Identifying Loci Under Positive Selection in Yellow Korean Cattle (Hanwoo)

Evolutionary Bioinformatics
Volume 15: 1–8

DOI: 10.1177/1176934319859001

Yi Li[1], Yun-Mi Lee[2], You-Sam Kim[2], Se-Pil Park[3] and Jong-Joo Kim[2]

[1]School of Statistics, Shanxi University of Finance and Economics, Taiyuan, China.
[2]Department of Biotechnology, Yeungnam University, Gyeongsan, Republic of Korea.
[3]Faculty of Biotechnology, Jeju National University, Jeju, Republic of Korea.

ABSTRACT: Jeju Black cattle is one of the aboriginal Korean cattle breeds that has been isolated in Jeju island for a long time, while Yellow Hanwoo cattle has been extensively selected for beef production traits for the last several decades. Aiming to investigate broader patterns of selection, we genotyped 352 Yellow Hanwoo and 169 Jeju Black cattle using a customized 150 K bovine chip. Our composite selection signals' analysis to identify selection signatures (cross-population extended haplotype homozygosity [XP-EHH], ΔSAF, and F_{ST}) identified recent and strong signature of selection near many loci with mutations affecting the traits under strong selection as outlier in Yellow Hanwoo, including SCP2 ($P = 8.41 \times 10^{-10}$) that may be involved in the meat quality. We found nine candidate regions with significant clusters of selection signals, and further bioinformatics analyses of the genes located within these regions revealed mainly genes involved in G-protein coupled receptor signaling pathway (GO:0007186) or olfactory transduction (bta04740), which may be due to adaptation to natural environments in Jeju island. Based on the stronger correlation of Ne_{10}/Ne_{100} ratio between Yellow Hanwoo (0.61) and Jeju Black (0.66) cattle, our results suggest that the difference of chromosomal regions of selection signature between the 2 cattle breeds was due to a consequence of selection processes to adapt to environmental differences between Jeju island and the main inland, Korean peninsula.

KEYWORDS: Yellow Hanwoo, Jeju black cattle, positive selection, effective population size

RECEIVED: May 5, 2019. **ACCEPTED:** May 23, 2019.

TYPE: Molecular Diversity and Conservation - Original Research

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by Program for the Outstanding Innovative Teams of Higher Learning Institutions of Shanxi (Yi Li), the National Natural Science Foundation of China (31501002; Yi Li), and the Fund for Shanxi "1331 Project" Key Innovative Research Team (Yi Li). Jong-Joo Kim was financially supported by a grant (715003-07) from the Research Center for Production Management and Technical Development for High-Quality Livestock Products through Agriculture, Food and Rural Affairs Research Center Program, Ministry of Agriculture, Food and Rural Affairs)

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHORS: Jong-Joo Kim, Department of Biotechnology, Yeungnam University, Gyeongsan 38541, Republic of Korea. Email: kimjj@ynu.ac.kr; Yi Li, School of Statistics, Shanxi University of Finance and Economics, Taiyuan 030006, China. Email: liyi@sxufe.edu.cn

Introduction

Recent evidence of selection on standing genetic variations is even more probable in domesticated cattle than in other natural outbred populations.[1,2] The bovine genome contains hundreds of regions whose patterns of genetic variation indicate adaptation for important traits, which might be due to strong selection pressures either by artificial selection programs or natural selection for adaptation in a specific environment.[3] These regions are broadly referred as signatures of selection, which has been focused on many recent studies with several cattle breeds. Their main goal was to unravel the causal mechanisms related to traits of interest in the genome.[4,5] The classic interpretation is that these genetic variations conferred differential adaptations to the cattle, causing many beneficial alleles of such genes to rapidly accumulate and fix within a population over time.[6]

Korean cattle are native taurine (*Bos primigenius taurus*) breeds that have been adapted to Korean peninsula long time ago.[7] In Korea, there are 3 native taurine-type cattle breeds, namely, Yellow Hanwoo, Chikso, and Jeju Black. These independent breeds mainly differ from each other in the coat colors and levels of nose darkness. Geographically, a large of Yellow Hanwoo live in the mainland while Chikso and Jeju Black exist on small areas on the mainland as well as 2 islands. While one of these breeds, Yellow Hanwoo, is a part of the national breeding and selection program for beef production purpose (carcass weight and marbling) and therefore undergoing artificial selection, the Chikso and Jeju Black are given little attention and adapted to the natural environment of their island. In the recent decades, the South Korean government took initiative for conservation and genetic improvement of Jeju Black in light of the need of improving the lineage and disease control measures, as well as raising awareness about the historical significance of the breed.[8–10]

With the rapid development of large-scale catalogs of genetic variation, uncovering the genetic footprints of this strong recent artificial selection could provide an insight into the mechanisms of selection in general and could, moreover, facilitate functional annotation of the genes related to important physiological and economical traits. The Korean government has selected Yellow Hanwoo for meat yield and quality using modern reproduction techniques; a completely different breeding history against Jeju Black remained isolated from the effects of artificial selection and will reveal the greatest selection pressure. Therefore, this study explores the selection signal by applying Randhawa composite selection signals (CSS)[11] statistics combining with all 3 single statistics (cross-population extended haplotype homozygosity [XP-EHH], ΔSAF, F_{ST}) using a custom-made single-nucleotide polymorphism (SNP) chip (147 792 SNPs) genotyping data of the Yellow Hanwoo and Jeju Black populations. Furthermore, we estimate the effective population size at the between-population level. Our study aimed to advance our understanding of mechanisms

underlying Yellow Hanwoo and elucidate the signatures of selection that have contributed to phenotypic appearances.

