

# Comparison of SNP Variation and Distribution in Indigenous Ethiopian and Korean Cattle (Hanwoo) Populations

Zewdu Edea<sup>1,2</sup>, Hailu Dadi<sup>3</sup>, Sang Wook Kim<sup>2</sup>, Tadelle Dessie<sup>1</sup>, Kwan-Suk Kim<sup>2\*</sup>

<sup>1</sup>International Livestock Research Institute (ILRI), P.O. Box 5689, Addis Ababa, Ethiopia, <sup>2</sup>Chungbuk National University, Cheongju 361-763, Korea, <sup>3</sup>Department of Animal Biotechnology, College of Animal Bioscience and Technology, Konkuk University, Seoul 143-701, Korea

Although a large number of single nucleotide polymorphisms (SNPs) have been identified from the bovine genome-sequencing project, few of these have been validated at large in *Bos indicus* breeds. We have genotyped 192 animals, representing 5 cattle populations of Ethiopia, with the Illumina Bovine 8K SNP BeadChip. These include 1 Sanga (Danakil), 3 zebu (Borana, Arsi and Ambo), and 1 zebu × Sanga intermediate (Horro) breeds. The Hanwoo (*Bos taurus*) was included for comparison purposes. Analysis of 7,045 SNP markers revealed that the mean minor allele frequency (MAF) was 0.23, 0.22, 0.21, 0.21, 0.23, and 0.29 for Ambo, Arsi, Borana, Danakil, Horro, and Hanwoo, respectively. Significant differences of MAF were observed between the indigenous Ethiopian cattle populations and Hanwoo breed ( $p < 0.001$ ). Across the Ethiopian cattle populations, a common variant MAF ( $\geq 0.10$  and  $\leq 0.5$ ) accounted for an overall estimated 73.79% of the 7,045 SNPs. The Hanwoo displayed a higher proportion of common variant SNPs (90%). Investigation within Ethiopian cattle populations showed that on average, 16.64% of the markers were monomorphic, but in the Hanwoo breed, only 6% of the markers were monomorphic. Across the sampled Ethiopian cattle populations, the mean observed and expected heterozygosities were 0.314 and 0.313, respectively. The level of SNP variation identified in this particular study highlights that these markers can be potentially used for genetic studies in African cattle breeds.

**Keywords:** Ethiopian cattle, Hanwoo cattle, minor allele frequency, SNP variation

## Introduction

Ethiopian cattle breeds are the core to sustain the livelihoods of most of the farming and pastoral communities. Indigenous breeds are known for their better adaptation for low input systems, heat tolerance, disease resistance, local preferences, and cultural and aesthetic values. The unique genetic background of Ethiopian cattle and their adaptation to the low input system support their conservation and breed development. However, common for developing countries, the genetic diversity and the genetic merits of most Ethiopian indigenous cattle populations are not yet well understood and exploited. Consequently, some of the indigenous cattle populations are already extinct and

endangered, while the risk status of many of them is unknown. A recent microsatellite based study revealed that most Ethiopian cattle breeds are under threat from uncontrolled breeding practices and are at high risk of becoming genetically homogenous [1].

Studies on genetic diversity of Ethiopian cattle have focused on limited cattle breeds, and all of these studies employed low-density microsatellite, mitochondrial, or Y-chromosome markers [1, 2]. A wide range of molecular markers has been developed and used for molecular characterization of populations. Among these, single nucleotides polymorphisms (SNPs) are the most abundant molecular markers in the genome. They are widely dispersed throughout genomes, with a variable distribution among

Received July 7, 2012; Revised August 17, 2012; Accepted August 20, 2012

\*Corresponding author: Tel: +82-43-261-2547, Fax: +82-43-273-2240, E-mail: kwanskim@chungbuk.ac.kr

Copyright © 2012 by The Korea Genome Organization

© It is identical to the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>).

