



Diversifying selection signatures among divergently selected subpopulations of Polish Red cattle

Artur Gurgul¹ · I. Jasielczuk¹ · E. Semik-Gurgul¹ · T. Szmatoła¹ · A. Majewska² · E. Sosin-Bzducha³ · M. Bugno-Poniewierska^{1,4}

Received: 3 September 2018 / Revised: 11 January 2019 / Accepted: 15 January 2019 / Published online: 26 January 2019
© The Author(s) 2019

Abstract

Polish Red cattle is one of the few indigenous breeds of European red cattle which is characterized by several desired features, such as high disease resistance, good health, longevity, good fertility, and high nutritional value of milk. Currently, Polish Red cattle population is a subject of two independent breeding programs: (i) improvement program and (ii) genetic resources conservation program. The aim of the improvement program is the genetic progress in terms of milk production and body conformation traits, while the conservation program mainly focuses on protection of the genetic resources of Polish Red cattle and preservation of the existing, original gene pool. By the analysis of F_{ST} genetic distances across genome-wide SNP panel, we detected diversifying selection signatures among these two subpopulations and indicated (among others) the significance of *DGAT1* and *FGF2* genes for milk production traits in these cattle. We also found that among genes being presumably under selection in terms of milk production, there are genes responsible, for example, for mammary gland development (e.g., *SOSTDC1*, *PYGO2*, *MED1*, and *CCND1*) and immune system response (e.g., *IL10RA*, *IL12B*, and *IL21*). The most important finding of this study is that the most pronounced genetic differences between the analyzed populations were associated with β -defensin genes (e.g., *DEFB1*, *DEFB4A*, *DEFB5*, *DEFB7*, *DEFB10*, *DEFB13*, *EBD*, *BNBD-6*, and *LAP*) located within so-called bovine cluster D on BTA27. The β -defensins are expressed mainly in the mammary gland and are antimicrobial peptides against the Gram-negative and Gram-positive bacteria, viruses, and other unicellular parasites. This suggests that antimicrobial resistance of mammary gland is of high importance during selection towards increased milk production and that genes responsible for this process are selected together with increasing levels of productivity.

Keywords β -Defensins · Milk production · Polish Red cattle · Selection signatures

Communicated by: Maciej Szydłowski

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13353-019-00484-0>) contains supplementary material, which is available to authorized users.

✉ Artur Gurgul
artur.gurgul@izoo.krakow.pl

- ¹ Department of Animal Molecular Biology, National Research Institute of Animal Production, Krakowska 1, 32-083 Balice, Poland
- ² Department of Cattle Breeding, National Research Institute of Animal Production, Krakowska 1, 32-083 Balice, Poland
- ³ Department of Nutrition Physiology, National Research Institute of Animal Production, Krakowska 1, 32-083 Balice, Poland
- ⁴ University of Agriculture in Krakow, Institute of Veterinary Sciences, Mickiewicza 24/28, 30-059 Kraków, Poland

Introduction

Polish Red cattle is one of the few indigenous breeds of European red cattle. It is characterized by several features typical for primitive populations, such as high disease resistance, good health, longevity, very good fertility, easy births, ease of calf rearing, and high biological value of milk. The Polish Red cattle is also characterized by a good adaptation to harsh environmental conditions, which is especially visible in the ability to limit the efficiency (enabling survival of seasonal feed deficiencies), as well as the relatively quick regeneration after condition loss. These features make the cattle of this breed well adapted to mountainous and submountainous living and production conditions (Szarek et al. 2004).

Currently, Polish Red cattle population is a subject of two independent breeding programs: (i) improvement program and (ii) genetic resources conservation program (Adamczyk

et al. 2008). The aim of the improvement program is the genetic progress in terms of milk production and body conformation traits, leading to the refinement of the economic aspects of breeding and preservation of the existing beneficial functional features. Specialization is directed at the characteristics that have a fundamental impact on improving the profitability of milk production, like milk yield, protein yield, fat yield, and functional features with special emphasis on the udder morphology and leg health. The aim of the conservation program is to protect the genetic resources of Polish Red cattle used for both meat and dairy production, to preserve the existing, original gene pool, genetic variability of the population, and to maintain the productivity of Polish Red cattle at an acceptable level. Maintaining the largely primitive character of this breed is one of the major breeding goals for this part of the population (Adamczyk et al. 2008).

Despite both these subpopulations originate from the same local red cattle population (Szarek et al. 2004), the application of the two different selection programs could have generated genetic variability among the studied groups. Therefore, in this study, we attempt to identify diversifying selection signatures among these two Polish Red cattle populations. The obtained results presumably should allow us to identify genome regions with differentially selected variants which are responsible for milk traits and functional features selected in the analyzed populations. To this end, we used the well-established F_{ST} -based approach which measures the genetic differentiation due to locus-specific allele frequencies variation between populations. With this approach, we identified numerous candidate genes for traits selected in both populations with especially strong reference to the udder health and udder developmental processes.