Materials and Methods

Studied samples, quality control, and population structure

A total of 352 Yellow Hanwoo and 169 Jeju Black cattle were genotyped using a customized 150K bovine chip, which was constructed by combining (1) the 55 491 SNPs that were already in Illumina BovineSNP50 BeadChip (version 3); (2) the 49 710 SNPs among the 648K SNPs in commercial Affymetrix bovine 648 K chips; (3) the 2603 SNPs in the hair or skin color genes including MC1R that were related to hair color specific in Hanwoo and Black cattle; (4) the 4665 SNPs in the candidate genes that were previously reported to be associated with economic traits in Hanwoo; (5) the 40 134 SNPs that were evenly positioned to cover the whole bovine chromosomes, which was selected after whole genome sequencing of 10 Jeju Black cattle sires. This array includes a total of 147 792 SNPs located throughout the genome with known mapping positions on the UMD 3.1 bovine reference genome,[12] which may alleviate ascertainment bias of genome analysis due to nonrandom distribution of SNPs across genomes, compared with the Illumina BovineSNP50 BeadChip. Genomic DNA purification and genotyping were accomplished by DNALink, a commercial genome analysis service provider in Korea. Single-nucleotide polymorphisms quality filtering was performed for each population using PLINK v1.9.[13] Only autosomal markers with unique genomic coordinates presenting call rates of at least 95%, minor allele frequency greater than 5%, and Fisher exact test for Hardy-Weinberg equilibrium greater than 0.001 were considered for linkage disequilibrium (LD) analyses. This left SNPs passing quality control (QC) across the breeds were then overlapped. To detect relatedness or duplicates in the dataset, individuals were further filtered using PLINK v1.9. Inferring the haplotype phase and imputing missing alleles were performed for the whole set of cattle populations with *BEAGLE* v5.0 using default options.[14]

Population stratification study was performed to assess the between breeds genetic differentiation using principal component (PC) analysis. Principal component uses SNP data generated by the R package *argyle*. The output files were graphically displayed using R script.

Estimation of effective population size

The prediction of Ne was calculated based on calculated LD, measured by r^2, using Corbin et al,[15] approach. The relevant formula is

$$N_{T(t)} = \left(4f\left(c_t\right)\right)^{-1}\left(E\left[r^2_{adj}|c_t\right]^{-1} - 1\right)$$

where N_T is the effective population size estimated t generations ago calculated as $(2f(c_t))^{-1}$,[16] r^2_{adj} is the LD value corrected for sample size, and c_t is the recombination rate using the physical distance between SNPs and corrected with the mapping functions, which is assumed as 1 Mb ≈ 1cM. *SNeP* v1.1 program was used to perform Ne estimation for each breed.[17]

Analysis of signatures of positive selection

Three single-test statistics (XP-EHH, ΔSAF, F_{ST}), which capture the augmentation in haplotype homozygosity, or the transformation in the allele frequency spectrum (ΔSAF), or the population differentiation (F_{ST}), were identified genome-wide selection signatures on Yellow Hanwoo cattle.

Haplotype-based method. Across population, EHH (XP-EHH) analysis was used to compare the extent of haplotype homozygosity between Yellow Hanwoo with Jeju Black cattle using the software *rehh* v2.0.[18] The integrated site-specific EHH of each SNP in each population was computed based on Sabeti et al.[19] The natural logarithm of the ratio between *iES*pop1 and iES-pop2 was used to compute the unstandardized XP-EHH score, which was then normalized by using median and standard deviation values. One-sided P values were derived as $-\log_{10}(1\text{-}2|\varphi[\text{XP-EHH}]|-0.5)$, where φ(XP-EHH) denotes the Gaussian distribution function.

Allele frequency spectrum-based method. ΔSAF was calculated as the directional change in the selected allele frequency between populations, so that $\Delta SAF = f_{Ai} - f_{Aj}$, where f_{Ai} is the major allele frequency in the assumed selected Yellow Hanwoo i; likewise, f_{Aj} is the A allele frequency in the control Jeju Black j. ΔSAF scores were standardized to $Z \sim N(0,1)$. The ΔSAF statistic[11] is similar to ΔDAF[20] except that the information about the ancestral and derived status of alleles at the focal SNP is not needed. Randhawa et al[11] reported on ΔDAF and ΔSAF and have found a very close relation ($r^2 > 0.8$) in cattle data with genome-wide population analysis.

Allelic frequency differentiation-based methods. The fixation index (F_{ST}) is a widely used approach for detecting highly differentiated loci between populations. The F_{ST} employed the unbiased estimator of Weir and Cockerham[21] using PLINK v1.9 program. Smaragdov et al[22] found F_{ST} assessment obtained using PLINK program appears to be unbiased.

Composite selection signals

As all applied methodologies (XP-EHH, ΔSAF, F_{ST}) are confined to test a small limit of selective hypotheses, a combination of multiple statistics is to detect sites bring into correspondence with rejection of the neutral locus among the different approaches and can be a robust test in detecting causal variants of positive selection.[23] Therefore, identifying such selection signals will be expected to be enrichment for highly frequent long-range haplotypes, derived alleles, and population

differentiation. The following process was used to compute CSS from integrating the 3 single test using Randhawa et al[11] method. For each test statistic i calculated at SNPj, to get the order of each observed test all the SNPs, the respective R_{ij} was transformed to a fractional ranking by scaling them into the interval between 0 and 1, R'_{ij}. Next, the fractional rank was converted into a Z-score by $Z_{ij} = \phi^{-1}(R'_{ij})$. Then, the mean of z-values was computed at each SNP position, $\bar{Z}_j$, and then, P values were calculated as $p = 1 - \Phi(m^{1/2}\bar{Z}_j)$. A consistent signal among the 3 tests would be regarded as an extreme P value at the certain sites, and each test can reject the common null hypothesis. The CSS with Bonferroni-adjusted P value $<.05$ on genome-wide level was considered significant.

Genomic regions and genes under selection

The identification of locus was performed by the selection pressure on neighboring alleles owing to hitchhiking; thus, these loci are detected to get together.[24] To reduce the noise from single tests to cause false-positives results, scan statistic was used to identify the selective regions harboring a greater than expected number of SNPs. Different from the moving window approaches, scan statistics overcome the weakness include subjective selection of the moving window size and arise naturally detecting clusters of events. In this study, the ChromoScan software (version 1.0)[25] was used to obtain statistical evidence of both a clustering of the SNP locations and clustering of extreme CSS P values for each chromosome. Consecutive clusters separated by less than 1 Mb were combined into a single cluster.

The selected regions of genomic features were then assessed on the Ensembl Genes 81 using software BioMart.[26] Genes were mapped to each genomic region under selection using its gene start and stop positions based on their Ensembl ID and associated gene name. Enrichment of genes (Ensembl gene ID) on biological processes and molecular pathways were then detected based on DAVID software 6.8.[27] Only pathways or annotations with Bonferroni-adjusted $P < 10^{-2}$ were retained.

