Iranian Journal of Applied Animal Science

Genomic Selection Signatures in Two French and Swedish Holstein Cattle Breeds Provide Evidence for Several Potential Candidate Genes Linked to Economic Traits

Research Article

R. Salehi[1], A. Javanmard[1*], M. Mokhber[2*] and S. Alijani[1]

[1] Department of Animal Science, Faculty of Agriculture, University of Tabriz, Tabriz, Iran
[2] Department of Animal Science, Faculty of Agriculture, Urmia University, Urmia, Iran

Received on: 30 Oct 2022
Revised on: 21 Feb 2023
Accepted on: 1 Mar 2023
Online Published on: Dec 2023

*Correspondence E-mail: a.javanmard@tabrizu.ac.ir; m.mokhber@urmia.ac.ir

ABSTRACT

The genomes of domestic animals have been artificially selected during long periods of domestication and have led to many significant changes in their economic traits. Exploring these changes is useful to expand our knowledge about the domestication process, understand the complexity of genetic diversity in livestock, and make an appropriate breeding decision. The genotypic information of 23 individuals representing French and Swedish Holsteins was used to detect signature signs between these two breeds. Quality control and data filtering were performed using PLINK software. The principal component analysis was performed to determine genetic diversity. In addition, linkage disequilibrium (LD), ancestral, and recent effective population size (Ne) were separately estimated for each breed. The signature of selection was determined by the fixation index (F_{ST}) statistic. The corrected r^2 (calculated statistics for LD) between single nucleotide polymorphisms (SNPs) decreased with increasing the physical distance from 100 Kb to 7.5 Mb (from about 0.3 to 0.09 in both breeds). These values were slightly smaller for the French breed. The effective population size was estimated at 1775 and 2120 individuals for French and Swedish Holsteins 900 generations ago to 31 and 39 in the recent generation, respectively. Overall, three regions with outlier F_{ST} values were identified as the signature of selection. The flanking SNP was mostly located on BTA2, BTA16, and BTA19. Candidate genes were found to be associated with SERTAD (liver metabolism and health), SYT14 (meat yield and marble levels), IRF6 (immune system), HSD11B19 (heat stress), LAMB3, and (fat deposition). Therefore, these potential genes can be considered invaluable genetic resources for future research to attempt to create generations of commercial breeds of cattle.

KEY WORDS annotated regions, candidate genes, genomics, selection sweep, signature of selection.

INTRODUCTION

Domestic animals are the product of selection, improvement, and domestication processes and have undergone the effects of genetic drift, mutation, and artificial selection (Giacomoni *et al.* 2008). The results of studies on domestic livestock have demonstrated that the genetic diversity of domestic animals has decreased during the domestication process due to various genetic factors such as genetic drift, and natural and artificial selection of breeds, as well as the use of a small number of breeds for more economic production. The decrease in the diversity of these animals has caused global concern (Holsinger and Weir, 2009). In this regard, using genome information is highly beneficial for identifying the population structure and the evolutionary history and developing genetic diversity conservation strategies. The identification of the signatures of selection is regarded as one of the important and practical aspects of population genomic analysis. The signatures of selection refer to patterns on the genome that are formed in specific regions of the genome by selective forces (Simianer *et al.* 2014). It is of interest to know how domestication is related to the corresponding traits or genes and what mechanism controls the divergence in the genome sequence for each particular breed of cattle (Johnsson, 2018). In addition, exploring the genetic signature will provide insights into the importance of signature sweeps and the genome footprint of previous selections related to randomly desirable single nucleotide polymorphisms (SNPs) and the history of introgression (Kim *et al.* 2013; Taye *et al.* 2017).

To our best knowledge, cattle are the primary species that provide a variety of human needs, such as meat, milk, and leather, in addition to being used for work and fertilization. Further, this livestock species is divided into two specific classes: *Bos taurus* and *Bos indicus*. In particular, the domestication of cattle has been estimated to have separately occurred around 11000 years ago for both subspecies of West Asian cattle (Troy *et al.* 2001; Orozco-ter Wengel *et al.* 2015). Holstein cattle are globally known as high milk producers and are the result of various artificial selection constraints. Distributed in over 150 countries worldwide, this breed has undergone intense selection, particularly in the last few decades and following the introduction of progeny testing-based breeding programs (Groeneveld *et al.* 2010; Decker *et al.* 2014).

So far, genome-wide association studies have mostly focused on the relative importance of signature recognition tools concerning the act of selection. In response to the selective force, chromosome fragments surrounded by favorable alleles are swept close or use the mechanism of the hitchhiker effect and consequently introduce a high rate of homozygosity patterns into the genome of the examined breed (Decker *et al.* 2014). Therefore, nucleotide variation and inheritance patterns within these types of specific loci in the genome largely affect crossover and recombination, effective population sizes, number of generations, relative ages of neutrally linked alleles, bottleneck and founder effects, type and intensity, and the selection coefficients. They also influence the length of chromosome fragments, over which the presence of a positive allele is inversely swept in proportion to the number of generations of selection, thereby influencing selection processes (Upadhyay *et al.* 2017).

