



## Evaluation of Single Nucleotide Polymorphisms (SNPs) Genotyped by the Illumina Bovine SNP50K in Cattle Focusing on Hanwoo Breed

Hailu Dadi, Jong-Joo Kim<sup>1</sup>, Duhak Yoon<sup>2</sup> and Kwan-Suk Kim\*

Department of Animal Science, Chungbuk National University, Cheongju, Chungbuk 361-763, Korea

**ABSTRACT :** In the present study, we evaluated the informativeness of SNPs genotyped by the Illumina Bovine SNP50K assay in different cattle breeds. To investigate these on a genome-wide scale, we considered 52,678 SNPs spanning the whole autosomal and X chromosomes in cattle. Our study samples consists of six different cattle breeds. Across the breeds approximately 72 and 6% SNPs were found polymorphic and fixed or close to fix in all the breeds, respectively. The variations in the average minor allele frequency (MAF) were significantly different between the breeds studied. The level of average MAF observed in Hanwoo was significantly lower than the other breeds. Hanwoo breed also displayed the lowest number of polymorphic SNPs across all the chromosomes. More importantly, this study indicated that the Bovine SNP50K assay will have reduced power for genome-wide association studies in Hanwoo as compared to other cattle breeds. Overall, the Bovine SNP50K assay described in this study offer a useful genotyping platform for mapping quantitative trait loci (QTLs) in the cattle breeds. The assay data represent a vast and generally untapped resource to assist the investigation of the complex production traits and the development of marker-assisted selection programs. (**Key Words :** Hanwoo Breed, Polymorphic SNPs, Fixed SNPs, Minor Allele Frequency)

### INTRODUCTION

The application of high density throughput genotyping array opens an opportunity for the discovery of genetic variants, known as SNPs. SNPs are nucleotide variations in the DNA sequence of individuals in a population and constitute the most abundant molecular markers in the genome. There are several millions of SNPs embedded in the genome of domestic livestock (Bovine Genome Sequencing and Analysis Consortium, 2009; Zimin et al., 2009). However, only a very small subset of these have been developed into genetic markers for population genetics, genomic selection, linkage disequilibrium pattern and genome-wide association studies.

SNPs are very important due to their quantity and their association to QTLs or in reshaping the proteins produced by QTLs that determine cellular functions (Falconer and Mackay, 1997; Graur and Li, 2000; Micklos et al., 2003).

With a close proximity to QTLs, SNPs are known either as a single or possibly as multiple in combination to reflect a QTL (Graur and Li, 2000; Micklos et al., 2003). Identification of SNP markers linked to QTLs, and implementing marker assisted selection become a contemporary cost-effective breeding programme. The prediction of accurate genomic breeding values from dense SNP markers dramatically changed the landscape of genetic improvement strategy (Goddard, 2008). For instance, genomic selection is underway in at least four dairy breeding programs around the world (Hayes et al., 2009). Hence, it seems that SNPs are increasingly becoming the dominant marker of choice in the field of genetics (Gupta et al., 2001).

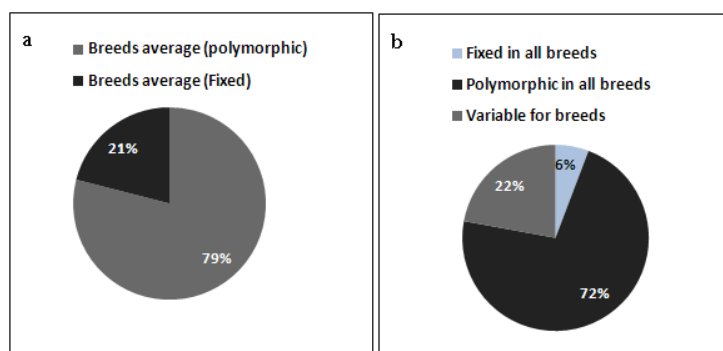
We assume that at some future time genomic selection will become practical in several breeds around the world. Therefore, it appears that understanding the extent of informativeness of SNPs in different cattle breeds is an entry to animal breeding. For example, it was suggested that the Bovine SNP50 assay will have reduced power in indicine and African cattle breeds in revealing genome-wide association studies (Matukumalli et al., 2009). Hence, we evaluated the informativeness of SNPs genotyped by the Bovine SNP50K in different cattle breeds, focusing on the Korean Hanwoo breed.

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

<sup>1</sup> School of Biotechnology, Yeungnam University, Gyeongsan, Gyeongbuk 712-749, Korea.

<sup>2</sup> Department of Animal Science, Kyungbook National University, Sangju, Kyungbook 742-711, Korea.