Known Quantitative Trait Loci (QTLs) that overlapped with each identified selective regions were obtained from the CattleQTL release 35 from the Animal Quantitative Trait Loci Database. Trait QTL were regarded as a trait class of "Meat and Carcass," "Production," "Health," and "Reproduction," which consisted of 64 515 QTLs.[28]

Results and Discussion

Data description

After SNP QC and filtering across the Yellow Hanwoo and Jeju Black populations, there were 92 922 autosomal SNPs from 260 Yellow Hanwoo cattles and 79 849 autosomal SNPs from 132 Jeju Black cattles. The mean of estimated genomic inbreeding coefficients were 0.02 ± 0.02 for Yellow Hanwoo (ranging from 0.0001 to 0.1529) and 0.06 ± 0.04 for Jeju Black (ranging from 0.0007 to 0.2577). Overlaps of the 2 SNP data sets obtained the final set of 73 489 markers for further analysis. The inter-marker distance mean was 34.05 kb, overlapping the average interval of 70.57 kb reported on the BovineSNP50K BeadChip.[29] These results suggested that the SNPs coverage would be satisfactory.

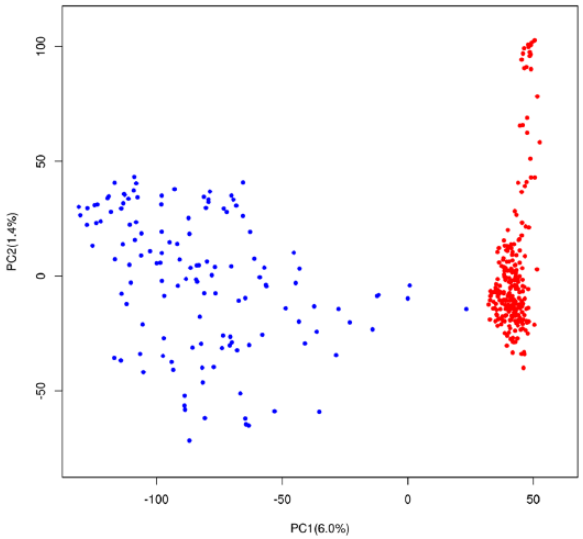

Figure 1. Principal component analysis between Hanwoo (red circles) and Jeju Black (blue circles) cattle breeds, 392 animals plotted.

As shown in Figure 1, PC analysis was performed. The first PC tended to separate Jeju Black from Hanwoo and explained 6% of the total variance. The genetic evidence further confirmed that Hanwoo cattle were traditionally considered to be distinct from Jeju Black cattle and shown some level of genetic difference in accordance with their geographic origin.[9,30] Based on PC analysis results, all 392 animals from the 2 populations were used in the subsequent analysis.

Historical effective population size

The Ne trends of Yellow Hanwoo and Jeju Black cattle were calculated for an SNP interval of 0.05 to 50 Mb corresponding to the last 1000 generations. The "demographic fingerprints"[17] estimated by Ne of the 2 cattle populations are shown in Figure 2. Values of Ne for both populations showed decreasing trend over the last 1000 generations, with an increasingly steeper slope since approximately 100 generations ago but a special period of increase from 950 to 850 generations ago (Figure 2), which is similar to the past Ne trends of other cattle breeds (Kerry, Belted Galloway, Angus, Jersey, and Holstein) reported by Browett et al.[31] These concordant patterns suggest that these cattle breeds might share similar demographic histories and that the evolution of these cattle was strongly affected by severe climate exchange. The estimated effective population size in the most recent generation varied from 86 for Jeju Black cattle to 374 for Yellow Hanwoo cattle. The Ne of Hanwoo was consistent with previous findings of Li and Kim.[29] To better understand different histories of migration, genetic drift, and artificial selection on the formation of cattle breeds, both N_{e10}/N_{e100} ratio and the N_{e100}/N_{e1000} ratio were estimated in

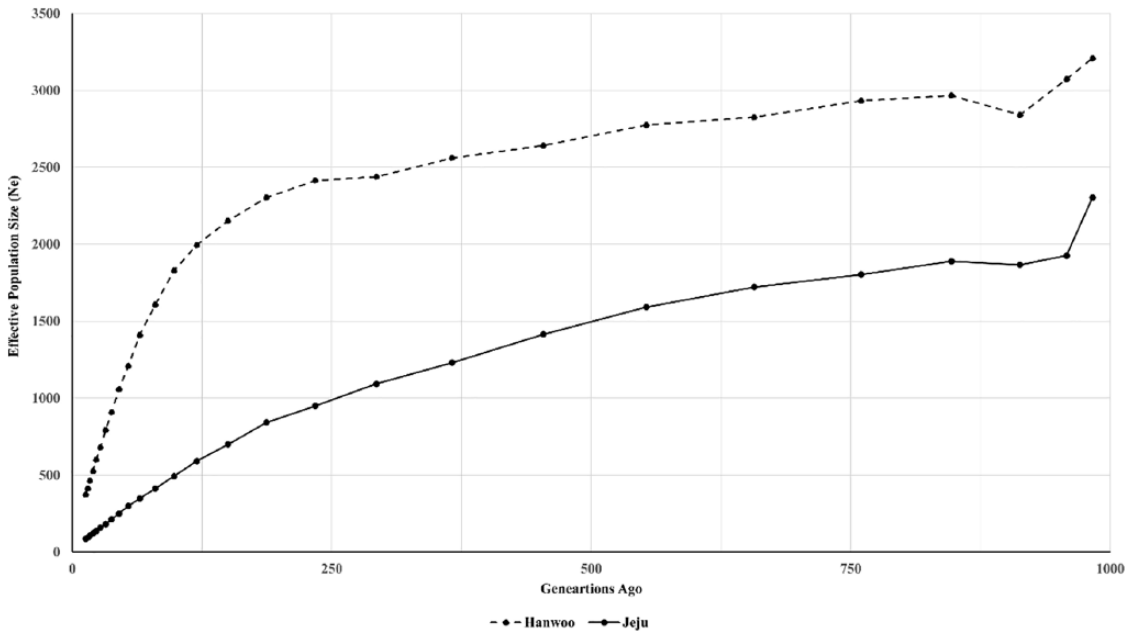

Figure 2. Genetic effective population size trends for Hanwoo and Jeju Black cattle in the last 1000 generations.

