**ANIMAL BIOSCIENCE**

# Genetic characteristics of Korean Jeju Black cattle with high density single nucleotide polymorphisms

**M. Zahangir Alam[1,2,a], Yun-Mi Lee[1,a], Hyo-Jung Son[1], Lauren H. Hanna[3], David G. Riley[4], Hideyuki Mannen[5], Shinji Sasazaki[5], Se Pill Park[6], and Jong-Joo Kim[1,*]**

* **Corresponding Author:** Jong-Joo Kim
**Tel:** +82-53-810-3027, **Fax:** +82-53-801-3027,
**E-mail:** kimjj@ynu.ac.kr

[1] Department of Biotechnology, Yeungnam University, Gyeongsan 38541, Korea
[2] Department of Genetic Engineering and Biotechnology, Shahjalal University of Science and Technology, Sylhet 3114, Bangladesh
[3] Department of Animal Sciences, North Dakota State University, Fargo, ND 58105, USA
[4] Department of Animal Sciences, Texas A&M University, College Station, TX 77843, USA
[5] Graduate School of Agricultural Science, Kobe University, Kobe 657-8501, Japan
[6] Faculty of Biotechnology, Jeju National University, Jeju 13557, Korea

[a] These authors contributed equally to this work.

ORCID
M. Zahangir Alam
https://orcid.org/0000-0003-0483-3256
Yun-Mi Lee
https://orcid.org/0000-0003-3023-7474
Hyo-Jung Son
https://orcid.org/0000-0003-0976-4456
Lauren H. Hanna
https://orcid.org/0000-0002-5238-7169
David G. Riley
https://orcid.org/0000-0002-3636-4418
Hideyuki Mannen
https://orcid.org/0000-0001-5627-7082
Shinji Sasazaki
https://orcid.org/0000-0003-2090-2653
Se Pill Park
https://orcid.org/0000-0002-6965-7831
Jong-Joo Kim
https://orcid.org/0000-0001-9687-0075

Submitted Nov 15, 2019; Revised Jan 1, 2020; Accepted Jun 29, 2020

**Objective:** Conservation and genetic improvement of cattle breeds require information about genetic diversity and population structure of the cattle. In this study, we investigated the genetic diversity and population structure of the three cattle breeds in the Korean peninsula. **Methods:** Jeju Black, Hanwoo, Holstein cattle in Korea, together with six foreign breeds were examined. Genetic diversity within the cattle breeds was analyzed with minor allele frequency (MAF), observed and expected heterozygosity ($H_O$ and $H_E$), inbreeding coefficient ($F_{IS}$) and past effective population size. Molecular variance and population structure between the nine breeds were analyzed using a model-based clustering method. Genetic distances between breeds were evaluated with Nei's genetic distance and Weir and Cockerham's $F_{ST}$. **Results:** Our results revealed that Jeju Black cattle had lowest level of heterozygosity ($H_E = 0.21$) among the studied taurine breeds, and an average MAF of 0.16. The level of inbreeding was –0.076 for Jeju Black, while –0.018 to –0.118 for the other breeds. Principle component analysis and neighbor-joining tree showed a clear separation of Jeju Black cattle from other local (Hanwoo and Japanese cattle) and taurine/indicine cattle breeds in evolutionary process, and a distinct pattern of admixture of Jeju Black cattle having no clustering with other studied populations. The $F_{ST}$ value between Jeju Black cattle and Hanwoo was 0.106, which was lowest across the pair of breeds ranging from 0.161 to 0.274, indicating some degree of genetic closeness of Jeju Black cattle with Hanwoo. The past effective population size of Jeju Black cattle was very small, i.e. 38 in 13 generation ago, whereas 209 for Hanwoo. **Conclusion:** This study indicates genetic uniqueness of Jeju Black cattle. However, a small effective population size of Jeju Black cattle indicates the requirement for an implementation of a sustainable breeding policy to increase the population for genetic improvement and future conservation.

**Keywords:** Jeju Black Cattle; Hanwoo; Genetic Diversity; Population Structure; Single Nucleotide Polymorphism (SNP) Chip

## INTRODUCTION

Cattle are an integral part of animal agriculture since 8000 BC, when it is thought that they become domesticated in different parts of the world such as India, Middle East and North Africa [1]. Different cattle breeds have been domesticated and adapted throughout the world due to variable geographical and climatic conditions. Jeju Black cattle (JJBC; Jeju Heugu) is one of the indigenous cattle breeds in the Jeju Island, south of Korean peninsula. The JJBC are thought to be originated from the native cattle in Korea main land according to island model of speciation [2]. The evidence of ancient cattle bones from the archaeological sites in Jeju Island suggests the existence of the breed approximately 1,100 to 2,000 years ago. DNA analysis of bones recovered from Gonaeri and Gwakji-ri in Aewaleup, Jeju city, indicates that ancestors of the present JJBC had been raised by humans since prehis-

toric times [3]. Also, the historical documents (Annals of the Choseon Dynasty) and the paintings found in the mural (Anak Tomb no. 3, during Goguryeo Dynasty in 357 AD) support cattle existence in Korean peninsula. However, there are several reports about cattle originating in a controversial way [4]. Whatever their origin, JJBC has been categorized as an endangered species due to a substantial shrinkage in population size until 1980s. It is reported that registered JJBC comprise approximately 619 individuals (Korea Seed Stock Database), while other sources indicated that the population size might be 400 to 500 [5,6]. Due to the small size of the JJBC population, it is essential to evaluate and monitor the level of inbreeding, which is an important parameter to assess the genetic diversity of the breed.

JJBC are adapted to the subtropical environment of the Jeju Island. Also, beef of JJBC is rich in oleic acid, linoleic acid, and unsaturated fatty acids, which make it a premium quality for Korean consumers. However, in the past decades, this indigenous breed was paid little attention by beef cattle producers due to their slow growth and thus were was not competitive compared to 'Hanwoo', a breed in mainland of Korea that has been extensively bred for superior meat quality.

Genomic studies using high throughput whole genome sequencing data have become popular in recent years. Single nucleotide polymorphisms (SNP) are one of the common genetic variants for any organism, and genotyping with a high density SNP microarray chip provides genome information in an efficient and cost effective manner. Many useful genetic parameters such as linkage disequilibrium (LD), effective population size ($N_e$), inbreeding coefficient, levels of heterozygosity, etc. can be estimated with high density SNP chips [7-9], so as to provide powerful tools to study evolutionary and conservation biology.

Development of sustainable breed improvement strategies depends on the precise characterization of animal genetic diversity [10]. Although several studies have investigated the diversity pattern of Korean cattle along with JJBC [2,5,11], it is still controversial whether Jeju Black are a separate breed. Thus, an accurate definition of breed origin for JJBC is necessary for future conservation and improvement programs. Genetic diversity study with microsatellite markers often causes estimated values of greater genetic differentiation than SNP markers [12,13]. Moreover, SNP markers give more accurate estimates in population admixture analysis than pedigree information [14]. In this study, genetic characteristics and population structure of JJBC breed were investigated with high density SNP chips, and were compared with Hanwoo, seven *Bos indicus* and *Bos taurus* breeds, by analyzing allelic richness ($A_R$), level of inbreeding ($F_{IS}$), effective population size ($N_e$), and genetic distances between the breeds.