It is a well-documented fact that the results of the selection signature study can aid in the identification of responsible desirable SNPs within candidate genes that control economic traits in cattle breeds and raw material for increasing the frequency of favorable alleles during post-domestication in response to the next-generation breeding program (Decker *et al.* 2014).

The majority of previous reports have argued that a selection sweeps for causative mutations within the candidate gene controlling certain economic traits is described in different cattle breeds, including *GDF-8* and *ABCG2* genes responsible for muscle development and milk composition (Ron and Weller, 2007), as well as CHCHD7, PLAG1, and SMAD2 for body size (Zhao *et al.* 2015). The other breeds were the *R3HDM1* gene that controls feed intake (Zhao *et al.* 2015) the SIGLEC5 gene as the possible candidate for productive lifespan and fertility trait variation (Hardie *et al.* 2017) DGAT1 for milk production (Cruz *et al.* 2019) and the XKR4 gene for growth trait (Porto Neto *et al.* 2012). The remaining breeds included the SF2 gene responsible for reproduction, *MC1R* and *KIT* genes responsible for coat color (Petersen *et al.* 2020), and recognition of good knowledge to understand the mechanism of genetic diversity in each breed of cattle (Pan *et al.* 2016).

Although varying reports exist on the identification of GWAS-based evidence of signature sweep in cattle breeds, limited studies have systematically examined this phenomenon in Swedish cattle breeds, which alone may not be sufficient. Accordingly, understanding the genetic structure and factors influencing genome variations in well-adapted native cattle breeds is the first critical step in developing an appropriate breeding program. Therefore, the present study intends to use the high-density SNP genotype information of two French and Swedish Holstein cattle breeds to characterize the population structure and effective population size (Ne) and identify the regions related to the genome of the studied cattle breeds that have been affected by natural or artificial selection forces. It is expected that the results of the study can be useful for analyzing the differences between the investigated breeds, implementing breeding programs, and preserving their biodiversity.

MATERIALS AND METHODS

Sample collection and breeds

The study group consisted of 46 individuals (French [n=23] and Swedish Holstein [n=23] cattle populations obtained from the publicly available database SNPchip. The genomic data related to Swedish and French Holstein cattle were obtained from the Dryad information repository (https://datadryad.org/) and the WIDDE repository (Widde.toulouse.inra.fr/widde), respectively.

Quality control process

Quality control and filtering were separately performed using PLINK software (Purcell *et al.* 2007) for each population. The informative and high-quality genotypes, including individuals and SNPs with a call rate of more than 95%, SNPs with minor allele frequency (MAF) of more than 1%, and Hardy-Weinberg Equilibrium (HWE=0.000001) on autosomal chromosomes, were used for further analyses.

Principal component analysis (PCA)

The PCA was performed to examine genetic diversity and determine genetic groups using Plink in R (R Development Core Team, 2013). The first two principal components (PC1 & PC2) were plotted in a two-dimensional graph.

Signature selection analysis using the fixation index (Fst) approach

To better understand the genetic divergence among all breeds, F_{ST} was calculated using the HierFstat R package with the unbiased estimator proposed by Weir and Cockerham (1983).

The Fst index explains the rate of heterogeneity and chromosomal localization of this event between two studied breeds. The plink Fst within the script was employed for such a calculation. Moreover, the Runnmed script and the Manhattan chart were employed to smooth the nose and summarize the statistics for each examined chromosome, respectively. The threshold line for the p-value was created with the z-score and the Lower.tail script in Plink. Carrying a P-value of 0.0001, the output was listed as significant loci for the first value. F_{ST} values less than 0 were converted to zeros, and the statistics were then smoothed using the Runmed function.

Linkage disequilibriums (LD) and population (Ne)

The haplotype phase was determined by BEAGLE (version 3.3.2,) and then Haploview software was used to visualize LD (Browning and Browning, 2011). Additionally, the Ne was calculated with SNeP, version 1.1 (Corbin *et al.* 2012; Barbato *et al.* 2015).

Downstream bioinformatics analysis for candidate gene annotation

Genes located in genomic regions, including the consecutive SNPs with a q-value of 0.05, were considered statistically significant intervals. The boundary for each interval was set such that the first flanking SNP was observable (the one q value greater than 0.1). Then, protein-coding genes were extracted from significant regions using the UMD3.1 cow reference genome assembly. Next, an extensive literature review was conducted to comment on the functions of the identified genes. Finally, biological pathways and gene networks were determined by DAVID 6.8 (Huang *et al.* 2009).

RESULTS AND DISCUSSION

Of all available SNPs (N=40823) after merging data from two studied breeds, 1667 and 512 SNPs were excluded due to missing genotypes and MAF, respectively, and finally, 38643 were retained for subsequent analyses.