Received July 21, 2011; Accepted September 18, 2011



**Figure 1.** (a) average polymorphic and fixed SNPs in the breeds, (b) degree of polymorphic and fixed SNPs sharing among the six cattle breeds; SNP with any missing frequency was excluded from the analysis.

## MATERIALS AND METHODS

Blood samples were obtained from Hanwoo and Holstein Nonghyup Animal Genetic Improvement Programs. For each breed, 50 samples were collected and DNA was extracted according to the laboratory procedure. Technical details of the experiment were previously described by Melka et al. (2011). Additionally, the available SNPs frequency data of Charolais, Angus, Simmental and Limousine breeds were obtained from Decker et al. (2009).

### Data analysis

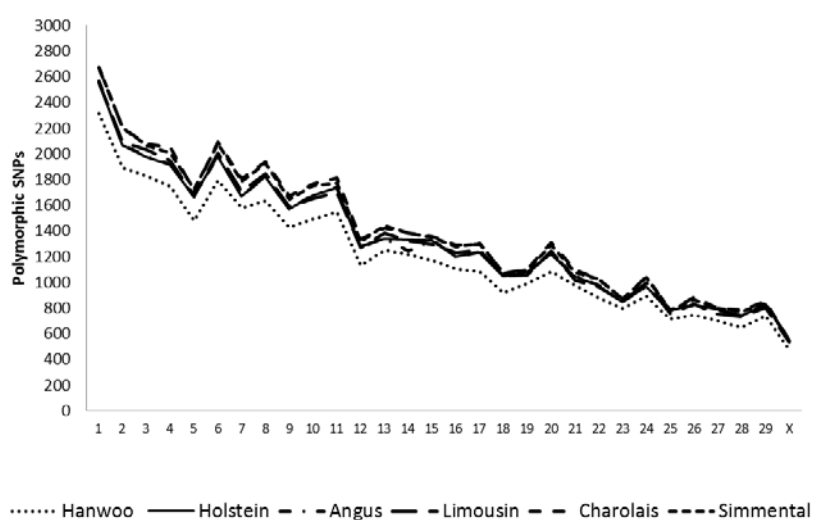
The Normalized bead intensity data for each sample were loaded into the Beadstudio 3.0 software (Illumina Inc., San Diego, CA, USA), which converted fluorescent intensities into SNP genotypes. SNPs clusters for genotype calling were examined for all SNPs using Beadstudio 3.0 software. SNP frequency variations were examined by using SPSS package version 17.0 program. Breed average, rare, intermediate and common allele frequencies were

computed and significance was tested by Chi-square.

## RESULTS

### Fixed and polymorphic SNPs in cattle

The rate of polymorphic and fixed SNPs observed in six cattle breeds are summarized in Figure 1a and b. On the average, 79% of the SNPs were polymorphic ( $\geq 0.05$ - $\leq 0.95$ ) for six breeds. Among the polymorphic, 72% (37,370) SNPs were found to be polymorphic across the whole breeds with variable or similar allele frequencies for SNPs. On the other hand, an average of 21% of the SNPs were fixed or close to fix ( $< 0.05$  or  $> 0.95$ ) in the studied breeds; among fixed SNPs, 6% (2,990) was common to all breeds. About, 22% (11,454) of SNPs were variable for breeds, which mean that some SNPs were fixed in some breeds but polymorphic in the other breeds. Figure 2 depicts chromosomal level variations of SNPs, Hanwoo breed displayed the lowest number of polymorphic SNPs, but had the highest proportion of fixed SNPs (data not shown).



**Figure 2.** Polymorphic SNPs dynamics across the chromosomes of different cattle breeds, unknown chromosomes were excluded (Polymorphic,  $\geq 0.05$ - $\leq 0.95$ ).