## MATERIALS AND METHODS

### Animals and genotypes

Animal ethics approval statement was not applicable as the Hanwoo DNA was extracted either from commercial bull semen straws or from tail hair samples obtained from different farmers with the permission of the owners. A total of 373 animals from nine cattle breeds were chosen to study, including two indicine breeds (Brahman, 15; Nelore, 26), seven taurine breeds (Angus, 27; Holstein, 76; Hereford, 25; Hanwoo, 66; Jeju Black, 78; Brown Wagyu, 10; and Black Wagyu, 50). Hanwoo DNAs were extracted either from AI bull semen straws or tail hair samples obtained from different farmers with the permission of the owners, and JJBC DNAs were prepared either from AI bull semen straws or tail hair in Jeju Island, Korea. Holstein DNAs were collected from semen straws provided by Nonghyup Dairy cattle improvement center. All other cattle DNAs were provided by Texas A & M University, USA, except the two Japanese Wagyu breeds. No ethics statement was required for the collection of DNA samples from the Brahman, Nelore, Angus, and Hereford cattle, because DNA samples were provided by the two authors in USA under their rules and regulations. The data of two Japanese Wagyu breeds were provided by the authors in Japan, which was previously used and published. Genomic DNA purification and genotyping was accomplished by DNA Link Inc. (Seoul, Korea), a commercial genome analysis service provider in Korea.

### Single nucleotide polymorphism genotyping and assembly of data sets

The samples of all breeds in this study except the two Wagyu breeds were genotyped using a customized Affymetrix 150K SNP Axiom array by DNA Link Inc. (Seoul, Korea), which included the 50K SNPs from the Bovine 50K v.3 Bead Chip (Illumina, San Diego, CA, USA). The samples of the two Wagyu breeds were genotyped with the Illumina 50K SNP chip. All the genotyped data were then merged and the common SNP markers on autosomal chromosomes were selected, resulting in 45,526 SNP markers in the final dataset across the breeds.

### Quality control and filtering of single nucleotide polymorphism markers

SNP quality control and filtering were performed across the nine cattle breeds to remove SNP markers with less than 95% call rate and animals with less than 95% call rate with PLINK software program [15]. This process resulted in 41,186 SNPs in 372 animals across the breeds. SNP quality filtering was also performed using PLINK version 1.9, and SNP markers with high LD were pruned using the following parameter option; -indep pair wise command--*indep-pairwise 50 5 2*

(SNP window size, 50; SNP markers shifted per step, 5; $r^2$ threshold, 2), because pruning of SNP markers in high LD can counter the effect of ascertainments bias and generate meaningful comparison between breeds [16]. After pruning, a total of 18,524 SNP markers were finally selected for analysis.

### Estimation of genetic diversity within breed and population differentiation

Genetic diversity within cattle breeds can be estimated by expected heterozygosity ($H_E$), observed heterozygosity ($H_O$), and inbreeding coefficient ($F_{IS}$), for which the parameters were calculated using R software package divRsity v1.9.9 [17]. Analysis of molecular variance (AMOVA) to determine the partition of genetic diversity was performed with ARLEQUIN v3.5 [18]. Population differentiation was calculated by pairwise $F_{ST}$ estimates according to the approach of Weir and Cockerham's [19] with R package, divRsity v1.9.9.

### Population structure analysis

Population structure of the studied cattle breeds were also carried out using the software ADMIXTURE v1.3. [20], which enables an unsupervised clustering of large numbers of samples and allows the incorporation of each individual cattle breed into a mixture of clusters. In ADMIXURE, a model-based estimation of individual ancestry was applied for a range of prior values of K defined by the user. To elicit the true number of genetic populations, i.e. K clusters among nine cattle breeds; a cross validation (CV) approach was used to determine the most likely number of populations (K) in the SNP data. The best possible number of ancestral populations (K) was inferred through 3 to 11 pre-assumed populations. For each tested value of K in ADMIXTURE, the proportion of each individual's genotype was estimated for clustering, showing a preferable value of K with a low CV error compared with other K values.

### Principal component analysis

Principal component analysis (PCA) determines breed relationships that are based directly on allele frequencies by using a multivariate method, in which the information from a large number of alleles and loci is condensed into a few synthetic variables known as principal components (PC) [21]. PCA was carried out to infer relationships between the nine cattle populations by using PLINK v1.9.

### Genetic distance

Phylogenic analysis on the basis of SNP data has become an important tool for studying evolutionary history of any organism. In this study, phylogenic tree was constructed by calculating Nei's genetic distances (DA) [22] with Poptree 2 [23] program. Neighbor-Joining (NJ) method was applied

for measuring Nei's genetic distance ($D_A$), which can be defined as

$$D_A = 1 - \frac{1}{r} \sum_{j}^{r} \sum_{i}^{m_j} \sqrt{x_{ij} y_{ij}}$$

where $x_{ij}$ and $y_{ij}$ are the frequencies of the $i$-th allele at the $j$-th locus in populations X and Y, respectively, $m_j$ is the number of alleles at the $j$-th locus, and $r$ is the number of loci used.

### Linkage disequilibrium

To measure the extent of LD, Lewontin's D′ [24] and Hill's $r^2$ [25] are widely used. However, $r^2$ is preferred for association studies due to its robustness, simplicity and not sensitivity to changing gene frequency and effective population size [26]. The $r^2$ estimator represents a squared correlation coefficient (r) between two variables (alleles) at two separate SNP marker loci [27]. PLINK software [15] was used for estimation of $r^2$ parameter:

$$r^2(P_a, P_b, P_{ab}) = \left( \frac{(P_{ab} - P_a P_b)^2}{P_a(1 - P_a)P_b(1 - P_b)} \right)$$

where, $P_{ab}$ represents the frequency of haplotypes consisting of allele $a$ at the first locus and allele $b$ at the second locus [28]. The $r^2$ values were calculated for each chromosome in each breed, as well as across breeds, so that correlations between all possible SNPs were tested and there was no $r^2$ threshold, in order to encapsulate all possible linkage interaction between SNPs per chromosome. To display the decay of LD, distances of pair-wise SNPs were binned into twenty types of intervals (0 to 1 kb, 10 kb intervals starting from 1kb up to 100 kb and 100 kb interval starting from 100 kb up to 1 Mb). For each chromosome, $r^2$ values were then sorted by inter-SNP distance, and were averaged across the afore-mentioned intervals to observe possible $r^2$ patterns with increasing inter-SNP distances.

### Effective population size ($N_e$)

$N_e$ was estimated using SNeP v1.1 by Barbato et al [29], which was based on LD data with the following formula suggested by Corbin et al [30],

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

where $N_{T(t)}$ represents the past effective population size with $t$ generations ago, $C_t$ represents the recombination rate, $t$ generations ago in the past for specific physical distance between markers calculated by the SNeP tool, $r^2_{adj}$ represents LD value adjusted for sample size, and $\alpha = (1, 2, 2.2)$ is the correction factor (constant) for the occurrence of mutation. The recombination rate was calculated using the following

equation suggested by Sved [31],

$$f(c) = c[(1-c/2)/(1-c)^2]$$

The data sets for each sub-population, as well as the merged dataset, were grouped into 20 distance bins of 10 to 100 kb each. $N_e$ estimates were subsequently obtained from the $r^2$ values for the average distance of each distance bin.

## RESULTS

### Within breed genetic diversity
Table 1 presents three measures of within breed diversity across the studied population. The minor allele frequency (MAF) ranged from 0.11 (Nelore) to 0.21 (Hanwoo, Angus, and Holstein), while the MAF was in middle (0.16) in JJBC. Nelore cattle had the lowest level of expected heterozygosity ($H_E$ = 0.15), while Hanwoo, Angus, and Holstein had the highest level of genetic diversity ($H_E$ = 0.28). JJBC had a middle range value of $H_E$ (0.21) in the studied breeds.