Material and methods

The study material comprised 60 samples of ear tissue collected from cows belonging to Polish Red breed and included in “conservation” (RP; $n = 37$) and “improvement” (RE; $n = 23$) breeding programs. The animals were randomly selected and verified to be unrelated for at least two generations. The animals were coming from at least three different herds. All animal procedures were approved by the Local Animal Care Ethics Committee No. II in Kraków—permission number 1293/2016 in accordance with EU regulations. The genomic DNA was purified using Sherlock AX kit (A&A Biotechnology, Poland) and after quality control was genotyped with the use of BovineSNP50v2 BeadChip assay (Illumina, San Diego, CA) according to the standard Infinium Ultra protocol. Only samples with call rate > 0.95 were retained for analysis. The initial markers set included 54,609 SNPs and were further filtered to remove markers without known chromosomal position or located on sex chromosomes. SNPs that had minor allele frequency (MAF) < 0.05 , genotyping rate < 0.8 (in joint populations), and deviated from Hardy-Weinberg

equilibrium with $p < 0.0001$ (in each population separately) were also removed. This filtering step allowed retaining for further analysis a common set of 43,165 SNPs with average inter-marker distance of 57.9 kb in UMD3.1 genome assembly.

The across-genome breeds genetic differentiation was evaluated using pairwise F_{ST} distances (Weir and Cockerham 1984) measuring locus-specific allele frequencies variation between populations. F_{ST} values obtained using Plink software were further averaged in 10-SNP sliding windows to account for stochasticity in locus-by-locus variation. The window-averaged F_{ST} values were then ranked and top 1% of the observations pointed to windows with the most pronounced selection signals. Overlapping windows with the top F_{ST} values were subsequently merged and genomic regions spanned by this merged windows were extended on both ends by 25 kb to account for extended linkage. The resulting genome regions spanning the strongest diversifying selection signals were finally analyzed in details to identify encoded genes and their associated biological processes using UCSC Genome Browser (Karolchik et al. 2004), ENSEMBL database, and KOBAS Web Server (Xie et al. 2011). The gene overrepresentation tests (in GO categories) were performed according to all annotated bovine genes with correction for multiple testing.

The linkage disequilibrium (LD) and haplotype block structure at the most divergently selected regions among the studied populations were analyzed using HaploView 4.2 (Barrett et al. 2005) software examining pairwise LD on the distance up to 500 kb and detecting blocks based on a method proposed by Gabriel et al. (2002).

Global population differentiation was analyzed using principal component analysis (PCA) based on individual genotypes and mean or weighted F_{ST} distances.

Data availability The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Results

The used SNP panel showed sufficient polymorphism parameters in both populations with average MAF in RP of 0.275 and 0.272 in RE. The mean observed heterozygosity was comparable for both populations and was 0.369 for RP and 0.377 for RE (Table 1). The general population differentiation was rather low with mean and weighted F_{ST} distances of 0.0226 and 0.0242, respectively. The PCA showed, however, that both populations form clearly separated clusters of observations with visibly higher level of genetic variation observed in RE population (Fig. 1). RE population was additionally subdivided into two visible clusters. A local genetic differentiation of the populations was predominantly concentrated in 76 separate genomic regions located on 25 different

Table 1 Marker statistics and basic population indexes

Population	<i>n</i>	SNPs (<i>n</i>)	Mean SNPs distance (kb)	Mean MAF	Ho	He	F_{ST} (mean/weighted)	<i>F</i>
RP	37	43,165	57.97	0.275	0.369	0.363	0.0226/0.0242	− 0.018
RE	23			0.272	0.377	0.359		− 0.049

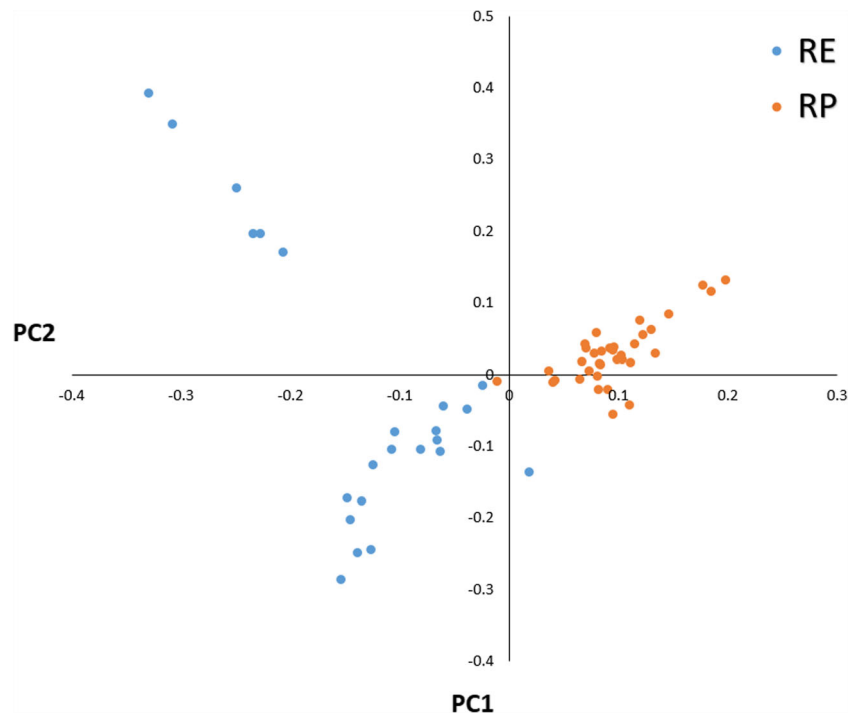
RP, cows under “conservation program”; *RE*, cows under “improvement program”; *Ho*, observed heterozygosity; *He*, expected heterozygosity; *MAF*, minor allele frequency; *F*, inbreeding coefficient based on the observed versus expected number of homozygous genotypes