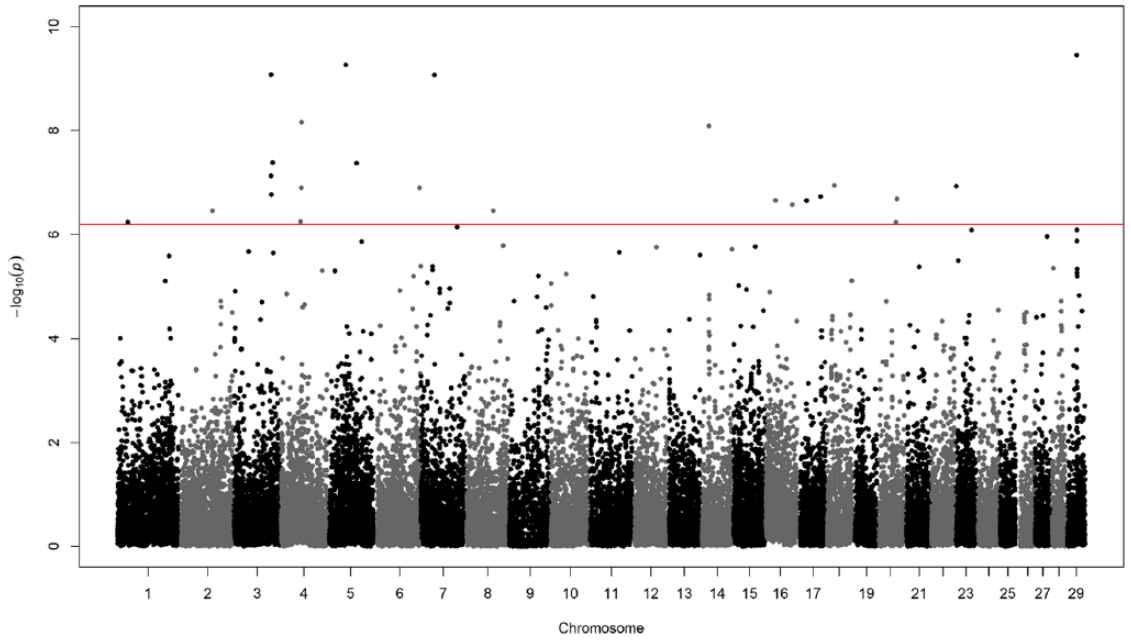

Figure 3. Manhattan plot of genome-wide autosomal composite selection signal (CSS) analyses of Hanwoo cattle. Red line in the CSS plots represents the genome-wide 0.5% thresholds based on the Bonferroni threshold correction.

each breed. Generally, N_{e10}/N_{e100}, representing the recent decline in Ne from ~100 to ~10 generations ago, shows the breed formation and artificial breeding period. Strucken et al[30] stated that Korea settle a cattle breeding program in the 1930s, whereas the N_{e100}/N_{e1000} ratio denoted the past decline in Ne from ~1000 to ~100 generations ago, which indicates an intensification of population isolation.[32] Strucken et al[30] reported that progenitors of Jeju Black settled the island 1100-2000 years ago. The stronger correlation with N_{e10}/N_{e100} ratio between Yellow Hanwoo (0.61) and Jeju Black (0.66) cattle may indicate a similar effect of artificial selection in each breed. In contrast, the N_{e100}/N_{e1000} ratio for Yellow Hanwoo (0.21) is approximately 3 times lower than Jeju Black (0.57), which may correspond to environmental differences in these breeds. These findings indicate the impact of environmental changes like bottlenecks, migration, and introgression, rather than artificial selection pressures on the selection signatures signals, owing to the different history of the cattle breeds detected here.

Genome-wide signatures of selection in the Yellow Hanwoo

The CSS findings using genome-wide scan results in the Yellow Hanwoo are illustrated in Figure 3; 24 distinct SNP crossing