In a structured population, the fixation index, F represents the degree of reduction in heterozygosity relative to Hardy-Weinberg expectation. $F_{IS}$ measures the heterozygosity of individuals (I) relative to the subpopulation (S) represented by non-random mating (inbreeding), so that negative $F_{IS}$ value means less inbreeding. The $F_{IS}$ ranged from –0.018 in Black Wagyu to –0.118 in Brown Wagyu cattle. JJBC had the $F_{IS}$ value of –0.076, while –0.025 in Hanwoo (Table 1).

### Analysis of molecular variance and population differentiation
The average Wright's F-statistics that were estimated with 20,000 bootstraps over loci were of values, $F_{ST}$ = 0.173, $F_{IS}$ = –0.030 and $F_{IT}$ = 0.148 (Table 2). $F_{IT}$ measures the genetic differentiation within individuals of the total population. $F_{ST}$ is the measure of the genetic differentiation between breeds, for which the value of 0.173 indicates population differentia-

**Table 1.** Sample sizes and measurements of genetic diversity for nine cattle breeds

| Breed | N | MAF (SD) | $H_O$ (SD) | $H_E$ (SD) | $F_{IS}$ |
|---|---|---|---|---|---|
| Angus | 27 | 0.21(0.16) | 0.30(0.20) | 0.28(0.18) | –0.056 |
| Brahman | 15 | 0.12(0.15) | 0.18(0.22) | 0.17(0.19) | –0.073 |
| Brown Wagyu | 10 | 0.16(0.17) | 0.25(0.24) | 0.22(0.20) | –0.118 |
| Black Wagyu | 50 | 0.19(0.16) | 0.26(0.20) | 0.26(0.19) | –0.018 |
| Hereford | 25 | 0.19(0.16) | 0.27(0.21) | 0.26(0.19) | –0.045 |
| Holstein | 75 | 0.21(0.16) | 0.28(0.19) | 0.28(0.18) | –0.026 |
| Hanwoo | 66 | 0.21(0.16) | 0.29(0.19) | 0.28(0.18) | –0.025 |
| Jeju Black | 78 | 0.16(0.16) | 0.23(0.21) | 0.21(0.19) | –0.076 |
| Nelore | 26 | 0.11(0.15) | 0.16(0.21) | 0.15(0.19) | –0.059 |

N, sample size; MAF, means of minor allele frequency; SD, standard deviation; $H_O$, observed heterozygosity; $H_E$, expected heterozygosity under Hardy-Weinberg equilibrium; $F_{IS}$, inbreeding coefficient.

**Table 2.** Average F-Statistics overall loci according to Weir and Cockerham [20]

| Fixation indices | Ninebreeds | p-values |
|---|---|---|
| $F_{ST}$ | 0.173** | 0.00000 |
| $F_{IS}$ | –0.030[NS] | 1.00000 |
| $F_{IT}$ | 0.148** | 0.00000 |

$F_{ST}$, genetic differentiation among breeds; $F_{IS}$, within population inbreeding; $F_{IT}$, total inbreeding; NS, not significant.
** p < 0.001 significant levels were obtained after 1,000 permutations.
Fixation indices were obtained with over 20,000 bootstraps.

tion with statistical significance (p<0.01). AMOVA indicated that almost 17% of the variation was estimated for variation among the populations, while 85% of the variation was accounted due to within individual variation (Table 3). Genetic differentiations between the nine cattle breeds that were based on pairwise $F_{ST}$ are displayed in Figure 1. The $F_{ST}$ ranged from 0.085 (Hanwoo and Brown Wagyu) to 0.376 (Nelore and Brown Wagyu). The genetic distance ($D_A$) between Hereford and Nelore (0.107) was relatively greater than between JJBC and Hanwoo (0.028). JJBC had great distances with Nelore (0.084) and Brahman (0.078), while small distances with Hanwoo (0.028) and Black Wagyu (0.042) cattle breeds.

The Nei's genetic distances ($D_A$) matrix of the nine cattle breeds were also used to construct phylogenetic trees with NJ method [22]. Our results showed clear separation of taurine and indicine cattle into two groups (branches). After diverging into two branches of *Bos Taurus* and *Bos Indicus*, the NJ trees showed two sub-branches within the taurine branch, i.e. one for JJBC and the other for Hanwoo, two Japanese breeds and three European taurine breeds (Figure 2).

### Principal component analysis
The first and second PC accounted for 27.9% and 24.9% of the total variation respectively, while the third and fourth PC accounted for 19.0% and 16.7% of the total variation, respectively. Thus, the first five PC accounted almost 100% variation across the breed populations. JJBC were uniquely located, even if Hanwoo breed is more closely positioned to JJBC than the other cattle breeds (Figure 3). Nelore and Brahman formed a distinct cluster due to large variation between taurine and indicine cattle. Hanwoo cattle formed a closer cluster with Brown Wagyu and Black Wagyu than Holstein

**Table 3.** Analysis of molecular variance among the nine cattle breeds

| Data set | Variance components (%) | | |
|---|---|---|---|
| | Among populations | Among individuals within populations | Within individuals |
| All nine cattle breeds | 17.31** | –2.48[NS] | 85.17** |

** p < 0.001 significant levels were obtained after 1,000 permutations; NS, not significant.
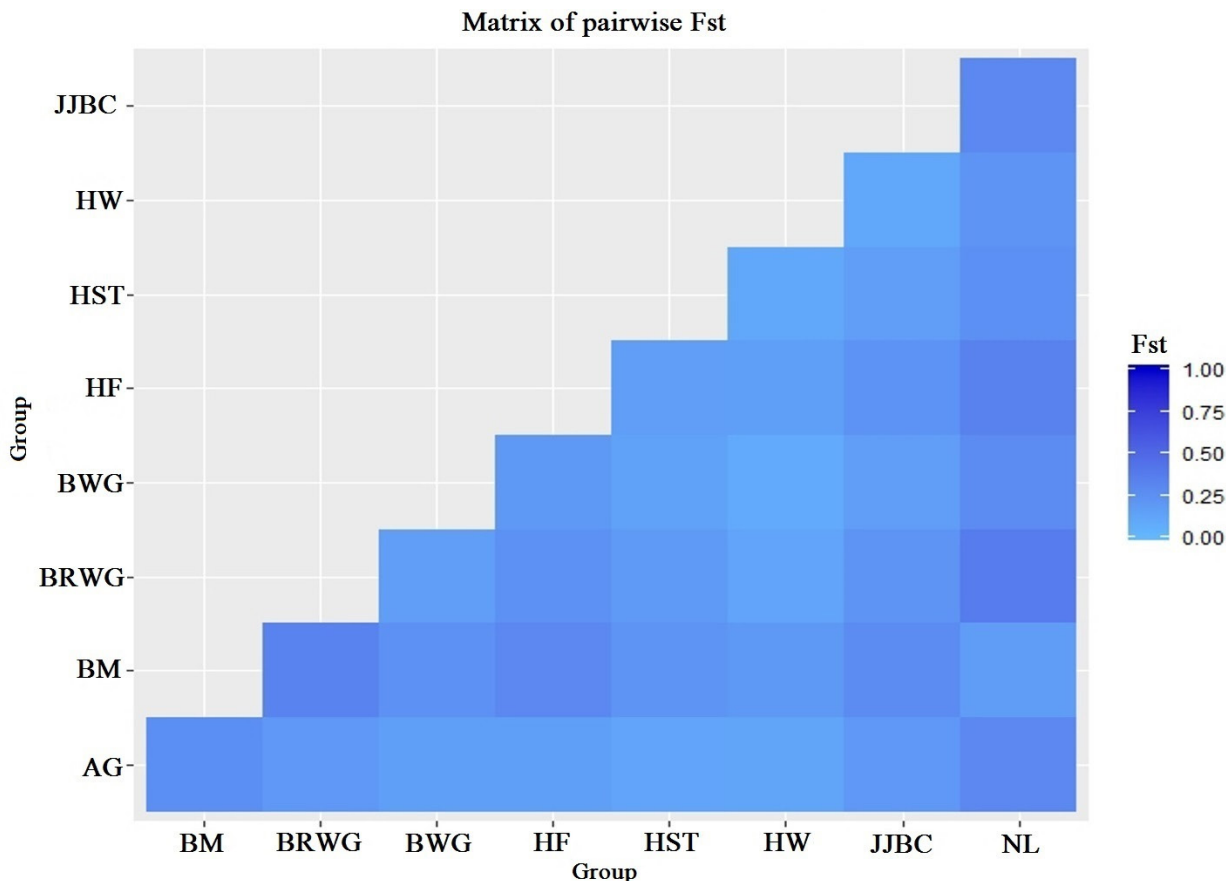
## Matrix of pairwise Fst