autosomes (Supplementary File 1). No strong selection signals were identified on chromosomes 18, 23, 26, and 28. The most pronounced differential selection signals were observed on BTA1, BTA3, BTA6, BTA10, BTA17, and BTA27 (Fig. 2, Table 2). The genome regions overlapped by the detected selection signals spanned from 26.4 kb to 4.6 Mb and encompassed 610 different genes (Supplementary Files 1 and 2). The analysis of biological processes associated with the genes allowed for detection of very general GO categories including primary metabolic process (143 genes), regulation of protein phosphorylation (34 genes), regulation of protein kinase activity (21 genes), phosphorylation (46 genes), growth factor activity (8 genes), mammary gland development (5 genes), and immune system processes (28 genes). Among the detected genes, we also identified previously described candidate genes for milk traits like *DGAT1* and *FGF2* (Ogorevc et al. 2009) or disease resistance-associated genes, like *IL10RA*, *IL12B*, and *IL21* (Malkovský et al. 1988). The detailed analysis of the genome regions spanned by the selection signatures with the top 0.1% of the F_{ST} values (strongest diversifying selection signals) allowed for detection of 46

overlapped genes. The overrepresentation test performed for those genes showed single enriched GO category connected with defense response to bacterium ($adjP = 0.02$). The genes associated with this process belonged mainly to the β -defensins family (e.g., *DEFB1*, *DEFB4A*, *DEFB5*, *DEFB7*, *DEFB10*, *DEFB13*, *EBD*, *BNBD-6*, *LOC783012*, *LAP*). In the analyzed genes set, we also detected other immune system functioning-related genes, like *CD53* and *TRATI* (Zhang and Samelson 2000; Todros-Dawda et al. 2014).

Two selection signals found in this study included, inter alia, the strongest detected signal on BTA3 (32.5–32.9 Mb) and the signal spanning the largest genomic region (1.3 Mb) among the signals with top 0.1% of F_{ST} values (BTA27; 4.9–6.2 Mb). These two selection signals were analyzed in details to identify their possible meaning and evaluate haplotype structure at these loci. The analysis of the divergently selected region on BTA3 showed that the detected selection signal colocalized with the large chromosomal region of strong LD (31.6–33.5 Mb) and encompassed large haplotype block with high-frequency haplotypes in both studied populations (> 0.5). The detailed analysis of haplotypes showed that in

Fig. 1 Principal component analysis for the studied individuals. *RP*, conserved population; *RE*, improved population



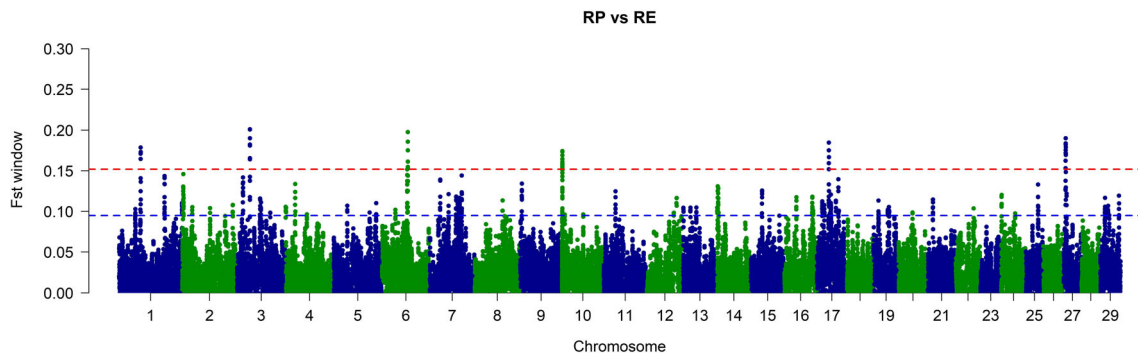


Fig. 2 Plot of window-averaged F_{ST} values across all bovine autosomes. The 10 SNP sliding window-averaged F_{ST} values are plotted against centered genomic positions of windows. *Blue dashed line* shows threshold of 1% of the top F_{ST} observations. *Red dashed line* represents 0.1% of the top F_{ST} observations. *RP*, conserved population; *RE*, improved population

Table 2 Genomic regions spanned by the strongest detected selection signals (with top 0.1% of F_{ST} values)