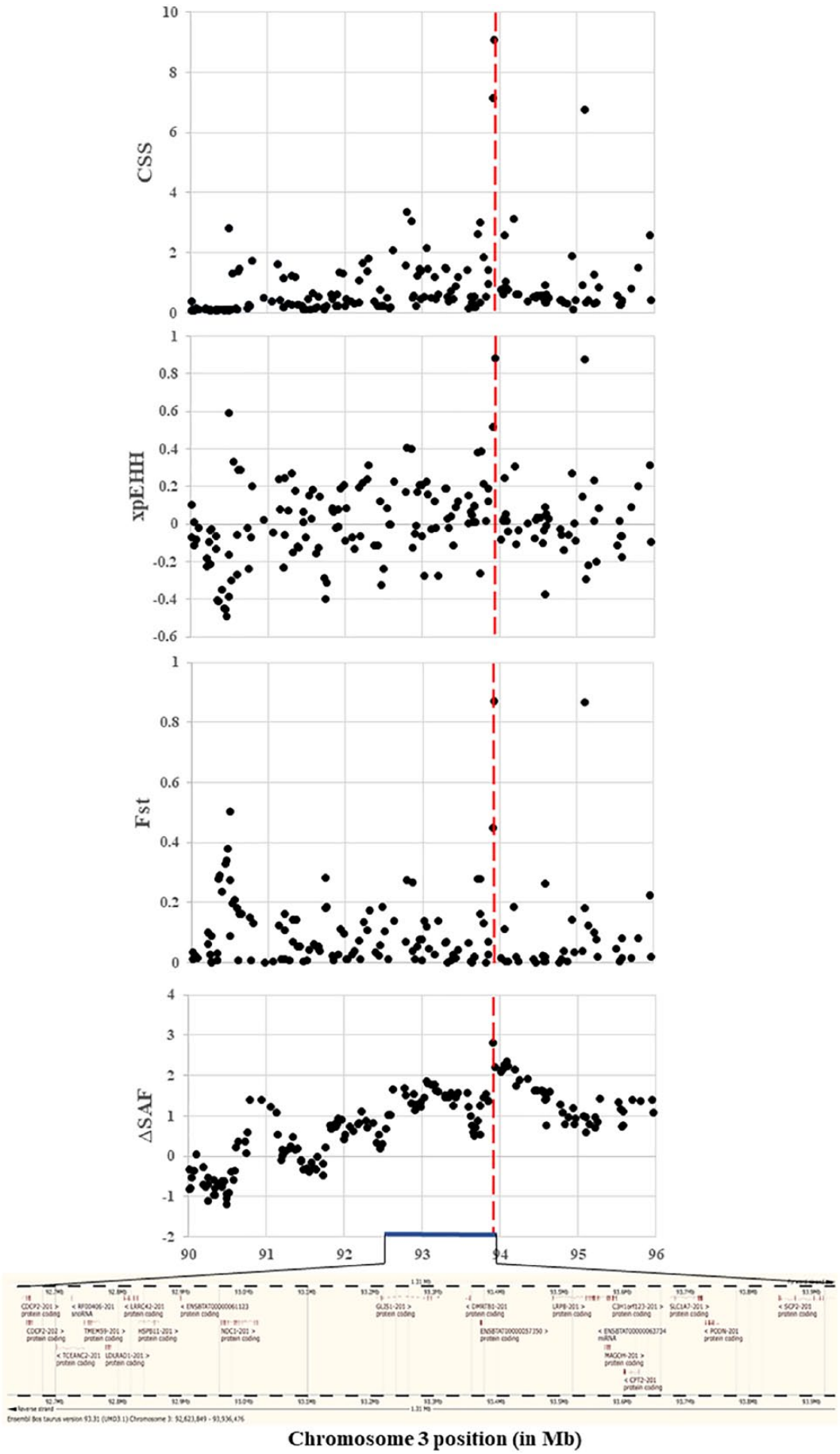

Figure 4. Composite selection signals, XP-EHH, ΔSAF, and F_{ST} in Hanwoo. Vertical dashed red lines represent the position of the SNP detected as highly significant in Hanwoo (Chr3: 93936476 bp, $P = 8.41 \times 10^{-10}$). The horizontal blue line marks the selective region (Chr3: 92.62-93.94 Mb).

Table 1. Genomic regions under selection in Yello Hanwoo cattle.

SELECTIVE SWEEPS	CHR	START (MB)	END (MB)	HIGHER CSS	SNP	TOTAL GENES	TRAITS	REFERENCE
1	3	92.62	93.94	9.07	43	20	FA-C14:1; SCS; LUTACT; CONCRATE; DPR; PUBAGE; NM; PL; CALEASE; SB	Lemos et al,[34] Cole et al,[35] Hawken et al,[36] Tenghe et al,[37] Duran Aguilar et al,[38] and Parker Gaddis et al[39]
2	4	51.69	52.24	8.17	23	5	PELAR; ADG	Michenet et al[40] and Peters et al[41]
3	5	45.67	45.96	9.27	11	4	EEF; FATTH; INHIB; SCRCIR	Peters et al,[41] Fortes et al,[42] Buzanskas et al[43] and Psaros et al[44]
4	6	112.33	113.54	6.91	33	3	RFI; CALEASE; NM; SB	Seabury et al[45] and Cole et al,[35]
5	14	21.05	23.02	8.09	60	13	FATTH; YGRADE; YIELD; CWT; MARBL; FATWT; SCFR; PL; BW; BTBS; ADG; TWIN; AGEFC; TICKR; IGFI; CALEST; PUBAGE; CALEASE; NM; SB; SCRCIR; W365	de Oliveira Silva et al,[46] Magalhaes et al,[47] Rempel et al,[48] Ali et al,[49] Saatchi et al,[50] Nayeri et al,[51] Snelling et al,[52] Richardson et al,[53] Sorbolini et al,[54] Mota et al,[55] Mapholi et al,[56] Hawken et al,[36] Fortes et al,[57] Cole et al,[35] Martinez et al,[58] Yurchenko et al,[59] and Michenet et al[40]
6	17	23.29	25.05	6.65	66	5	FA-C18:3; CALEST	Cesar et al[60] and Hawken et al[36]
7	17	58.26	59.24	6.73	26	5	FATTH; CANCIR; RFI; ADG	Santana et al,[61] Sorbolini et al,[54] Seabury et al,[45] and Olivieri et al[62]
8	18	19.49	20.10	6.95	15	2	FA-n3; SCRCIR; W365; LACTPER	Saatchi et al,[63] Buzanskas et al,[43] and Pryce et al[64]
9	29	27.10	29.77	9.45	110	81	SUBFAT; FATTH; W365; BWG; SCS; CALEST	Li and Kim,[65] Nalaila et al,[66] Aguilar et al,[38] Wang et al,[67] and Hawken et al[36]

Abbreviations: CSS, composite selection signals; SNP, single-nucleotide polymorphism.
Meat and carcass traits: (1) FA-C14:1, myristoleic acid content; (2) PELAR, pelvic area; (3) EEF, intramuscular fat; (4) FATTH, fat thickness at the 12th rib; (5) SF, shear force; (6) FA-C18:3, linolenic acid content; (7) FA-n3, omega-3 unsaturated fatty acid content; (8) SUBFAT, subcutaneous fat; (9) YGRADE, yield grade; (10) YIELD, retail product yield; (11) CWT, carcass weight; (12) MARBL, marbling score; (13) FATWT, fat weight; (14) SCFR, subcutaneous rump fat thickness.
Production traits: (1) ADG, average daily gain; (2) NM, net merit; (3) PL, length of productive life; (4) RFI, residual feed intake; (5) BW, body weight (birth); (6) W365, body weight (yearling); (7) CANCIR, cannon bone circumference; (8) LACTPER, lactation persistency; (9) BWG, body weight gain.
Health traits: (1) SCS, somatic cell score; (2) BTBS, bovine tuberculosis susceptibility; (3) TICKR, tick resistance; (4) IGFI, insulin-like growth factor 1 level.
Reproductive traits: (1) LUTACT, luteal activity; (2) CONCRATE, conception rate; (3) DPR, daughter pregnancy rate; (4) PUBAGE, age at puberty; (5) CALEASE, calving ease; (6) SB, still birth; (7) INHIB, inhibin level; (8) SCRCIR, scrotal circumference; (9) PREGRATE, pregnancy rate; (10) TWIN, twinning; (11) AGEFC, age at first calving; (12) CALEST, interval to first estrus after calving.