**Figure 1.** Graphical representation of pairwise $F_{ST}$ distance matrix. AG, Angus; BM, Brahman; BRWG, Brown Wagyu; BWG, Black Wagyu; HF, Hereford; HST, Holstein; HW, Hanwoo; JJBC, Jeju Black cattle; NL, Nelore.

or other European taurine breeds (Figure 3).

### Population structure analysis between eight cattle breeds

The results of proportion of individuals into each of the nine breeds that were inferred by the ADMIXTURE are presented in Figure 4. The lowest CV error values were expected when K = 9. However, the lowest CV error estimator was found when K = 11 (data not shown). Thus, K = 11 was taken as the most probable number of inferred populations.

### Effective population size ($N_e$) over the past generations

$N_e$ is needed to determine the accuracy of genomic selection [28], and the $N_e$ estimates in the nine cattle breeds are shown
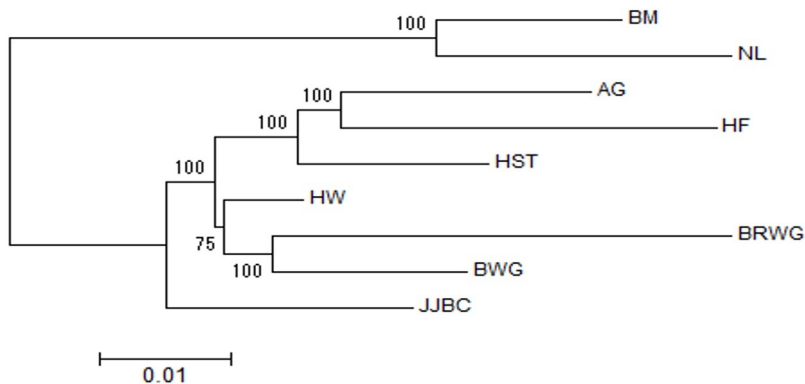


**Figure 2.** Unrooted consensus tree showing the genetic relationships among the nine breeds using the neighbor-joining method and the unbiased Nei's $D_A$ genetic distance. The values at the nodes are the percentages of bootstrap values from 1,000 replications of resampling. AG, Angus; BM, Brahman; BRWG, Brown Wagyu; BWG, Black Wagyu; HF, Hereford; HST, Holstein; HW, Hanwoo; JJBC, Jeju Black cattle; NL, Nelore.
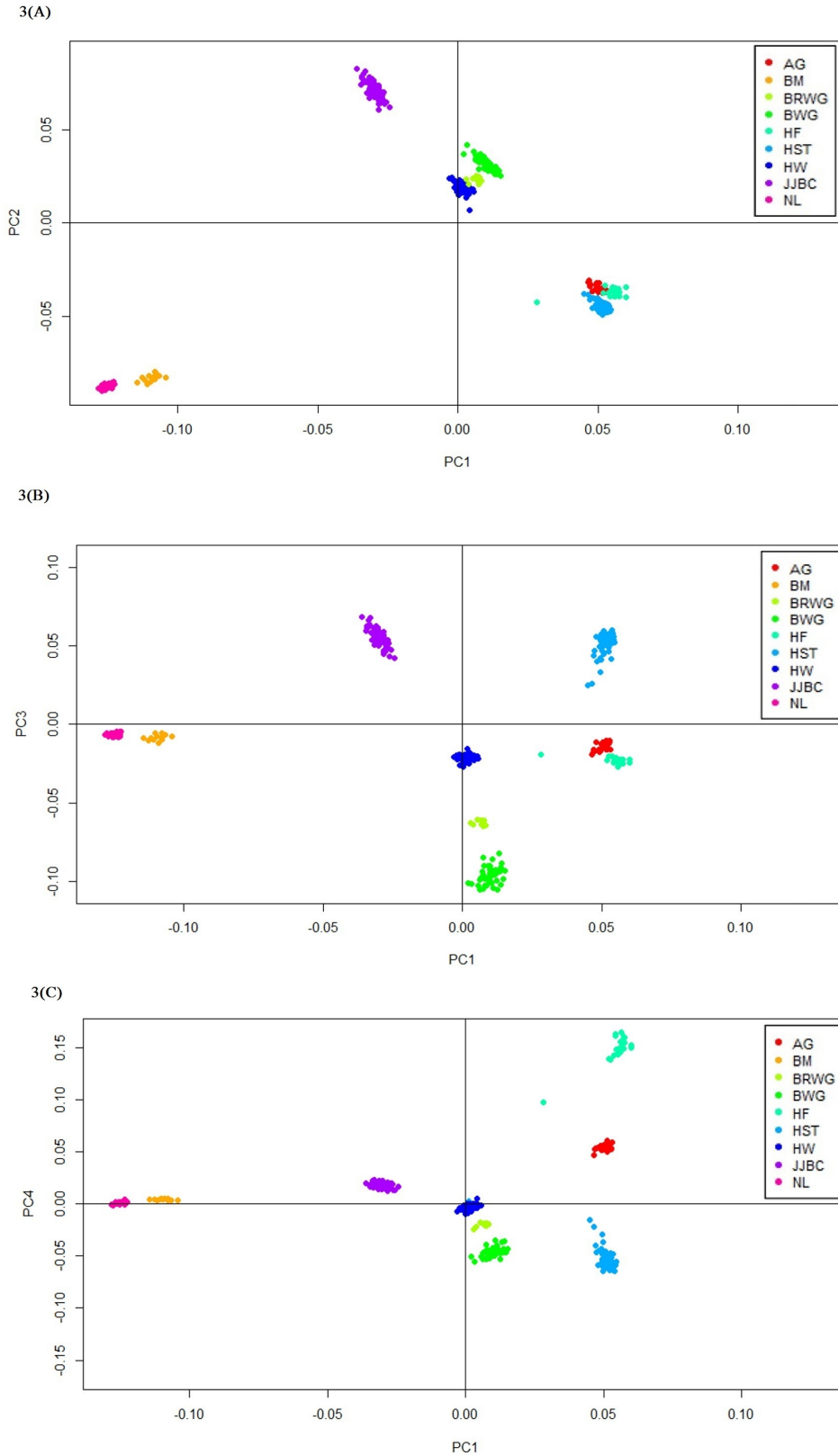
3(A)