Chr	Start	End	Size	Genes*
1	54,052,439	54,625,163	572,724	<i>TRAT1—T Cell Receptor Associated Transmembrane Adaptor 1</i> <i>MORC1—MORC Family CW-Type Zinc Finger 1</i> <i>DPPA2—Developmental Pluripotency Associated 2</i>
3	32,527,488	32,883,278	355,790	<i>LRIF1—Ligand-Dependent Nuclear Receptor Interacting Factor 1</i> <i>CD53—Cell Surface Glycoprotein CD53</i> <i>KCNA3—Potassium Voltage-Gated Channel Subfamily A Member 3</i>
6	63,018,446	63,186,617	168,171	<i>ATP8A1—ATPase Phospholipid Transporting 8A1</i>
6	64,418,457	64,662,856	244,399	–
10	766,191	832,105	65,914	<i>MCC - MCC, WNT Signaling Pathway Regulator</i>
17	29,760,863	30,508,078	747,215	<i>LOC101903064—Uncharacterized protein</i> <i>PGRMC2—Progesterone Receptor Membrane Component 2</i> <i>LARP1B—La Ribonucleoprotein Domain Family Member 1B</i> <i>ABHD18—Abhydrolase Domain Containing 18</i> <i>MFSB8—Major Facilitator Superfamily Domain Containing 8</i> <i>PLK4—Polo Like Kinase 4</i> <i>HSPA4L—Heat Shock Protein Family A (Hsp70) Member 4 Like</i> <i>SLC25A31—Solute Carrier Family 25 Member 31</i> <i>INTU—Inturned Planar Cell Polarity Protein</i>
27	4,947,815	6,243,961	1,296,146	<i>SPAG11—Sperm-associated Antigen 11</i> <i>ZNF705A—Zinc Finger Protein 705A</i> <i>LAP—Laryngeal Adductor Paralysis</i> <i>DEFB7—Beta-defensin 7</i> <i>BNBD-6—Dfensin Beta 6</i> <i>DEFB4A—Defensin Beta 4A</i> <i>DEFB—Defensin, Beta</i> <i>DEFB13—Defensin Beta 113</i> <i>EBD—Enteric Beta-defensin</i> <i>DEFB5—Defensin, Beta 5</i> <i>DEFB10—Defensin, Beta 10</i> <i>DEFB1—Defensin, Beta 1</i> <i>GPM6A—Glycoprotein M6A</i>

*Genes with no assigned name or pseudogenes were not included in the table

separate populations, the most common haplotypes were composed of alternative SNP variants. Additionally, only haplotypes found in RP breed showed some minor signs of recombination (Fig. 3). The analyzed region on BTA3 overlapped with four genes (*LRIF1*—*Ligand-Dependent Nuclear Receptor Interacting Factor 1*, *ENSBTAG00000042091*, *CD53*—*Cell Surface Glycoprotein CD53*, *KCNA3*—*Potassium Voltage-Gated Channel Subfamily A Member 3*) including one non-coding small nuclear RNA.

The genomic region on BTA27 overlapping with the detected large and strong diversifying selection signal was also characterized by strong LD (unique on the chromosome scale) and encompassed three haplotype blocks showing common variants with haplotype frequency above 0.5. The detailed analysis of haplotypes again showed signs of diversifying selection among the analyzed populations, represented by haplotype variation and divergent selection of separate variants (Fig. 4). The signal on chromosome 27 overlapped with 21 different genes, mainly belonging to β -defensins family (Table 2) grouped in bovine cluster D (Gurao et al. 2017).

Discussion

In this study, we attempt to identify diversifying selection signatures among two subpopulations of Polish Red cattle formed by animals subjected to conservation program (selected mainly

towards primitive type and breed standard) and animals used for milk production (selected for milk and functional traits). These two populations visibly diversified over time and the phenotypical differences between them mainly refer to milk yield traits as well as exterior features which were strongly shifted towards milk production type in the improved cattle (Fig. 5).

The low average level of genetic differentiation among the studied populations (as shown by global polymorphism parameters) confirms their common origin from local population (Szarek et al. 2004) and suggests that major genetic differences between them are concentrated over the specific genomic *loci* which were divergently selected since the implementation of the improvement program. The performed PCA showed that the genomes of both these populations bear signs of differentiation with fraction of individuals being still genetically close. RE population was, however, subdivided into two visible clusters due to the different level of historical crossbreeding present in the pedigrees of the analyzed animals. The fact that both subpopulations originate from the same native population created an opportunity to trace allele frequency differences among them with high probability that they result from application of selection directed to improvement of milk traits in commercial population. This allowed for identification of candidate genes affecting presumably both primitive traits found in conserved population (being lost due to the intensive selection) (Lindhé and Philipsson 1998) but also *loci* of genes responsible for traits selected in the improved subpopulation.

