

Identification of selection footprints in a brown Hanwoo (Korean cattle) population for production traits

**Dajeong Lim¹, Seung Hwan Lee², Cedric Gondro³, Bong Hwan Choi¹, Han Ha Chai¹, Eva Strucken³, Yong Min Cho¹,
Gul Won Jang¹, Seong Koo Hong¹, Tae-Hun Kim¹, Seung Hwan Lee³**

¹Division of Animal Genomics & Bioinformatics, ²Hanwoo Experiment Station, National Institute of Animal Science, RDA, Republic of Korea, ³School of Environmental and Rural Science, University of New England, Armidale, NSW Australia

ABSTRACT: Korean cattle (Hanwoo) are categorized into three types of breeds according to their color, brown, brindle and black. Among the three breeds, the brown Hanwoo has been subjected to intensive artificial selection over the past seventy years to improve meat production traits such as carcass weight and marbling. In this study, Rsb method was applied to identify recent selection signals for production traits in brown Hanwoo.

To identify traces driven by the artificial selection process within the Hanwoo breeding program, we used 3 types (breeds) of Hanwoo (brown Hanwoo; HW, brindle Hanwoo; BR, and Jeju black Hanwoo; JB) and Chinese Yanbian cattle (YB). The HW population was under selection pressure in the Hanwoo breeding program whilst BR, JB, and YB were unselected. We investigated each of the four comparisons (HW/BR, HW/JB, HW/YB cattle and HW/all three breeds; UN). As a result, we obtained 41, 38, 105, 87 possible candidate regions by the four comparisons, respectively. Finally, we identified 31 candidate regions that seem to be recently artificially selected in brown Hanwoo compared to other breeds. Moreover, 18 significant SNPs were detected in all four comparisons, 16 genes were located in near the significant the SNPs such as CAPN1, POMC or ADRB2.

Keywords: Hanwoo (Korean cattle); Chinese Yanbian cattle; Production traits; Selection signals

Introduction

Hanwoo is an indigenous cattle breed of Korea and has three strains showing phenotypic and genetic differences: brown Hanwoo (HW), brindle Hanwoo (BR), and Jeju black Hanwoo (JB) (Seo et al. (2007)). BR and JB have very small populations (less than 1,000 individuals) in Korea with low genetic variation and they might be considered as typical island populations in which a founder effect and genetic drift could have contributed to a loss of variation. Chinese Yanbian (YB) cattle are closer related to HW, with which they share phenotypic features, than BR or JB according to the genomic relation matrix (GRM) (Porto-Neto et al. (2014)). In contrast to HW, which are part of a breeding program, the other three breeds (BR, JB, and YB) are unselected native breeds. This intensive selection will likely have increased the frequency of favorable alleles at loci affecting the selected production traits in Hanwoo.

The HW breeding program was started in the 1930s and HW have been intensively selected based on production traits such as carcass weight (CWT), eye muscle area (EMA) and marbling (intramuscular fat) through progeny testing.

Searches for signatures of selection successfully revealed many genes that have been important in cattle. For examples, the International Bovine HapMap project described a range of breeds that have been historically selected for different phenotypic traits. Hayes et al. (2009) also proposed to identify divergently selected regions of the genome between dairy cattle and beef cattle breeds within *Bos taurus* (Hayes et al. 2009). Qanbari et al. (2010) performed a whole genome scan for selection signatures in the Holstein genome and detected a set of 10 candidate regions including diacylglycerol O-acyltransferase 1 (DGAT1), Casein cluster, Growth Hormone Receptor (GRH), Somatostatin and Leptin receptor (Qanbari et al. (2010)) based on extended haplotype homozygosities (EHHs). The aim of this study is to identify selection signatures in Hanwoo populations through pairwise comparisons between HW and the other unselected breeds BR, JB, and YB. We tried to find common selection signals from the results of the different comparisons. Finally, the putative candidate genes that are in proximity to the significant genome-wide markers are determined to underline the potential selection pressure on the identified genomic region.

Materials and Methods

Animals and genotype data. The Hanwoo data were collected from 160 steers in three populations of HW, BR, and JB; a) steers of candidate bulls for progeny testing in the Hanwoo Improvement Center of National Agricultural Cooperative Federation in Seosan (N = 100), b) steers raised in Hankyung University and Gyeonbuk province (N = 20), and c) steers raised on Jeju island (N = 20), respectively. YB (N=39) were obtained from the northeast of China. Genomic DNA for genotyping assays was extracted from blood samples and the single nucleotide polymorphism (SNP) genotyping was performed using the Illumina Bovine SNP 50K Bead chip (Illumina, San Diego). Stringent filtering criteria were applied to the genotype data. Briefly, SNPs were excluded from the analysis if they failed in over 5% of the genotypes, had median GC scores below 0.6, had GC scores under 0.6 in less than 90% of the samples, deviated in heterozygosity more than 3 standard devia-

tions from the other SNPs and were out of Hardy-Weinberg equilibrium for a cutoff p-value of 1-15. Unmapped SNPs and SNPs on sex chromosomes were also excluded. Individual genotypes with GC scores under 0.6 were treated as missing. After this SNP filtering, genotypes were tested for Hardy-Weinberg equilibrium (HWE) to identify possible genotyping errors using a chi-square test in R/SNPassoc Package. Single nucleotide polymorphisms not in HWE ($P < 0.05$), monomorphic SNPs, and with minor allele frequency $< 1\%$ were removed from this study. Finally, from a total of 55,704 SNPs, genotype data of 38,266 SNP remained after quality control, all of which were previously mapped to a chromosome in the Btau4.0 assembly