3(B)

3(C)

**Figure 3.** Principal component analysis (PCA) analyses. First and second principal component (A), first and third principal component (B), and first and fourth principal component (C) analysis resulted from the 18,524 single nucleotide polymorphisms (SNPs) in the nine cattle breeds. AG, Angus; BM, Brahman; BRWG, Brown Wagyu; BWG, Black Wagyu; HF, Hereford; HST, Holstein; HW, Hanwoo; JJBC, Jeju Black cattle; NL, Nelore.
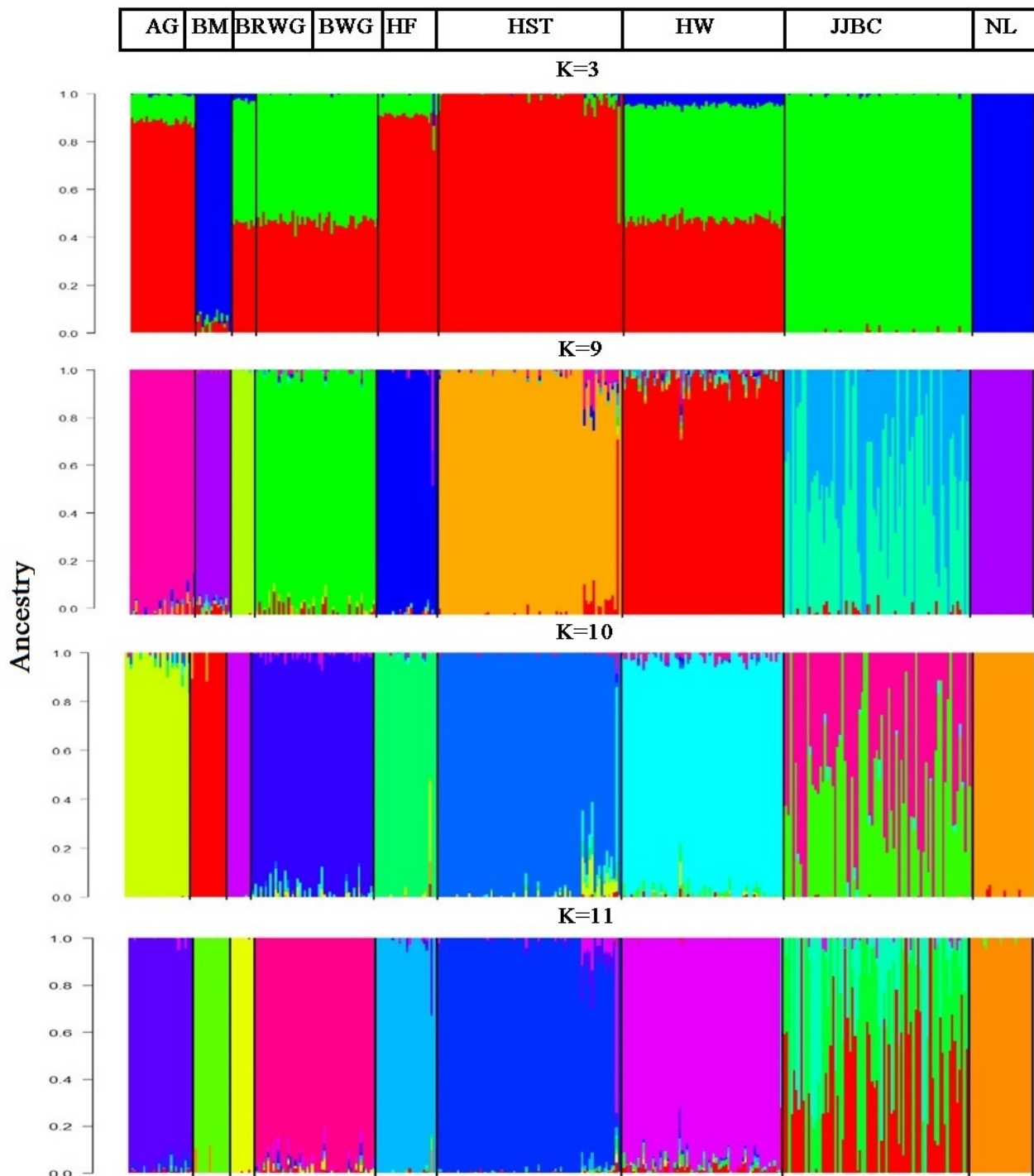
**Figure 4.** Clustering assignments of individuals into the nine cattle populations. ADMIXTURE analysis were performed with inferred K values ranging from 2 to 15, while the clustering results were shown when k = 3, 9, 10 and K = 11. AG, Angus; BM, Brahman; BRWG, Brown Wagyu; BWG, Black Wagyu; HF, Hereford; HST, Holstein; HW, Hanwoo; JJBC, Jeju Black Cattle; NL, Nelore. Each individual was represented by a single vertical line, which was divided into K colored segments. K is thenumber of the clusters assumed to have proportional length to each of the K inferred clusters. Black color separates the populations. Breeds are labeled by abbreviation at the top of figure.

in Figure 5 and Table 4 at *t* generation ago. In general, all the nine breeds showed a marked decrease over time as expected. The Ne estimate of Hanwoo cattle in the most recent 13 generation ago was 209, which was the greatest value across the

nine breeds, while JJBC has an Ne size of 38 at the 13 generation ago, which was small compared with other cattle breeds, e.g. 78, 37, 39, 51, 29, 93, and 91 for Angus, Brahman, Nelore, Hereford, Brown Wagyu, Black Wagyu, and Holstein, re-
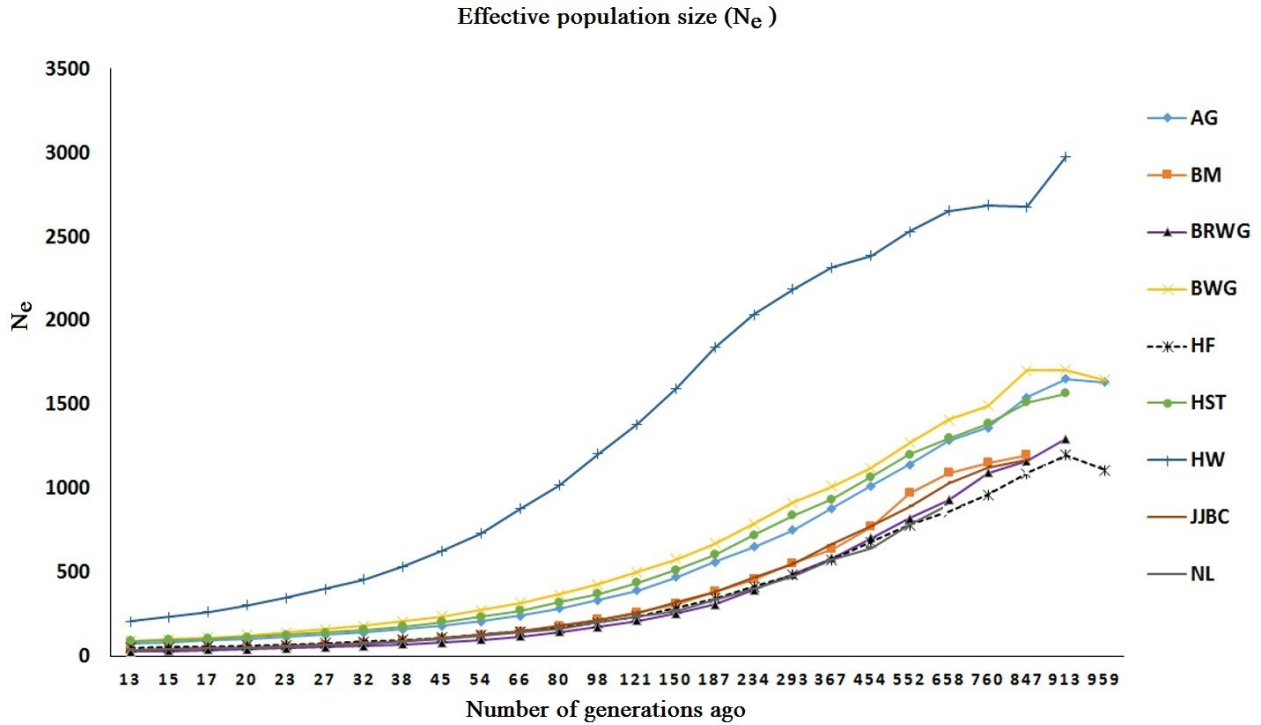
**Effective population size (N$_e$)**