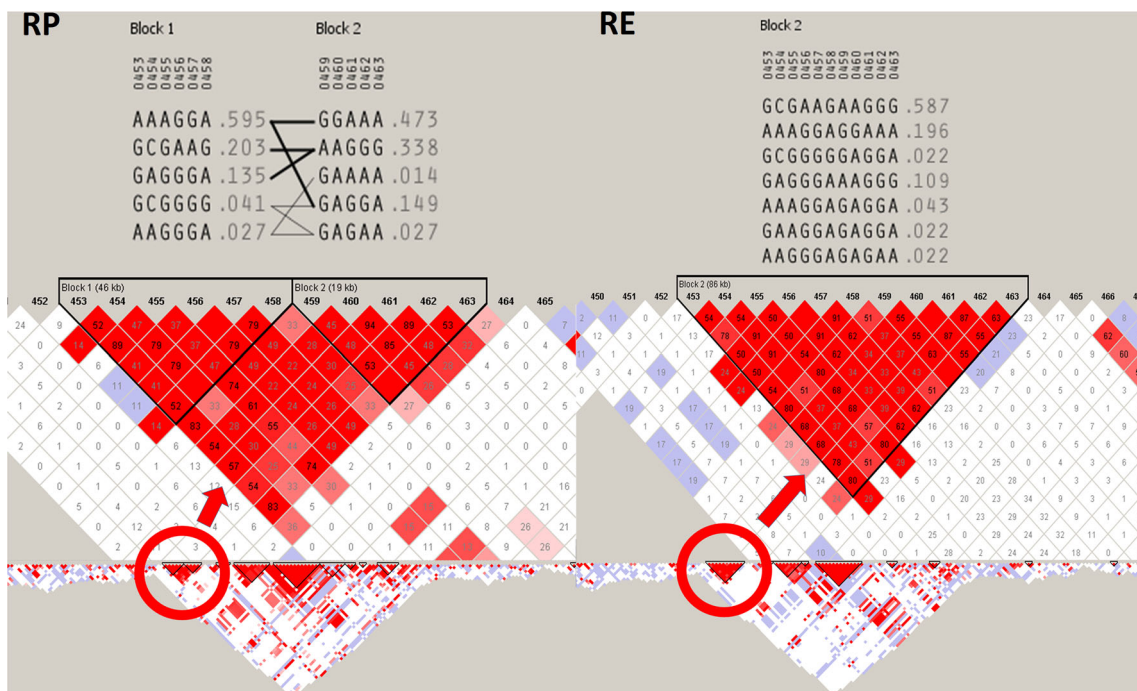


Fig. 3 Linkage disequilibrium and haplotype block structure at the strongest detected diversifying selection signal on BTA3 between 32.5 and 32.9 Mb of genomic sequence. RP, population under “conservation program”; RE, population under “improvement program”. LD nodes

were colored according to D' values and R^2 values were plotted inside the nodes. Red circle marks the haplotype blocks spanned by the detected selection signal. Haplotypes sequences and frequencies are presented on the top

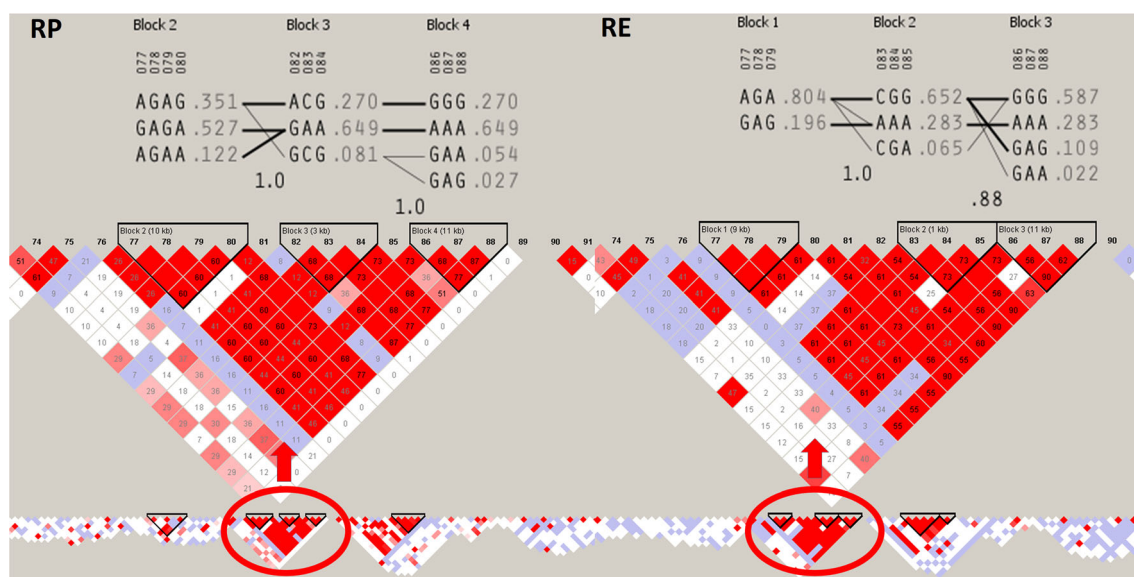


Fig. 4 Linkage disequilibrium and haplotype block structure at the largest strong diversifying selection signal on BTA27 between 4.9 and 6.2 Mb of genomic sequence. *RP*, conserved population; *RE*, improved population. LD nodes were colored according to D' values and R^2 values were plotted

inside the nodes. *Red circle* marks the haplotype blocks spanned by the detected selection signal. Haplotypes sequences and frequencies are presented on the top

The F_{ST} -based approach applied in this study allowed us to identify several selection signals of which the most pronounced were localized on chromosomes 1, 3, 6, 10, 17, and 27. Among the genes associated with the detected signals, we found well-established fat yield and fertility markers like *DGAT1* (*Diacylglycerol O-Acyltransferase 1*) and *FGF2* (*Fibroblast Growth Factor 2*) (Ogorevc et al. 2009). Large portion of genes detected in this study was associated with immune system functioning or antimicrobial resistance and included several members of interleukins and β -defensins family. We also detected some genes responsible for mammary gland development (*SOSTDC1*, *PYGO2*, *MED1*, *CCND1*, *FGF2*), which may be of high importance for the improved population and its milking abilities (Gu et al. 2009; Närhi et al. 2012; Casimiro et al. 2013; Marete et al. 2018). A revision of the gene functions showed that *SOSTDC1* (*Sclerostin Domain Containing 1*) controls the size and shape of mammary buds in mice (Närhi et al. 2012), while *PYGO2* (*Pygopus Family PHD*

Finger 2) ablation results in defective mammary morphogenesis and regeneration (Gu et al. 2009). Moreover, the *FGF2* gene (in addition to its effect on fat yield) was shown to play a role in development and reorganization of the mammary gland (Marete et al. 2018).