Analysis of population structure. To identify hierarchical clustering of individuals based on the SNP genotyping data was performed using the maximum likelihood method implemented in the program Admixture 1.22 (Alexander et al. (2009)). We also performed a PCA based on the genomic relationship matrix (GRM) according to van Raden (Van Raden 2008). F_{ST} statistics were also estimated between pairs of breeds. F_{ST} These distance matrices were then used for hierarchical clustering using Ward's minimum variance algorithm. To clear visible signs of population structure, we also added European Angus samples ($N=20$) as an out-group.

Detecting footprints of selection using *Rsb* score. Haplotypes were estimated with fastPHASE (Scheet and Stephens 2006) for the HW, BR, JB and YB populations. To compare the selected population (HW) with the other unselected populations (BR, JB and YB), we combined genotypes of the three unselected breeds (UN) and created haplotypes based on the pooled genotypes. Using the haplotype information, we further computed *Rsb* scores according to Tang (Tang et al. (2007)) using the 'rehh' package (Gautier and Vitalis 2012). *Rsb* is the standardized log-ratio of the integrated EHHS (*iES*) between pairs of populations. These statistics are designed to detect regions with high levels of haplotype homozygosity over an unexpectedly long distance (relative to neutral expectations) across populations as,

$$Rsb = \ln(iES_{pop_1}/iES_{pop_2})$$

Therefore, genomic regions with unusually high *Rsb* referred as signals of positive selection. For four comparisons, HW/BR, HW/JB, HW/YB, and HW/UN, *Rsb* statistics at per SNP were calculated by standardizing the ratio of the corresponding *iES* of the unselected populations (BR, JB, YB and UN as pop_2) and that of the selected population (HW as pop_1) by breeding program. To identify candidate regions across the genome, we first counted the number of SNP with $P_{Rsb} > 2$ ($P < 0.01$) for each 1 Mb window (with a 0.5 Mb overlap) for each of the comparisons

Results and Discussion

Population structure. The Principal Component Analysis (PCA) based on the GRM separated and grouped individuals of all five breeds in agreement to their origins (Table 1). Pairwise F_{ST} also confirmed that YB are genetically very close to HW ($F_{ST} = 0.004$ for HW/YB pair) relative to the pair of HW/BR ($F_{ST} = 0.011$) or HW/JB ($F_{ST} = 0.010$). This is probably mainly due to the small population sizes of island varieties BR and JB as founder effect.

Table 1. Pairwise F_{ST} estimates between five cattle breeds (below the diagonal) and pairwise distances based on the genomic relationship matrix (above the diagonal). AG: Angus; JB: Jeju Black Hanwoo; BR: Brindle Hanwoo; HW: Brown Hanwoo; YB: Chinese Yanbian cattle

	AG	JB	BR	HW	YB
AG	NA	0.032132	0.035135	0.028352	0.022256
JB	0.167763	NA	0.017421	0.010771	0.008228
BR	0.166517	0.001246	NA	0.01149	0.009939
HW	0.23765	0.069886	0.071132	NA	0.004175
YB	0.253095	0.085332	0.086578	0.015445	NA

The result of Admixture has also almost same patterns with PCA and pairwise F_{ST} (Figure 1). $K = 4$ was found in good agreement with PCA result. A first cluster (represented in blue) could be unambiguously interpreted as HW clusters (HW and YB). Interestingly, YB still has a characteristic as cross-breed in this result. YB individuals had on average 67.2% (ranging from 52.6 % to 75.9%) of HW ancestry, 19.8% (ranging from 11.1% to 26.6%) of BR and JB ancestry (island cluster of Korean cattle) and 12.8% (ranging from 3.7% to 32.7%) of AG ancestry. Moreover, All Korean cattles (HW, JB and BR) showed very lower admixture from the AG than the YB.

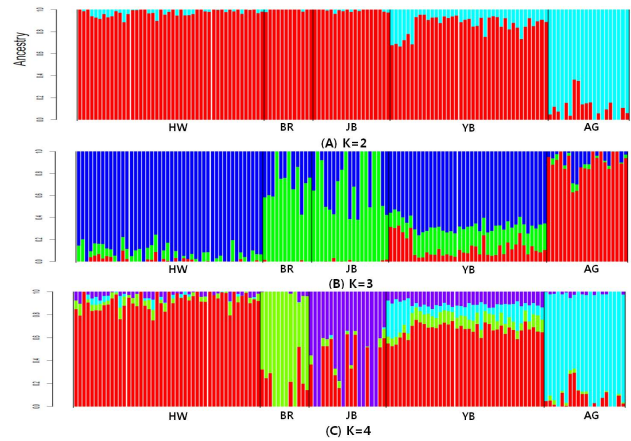


Figure 1. Unsupervised hierarchical clustering results of the individuals genotyped for 38,404 SNPs with an inferred number of clusters $K = 2$ (a), $K = 3$ (b) and $K = 4$ (c).