**Figure 5.** Effective population size (N$_e$) in the nine cattle breeds. AG, Angus; BM, Brahman; BRWG, Brown Wagyu; BWG, Black Wagyu; HF, Hereford; HST, Holstein; HW, Hanwoo; JJBC, Jeju Black cattle; NL, Nelore.

**Table 4.** Effective population size (N$_e$) across the nine cattle breeds

| Generation Ago | AG | BM | BRWG | BWG | HF | HST | HW | JJBC | NL |
|---|---|---|---|---|---|---|---|---|---|
| 13 | 78 | 37 | 29 | 93 | 51 | 91 | 209 | 38 | 39 |
| 15 | 85 | 42 | 33 | 102 | 55 | 97 | 235 | 42 | 43 |
| 17 | 94 | 47 | 37 | 113 | 59 | 106 | 264 | 47 | 49 |
| 20 | 102 | 53 | 41 | 126 | 64 | 115 | 306 | 53 | 53 |
| 23 | 114 | 59 | 47 | 142 | 70 | 128 | 347 | 61 | 60 |
| 27 | 128 | 67 | 54 | 161 | 77 | 142 | 401 | 68 | 68 |
| 32 | 142 | 78 | 61 | 183 | 87 | 158 | 459 | 79 | 78 |
| 38 | 161 | 93 | 71 | 208 | 97 | 179 | 536 | 90 | 88 |
| 45 | 182 | 108 | 83 | 238 | 111 | 205 | 628 | 107 | 105 |
| 54 | 210 | 125 | 99 | 275 | 128 | 236 | 732 | 127 | 121 |
| 66 | 242 | 150 | 120 | 318 | 147 | 273 | 880 | 151 | 141 |
| 80 | 284 | 178 | 144 | 369 | 170 | 321 | 1,020 | 181 | 166 |
| 98 | 336 | 217 | 174 | 427 | 202 | 371 | 1,205 | 216 | 201 |
| 121 | 393 | 263 | 212 | 503 | 239 | 438 | 1,383 | 260 | 238 |
| 150 | 471 | 313 | 255 | 578 | 290 | 515 | 1,595 | 326 | 271 |
| 187 | 562 | 385 | 308 | 671 | 344 | 604 | 1,839 | 386 | 336 |
| 234 | 652 | 460 | 394 | 789 | 416 | 723 | 2,035 | 467 | 402 |
| 293 | 753 | 554 | 488 | 917 | 484 | 840 | 2,188 | 551 | 474 |
| 367 | 881 | 639 | 586 | 1,014 | 574 | 937 | 2,318 | 668 | 575 |
| 454 | 1,016 | 771 | 707 | 1,121 | 683 | 1,068 | 2,386 | 775 | 646 |
| 552 | 1,141 | 971 | 823 | 1,272 | 781 | 1,203 | 2,532 | 891 | 787 |
| 658 | 1,286 | 1,092 | 933 | 1,408 | 863 | 1,300 | 2,653 | 1,031 | 903 |
| 760 | 1,364 | 1,151 | 1,093 | 1,492 | 962 | 1,386 | 2,686 | 1,125 | - |
| 847 | 1,544 | 1,200 | 1,162 | 1,703 | 1,088 | 1,514 | 2,678 | 1,168 | - |
| 913 | 1,653 | - | 1,297 | 1,708 | 1,200 | 1,567 | 2,978 | - | - |

AG, Angus; BM, Brahman; BRWG, Brown Wagyu; BWG, Black Wagyu; HF, Hereford; HST, Holstein; HW, Hanwoo; JJBC, Jeju Black cattle; NL, Nelore.

spectively.

## Linkage disequilibrium

LD measurements with $r^2$ in the nine cattle breeds are presented in the Supplementary Table S1. In general, cattle breeds showed different pattern of LD at different inter-SNP distances (Supplementary Table S2). The average $r^2$ value for JJBC was 0.71 at 0 to 1 kb distance bin, which was very high compared with other cattle breeds, i.e. Hanwoo, Brown Wagyu, Black Wagyu, Brahman, Nelore, Angus, Hereford, and Korean Holstein had the $r^2$ values of 0.63, 0.65, 0.54, 0.55, 0.58, 0.71, 0.63, and 0.67, respectively. The highest LD within close proximity of 0 to 1 kb interval distance decreased rapidly with increasing distance between pairs of SNP markers in all populations.

## DISCUSSION

Genetic characterization of breeds or animals based on genomic data has become an attractive method due to easy access of high throughput data derived from microarray SNP chip technology. In genetic diversity analysis, SNP markers have many advantages over microsatellite markers due to higher level of resolution, despite a set of microsatellites being suggested by the FAO to assess genetic diversity of farm animals and endangered species [10,32]. In the most recent years, genomic characterization using SNP markers have been studied in a variety of cattle breeds such as Irish Carry cattle [33], Tyrol Grey [34], Spanish beef cattle breeds [35], Canchim [36], Chinese Yiling yellow cattle [37] and many other indigenous and exotic cattle breeds raised in different countries worldwide [38-44].

Among Korean cattle breeds, Hanwoo was paid much more attention due to its incorporation into the national breeding program since 1970s [2]. In this study, we emphasized on the genetic characterization of JJBC, because these cattle are raised in Jeju Island in Korea and thus have unique characteristics that are different from the breeds on the mainland of Korea.

JJBC showed lower level of genetic variability ($H_E = 0.21$) than Hanwoo and Holstein in Korea ($H_E = 0.218$). Sharma et al [11,45] demonstrated a different level of heterozygosity in JJBC ($H_E = 0.39$ and 0.25), while Struken et al [2] reported $H_E = 0.29$. Heterozygosity level in our study is close to Sharma et al [11,45]. However, different results might be due to the use of various genotyping platforms, markers, and quality control criteria [2]. The lower level of heterozygosity in JJBC than Hanwoo and Holstein might be due to small population sizes, or few breeding males with more chance for increased inbreeding. However, Makina et al [40] stated that allele frequencies might be a poor estimate of inbreeding. Inbreeding level ($F_{IS}$) in JJBC was estimated to be –0.076,