Within the selection signal on BTA19, we detected *LASP1* (*LIM and SH3 Protein 1*) gene. This gene was previously shown to have effect on body size regulation in horses (Makvandinejad et al. 2012). It was also demonstrated that its expression level is 18-fold higher in lactating mammary tissue relative to non-lactating tissue in cows (Suchyta et al. 2003).

The strongest detected selection signal on BTA3

The strongest detected diversifying selection signal found in this study was localized on BTA3 between 32.5 and 32.9 Mb of genomic sequence. This region was found to be closely positioned with previously described quantitative trait loci

Fig. 5 Two champion cows belonging to the “conservation program” (on the left; *RP*) and “improvement program” (on the right; *RE*). Clear differences in body constitution and udder morphology traits can be seen between conserved and improved animals (picture by Majewska A)



(QTL; 30,706,774–31,080,623 and 34,065,565–36,461,414) affecting milk yield, fat yield, and protein yield in Holstein and other cattle populations (Rodriguez-Zas et al. 2002; Boichard et al. 2003; Viitala et al. 2003; Ashwell et al. 2004; Ihara et al. 2004). In the study of Cohen-Zinder et al. (2011), genes such as *RAP1A*, *ADORA3*, and *C3H1orf88* were proposed as having a major effect on the milk traits within this QTL. The divergently selected region found in this study did not directly overlap with those genes; however, it was in a large region of high LD (unique in chromosome scale), spanning over 1.1 Mb (32.4–33.6 Mb) in both analyzed populations. Within the detected selection signal, we found only three genes, namely *LRIF1*, *CD53*, and *KCNA3*. These gene functions could be connected with both male fertility and immune system response. Exemplary, *LRIF1* product represses the ligand-induced transcriptional activity of retinoic acid receptor alpha (RARA) and disruption of retinoid signaling was shown to impair mammalian spermatogenesis and fertility (Chung et al. 2011). The region also displayed high marker linkage and extensive haplotype structure, representing possibly reduced genetic variation resulting from selection pressure directed on the analyzed locus.

Extensive selection signal on BTA27

Among the strongest selection signals detected in this study (0.1% of highest F_{ST} values), the largest genomic region with signs of divergent selection was identified on BTA27 (1.29 Mb). This signal overlapped with several genes coding proteins belonging mainly to the β -defensins family (Table 2). β -Defensins are amphipathic cationic peptides that have been reported to function as antimicrobial peptides (AMPs) for the Gram-negative and Gram-positive bacteria, viruses, fungi, and other unicellular parasites (Brogden 2005). Their antimicrobial activity is mediated mainly through permeabilization of cell membrane (Huang 2000) or stimulation of hydrolases, therefore degradation of the cell wall (Bierbaum and Sahl 1985). The first β -defensin was isolated from bovine respiratory tract and was denoted as tracheal AMP (TAP—tracheal antimicrobial peptide) (Diamond et al. 1991). Currently, 58 β -defensin genes located within four clusters have been described in bovine genome on chromosomes 8, 13, 23, and 27 and were designated as cluster A, cluster B, cluster C, and cluster D, respectively (Meade et al. 2013). It was found that the genes important for resistance against intramammary infections are located on BTA27 in the cluster D (Gurao et al. 2017). These β -defensins are expressed both in the mammary gland and in the milk somatic cells, thus have a potential to prevent the intramammary infection (Gurao et al. 2017). In this study, among β -defensin genes overlapped by selection signal on BTA27, we detected, inter alia, *LAP*, *DEFB1*, *DEFB5*, *DEFB10*, and *DEFB4A*. Constitutive expression of *LAP* has been reported in the mammary gland of juvenile,

lactating (both healthy and infected), and non-lactating cows (Roosen et al. 2004). Another study has shown that the *LAP* is expressed only in infected cattle (Swanson et al. 2004). The direct relationship between *LAP* and somatic cell count (SCC) has been also presented, where higher concentration of *LAP* in the milk of cattle infected with *Staphylococcus aureus*, *Streptococcus bovis*, *Streptococcus dysgalactiae*, and *E. coli* was observed than in uninfected cows (Kazuhiro et al. 2013). Similarly, *DEFB1* gene was found to be inducibly expressed during intramammary gland infection (Roosen et al. 2004), while *DEFB5* has been reported in mammary gland infected with coagulase-positive staphylococci and coagulase-negative staphylococci (Kościuczuk et al. 2014). Additionally, the expression of *DEFB10* was high in early lactation stages accompanied by infection by coagulase-positive staphylococci (Kościuczuk et al. 2014). What is important, β -defensins were also shown to have effect on milk traits, where β 4-defensin (*DEFB4A*) polymorphisms were significantly associated with protein yield and fat or protein contents (Bagnicka et al. 2008) and milk yield (Krzyszewski et al. 2008) in Polish Holstein-Friesian cattle.