Identification of footprints of selection. Several regions appear to display highly significant scores. In the comparison of HW/BR, marker Hapmap35883-POMC_121F2-SNP1 showed the highest *Rsb* score (*Rsb* score = 4.53 and *pRsb* = 5.23); the corresponding gene at this position is the proopiomelanocortin (*POMC*) gene. Further, we created 5103 windows (1 Mb across the genome) with 14.96 SNPs on average per window. From this step, we obtained 41, 38, 105, 87 regions that differed in their *pRsb* value in the comparisons of HW/BR, HW/JB, HW/YB and HW/UN, respectively. Candidate regions were considered those windows containing at least two SNPs with *pRsb* >2 in at least three tests which left 31 candidate regions. From the comparisons of HW/BR, HW/JB, HW/YB, and HW/UN, 221, 196, 516 and 487 SNPs, respectively, were detected according to the counts of significant SNPs from the *Rsb* test (Supplementary Figure 32). We found that 18 significant SNPs were commonly shared between all four breeds (Figure 2). Among these 18 significant SNPs, 16 genes were located near the SNPs. The genes are *POMC*, *NOV*, *ENDOD1*, *CAPN1*, *LHFP3*, *CNTLN*, *PPMIH*, *PACS2*, *CWC15*, *ADRB2*, *PTBP2*, *CNOT1*, *CNTN6*, *SLC25A45*, *ACCN1* and *TTH1*. Some of the genes were previously reported for the production traits such as marbling, carcass or growth traits. In particular, *CAPN1* is already used in the commercial cattle industry for marker assisted selection and significant (Cheong et al. (2008)). *CAPN1* is located on chromosome 29 at 44.5 - 47Mb and was detected in all breed comparisons as evidence of selection in our study. The marker CAPN_1 has significant *Rsb* scores for the comparisons of HW/BR (3.73), HW/JB (4.17), HW/YB (3.08) and HW/UN (4.28).

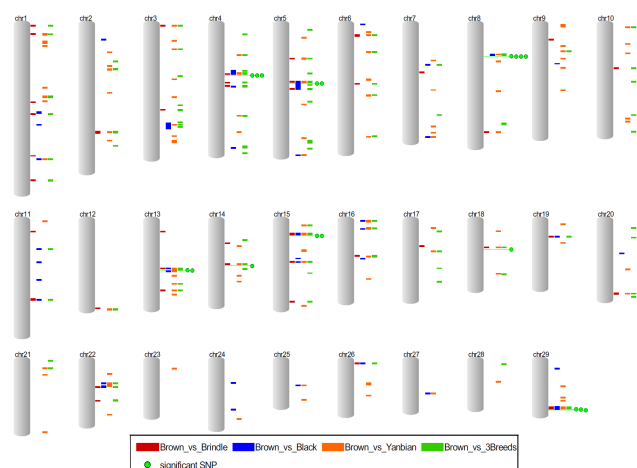


Figure 2. The map of the selection signals in the Brown Hanwoo and 18 common significant SNPs from the four comparisons using *Rsb* statistic. The box is indicated as candidate regions. The circle is genomic position of the significant SNP.

Conclusion

In this study, we were able to trace selection signatures by comparing a selected and three un-selected breeds with each other. Such a comparison has the potential to identify genomic regions influencing production traits. The regions can contribute to the identification of candidate genes and to the understanding of the biological mechanisms controlling production traits. Some regions harboring genes that are associated with traits of economic importance for livestock have been detected. In particular, the fatness and growth metabolism have been highlighted. These results suggest that although brown Hanwoo have experienced only recent selective pressure with short divergence time, signals of selection can be observed.

Literature Cited

- Seo, K., T. R., Mohanty. et al. (2007). Vet. Dermatol. 18(6): 392-400.
- Porto - Neto, L. R., Lee, S. H., Sonstegard, T. S., et al. (2014). Anim. Genet.
- Hayes, B. J., Chamberlain, A. J., Maceachern, S., et al. (2009). Anim. Genet. 40(2): 176-184.
- Qanbari, S., Pimentel, E. C. G., et al. (2010). Anim Genet 41(4): 346-356
- Alexander, D. H., Novembre, J., and Lange, K. (2009). Genome Res. 19(9): 1655-1664
- VanRaden, P. M. (2008). J Dairy Sci 91(11): 4414-4423
- Scheet, P., and Stephens, M. (2006). Am J Hum Genet 78(4): 629-644
- Tang, K., Thornton, K. R., & Stoneking, M. (2007). PLoS biol. 5(7): e171.
- Gautier, M., & Vitalis, R. (2012). Bioinformatics 28(8): 1176-1177
- Cheong, H. S., Yoon, D. H., Park, B. L., et al. (2008). BMC genet. 9(1): 33.