The first and second PCAs explained 12.87% and 5.44% of the variance, respectively (Figure 1). Based on the PCA analysis results, despite the small differentiation, no significant or outstanding differentiation was found between the two studied breeds (Figure 1).

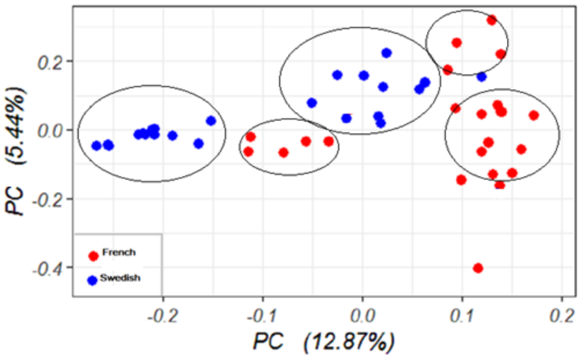

Figure 1 Principal component analysis (PCA) related to Swedish and French Holstein breeds

As expected, the average r^2 values decreased with increasing the distance between pairwise SNPs for all the studied populations (Figure 2). Differences among French and Swedish Holstein breeds were negligible.

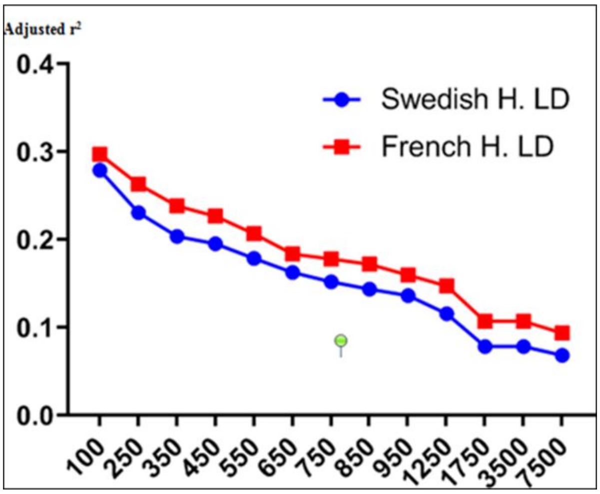

Figure 2 Linkage disequilibrium (LD) decay for increasing distance (Kb) for Swedish and French Holstein populations

Ne was estimated from the last 900 to recent generations in the present study. Based on the data (Figure 3), Ne decreased from the past (900) to the recent generation (4). Ne was clear in both Swedish and French Holstein breeds due to intensive direct selection for economical traits.

This analysis represented evidence for the F_{ST} index, and the mean and weighted Fst between Swedish and French Holstein populations were 0.018227 and 0.02093 (no differentiation and high closeness), respectively.

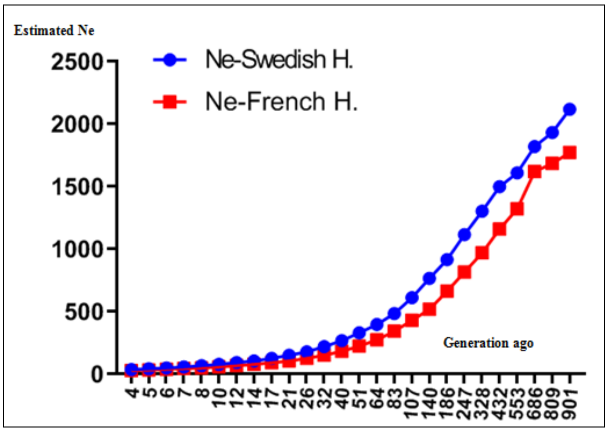

Figure 3 Estimates of effective population size (Ne) over time across the last generations based on linkage disequilibrium calculations from 29 autosomal chromosomes

Significant genomic regions were identified, and it was found that the flanking SNP was located mostly on BTA2, BTA16, and BTA19, indicating that the annotation of genes in these regions revealed several verified and novel candidate genes. The identified genes relating to detected regions as a signature of selections were associated with a variety of traits, including SERTAD (liver metabolism and health), SYT14 (meat yield and marble levels), IRF6 (immune system), HSD11B19 (heat stress), andLAMB3 (fat deposition). Figure 4 illustrates the Manhattan plot of the genomic regions detected by the Fst statistic to be under putative selection. The dashed lines represent the significant threshold level at an FDR of 5% (q-.value<0.05).

Gene ontology and study of the biological function of the highlighted candidate gene illustrated the role of the identified gene in different cell pathways and functionally, the MAPK signaling pathway, cardiac muscle contraction, adrenergic signaling in cardiomyocytes, oxytocin signaling pathway, voltage-gated cation channels, hypertrophic cardiomyopathy, and steroid hormone biosynthesis (Figure 5).

Most livestock were artificially selected during human domestication and, consequently, were found to have varied desirable traits, behaviors, body sizes, and production. A signature sweep is a suitable tool for the identification of footprint patterns on their genomes during artificial and natural selection.