All these findings suggest that the selection signal detected on BTA27 in this study is associated with selection towards increased resistance against mastitis. Based on LD and haplotype block structure analysis, we presume that the observed allelic differences were created under the influence of relatively recent selection events. This is supported by the fact that genome regions being under strong ongoing selection are characterized by strong LD and extensive haplotype structure (Qanbari et al. 2014) while recombination mechanisms acting over several generations result in LD decay around variants selected a relatively long time ago in the population history (Slatkin 2008). Thus, we propose that both natural and artificial selection are mechanisms acting on BTA27 locus; however, in extensive and intensive breeding conditions, different microbial and environmental factors are of major importance for mammary health, and thus different variants are selected in the studied cattle populations. It is also a known fact that together with increased milk yield and milk synthesis intensity, the immune system activity is being altered. It was shown that selection towards increased milk yield has a detrimental effect on the health condition of the mammary gland (Rupp and Boichard 2003); thus, the mastitis is a problem mainly in high-production herds raised in intensive farming conditions.

Conclusions

By the comparative analysis of divergently selected subpopulations of Polish Red cattle, we again indicated the importance of *DGAT1* and *FGF2* genes for milk production traits in cattle. We also found that among genes being under selection in terms of milk production in red cattle, there are genes

responsible for mammary gland development such as *SOSTDC1*, *PYGO2*, *MEDI1*, and *CCND1* and immune system response like *IL10RA*, *IL12B*, and *IL21*. The most pronounced differences between the analyzed populations were, however, associated with β -defensin genes located within bovine cluster D on BTA27 which shows that antibacterial resistance of mammary gland is of high importance during selection towards increased milk production.

Funding The study was financed from funds of the project: “Directions for use and conservation of livestock genetic resources in sustainable development” co-financed by the National Research and Development Center (Poland) under the Strategic Research and Development Program: “Environment, Agriculture and Forestry” – BIOSTRATEG, the decision number BIOSTRATEG2/297267/14/NCBR/2016.

Compliance with ethical standards

Ethics approval and consent to participate All animal procedures were approved by the Local Animal Care Ethics Committee No. II in Kraków – permission number 1293/2016 in accordance with EU regulations.

Consent for publication Not applicable.

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Adamczyk K, Felenczak A, Jamrozy J, Szarek J, Bulla J (2008) Conservation of Polish Red cattle. *Slovak J Anim Sci* 41:72–76
- Ashwell MS, Heyen DW, Sonstegard TS, Van Tassell CP, Da Y, VanRaden PM, Ron M, Weller JI, Lewin HA (2004) Detection of quantitative trait loci affecting milk production, health, and reproductive traits in Holstein cattle. *J Dairy Sci* 87:468–475
- Bagnicka E, Strzalkowska N, Szreder T, Prusak B, Jozwik A, Kosciuzuk E, Krzyzewski J, Zwierzchowski L (2008) A/C polymorphism in the beta-4 defensin gene and its association with phenotypic and breeding values of milk production traits in Polish-Friesian cows. *Anim Sci Pap Rep* 26:239–250
- Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21:263–265
- Bierbaum G, Sahl HG (1985) Induction of autolysis of staphylococci by the basic peptide antibiotics pep5 and nisin and their influence on the activity of autolytic enzymes. *Arch Microbiol* 141:249–254
- Boichard D, Grohs C, Bourgeois F, Cerqueira F, Faugeras R, Neau A, Rupp R, Amigues Y, Boscher MY, Levéziel H (2003) Detection of genes influencing economic traits in three French dairy cattle breeds. *Genet Sel Evol* 35:77–101
- Brogden KA (2005) Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? *Nat Rev Microbiol* 3:238–250
- Casimiro MC, Wang C, Li Z, Di Sante G, Willmart NE, Addya S, Chen L, Liu Y, Lisanti MP, Pestell RG (2013) Cyclin D1 determines estrogen signaling in the mammary gland in vivo. *Mol Endocrinol* 27:1415–1428
- Chung SSW, Wang X, Roberts SS, Griffey SM, Reczek PR, Wolgemuth DJ (2011) Oral administration of a retinoic acid receptor antagonist reversibly inhibits spermatogenesis in mice. *Endocrinology* 152:2492–2502
- Cohen-Zinder M, Donthu R, Larkin DM, Kumar CG, Rodriguez-Zas SL, Andropolis KE, Oliveira R, Lewin HA (2011) Multisite haplotype on cattle chromosome 3 is associated with quantitative trait locus effects on lactation traits. *Physiol Genomics* 43:1185–1197
- Diamond G, Zasloff M, Eck H, Bresseur M, Maloy WL, Bevins CL (1991) Tracheal antimicrobial peptide, a novel cysteine-rich peptide from mammalian tracheal mucosa: peptide isolation and cloning of a cDNA. *Proc Natl Acad Sci U S A* 88:3952–3956
- Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J et al (2002) The structure of haplotype blocks in the human genome. *Science* 296:2225–2229
- Gu B, Sun P, Yuan Y, Moraes RC, Li A, Teng A, Agrawal A, Rhéaume C, Bilanchone V, Veltmaat JM, Takemaru K, Millar S, Lee EY, Lewis MT, Li B, Dai X (2009) Pygo2 expands mammary progenitor cells by facilitating histone H3 K4 methylation. *J Cell Biol* 185:811–826
- Gurao A, Kashyap SK, Singh R (2017) β -defensins: an innate defense for bovine mastitis. *Veterinary World* 10:990–998
- Huang HW (2000) Action of antimicrobial peptides: two-state model. *Biochemistry* 39:8347–8352
- Ihara N, Takasuga A, Mizoshita K, Takeda H, Sugimoto M, Mizoguchi Y, Hirano T, Itoh T, Watanabe T, Reed KM, Snelling WM, Kappes SM, Beattie CW, Bennett GL, Sugimoto Y (2004) A comprehensive genetic map of the cattle genome based on 3802 microsatellites. *Genome Res* 14:1987–1998
- Karolchik D, Hinrichs AS, Furey TS, Roskin KM, Sugnet CW, Haussler D, Kent WJ (2004) The UCSC Table Browser data retrieval tool. *Nucleic Acids Res* 32(Database issue):D493–D496
- Kazuhiro K, Akamatsu H, Obayashi T, Nagahata H, Higuchi H, Iwano H, Oshida T, Yoshimura Y, Isobe N (2013) Relationship between concentration of lingual antimicrobial peptide and somatic cell count in milk of dairy cows. *Vet Immunol Immunopathol* 153:298–301
- Koścuzuk EM, Lisowski P, Jarczak J, Krzyzewski J, Zwierzchowski L, Bagnicka E (2014) Expression patterns of β -defensin and cathelicidin genes in parenchyma of bovine mammary gland infected with coagulase-positive or coagulase-negative staphylococci. *BMC Vet Res* 10:1
- Krzyzewski J, Bagnicka E, Strzalkowska N, Jozwik A, Pyzel B, Zwierzchowski L (2008) Association between the polymorphism of bovine beta4-defensin gene and milk traits in Holstein-Friesian cows as computed for standard [305 days] and the whole lactation. *Anim Sci Pap Rep* 26:191–198
- Lindhé B, Philipsson J (1998) Genetic correlations between production with disease resistance and fertility in dairy cattle and consequences for total merit selection. *Acta Agric Scand* 48:216–221
- Makvandi-Nejad S, Hoffman GE, Allen JJ et al (2012) Four loci explain 83% of size variation in the horse. *PLoS One* 7(7):e39929
- Malkovský M, Sondel PM, Strober W, Dalgleish AG (1988) The interleukins in acquired disease. *Clin Exp Immunol* 74:151–161
- Marete A, Lund MS, Boichard D, Ramayo-Caldas Y (2018) A system-based analysis of the genetic determinism of udder conformation and health phenotypes across three French dairy cattle breeds. *PLoS One* 13:e0199931
- Meade KG, Cormican P, Narciandi F, Lloyd A, O’farrelly C (2013) Bovine β -defensin gene family: opportunities to improve animal health? *Physiol Genomics* 46:17–28