**Table 1.** Summary of sampled populations and their characteristics with respect to breed group, distribution over ecological zones, and affinity to ethnic communities

Breed/Population	Breed group	Agro-ecology	Community	Production system
Ambo	Small East African zebu	III	Oromo	Mixed crop-livestock
Arsi	Large East African zebu	III	Oromo	Mixed crop-livestock
Borana	Large East African zebu	IV	Oromo	Pastoral
Danakil	Sanga	IV	Afar	Pastoral
Horro	Sanga × zebu	III	Oromo	Mixed crop-livestock

Agro-ecologies: III. Wet highland: tepid to cool wet highlands, very high vegetation, 2,091 m a.s.l, 1,437 mm rain, maximum 24.8°C, minimum 10.1°C; IV. Arid lowland: hot arid lowland plain, very low vegetation, 894 m a.s.l, 404.5 mm rain, maximum 33.2°C, minimum 17.4°C.

species; genotyping results are easy to reproduce in different laboratories and are simple to score. The usefulness of SNPs in population diversity and structure studies has been demonstrated in several studies in livestock species [3-6].

Although a large number of SNPs have been identified from the bovine genome-sequencing project, few of these have been validated at large in *Bos indicus* breeds. Yet, the level of polymorphism in Ethiopian cattle population has not been known and analyzed. Prior to this study, no SNPs have been described for these populations. Hence, we report the first preliminary findings on SNP variation in Ethiopian indigenous cattle populations and Hanwoo breed. This study is part of the collaborative project between the Korean Government, the Ethiopian Institute of Agricultural Research, and the International Livestock Research Institute, designed with the objectives to estimate genetic diversity between and within populations using SNP markers. We also evaluated the level of SNP polymorphisms in selected Ethiopian cattle populations and compared to the Hanwoo cattle breed.

## Methods

### Sampling and DNA sample collection

Nasal samples were collected using Performagene LIVESTOCK's nasal swab DNA collection kit (DNA Genotek Inc., Kanata, ON, Canada) from 5 cattle populations-Ambo, Borana, Arsi, Horro, and Danakil-inhabiting different agro-ecologies. Sampling was carried out by considering potential geo-environmental gradients (highland, lowlands), production systems (mixed-crop livestock, pastoral/agro-pastoral), and ethnic groups raising predominantly the populations (Table 1, Fig. 1). Representative sites were selected based upon secondary data obtained from livestock departments and expertise. Herdsmen and owners of the animals were contacted to ensure that the animals were not genetically related. Animals were randomly sampled from multiple herds. Genomic DNA was extracted from nasal

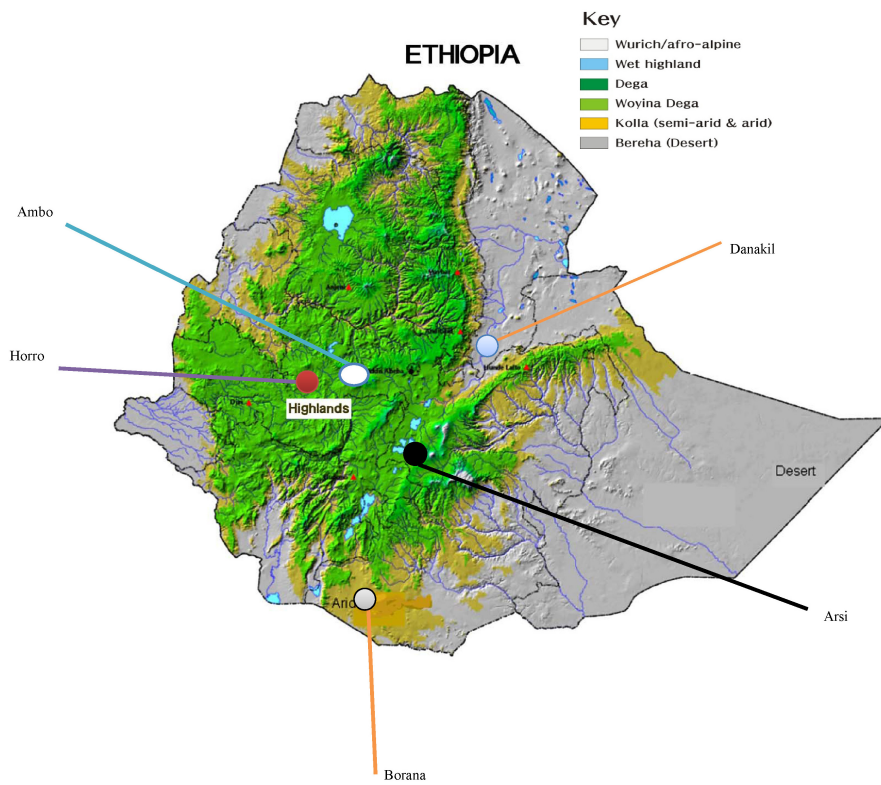
samples based on laboratory protocols. In addition, DNA samples from Hanwoo cattle were included for comparison purposes.