PCA is a suitable tool for the recognition of admixture and relatedness among the studied Holstein cattle. Our finding (with more than 18% explained variance) demonstrated no clear differentiation between the studied populations. This is probably because of intensive selection in the Holstein cattle. Kim *et al.* (2013) reported that intensive selection programs for milk have been facilitated by mass artificial insemination since the 1960s, increasing the similarity of genomes among North American Holsteins.

Bodenmüller Filho *et al.* (2010) evaluated the differences between production systems in the Northern region of Paraná using seven characteristics of milk production and quality and verified that three main components were sufficient to explain 70.52% of the total variance of the characteristics.

Moreover, Fraga *et al.* (2016) assessed the relationship between the productive characteristics and the genotypic proportions of Holstein and Zebu crossbred dairy cattle and observed that the first two principal components accounted for 89.4% of the total variance, representing production and genotypic components. Likewise, Upadhyay *et al.* (2017) identified genomic relatedness and diversity of Swedish native cattle breeds. They found that the first principal component (EV1), which explains 5.76% of the total variance, clearly separated Fjäll, Swedish Polled, Fjällnära, and Bohus Polled breeds from the Swedish Red, Väne, and Ringamåla breeds. Zhang *et al.* (2022) identified genomic diversity and concludedthat the first and second PCs explained 6.23% and 2.58% of the variation in the entire genomic data, respectively.

During positive selection, allele frequency tends to change; therefore, some specific alleles relatively increase and then are measurable through long haplotypes and low recombination situations. LD is an informative index. It can provide useful information about variations related to quantitative traits, various evolutionary forces, genetic diversity, and population structure and reveal genomic regions affected by selection. The first reports on the extent of LD in the bovine genome described a large LD range (e.g., up to 20 cM) (Farnir *et al.* 2000; Tenesa *et al.* 2003). Further analyses with denser markers confirmed extensive LD but generally represented lower values (Khatkar *et al.* 2006). In some other studies, two genome-wide studies based on 10K SNP data demonstrated that the level of LD is lower than what was previously thought (e.g., Sargolzaei *et al.* 2008; Kim and Kirkpatrik, 2009).

Sargolzaei *et al.* (2008) for North American Holstein cattle and Kim and Kirkpatrik (2009) reported a strong LD (r^2>0.8) in the genomic regions of the proportion of SNP pairs of the LD patterns of approximately 50 Kb or less, which is extremely larger than our results (r^2<0.3 for distances less than 200Kb).

As mentioned earlier, our findings demonstrated that LD decreased with increasing the genomic distance for Swedish and French Holstein populations. It was found that the r^2 value decayed below 0.1 after a 7500Kbp distance between adjacent SNPs (Figure 2). The low levels of LD have been previously reported by Edea *et al.* (2015) and by Espigolan *et al.* (2013) for Ethiopian (0.14) and Nellore (0.17) cattle at a similar genetic distance, respectively.

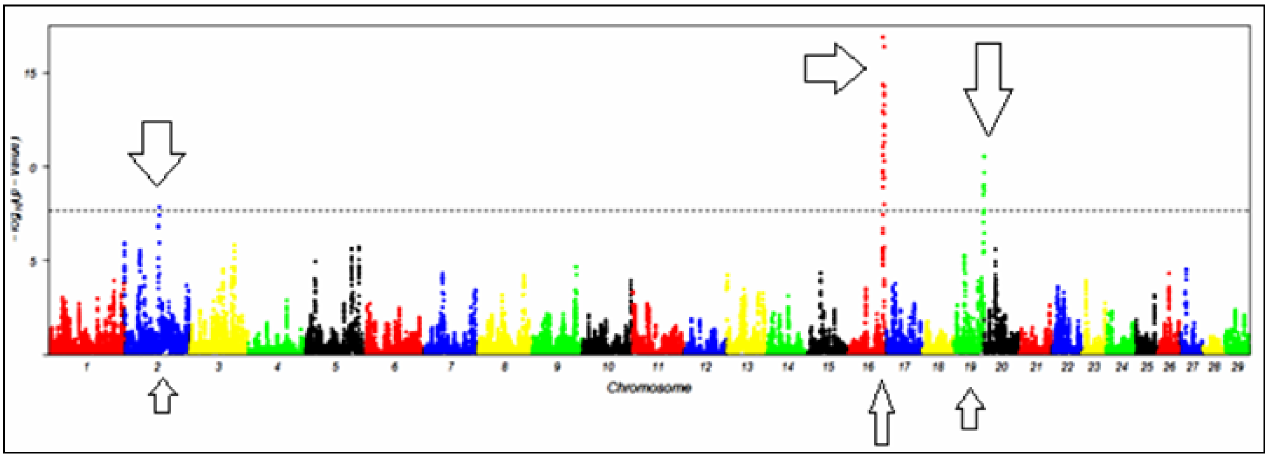

Figure 4 Manhattan plot of the genomic regions detected by the fixation index (Fst) statistic as being under putative selection
The dashed lines demonstrate the significant threshold level at an FDR of 5% (q-value<0.05)