which was lower than Hanwoo (–0.025) and Holstein (–0.026) in Korea. Genetic variability was lowest in Nelore ($H_E = 0.15$) and low in Brahman ($H_E = 0.17$). Indicine breeds might have less genetic variability than taurine breeds as reported by Lin et al [46]. Analysis of molecular variance that enables partitioning of genetic variation into overall fixation indices ($F_{ST}$), within population inbreeding ($F_{IS}$) and total inbreeding ($F_{IT}$), showed that 85% of total genetic variation was due to within populations across nine cattle breeds. This value was lower than the within populations genetic variation observed in South African cattle populations (92%) [40], but higher than those for Iranian cattle (82.9%) [42] and Ethiopian cattle populations (84.0%) [39]. Total inbreeding estimate ($F_{IT}$) and estimate of population differentiation ($F_{ST}$) was 0.148 and 0.173 respectively, with statistical significance (p<0.001) (Table 2). JJBC and Hanwoo were found to be least differentiated ($F_{ST} = 0.106$) compared to other breed pairs. Sharma et al [45] calculated $F_{ST}$ values for Korean cattle breeds, ranging 0.02 to 0.06, and reported that JJBC was differentiated from another Korean cattle breed, Brindle Hanwoo (Chikso). Struken et al [2] estimated $F_{ST}$ value of 0.024 for JJBC-Hanwoo, 0.038 for JJBC-Chikso and 0.023 for Hanwoo-Chikso. Both studies confirmed that JJBC was least differentiated with Hanwoo, which was in good agreement with our study.

Japanese breed Black Wagyu breeds showed less differentiation with Hanwoo (0.085) than JJBC (0.164), indicating a closer relation with Hanwoo. Based on $F_{ST}$ value and phylogenetic tree, we are in agreement with the concept that Hanwoo and Japanese Wagyu cattle breeds were genetically much closer to each other than other indicine and European taurine breeds (Table 1; Figure 2). PCA analysis also showed that Hanwoo and Wagyu breeds are much closer than JJBC (Figure 3A).

Unsupervised hierarchical clustering of our data implemented with ADMIXTURE analysis revealed that 95% of Hanwoo individuals were assigned to cluster ten, whereas 39%, 38%, 23%, and 1% of JJBC individuals were assigned to four different clusters, i.e. five, one, six, and ten, respectively. This means that JJBC shared its genome only with Hanwoo cattle. Japanese Brown Wagyu (100%) stands alone in cluster three, while Black Wagyu (96%) individuals were assigned in cluster eleven and the rest of the 4% were assigned to cluster three, nine and ten (Table 4). This level of admixture between JJBC and Hanwoo and between Brown Wagyu and Black Wagyu might be due to co-ancestry regarding the origin of these two breeds.

Phylogenic tree analysis also showed a distinct branch for JJBC, while Hanwoo and Wagyu breeds shared the same branch more recently than JJBC (Figure 2), suggesting that Hanwoo and Wagyu breeds descended from a more close common ancestors than JJBC, and that JJBC was evolutionally diverged for a longer time than the breeds in inland Korea and Japan. This result supports that JJBC has been adapted

in Jeju island, to have genetically unique characteristics as one cattle breed. PCA results also showed that Hanwoo individuals formed close clusters with those of the two Wagyu breeds than the JJBC individuals (Figure 3A), supporting the results of phylogenic tree analysis (Figure 2).

JJBC have an effective population size of 38 in the nearest 13 generation ago, whereas Sharma et al [11] reported the $N_e$ estimate of 67. Sudrajad et al [5] reported that $N_e$ in JJBC was estimated to be 60 until 11 generation ago, and Struken et al [2] reported the $N_e$ estimate of 11 at nearest generation. These $N_e$ estimates were much smaller than the $N_e$ of Hanwoo, i.e. 209. Differences of $N_e$ estimates in the various reports might be caused by many factors such as sample size, SNP quality control measures and models used to study LD and $N_e$ [5].

The small $N_e$ value of JJBC seems to be sufficient for maintaining genetic diversity for short term species management as suggested by Frankham et al [47], but not enough at long term level. This result indicates that careful implementation of breed conservation is needed to keep genetic diversity of JJBC while increasing the population, as well as in a breeding program for genetic improvement of economically important traits in beef cattle.

## CONCLUSION

This study supports genetic uniqueness of JJBC breed that has been evolved differently from Hanwoo and Wagyu breeds, as well as from indicine and western taurine breeds, although a small amount of genetic components in terms of allele sharing exists between JJBC and Hanwoo. However, further in-depth study with whole genome sequencing and scanning using high density markers with larger samples would help us to accurately measure genetic parameters for establishment of JJBC as a unique breed.

## AUTHOR CONTRIBUTIONS

JJK conceived, designed the experiment and revised the manuscript critically for important intellectual content. HJS and YML carried out the experiments, laboratory analyses, statistical analyses and interpretation of the data, YML assisted statistical analysis and contributed as part of the laboratory team. MZA wrote the manuscript, equally contributed to the data analysis with HJS and YML. All authors read and approved the final manuscript. All other authors including LHH, DR, HM, SS and SPP contributed by providing DNA sample, carefully read and approved the final manuscript.

## CONFLICT OF INTEREST

We certify that there is no conflict of interest with any financial organization regarding the material discussed in the manu-

script.

## REFERENCES

1. Oldenbroek K,van der Waaij L. Textbook animal breeding: animal breeding and genetics for BSc students. Wageningen, Netherlands: Wageningen University and Research Centre; 2014.

2. Struken 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. https://doi.org/10.1186/s12862-015-0563-2

3. Kim JH, Oh JH, Song JH, et al. Molecular genetic analysis of ancient cattle bones excavated from archaeological sites in Jeju, Korea. Mol Cells 2005;20:325-30.

4. Lee SH, Park BH, Sharma A, et al. Hanwoo cattle: origin, domestication, breeding strategies and genomic selection. J Anim Sci Technol 2014;56:2. https://doi.org/10.1186/2055-0391-56-2

5. Sudrajad P, Seo DW, Choi TJ, et al. Genome-wide linkage disequilibrium and past effective population size in three Korean cattle breeds. Anim Genet 2017;48:85-9. https://doi.org/10.1111/age.12488

6. Kim EY, Song DH, Park MJ, et al. Post-death cloning of endangered Jeju Black cattle (Korean native cattle): fertility and serum chemistry in a cloned bull and cow and their offspring. J Reprod Dev 2013;59:536-43. https://doi.org/10.1262/jrd.2013-047

7. Zanella R, Peixoto JO, Cardoso FF, et al. Genetic diversity analysis of two commercial breeds of pigs using genomic and pedigree data. Genet Sel Evol 2016;48:24. https://doi.org/10.1186/s12711-016-0203-3

8. Li Y, Kim JJ. Effective population size and signatures of selection using bovine 50K SNP chips in Korean native cattle (Hanwoo). Evol Bioinform 2015;11:143-53. https://doi.org/10.4137/EBO.S24359

9. Tenesa A, Navarro P, Hayes BJ, et al. Recent human effective population size estimated from linkage disequilibrium. Genome Res 2007;17:520-6. https://doi.org/10.1101/gr.6023607

10. Groeneveld LF, Lenstra JA, Eding H, et al. Genetic diversity in farm animals – a review. Anim Genet 2010;41:6-31. https://doi.org/10.1111/j.1365-2052.2010.02038.x

11. Sharma A, Lim D, Chai HH, Choi BH, Cho Y. Demographic

trends in Korean native cattle explained using bovine SNP50 beadchip. Genomics Inform 2016;14:230-3. https://doi.org/10.5808/GI.2016.14.4.230

12. Ciani E, Cecchi F, Castellana E, et al. Poorer resolution of low-density SNP vs. STR markers in reconstructing genetic relationships among seven Italian sheep breeds. Large Anim Rev 2013;19:236-41.