the genome-wide significance threshold ($P < 6.28 \times 10^{-7}$) were detected on 15 *Bos Taurus* chromosomes (ie, BTA 1, 2, 3, 4, 5, 6, 7, 8, 14, 16, 17, 18, 20, 23, and 29). The most significant SNP was found ($P = 3.52 \times 10^{-10}$) in transmembrane protein 225 (TMEM225, ENSBTAG00000033502) gene, which is located at position BTA29: 27749740. Supplementary Table S1 provides detailed information for the selective signals. One gene detected by 2 significant SNPs is a sterol carrier protein (SCP2, ENSBTAG00000003746), which is known to be an intracellular protein potentially including lipid transfer in organs involved in lipid metabolism, including mammary tissue (Figure 4). Qin et al[33] reported that the *SCP2* gene may participate in the regulation and control of intramuscular fatty acid metabolism in cattle. Because intramuscular fat content is a key factor affecting meat quality, intramuscular fat content is the most important breeding goals in the national Hanwoo breeding progam.[4,8] Consequently, selective breeding might lead to some genomic footprint, which could show up as the CSS value. Unsurprisingly, SCP2 was detected to overlap with QTL for myristoleic acid content.[34] The detailed functions of genes in the positive selection are outside the scope of this study, notwithstanding our study present availability of genomic information in the further studies to elucidate genetic mechanisms underlying economical traits in cattle.

In the interest of explaining the CSS results from different tests, a selective regional (Chr3: 92.62-93.94 Mb) plot of *P* values for each individual test in Yellow Hanwoo is present in Figure 4. On the same position, 3 extra figures were performed: XP-EHH, F_{ST}, and ΔSAF analyses. The selective signals that overlap among the different approaches could be the convincing result that the site has signatures of selection. Besides, this study illustrates the relevance of combining signals, which might improve the resolution of quite extended signatures and the statistical power for identifying footprints of selection.[11,20] To further identify signals displaying strong footprints of selection, scan statistics were used to identify chromosomal selective regions harboring a greater than expected number of signals.[24] A total of nine statistically significant selective regions were

Table 2. Summary of the results from the enrichment analysis.

CATEGORY[A]	ID	ENRICHED TERM	ENSEMBL GENE COUNT	CORRECTED *P* VALUE[B]	ENRICHMENT
GO term	GO:0007186	G-protein-coupled receptor signaling pathway	35	1.46×10^{-19}	7.8
KEGG	bta04740	Olfactory transduction	39	5.25×10^{-18}	4.8

Abbreviations: GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.
[a]The GO and KEGG analysis performed with DAVID 6.8.
[b]*P* values are Bonferroni-corrected *P* values $\leqslant 10^{-2}$.

detected (Table 1 and Table S2), which led to a list of 138 candidate genes. These genes may be an outcome of selection that is closely related to traits of interest, presenting availability of genomic information to be applied on examination of the phenotypic differences between Yellow Hanwoo and Jeju Black. To test this further, the CattleQTL database was employed and a total of 339 QTL intersecting with the selective regions advanced the understanding of the mechanisms underlying Yellow Hanwoo characteristics. These QTLs are related to several traits, for instance, meat and carcass (13 traits), production (9 traits), health (4 traits), and reproduction (12 traits). Interestingly, all the selective regions detected here overlap with the QTL identified in previously published genetic study for cattle breeds (Table 1). This may be an outcome of selection processes to encourage adaptation to environmental change, which is coincident with the aims of the formation of the Yellow Hanwoo.

For providing a whole view of the types of genes or pathways that were performed among the 138 candidate Ensembl Genes (Table S2), the DAVID tool was used for their Gene Ontology (GO) categories and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Table 2). The KEGG pathway study found a major enrichment of genes (eg, OR10D3, OR8D2, OR8D4, and OR8D8) gathered in sensory perception (bta04740, olfactory transduction; Bonferroni-corrected $P < .01$; Table 1). The olfactory receptor interacts with odor molecules in the nose, initiating a neuronal response that triggers olfactory perception. For cattle, olfactory receptors are necessary to forage and to avoid eating toxic food.[68,69] Jeju Black were formatted in the absence of food and graze in the field. Yellow Hanwoo, however, have been artificially selected for intensive production. The cause of olfactory genes might therefore to adapt to these environments.

The GO study detected a significant over-representation of genes list involved in the G-protein-coupled receptor signaling pathway (GO:0007186). G-protein-coupled receptors (GPCR), also referred as free fatty acid receptors (FFAR), potentially affect immune response in cattle.[70] Moreover, we found SNP frequency differences between Yellow Hanwoo and Black Jeju cattle, which might be the reasonable cause to the superior resistance to metabolic disorders via an immune system response. Thus, these gene sets represent availability of genomic information to further study how these genes are genetically related to immune response in Yellow Hanwoo.

Acknowledgements

The authors gratefully acknowledge Imtiaz Randhawa for providing the R script of composite selection signals method.

Author Contributions

YL and JJK conceived and designed the experiment; YL, YML, and YSK analyzed the data; YL and JJK drafted and polished the paper; SPP provided the genotype data and samples of Jeju Black cattle. All contributing authors have reviewed the manuscript and approved the submitting of final manuscript.

Supplemental Material

Supplemental material for this article is available online.