### Breeds/populations descriptions

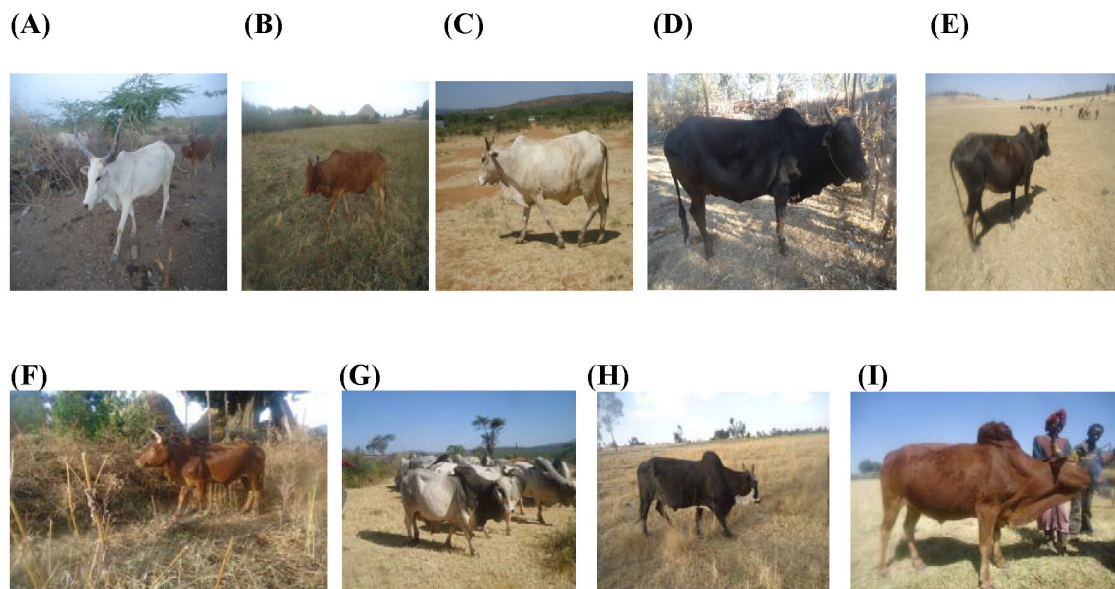
The Danakil cattle breed is grouped under Sanga and originated in East and Northeast Africa, where Sanga cattle evolved. The Horro breed was formed as a result of interbreeding between zebu-Sanga and Sanga-zebu admixture populations. The large-sized Borana breed descended from secondary cattle domestication in the arid areas of the 'Fertile Crescent' about 5000 BP [7]. The breed is predominantly distributed in arid and semi-arid agro-ecologies of the Borana plateau. The small-sized Arsi cattle probably developed from a group of small short-horned Abyssinian zebu by the highland Oromo people. The Ambo cattle is grouped under Small East African zebu and is believed to have descended from the recent introductions of zebu into Africa [8]. Pictures of sampled animals are depicted in Fig. 2.