13. Fischer MC, Rellstab C, Leuzinger M, et al. Estimating genomic diversity and population differentiation – an empirical comparison of microsatellite and SNP variation in *Arabidopsis halleri*. BMC Genomics 2017;18:69. https://doi.org/10.1186/s12864-016-3459-7

14. Frkonja A, Gredler B, Schnyder U, Curik I, Solkner J. Prediction of breed composition in an admixed cattle population. Anim Genet 2012;43:696-703. https://doi.org/10.1111/j.1365-2052.2012.02345.x

15. 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;81:559-75. https://doi.org/10.1086/519795

16. Kijas JW, Townley D, Dalrymple BP, et al. A genome wide survey of SNP variation reveals the genetic structure of sheep breeds. PLoS One 2009;4:e4668. https://doi.org/10.1371/journal.pone.0004668

17. Keenan K, McGinnity P, Cross TF, Crozier WW, Prodohl PA. diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. Methods Ecol Evol 2013;4:782-8. https://doi.org/10.1111/2041-210X.12067

18. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 2010;10:564-7. https://doi.org/10.1111/j.1755-0998.2010.02847.x

19. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. Evolution 1984;38:1358-70. https://doi.org/10.2307/2408641

20. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Res 2009;19:1655-64. https://doi.org/10.1101/gr.094052.109

21. Edea Z, Dadi H, Kim SW, et al. Genetic diversity, population structure and relationships in indigenous cattle populations of Ethiopia and Korean Hanwoo breeds using SNP markers. Front Genet 2013;4:35. https://doi.org/10.3389/fgene.2013.00035

22. Nei M, Tajima F, Tateno Y. Accuracy of estimated phylogenetic trees from molecular data. J Mol Evol 1983;19:153-70. https://doi.org/10.1007/BF02300753

23. Takezaki N, Nei M, Tamura K. POPTREE2: software for constructing population trees from allele frequency data and computing other population statistics with Windows interface. Mol Biol Evol 2010;27:747-52. https://doi.org/10.1093/molbev/msp312

24. Lewontin RC. The interaction of selection and linkage. I. General considerations; heterotic models. Genetics 1964;49:49-67.

25. Hill WG. Estimation of linkage disequilibrium in randomly mating populations. Heredity 1974;33:229-39. https://doi.org/10.1038/hdy.1974.89

26. Visser C, Lashmar SF, Van Marle-Köster E, Poli MA, Allain D. Genetic diversity and population structure in South African, French and Argentinian Angora goats from genome-wide SNP data. PLoS One 2016;11:e0154353. https://doi.org/10.1371/journal.pone.0154353

27. VanLiere JM, Rosenberg NA. Mathematical properties of the $r^2$ measure of linkage disequilibrium. Theor Popul Biol 2008;74:130-7. https://doi.org/10.1016/j.tpb.2008.05.006

28. Lee SH, Clark S, van der Werf JHJ. Estimation of genomic prediction accuracy from reference populations with varying degrees of relationship. PLoS One 2017;12:e0189775. https://doi.org/10.1371/journal.pone.0189775

29. 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. https://doi.org/10.3389/fgene.2015.00109

30. Corbin LJ, Liu AYH, Bishop SC, Woolliams JA. Estimation of historical effective population size using linkage disequilibria with marker data. J Anim Breed Genet 2012;129:257-70. https://doi.org/10.1111/j.1439-0388.2012.01003.x

31. Sved JA. Linkage disequilibrium and homozygosity of chromosome segments in finite populations. Theor Popul Biol 1971;2:125-41. https://doi.org/10.1016/0040-5809(71)90011-6

32. Vargas J, Landi V, Martinez A, et al. Molecular study of the Amazonian Macabea cattle history. PLoS One 2016;11:e0165398. https://doi.org/10.1371/journal.pone.0165398

33. Browett S, McHugo G, Richardson IW, et al. Genomic characterisation of the indigenous Irish Kerry cattle breed. Front Genet 2018;9:51. https://doi.org/10.3389/fgene.2018.00051

34. Mészáros G, Boison SA, Pérez O'Brien AM, et al. Genomic analysis for managing small and endangered populations: a case study in Tyrol Grey cattle. Front Genet 2015;6:173. https://doi.org/10.3389/fgene.2015.00173

35. Cañas-Álvarez JJ, González-Rodríguez A, Munilla S, et al. Genetic diversity and divergence among Spanish beef cattle breeds assessed by a bovine high-density SNP chip. J Anim Sci 2015;93:5164-74. https://doi.org/10.2527/jas.2015-9271

36. Buzanskas ME, Ventura RV, Seleguim Chud TCS, et al. Study on the introgression of beef breeds in Canchim cattle using single nucleotide polymorphism markers. PLoS One 2017;12:e0171660. https://doi.org/10.1371/journal.pone.0171660

37. Ling X, Zhang W, Li J, et al. Genetic background analysis and breed evaluation of Yiling yellow cattle. J Integr Agric 2017;16:2246-56. https://doi.org/10.1016/S2095-3119(17)61679-4

38. Edea Z, Bhuiyan MSA, Dessie T, Rothschild MF, Dadi H, Kim

KS. Genome-wide genetic diversity, population structure and admixture analysis in African and Asian cattle breeds. Animal 2015;9:218-26. https://doi.org/10.1017/S1751731114002560

39. Eusebi PG, Cortes O, Dunner S, Canon J. Genomic diversity and population structure of Mexican and Spanish bovine Lidia breed. Anim Genet 2017;48:682-5. https://doi.org/10.1111/age.12618

40. Makina SO, Muchadeyi FC, van Marle-Köster E, MacNeil MD, Maiwashe A. Genetic diversity and population structure among six cattle breeds in South Africa using a whole genome SNP panel. Front Genet 2014;5:333. https://doi.org/10.3389/fgene.2014.00333

41. Msalya G, Kim ES, Laisser ELK, et al. Determination of genetic structure and signatures of selection in three strains of Tanzania Shorthorn Zebu, Boran and Friesian cattle by genome-wide SNP analyses. PLoS One 2017;12:e0171088. https://doi.org/10.1371/journal.pone.0171088

42. Karimi K, Strucken EM, Moghaddar N, Ferdosi MH, Esmailizadeh A, Gondro C. Local and global patterns of admixture and population structure in Iranian native cattle. BMC Genet 2016;17:108. https://doi.org/10.1186/s12863-016-0416-z

43. Kelleher MM, Berry DP, Kearney JF, McParland S, Buckley F, Purfield DC. Inference of population structure of purebred dairy and beef cattle using high-density genotype data. Animal 2017;11:15-23. https://doi.org/10.1017/S1751731116001099

44. Signer-Hasler H, Burren A, Neuditschko M, et al. Population structure and genomic inbreeding in nine Swiss dairy cattle populations. Genet Sel Evol 2017;49:83. https://doi.org/10.1186/s12711-017-0358-6

45. Sharma A, Lee SH, Lim D, Chai HH, Choi BH, Cho Y. A genome-wide assessment of genetic diversity and population structure of Korean native cattle breeds. BMC Genet 2016;17:139. https://doi.org/10.1186/s12863-016-0444-8

46. Lin BZ, Sasazaki S, Mannen H. Genetic diversity and structure in *Bos taurus* and *Bos indicus* populations analyzed by SNP markers. Anim Sci J 2010;81:281-9. https://doi.org/10.1111/j.1740-0929.2010.00744.x

47. Frankham R, Ballou JD, Briscoe DA. Introduction to conservation genetics. Cambridge, UK: Cambridge University Press; 2002.