REFERENCES

1. Innan H, Kim Y. Pattern of polymorphism after strong artificial selection in a domestication event. *Proc Natl Acad Sci U S A*. 2004;101:10667-10672.
2. Rothammer S, Seichter D, Forster M, Medugorac I. A genome-wide scan for signatures of differential artificial selection in ten cattle breeds. *BMC Genomics*. 2013;14:908.
3. Utsunomiya YT, Perez O'Brien AM, Sonstegard TS, Solkner J, Garcia JF. Genomic data as the "hitchhiker's guide" to cattle adaptation: tracking the milestones of past selection in the bovine genome. *Front Genet*. 2015;6:36.
4. Lee HJ, Kim J, Lee T, et al. Deciphering the genetic blueprint behind Holstein milk proteins and production. *Genome Biol Evol*. 2014;6:1366–1374.
5. Chen NB, Cai YD, Chen QM, et al. Whole-genome resequencing reveals world-wide ancestry and adaptive introgression events of domesticated cattle in East Asia. *Nat Commun*. 2018;9:2337.
6. Boero F. From Darwin's origin of species toward a theory of natural history. *F1000Prime Rep*. 2015;7:49.
7. Lee SW, Park BH, Sharma A, et al. Hanwoo cattle: origin, domestication, breeding strategies and genomic selection. *J Anim Sci Technol*. 2014;56:2.
8. Chung KY, Lee SH, Cho SH, Kwon EG, Lee JH. Current situation and future prospects for beef production in South Korea—a review. *Asian-Australas J Anim Sci*. 2018;31:951–960.
9. Kim S, Cheong HS, Shin HD, et al. Genetic diversity and divergence among Korean cattle breeds assessed using a BovineHD single-nucleotide polymorphism chip. *Asian-Australas J Anim Sci*. 2018;31:1691–1699.
10. Sharma A, Lee SH, Lim D, Chai HH, Choi BH, Cho YM. A genome-wide assessment of genetic diversity and population structure of Korean native cattle breeds (vol 17, 139, 2016). *Bmc Genet*. 2017;18:12.
11. Randhawa IAS, Khatkar MS, Thomson PC, Raadsma HW. Composite selection signals can localize the trait specific genomic regions in multi-breed populations of cattle and sheep. *BMC Genet*. 2014;15:34.
12. Elsik CG, Tellam RL, Worley KC, et al. The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science*. 2009;324:522–528.
13. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559–575.
14. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet*. 2007;81:1084–1097.
15. Corbin LJ, Liu AY, Bishop SC, Woolliams JA. Estimation of historical effective population size using linkage disequilibria with marker data. *J Anim Breed Genet*. 2012;129:257–270.