### Genotyping and quality control

In total, 192 animals representing 5 cattle populations of Ethiopia were genotyped for 8,773 SNPs with the Illumina Bovine 8K SNP BeadChip (Illumina, San Diego, CA, USA) according to Illumina protocols [9]. Over the loci, the average GenTrain score, 10% GenCall (10% GC) score, and 50% GenCall (50% GC) score were 0.86, 0.87, and 0.83, respectively. In this experiment, about 99.91% of the markers identified had a GenTrain score greater than the minimum acceptable value (0.25) (<http://www.illumina.com>). All of the samples had at least a 0.20 or greater 10% GC score, while the 50% GC score ranged from 0.75 to 0.91, with the average call rate of 0.913. For Hanwoo cattle, about 51,162 SNPs were genotyped using the Illumina Bovine 50K SNP BeadChip [10]. 8K and 50K genotypes were merged using Golden Helix SNP Variation Suite software version 7 (Golden Helix, Inc., Bozeman, MT, USA). To ensure the highest quality of data from the 8K and 50K SNP BeadChips,



**Fig. 1.** Map of Ethiopia showing location of sampled populations.



**Fig. 2.** Pictures of sampled Ethiopian cattle populations. (A) Danakil adult female. (B) Horro adult female. (C) Borana adult female. (D) Arsi adult female. (E) Ambo adult female. (F) Horro adult male. (G) Borana adult males. (H) Arsi adult male. (I) Ambo adult male.

duplicate marker genotypes were identified and dropped. Finally about 7,045 SNPs that were common for both Bead-Chips were subjected to analysis. Markers selected for diversity analysis were all located on autosomal chromosomes and had at least an 80% call rate.

### Data analysis

Frequency of minor allele frequency (MAF), distribution, and deviation from Hardy-Weinberg equilibrium (HWE) were estimated using Golden Helix SNP Variation Suite ver-

sion 7 (Golden Helix, Inc.). The SAS GLM package version 9.1 (SAS Institute Inc., Cary, NC, USA) procedure was used to test the effect of breed on MAF variation. Observed and expected heterozygosities and hierarchical F-statistics were estimated using GENALEX software [11].

## Results and Discussion

### MAF variation

The MAF for each SNP was calculated from the genotypic data and is presented in Table 2. Analysis of 7,045 markers revealed that the mean MAF was 0.23, 0.22, 0.21, 0.21, 0.23, and 0.29 for Ambo, Arsi, Borana, Danakil, Horro, and Hanwoo cattle, respectively. A significant difference was observed between the Ethiopian indigenous cattle populations and Hanwoo ( $p < 0.001$ ). The average value for Hanwoo was 32% higher than the value for Ethiopian cattle populations. The overall MAF in the present finding for indigenous Ethiopian cattle breeds is higher than an earlier report for indicine breeds but lower than the average value reported for most taurine breeds [3]. The value found for Hanwoo cattle was comparable with the previous results for taurine breeds but higher than the reported average for the

same breed [10, 12]. This could be attributed to the different marker density used, and a higher density marker (Illumina Bovine 50K SNP BeadChip) was used in the previous study. The higher values for taurine breeds can be explained by the fact that SNP loci used in this study were detected in taurine breeds, and their average minor allele frequency was much lower in indicine breeds.