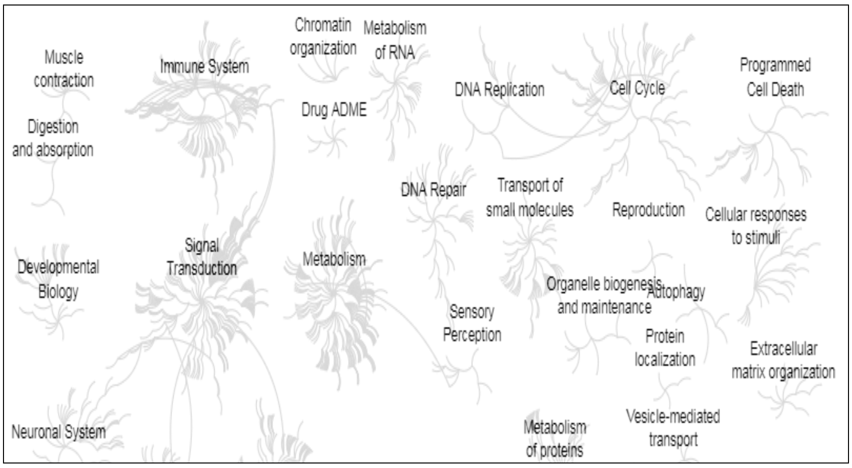

Figure 5 The mainly detected biological functions related to identified genes from the signature of selection

Jemaa *et al.* (2019) showed more similar LD patterns to LIM than the Tuscan breeds with r^2 values lower than 0.05 at larger distances (>0.5 Mbp). In another study, Makina *et al.* (2015) hypothesized that the decay of Tuscan breeds was slower than the other local breeds. They studied four local African cattle compared with Angus and Holstein breeds. In all six breeds, at 1 Mbp, r^2 values were lower than 0.1 and in the four local breeds (Afrikaner, Nguni, Drakensberger, and Bonsmara). In the Chinese Simmental breed, Zhu *et al.* (2013) in the window of 0.5-1 Mbp, the r^2 was 0.05. Similarly, Biegelmeyer *et al.* (2016) investigated LD patterns in Hereford and Braford and found average r^2 values equal to 0.07 and 0.06, respectively.

Our finding confirmed that the decrease in Ne was clear in both Swedish and French Holstein breeds due to intensive direct selection for economical traits.

The estimated Ne for French and Swedish Holstein cattle breeds was about 1775 and 2120 at 900 generations ago but decreased to 184 and 270, as well as 31 and 39, 40 generations ago and recent generations, respectively. Decreased trends in Ne could be explained by strong selection and the use of relatively few elite sires in the breeding process. The estimated Ne in the recent generations of this study was less than the critical threshold of Ne= 100 estimated for the long-term viability of discrete livestock breeds (Meuwissen, 2009). Kukuˇcková *et al.* (2017) evaluated the Ne of 15 European cattle breeds (Ne was equal to ~300 in the 60th generation ago. Mokhber *et al.* (2019) estimated Ne for Azari, Khuzestani, and Mazandarani breeds at 1530, 1375, and 1141, respectively, for 700 generations ago. Ne for the present generation in Azeri, Khuzestani, and Mazandarani breeds was estimated at 447, 226, and 35, respectively.

Moreover, Kour *et al.* (2022) investigated Ne for Arunachali yak animals, and the most recent generation (i.e., 13 generations ago) was estimated to be 83. The Nevalues for the Kholmogor breed declined over time, whereas the lowest *N*E value was observed for the Yaroslavl breed 5 generations ago (Ne=111) with a subsequent increase to 118 in 3 generations ago. The most recent Ne values detected for Yaroslavl (Ne5=111) and Kholmogor (Ne5=130) cattle were higher than those found in Holsteins (Ne5=95) (Zinovieva *et al.* 2020).

In this study, the selection signatures were recognized at autosomal regions with an F_{ST} value higher than 0.05 (Figure 4).