16. Hayes BJ, Visscher PM, McPartlan HC, Goddard ME. Novel multilocus measure of linkage disequilibrium to estimate past effective population size. *Genome Res*. 2003;13:635–643.
17. Barbato M, Orozco-terWengel P, Tapio M, Bruford MW. SNeP: a tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. *Front Genet*. 2015;6:109.
18. Gautier M, Klassmann A, Vitalis R. REHH 2.0: a reimplementation of the R package REHH to detect positive selection from haplotype structure. *Mol Ecol Resour*. 2017;17:78–90.
19. Sabeti PC, Varilly P, Fry B, et al. Genome-wide detection and characterization of positive selection in human populations. *Nature*. 2007;449:913.
20. Grossman SR, Shlyakhter I, Karlsson EK, et al. A composite of multiple signals distinguishes causal variants in regions of positive selection. *Science*. 2010;327:883–886.
21. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population-structure. *Evolution*. 1984;38:1358–1370.
22. Smaragdov MG, Saksa EI, Kudinov AA, Dement'Eva NV, Mitrofanova OV, Plemyashov KV. Genome-wide analysis of across herd Fst heterogeneity in holsteinized cattle. *Russian J Genet*. 2016;52:173–179.
23. Lin K, Li H, Schlotterer C, Futschik A. Distinguishing positive selection from neutral evolution: boosting the performance of summary statistics. *Genetics*. 2011;187:229–244.
24. Hoh J, Ott J. Scan statistics to scan markers for susceptibility genes. *Proc Natl Acad Sci U S A*. 2000;97:9615–9617.
25. Sun YV, Jacobsen DM, Kardia SLR. ChromoScan: a scan statistic application for identifying chromosomal regions in genomic studies. *Bioinformatics*. 2006;22:2945–2947.
26. Kinsella RJ, Kahari A, Haider S, et al. Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database (Oxford)*. 2011;2011:bar030.
27. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4:44–57.
28. Hu ZL, Park CA, Reecy JM. Developmental progress and current status of the Animal QTLdb. *Nucleic Acids Res*. 2016;44:D827–D833.
29. Li Y, Kim JJ. Effective population size and signatures of selection using bovine 50K SNP chips in Korean native cattle (Hanwoo). *Evol Bioinform Online*. 2015;11:143–153.
30. Strucken EM, Lee SH, Jang GW, Porto-Neto LR, Gondro C. Towards breed formation by island model divergence in Korean cattle. *BMC Evol Biol*. 2015;15:284.
31. Browett S, McHugo G, Richardson IW, et al. Genomic characterisation of the indigenous Irish Kerry cattle breed. *Front Genet*. 2018;9:51.
32. de Roos AP, Hayes BJ, Spelman RJ, Goddard ME. Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus cattle. *Genetics*. 2008;179:1503–1512.
33. Qin W, Liang CN, Guo X, et al. PPARα signal pathway gene expression is associated with fatty acid content in yak and cattle longissimus dorsi muscle. *Genet Mol Res*. 2015;14:14469–14478.
34. Lemos MVA, Chiaia HLJ, Berton MP, et al. Genome-wide association between single nucleotide polymorphisms with beef fatty acid profile in Nellore cattle using the single step procedure. *BMC Genomics*. 2016;17:213.
35. Cole JB, Wiggans GR, Ma L, et al. Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary US Holstein cows. *Bmc Genomics*. 2011;12:408.
36. Hawken RJ, Zhang YD, Fortes MRS, et al. Genome-wide association studies of female reproduction in tropically adapted beef cattle. *J Anim Sci*. 2012;90:1398–1410.
37. Tenghe AMM, Bouwman AC, Berglund B, Strandberg E, de Koning DJ, Veerkamp RF. Genome-wide association study for endocrine fertility traits using single nucleotide polymorphism arrays and sequence variants in dairy cattle. *J Dairy Sci*. 2016;99:5470–5485.
38. Duran Aguilar M, Roman Ponce SI, Ruiz Lopez FJ, et al. Genome-wide association study for milk somatic cell score in Holstein cattle using copy number variation as markers. *J Anim Breed Genet*. 2017;134:49–59.
39. Parker Gaddis KL, Null DJ, Cole JB. Explorations in genome-wide association studies and network analyses with dairy cattle fertility traits. *J Dairy Sci*. 2016;99:6420–6435.
40. Michenet A, Barbat M, Saintilan R, Venot E, Phocas F. Detection of quantitative trait loci for maternal traits using high-density genotypes of Blonde d'Aquitaine beef cattle. *BMC Genet*. 2016;17:88.
41. Peters SO, Kizilkaya K, Garrick DJ, et al. Bayesian genome-wide association analysis of growth and yearling ultrasound measures of carcass traits in Brangus heifers. *J Anim Sci*. 2012;90:3398–3409.
42. Fortes MRS, Reverter A, Kelly M, McCulloch R, Lehnert SA. Genome-wide association study for inhibin, luteinizing hormone, insulin-like growth factor 1, testicular size and semen traits in bovine species. *Andrology*. 2013;1:644–650.
43. Buzanskas ME, Grossi DD, Ventura RV, et al. Candidate genes for male and female reproductive traits in Canchim beef cattle. *J Anim Sci Biotechnol*. 2017;8:67.
44. Psaros KM, McDaneld TG, Kuehn LA, Snelling WM, Keele JW. Evaluation of single nucleotide polymorphisms in chromosomal regions impacting pregnancy status in cattle. *J Anim Sci*. 2015;93:978–987.
45. Seabury CM, Oldeschulte DL, Saatchi M, et al. Genome-wide association study for feed efficiency and growth traits in US beef cattle. *Bmc Genomics*. 2017;18:386.
46. de Oliveira Silva RM, Bonvino Stafuzza N, de Oliveira Fragomeni B, et al. Genome-wide association study for carcass traits in an experimental Nelore cattle population. *PLoS ONE*. 2017;12:e0169860.
47. Magalhaes AF, de Camargo GM, Fernandes GA Junior, et al. Genome-wide association study of meat quality traits in Nellore cattle. *PLoS ONE*. 2016;11:e0157845.
48. Rempel LA, Casas E, Shackelford SD, Wheeler TL. Relationship of polymorphisms within metabolic genes and carcass traits in crossbred beef cattle. *J Anim Sci*. 2012;90:1311.
49. Ali AA, Khatkar MS, Kadarmideen HN, Thomson PC. Additive and epistatic genome-wide association for growth and ultrasound scan measures of carcass-related traits in Brahman cattle. *J Anim Breed Genet*. 2015;132:187–197.
50. Saatchi M, Schnabel RD, Taylor JF, Garrick DJ. Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. *BMC Genomics*. 2014;15:442.
51. Nayeri S, Sargolzaei M, Abo-Ismail MK, et al. Genome-wide association study for lactation persistency, female fertility, longevity, and lifetime profit index traits in Holstein dairy cattle. *J Dairy Sci*. 2017;100:1246–1258.
52. Snelling WM, Allan MF, Keele JW, et al. Genome-wide association study of growth in crossbred beef cattle. *J Anim Sci*. 2010;88:837–848.
53. Richardson IW, Berry DP, Wiencko HL, et al. A genome-wide association study for genetic susceptibility to Mycobacterium bovis infection in dairy cattle identifies a susceptibility QTL on chromosome 23. *Genet Sel Evol*. 2016;48:19.
54. Sorbolini S, Bongiorni S, Cellesi M, et al. Genome wide association study on beef production traits in Marchigiana cattle breed. *J Anim Breed Genet*. 2017;134:43–48.
55. Mota RR, Guimaraes SEF, Fortes MRS, et al. Genome-wide association study and annotating candidate gene networks affecting age at first calving in Nellore cattle. *J Anim Breed Genet*. 2017;134:484–492.
56. Mapholi NO, Maiwashe A, Matika O, et al. Genome-wide association study of tick resistance in South African Nguni cattle. *Ticks Tick Borne Dis*. 2016;7:487–497.
57. Fortes MRS, Reverter A, Hawken RJ, Bolormaa S, Lehnert SA. Candidate genes associated with testicular development, sperm quality, and hormone levels of inhibin, luteinizing hormone, and insulin-like growth factor 1 in Brahman bulls. *Biol Reprod*. 2012;87:58.
58. Martinez R, Gomez Y, Rocha JFM. Genome-wide association study on growth traits in Colombian creole breeds and crossbreeds with Zebu cattle. *Genet Mol Res*. 2014;13:6420–6432.
59. Yurchenko AA, Daetwyler HD, Yudin N, et al. Scans for signatures of selection in Russian cattle breed genomes reveal new candidate genes for environmental adaptation and acclimation. *Sci Rep*. 2018;8:12984.
60. Cesar ASM, Regitano LCA, Mourao GB, et al. Genome-wide association study for intramuscular fat deposition and composition in Nellore cattle. *BMC Genet*. 2014;15.
61. Santana MHA, Ventura RV, Utsunomiya YT, et al. A genome wide association mapping study using ultrasound-scanned information identifies potential genomic regions and candidate genes affecting carcass traits in Nellore cattle. *J Anim Breed Genet*. 2015;132:420–427.
62. Olivieri BF, Mercadante MEZ, Cyrillo JNDG, et al. Genomic regions associated with feed efficiency indicator traits in an experimental Nellore cattle population. *PLoS ONE*. 2016;11:e0164390.
63. Saatchi M, Garrick DJ, Tait RG Jr, et al. Genome-wide association and prediction of direct genomic breeding values for composition of fatty acids in Angus beef cattle. *BMC Genomics*. 2013;14:730.
64. Pryce JE, Haile-Mariam M, Verbyla K, Bowman PJ, Goddard ME, Hayes BJ. Genetic markers for lactation persistency in primiparous Australian dairy cows. *J Dairy Sci*. 2010;93:2202–2214.
65. Li Y, Kim JJ. Multiple linkage disequilibrium mapping methods to validate additive quantitative trait loci in Korean native cattle (Hanwoo). *Asian-Australas J Anim Sci*. 2015;28:926–935.
66. Nalaila SM, Stothard P, Moore SS, Li C, Wang Z. Whole-genome QTL scan for ultrasound and carcass merit traits in beef cattle using Bayesian shrinkage method. *J Anim Breed Genet*. 2012;129:107–119.
67. Wang X, Ma PP, Liu JF, et al. Genome-wide association study in Chinese Holstein cows reveal two candidate genes for somatic cell score as an indicator for mastitis susceptibility. *BMC Genet*. 2015;16:111.
68. Mombaerts P. Genes and ligands for odorant, vomeronasal and taste receptors. *Nat Rev Neurosci*. 2004;5:263–278.
69. Niimura Y. Evolutionary dynamics of olfactory receptor genes in chordates: interaction between environments and genomic contents. *Hum Genomics*. 2009;4:107–118.
70. Agrawal A, Alharthi A, Vailati-Riboni M, Zhou Z, Loor JJ. Expression of fatty acid sensing G-protein coupled receptors in peripartal Holstein cows. *J Anim Sci Biotechnol*. 2017;8:20.