Across the Ethiopian cattle populations, common variant MAFs ( $\geq 0.10$  and  $\leq 0.5$ ) accounted for 74% of the total SNPs (Table 3). Hanwoo cattle displayed a higher proportion of common variant SNPs (90%). Rare variant MAFs were observed in about 9% and 4% of markers in the Ethiopian and Hanwoo breeds, respectively. The relatively higher proportion of fixed alleles in Ethiopia cattle populations could be attributed to the inbreeding as a result of uncontrolled breeding practices. A common variant MAF proportion of 83% was reported in sheep breeds [5]. The distribution of MAFs revealed that nearly 34.08% of the markers in Ethiopian cattle populations and 53% in Hanwoo cattle displayed a high degree of polymorphisms ( $0.3 \leq \text{MAF} \leq 0.5$ ). Yet, this value was higher than an earlier report both for Hanwoo (25%) and Holstein (28%) [12]. Much higher MAFs were detected in *Bos taurus* breeds than in *Bos indicus* breeds [3]. The proportion of fixed markers varied among the sampled populations. Among the Ethiopian cattle populations, the highest (8.96%) and lowest (6.30%) were obtained in the Borana and Ambo populations, respectively. The Hanwoo breed displayed a lower level of fixed SNPs (1.10%).

On average, 83.36% of the markers displayed polymorphisms ( $\text{MAF} > 0.05$ ) within Ethiopian cattle populations, and the remaining 16.64% of the SNPs considered as monomorphic. The Hanwoo breed showed the highest proportion of polymorphisms (94%) (Fig. 3). The level of polymorphic SNPs in the present study was higher than what was previously reported for taurine breeds [10]. But, it was in close agreement with the value reported for sheep breeds

**Table 2.** Mean MAF values in indigenous Ethiopian cattle populations and in Korea Hanwoo cattle

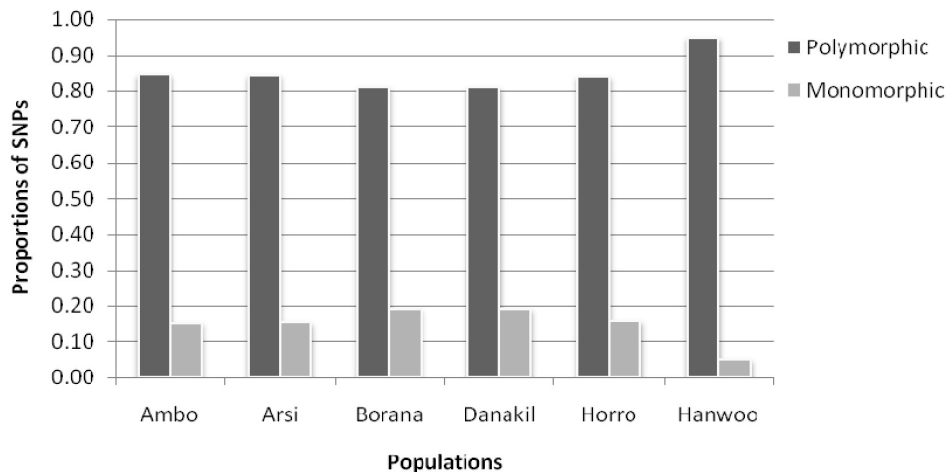
Breed/Population	No.	Mean $\pm$ SD
Ethiopian cattle		
Ambo	27	0.23 $\pm$ 0.148
Arsi	30	0.22 $\pm$ 0.148
Borana	35	0.21 $\pm$ 0.150
Danakil	38	0.21 $\pm$ 0.150
Horro	36	0.23 $\pm$ 0.148
Overall	<b>166</b>	<b>0.22 <math>\pm</math> 0.149</b>
<i>Bos taurus</i>		
Hanwoo	195	0.29 $\pm$ 0.135

MAF, minor allele frequency.