The selected genomic regions as a signature of selection were mostly located on chromosomes BTA2, BTA16, and BTA19, implying that the annotation of genes in these regions revealed several verified and novel candidate genes. Moravčíková *et al.* (2019) analyzed selection signatures in the beef cattle genome and concluded that the selection signatures were recognized as autosomal regions with an F_{ST} value higher than 0.2. The identified genes related to detected regions as a signature of selections were associated with a variety of traits, including muscle development and the MYLK (myosin light chain kinase). Strong selection signals were identified close to the genes coding the coat color of cattle on BTA6 (KIT & KDR) and BTA18 (MC1R). In cattle, the KIT gene is responsible for pied color patterns also assigned as “spotting loci” (Fontanesi *et al.* 2010). The average F_{ST} values obtained in the present study were close to those observed in different commercial cattle breeds. Alejandra *et al.* (2022) evaluated genetic architecture and signatures of selection with Fst in the Caqueteño Creole and found that four of the eight SNPs with significant values overlapped candidate genes previously associated with disease resistance. For example, the TMPRSS15 gene (BTA1) has been associated with mastitis resistance. Additionally, three F_{ST} peaks were within the genomic regions of BTA22, and the PGAM2 gene was associated with muscle fat deposition, body conformation, tenderness, carcass, and meat quality. In addition, the EGFR gene was related to adaptation and reproduction in cattle populations (Feitosa *et al.* 2021). Mokhber *et al.* (2018) studied the signatures of selection with F_{ST} in Azeri and Khuzestani buffalo breeds. Based on their results, 13 regions with outlier F_{ST} values (0.1%) were identified, and related genes were involved in milk production, cytoskeleton organization, growth, metabolic function, apoptosis, and domestication-related changes, including immune and nervous system development. Eventually, Zinovieva *et al.* (2020) investigated selection signatures in the two oldest Russian native cattle breeds and observed variations in genetic differentiation between breeds based on F_{ST} through the genome. Future investigations are necessary to validate the kinds of conclusions that can be drawn from this study.

CONCLUSION

Currently, SNP chip tools can be applied for identifying genomic regions under selection and highlighting the occurrence of selective sweep through the whole genome. The data related to the genomics of domestication on two cattle breeds from Swedish and French cattle were examined in the present research. Accordingly, to address the genetic structure and evolutionary connectedness of both cattle populations, the signature of the selection theory was used as a mirror of genome variation during historical changes and response for their population size and putative selection. Significant genomic regions were identified, and searching for the flanking SNP confirmed that the annotation of genes in these regions demonstrated several verified and novel candidate genes. The highlighted candidate genes were associated with a wide variety of traits such as SERTAD (liver metabolism and health), SYT14 (meat yield and marble levels), IRF6 (immune system), HSD11B19 (heat stress), and LAMB3 (fat deposition). The scanning of genomic signature selection produces invaluable knowledge about economic traits in Holstein cattle. Therefore, these potential genes can be considered an invaluable genetic resource for future research to attempt to create generations of Holstein breeds well suited to French, Swedish, and other countries with the same conditions. Future investigations are required to validate the outcomes.

ACKNOWLEDGEMENT

The authors would like to thank the reviewer for the constructive feedback.

REFERENCES

Alejandra M., Ana C., Wellington B., Gustavo P., Viviana H., Vallejo A. and Morea O. (2022). Genetic architecture and signatures of selection in the caqueteño creole (Colombian Native Cattle). *Diversity.* **14,** 828-835.

Barbato M., Orozco-terWengel P.A., Tapio M. and Bruford M.W. (2015). SNeP: A tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. *Front. Genet*. **6,** 1-6.

Biegelmeyer P., Gulias-Gomes C.C., Caetano A.R., Steibel J.P. and Cardoso F.F. (2016). Linkage disequilibrium, persistence of phase and effective population size estimates in Hereford and Braford cattle. *BMC Genet.* **17,** 32-45.

Bodenmüller Filho A. Damasceno J.C., Previdelli I.T.S., Santana R.G., Ramos C.E.C.D.O. and Santos G.T.D. (2010). Typology of production systems based on the milk characteristics. *Rev. Bras. Zootec.* **39,** 1832-1839.

Browning B.L. and Browning S.R. (2011). A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am. J. Hum. Genet.* **84,** 210-223.

Corbin L.J., Liu A.Y.H., Bishop S.C. and Woolliams J.A. (2012). Estimation of historical effective population size using linkage disequilibria with marker data. *Anim. Breed. Genet. 129, 257–270.*

Cruz V.A.R., Oliveira H.R., Brito L.F., Fleming A., Larmer S., Miglior F. and Schenkel F.S. (2019). Genome-Wide Association Study for Milk Fatty Acids in Holstein Cattle Accounting for the DGAT1 Gene Effect. *Animals (Basel). 9,997-1009.*

Decker J.E., McKay S.D., Rolf M.M., Kim J.W., Alcala´ A.M., Sonstegard T.S., Hanotte O., Go¨therstro¨ A., Seabury C.M., Praharani L., Babar M.E., Correia de Almeida Regitano L., Yildiz1 M.A., Heaton M.P., Liu W.S., Lei C.Z., Reecy J.M., Saif-Ur-Rehman M., Schnabel R.D. and Taylor J.F. (2014). Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. *PLoS Genet*. **10(3),** e1004254.

Edea Z., Dadi H., Dessie T., Lee S.H. and Kim K.S. (2015). Genome-wide linkage disequilibrium analysis of indigenous cattle breeds of Ethiopia and Korea using di_erent SNP genotyping BeadChips. *Genes Genom.* **37,** 759-765.