- Närhi K, Tummers M, Ahtiainen L, Itoh N, Thesleff I, Mikkola ML (2012) *Sostdc1* defines the size and number of skin appendage placodes. *Dev Biol* 364:149–161
- Ogorevc J, Kunej T, Razpet A, Dovc P (2009) Database of cattle candidate genes and genetic markers for milk production and mastitis. *Anim Genet* 40:832–851
- Qanbari S, Pausch H, Jansen S, Somel M, Strom TM, Fries R, Nielsen R, Simianer H (2014) Classic selective sweeps revealed by massive sequencing in cattle. *PLoS Genet* 10:e1004148
- Rodriguez-Zas SL, Southey BR, Heyen DW, Lewin HA (2002) Interval and composite interval mapping of somatic cell score, yield, and components of milk in dairy cattle. *J Dairy Sci* 85:3081–3091
- Roosen S, Exner K, Paul S, Schroeder JM, Kalm E, Looft C (2004) Bovine b-defensins: identification and characterization of novel bovine b-defensin genes and their expression in mammary gland tissue. *Mamm Genome* 15:834–842
- Rupp R, Boichard D (2003) Genetics of resistance to mastitis in dairy cattle. *Vet Res* 34:671–688
- Slatkin M (2008) Linkage disequilibrium — understanding the evolutionary past and mapping the medical future. *Nat Rev Genet* 9:477–485
- Suchyta SP, Sipkovsky S, Halgren RG, Kruska R, Elftman M, Weber-Nielsen M, Vandehaar MJ, Xiao L, Tempelman RJ, Coussens PM (2003) Bovine mammary gene expression profiling using a cDNA microarray enhanced for mammary-specific transcripts. *Physiol Genomics* 16:8–18
- Swanson K, Gorodetsky S, Good L, Davis S, Musgrave D, Stelwagen K, Farr V, Molenaar A (2004) Expression of a beta-defensin mRNA, lingual antimicrobial peptide, in bovine mammary epithelial tissue is induced by mastitis. *Infect Immunol* 72:7311–7314
- Szarek J, Adamczyk K, Felenczak A (2004) Polish Red cattle breeding: past and present. *AGRI* 35:21–35
- Todros-Dawda I, Kveberg L, Vaage JT, Inngjerdigen M (2014) The tetraspanin CD53 modulates responses from activating NK cell receptors, promoting LFA-1 activation and dampening NK cell effector functions. *PLoS One* 9:e97844
- Viitala SM, Schulman NF, de Koning DJ, Elo K, Kinos R, Virta A, Virta J, Mäki-Tanila A, Vilkki JH (2003) Quantitative trait loci affecting milk production traits in Finnish Ayrshire dairy cattle. *J Dairy Sci* 86:1828–1836
- Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38:1358–1370
- Xie C, Mao X, Huang J, Ding Y, Wu J et al (2011) KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res* 39:W316–W322
- Zhang W, Samelson LE (2000) The role of membrane-associated adaptors in T cell receptor signalling. *Semin Immunol* 12:35–41