**Table 3.** Minor allele frequency distribution of 8K BeadChip SNP in *Bos taurus* and in Ethiopian cattle breeds

Population	No.	Fixed (0)		Rare ( $>0$ and $<0.05$ )		Intermediate ( $\geq 0.05$ and $<0.10$ )		Common ( $\geq 0.10$ and $\leq 0.50$ )		$\geq 0.30$ and $\leq 0.5$	
		SNP	Prop.	SNP	Prop.	SNP	Prop.	SNP	Prop.	SNP	Prop.
Ambo	27	444	0.06	627	0.09	711	0.10	5,262	0.75	2,685	0.38
Arsi	30	479	0.07	608	0.09	708	0.10	5,249	0.76	2,392	0.34
Borana	35	631	0.09	692	0.09	643	0.90	5,078	0.72	2,298	0.33
Danakil	38	600	0.09	668	0.09	821	0.12	4,958	0.70	2,230	0.32
Horro	36	471	0.07	633	0.09	601	0.09	5,340	0.76	2,494	0.35
Overall	<b>166</b>	<b>525</b>	<b>0.08</b>	<b>645</b>	<b>0.09</b>	<b>696</b>	<b>0.10</b>	<b>5,177</b>	<b>0.74</b>	<b>2,419</b>	<b>0.34</b>
Hanwoo	195	79	0.01	247	0.04	381	0.05	6,340	0.90	3,742	0.53

SNP, single nucleotide polymorphism.



**Fig. 3.** Polymorphic and monomorphic dynamics of single nucleotide polymorphism (SNPs).

**Table 4.** Observed ( $H_O$ ) and expected heterozygosities ( $H_E$ ), mean number of different alleles ( $N_a$ ), mean number of effective alleles ( $N_e$ ), Shannon's information index ( $I$ ) and fixation index ( $F$ ) in indigenous Ethiopian and Korean cattle populations

Population	No.	$H_O$	$H_E$	$N_a$	$N_e$	$I$	$F$
Ambo	27	0.325	0.321	1.963	1.545	0.484	-0.010
Arsi	30	0.313	0.315	1.959	1.532	0.475	0.004
Borana	35	0.309	0.304	1.943	1.512	0.461	-0.016
Danakil	38	0.299	0.303	1.946	1.509	0.459	0.011
Horro	36	0.322	0.321	1.962	1.544	0.483	-0.006
Overall	166	0.314	0.313	1.955	1.529	0.472	-0.003
Hanwoo	195	0.399	0.394	2.00	1.700	0.567	-0.012

using genome-wide SNPs [5]. Similarly, Gautier *et al.* [13] reported a lower rate of polymorphisms using 696 SNPs localized in BTA03 from European and West African cattle breeds (94.6% and 81.8%), respectively. The higher observed polymorphisms in Hanwoo cattle could be explained by the fact that most of the bovine sequence data available in the database are from *Bos taurus* breeds.

### Genetic diversity among Ethiopian cattle populations

An examination of the variability within breeds was carried out to compare the level of heterozygosity between the populations, and the results are presented in Table 4. Across the sampled Ethiopian cattle populations, the mean observed and expected heterozygosities were  $0.314 \pm 0.001$  and  $0.313 \pm 0.001$ , respectively. The mean values were lower than the values reported from a microsatellite markers analysis [1]. However, they were in concurrence with the values reported using SNPs for the Angus breed (0.332) [14] and for sheep breeds [5].

The overall estimated  $F$ -statistics were  $F_{IS} = -0.003$  (within-population inbreeding estimates) and  $F_{IT} = 0.016$  (total inbreeding). Among the Ethiopian cattle populations, genetic differentiation ( $F_{ST}$ ) was estimated at 0.019. Ge-

netic characterization of Ethiopian cattle breeds using microsatellites reported within-population inbreeding of 0.071, total inbreeding of 0.083, and genetic differentiation of 0.013 [1]. The  $F_{ST}$  for Ethiopian cattle populations was low, which could be due to their common ancestral origin.

There were loci significantly deviating from HWE. Overall, significant deviation from Hardy-Weinberg proportions ( $p < 0.05$ ) was observed for 405 markers in Ethiopian cattle populations (Ambo, 441; Arsi, 475; Borana, 314; Danakil, 314; and Horro, 481) and in 549 of the markers in Hanwoo. The highest number of SNPs that departed from HWE was observed in Hanwoo cattle, whereas within the Ethiopia cattle populations, Borana displayed lower proportions of markers (SNPs) deviating from HWE. The variation in proportion of SNPs displaying deviation from HWE among the sampled populations could be explained by selection pressure and population structure (stratification, admixture).