Espigolan R., Baldi F., Boligon A.A. Souza F.R.P., Gordo D.G.M. and Tonussi R.L. (2013). Study of whole genome linkage disequilibriumin Nellore cattle. *BMC Genom.* **14,** 305-315.

Farnir F., Coppieters W., Arranz J.J., Berzi P., Cambisano N., Grisart B., Karim L., Marcq F., Moreau L., Mni M., Nezer C., Simon P., Vanmanshoven P., Wagenaar D. and Georges M. (2000). Extensive genome-wide linkage disequilibrium in cattle. *Genome Res.* **10,** 220-227.

Feitosa F.L.B., Pereira A.S.C., Mueller L.F., Fonseca P.A., Braz C.U., Amorin S., Espigolan R., Lemos M.A., Albuquerque L.G. and Schenkel F.S. (2021). Genome-wide association study for beef fatty acid profile using haplotypes in Nellore Cattle. *Livest. Sci*. **245,** 104396-104405.

Fontanesi L., Scotti E. and Russo V. (2010). Analysis of SNPs in the KIT gene of cattle with different coat colour patterns and perspectives to use these markers for breed traceability and authentication of beef and dairy products. Italian *J. Anim. Sci.* **9,** 42-52.

Fraga A.B., De Lima Silva F., Hongyu K., Da Silva Santos D., Murphy T.W. and Lopes F.B. (2016). Multivariate analysis to evaluategenetic groups and production tr*8*.aits of crossbred Holstein × Zebu cows. *Trop. Anim. Health Prod.* **48,** 533-542.

Giacomoni E.H., Fernández-Stolz G.P. and Freitas T.R.O. (2008). Genetic diversity in the Pantaneiro horse breed assessed using microsatellite DNA markers. *Genet. Mol. Res.* **7(1),** 261-270.

Groeneveld L.F., Lenstra J.A., Eding H., ToroM. A., Scherf B. and Pilling D. (2010). Genetic diversity in farm animals. *Anim. Genet.* **41,** 6-31.

Hardie L.C., VandeHaar M.J., Tempelman R.J., Weigel K.A., Armentano L.E. and Wiggans G.R. (2017). The genetic and biological basis of feed efficiency in mid-lactation Holstein dairy cows. *Dairy Sci.* **100,** 9061-9075.

Holsinger K.E. and Weir B.S. (2009). Genetics in geographically structured populations: defining, estimating and interpreting F_{ST}. *Nat. Rev. Genet.* **10(9),** 639-650.

Huang D.W., Sherman B.T. and Lempicki R.A. (2009). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37(1),** 1-13.

Jemaa S.B., Thamri N., Mnara S., Rebours E., Rocha D., Boussaha M., Jemaa S.B., Thamri N., Mnara S. and Rebours E. (2019). Linkage disequilibrium and past e_ective population size in native Tunisian cattle. *Genet. Mol. Biol.* **42,** 52-61.

Johnsson M. (2018). Integrating selection mapping with genetic mapping and functional genomics. *Front. Genet.* **9,** 603-610.

Kim J., Hanotte O., Mwai O.A., Dessie T., Bashir S. and Diallo B. (2013). The genome landscape of indigenous African cattle. *Genome Biol.* **18,** 34-45.

Khatkar M.S., Collins A., Cavanagh J.A.L., Hawken R., Hobbs M., Zenger K., Barris W., McClintock A.E., Thomson P.C., Nicholas F.W. and Raadsma H.W. (2006). A first generation metric linkage disequilibrium map of bovine chromosome 6. *Genetics.* **174,** 79-85.

Kim E.S. and Kirkpatrick B.W. (2009). Genome-widescan for bovine twinning rate QTL using linkage disequilibrium. *Anim. Genet.* **40,** 300-307.

Kour A., Niranjan S.K., Malayaperumal M., Surati U., Pukhrambam M., Sivalingam J., Kumar A. and Sarkar M. (2022). Genomic diversity profiling and breed-specific evolutionary signatures of selection in arunachali yak. *Genes.* **13,** 2-16.

Kukuˇcková V., Moravˇcíková N., Ferenˇcakovi´c M., Simˇciˇc M., Mészáros G., Sölkner J., Trakovická A., Kadleˇcík O., Curik I. and Kasarda R. (2017). Genomic characterization of Pinzgau cattle: Genetic conservation and breeding perspectives. *Conserv. Genet.* **18,** 893-910.

Makina S.O., Taylor J.F., van Marle-Köster E., Muchadeyi F.C., Makgahlela M.L., MacNeil M.D. and Maiwashe A. (2015). Extent of linkage disequilibrium and e_ective population size in four South African Sanga cattle breeds. *Front. Genet.* **6,** 11-23.

Meuwissen T. (2009) .Genetic management of small populations: A review. *Acta Agric. Scand. A Anim. Sci.* **59,** 71-79.

