A1</em>)</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>26.86</td>
<td>myosin IIIA (<em>MYO3A</em>)</td>
</tr>
<tr>
<td>23</td>
<td>1</td>
<td>25.59</td>
<td>major histocompatibility complex, class II (<em>BOLA-</em>)</td>
</tr>
</tbody>
</table>
Overall, the results indicate the use of very similar breeding schemes in both Slovak and Austrian herds. Nevertheless, the genomic response to several differences resulted in the changes of allele frequencies in loci located mainly in genomic regions responsible for muscle formation and beef production (CAPN, MSTN, MYO3A) as well as control of immunity system (IGF1, BOLA-DRA). The presence of strong selective footprints across the bovine genome were tested in multiple populations using various approaches mainly based on site frequency spectrum, population differentiation represented by FST statistic and haplotype length (extend of linkage disequilibrium) (Qanbari et al., 2011; Druet et al., 2013; Mancini et al., 2014; Zhao et al., 2015).

Conclusion

The utilization of high-density genomic data allowed obtaining more detailed view on genomic differences in two local, closely related populations. To protect as much diversity as possible in terms of sustainable genetic gain, breeders have to manage the strategy of animals’ selection. The analysis of genomic regions and loci under strong selection pressure is necessary to understand the significance of molecular variations involved in the biological process of any selection strategy. The study provided data that can be beneficial not only for increase of selection efficiency but also a case for the protection of the unique characteristics of Slovak Pinzgau population in the future.

Acknowledgement

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

List of References