In conclusion, a significant difference was detected for the rate of polymorphisms and MAFs between the Hanwoo and the Ethiopian cattle populations. The level of variation identified in this particular study highlights that these markers can be potentially used for genetic studies in African cattle populations. The high within-breed diversity observed

in Ethiopian indigenous cattle populations could be an opportunity and allow improvements through within-population selection schemes.

## Acknowledgments

This study was funded by the Africa & International Organization Programs under the National Research Foundation of Korea and the Ministry of Education and Human Resource Development. Participant framers and pastoralists are highly thanked for their willingness to allow us to collect samples from their animals and for their indisputable cooperation. We are grateful to officials from livestock departments in Ethiopia for their guidance and facilitation of the sample collection process. We would like to express our gratitude to ILRI-Addis for providing logistical support.

## References

1. Dadi H, Tibbo M, Takahashi Y, Nomura K, Hanada H, Amano T. Microsatellite analysis reveals high genetic diversity but low genetic structure in Ethiopian indigenous cattle populations. *Anim Genet* 2008;39:425-431.
2. Li MH, Zerabruk M, Vangen O, Olsaker I, Kantanen J. Reduced genetic structure of north Ethiopian cattle revealed by Y-chromosome analysis. *Heredity (Edinb)* 2007;98:214-221.
3. McKay SD, Schnabel RD, Murdoch BM, Matukumalli LK, Aerts J, Coppieters W, *et al.* An assessment of population structure in eight breeds of cattle using a whole genome SNP panel. *BMC Genet* 2008;9:37.
4. Williams JL, Dunner S, Valentini A, Mazza R, Amarger V, Checa ML, *et al.* Discovery, characterization and validation of single nucleotide polymorphisms within 206 bovine genes that may be considered as candidate genes for beef production and quality. *Anim Genet* 2009;40:486-491.
5. Kijas JW, Townley D, Dalrymple BP, Heaton MP, Maddox JF, McGrath A, *et al.* A genome wide survey of SNP variation reveals the genetic structure of sheep breeds. *PLoS One* 2009;4:e4668.
6. 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;81:281-289.
7. Payne WJA, Wilson RT. *An Introduction to Animal Husbandry in the Tropics*. Oxford: Blackwell Science Ltd., 1999.
8. Kemp S, Mamo Y, Asrat B, Dessie T. Domestic animal genetic resources information system. Addis Ababa: International Livestock Research Institute, 2007. Accessed 2012 Jun 20. Available from: <http://dagris.ilri.cgiar.org>.
9. Boichard D, Chung H, Dasonneville R, David X, Eggen A, Fritz S, *et al.* Design of a bovine low-density SNP array optimized for imputation. *PLoS One* 2012;7:e34130.
10. Dadi H, Kim JJ, Yoon D, Kim KS. Evaluation of single nucleotide polymorphisms (SNPs) genotyped by the Illumina Bovine SNP50K in cattle focusing on Hanwoo breed. *Asian-Aust J Anim Sci* 2012;25:28-32.
11. Peakall R, Smouse PE. GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 2006;6:288-295.
12. Melka HD, Jeon EK, Kim SW, Han JB, Yoon D, Kim KS. Identification of genomic differences between Hanwoo and Holstein breeds using the Illumina Bovine SNP50 BeadChip. *Genomics Inform* 2011;9:69-73.
13. Gautier M, Faraut T, Moazami-Goudarzi K, Navratil V, Foglio M, Grohs C, *et al.* Genetic and haplotypic structure in 14 European and African cattle breeds. *Genetics* 2007;177:1059-1070.
14. Carruthers CR, Plante Y, Schmutz SM. Comparison of Angus cattle populations using gene variants and microsatellites. *Can J Anim Sci* 2011;91:81-85.