Mokhber M., Moradi-Shahrbabak M., Sadeghi M., Moradi H. and Rahmani J. (2019). Estimation of effective population size of Iranian water buffalo by genomic data. *Iranian J. Anim. Sci.* **50,** 197-205.

Mokhber M., Moradi-Shahrbabak M., Sadeghi M., Moradi-Shahrbabak H., Stella A., Nicolzzi E., Rahmaninia J. and Williams J. (2018). A genome-wide scan for signatures of selection in Azeri and Khuzestani buffalo breeds. *BMC Genom.* **19,** 449-458.

Moravčíková N., Kasarda R., Vostrý L., Krupová Z., Krupa E., Lehocká K., Olšanská B., Trakovická A., Belej L. and Golian J. (2019). Analysis of selection signatures in the beef cattle genome. *J. Anim. Sci.* **12,** 491-503.

Orozco-ter Wengel P., Barbato M., Nicolazzi E., Biscarini F., Milanesi M. and Davies W. (2015). Revisiting demographic processes in cattle with genome-wide population genetic analysis. *Front. Genet.* **6,** 191-205.

Pan Z.Y., He X.Y., Wang X.Y., Guo X.F., Cao X., Hu W.P., Di R., Liu Q.Y. and Chu M.X. (2016). Selection signature in do mesticated animals. *Yi Chuan*. **38,** 1069-1080.

Petersen J.L., Kalbfleisch T.S., Parris M., Tietze S.M. and Cruickshank J. (2020). MC1R and KIT haplotypes associate with pigmentation phenotypes of North American Yak (Bosgrunniens). *J. Hered.* **111,** 182-193.

Porto Neto L.R., Bunch R.J., Harrison B.E. and Barendse W. (2012). Variation in the *XKR4* gene was significantly associated with subcutaneous rump fat thickness in indicine and composite cattle. *Anim. Genet.* **43,** 785-789.

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A.R., Bender D., Maller J., Sklar P., de Bakker P.I.W., Daly M.J. and Sham P.C. (2007). PLINK: A toolset for whole-genome association and population-based linkage analysis. *Am. J. Hum. Genet.* **81,** 559-575.

Ron M. and Weller J.I. (2007). From QTL to QTN identification in livestock—Winning by points rather than knock-out. *Anim. Genet.* **38,** 429-439.

Sargolzaei M., Schenkel F.S., Jansen G.B. and Schaeffer L.R. (2008). Extent of linkage disequilibrium in Holstein cattle in North America. *Dairy Sci.* **91,** 2106-2117.

Simianer H., Ma Y. and Qanbari S. (2014). Statistical problems in livestock population genomics. Pp. 25-31 in Proc. 10th Congr. Genet. Appl. Livest. Prod. Vancouver, Canada.

Taye M., Lee W., Jeon S., Yoon J., Dessie T. and Hanotte O. (2017). Exploring evidence of positive selection signatures in cattle breeds selected for different traits. *Mamm Genome*. **28,** 528-541.

Tenesa A., Knott S.A., Ward D., Smith D., Williams J.L. and Visscher P.M. (2003). Estimation of linkage disequilibrium in a sample of the United Kingdom dairy cattle population using unphased genotypes. *Anim. Sci.* **81,** 617-623.

Troy C.S., MacHugh D.E., Bailey J.F., Magee D.A., Loftus R.T. and Cunningham P. (2001). Genetic evidence for Near-Eastern origins of European cattle. *Nature*. **410,** 1088-1091.

Upadhyay M.R., Chen W., Lenstra J.A., Goderie C.R.J., MacHugh D.E., Park S.D.E. and Crooijmans R.P.M.A. (2017). Genetic origin, ad- mixture and population history of aurochs (*Bos primigenius*) and prim-itive European cattle. *Heredity*. **118,** 1-8.

Weir B.S. and Cockerham C. (1983). Estimating F-statistics for the analysis of population structure. *Evolution*. **38,** 1358-1370.

Zhao F., McParland S. and Kearney F. (2015). Detection of selection signatures in dairy and beef cattle using high-density genomic information. *Genet. Sel. Evol.* **47,** 25-33.

Zhang S., Yao Z., Li X., Zhang Z., Liu X., Yang P., Chen N., Xia X., Lyu S., Shi Q., Wang E., Ru B., Jiang Y., Lei C., Chen H. and Huang Y. (2022). Assessing genomic diversity and signatures of selection in Pinan cattle using whole-genome sequencing data. *BMC Genom.* **23,** 2-10.

Zinovieva N.A., Dotsev A.V., Sermyagin A.A., Deniskova T.E., Abdelmanova A.S., Kharzinova V.R., Sölkner J., Reyer H.,Wimmers K. and Brem, G. (2020). Selection signatures in two oldest Russian native cattle breeds revealed using high-density singlenucleotide polymorphism analysis. *PLoS One*. **15,** e0242200.