**Table 1.** Minimum, maximum and mean values for minor allele frequencies observed in each breed

Breeds	n (%)	Allele frequency			
		Minimum	Maximum	Average	Std. deviation
Hanwoo (MAF)	29,840(56)	0.000	0.500	0.191 <sup>a</sup>	0.161
Zero	4,837(9)				
Rare	3,767(7)	0.010	0.040	0.021	0.011
Intermediate	2,654(5)	0.052	0.098	0.069	0.014
Common	18,582(35)	0.100	0.500	0.292	0.119
Angus (MAF)	29,033(55)	0.000	0.500	0.212 <sup>b</sup>	0.159
Zero	3,694(7)				
Rare	2,934(5.5)	0.010	0.040	0.021	0.011
Intermediate	2,510(4.7)	0.05	0.090	0.070	0.014
Common	19,895(37.7)	0.100	0.500	0.297	0.114
Limousine (MAF)	29,241(55)	0.000	0.500	0.210 <sup>c</sup>	0.157
Zero	3,228 (6)				
Rare	3,195(6)	0.010	0.040	0.022	0.011
Intermediate	2,745(5)	0.050	0.090	0.069	0.014
Common	20,074(38)	0.100	0.500	0.293	0.115
Charolais (MAF)	29,558(56)	0.000	0.500	0.224 <sup>d</sup>	0.155
Zero	2,715(5)				
Rare	2,634(5)	0.010	0.040	0.023	0.011
Intermediate	2,691(5)	0.050	0.090	0.070	0.013
Common	21,518(40)	0.100	0.500	0.295	0.117
Simmental (MAF)	29,271(55)	0.000	0.500	0.223 <sup>e</sup>	0.151
Zero frequency	2,263(4)				
Rare	2,907(5.5)	0.010	0.040	0.022	0.011
Intermediate	2,720(5)	0.050	0.090	0.069	0.014
Common	21,381(40)	0.100	0.500	0.294	0.115
Holstein (MAF)	29,410(55.8)	0.000	0.500	0.215 <sup>f</sup>	0.160
Zero	1,620(3)				
Rare	4,861(9)	0.010	0.040	0.021	0.010
Intermediate	2,552(4.8)	0.050	0.099	0.069	0.013
Common	20,377(38.6)	0.100	0.500	0.298	0.117

\*\*\* Significantly different at  $p < 0.001$  within the column, Rare,  $>0 < 0.05$ ; intermediate,  $\geq 0.05 < 0.10$ ; Common,  $\geq 0.1 < 0.50$ .

These fixed SNPs accounted for the reduced percentage of informative markers within Hanwoo. The pattern of polymorphic SNPs distribution on chromosomes was homogenous across the breeds.

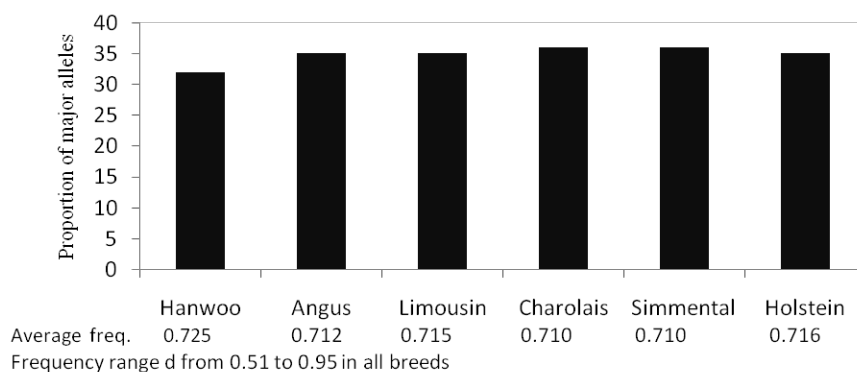
#### Minor allele frequencies (MAFs)

The lowest number of polymorphic SNPs observed across the chromosomes in Hanwoo was further investigated to determine if the Bovine SNP50K assay had low MAF within Hanwoo breed. The estimated values for different MAF classes for different breeds are presented in Table 1. The level of average MAF observed in Hanwoo was significantly lower than the other breeds. Both average MAF and the number of informative SNP markers were

reduced in Hanwoo breed. The proportion of SNPs with zero MAF was highest in Hanwoo, reflecting the presence of single fixed alleles. On the other hand, all the breeds had almost similar values for intermediate and common MAF classes.

The major alleles on the average accounted for 35% of the total SNPs, ranged from 32% in Hanwoo to 36% in Charolais and Simmental (Figure 3). The average allele frequency for Hanwoo was 0.725, with the average of 0.71 across the breeds. The proportion of major alleles in Hanwoo was the lowest, but allele frequency was slightly higher. Several markers tend to had higher allele frequencies in Hanwoo.

In comparison with each breed, the number of SNPs



**Figure 3.** Proportion of major alleles and their corresponding average allele frequencies for each cattle breed.

observed only in Hanwoo is depicted in Table 2. In total, about 530 SNPs were found only in Hanwoo when compared to each breed. The number of these markers varied from 15 for Simmental to 207 for Charolais. The number of Hanwoo specific SNPs detected from this assay data was very small (ARS-BFGL-NGS-28865, Hapmap43053-BTA-79375, ARS-BFGL-NGS-3599, ARS-BFGL-NGS-68165, Hapmap51464-BTA-51997, ARS-BFGL-NGS-48933, DIAS-188). Indeed, it needs more investigation in large sample sizes for robust inference. These variants were found at frequencies ranged from 1.1 to 21.3% in Hanwoo.

## DISCUSSION

The Illumina's chips are rich source of SNPs for different applications in livestock, plants and humans. The most promising applications in livestock could be genome-wide association studies and genomic selection for improvement of meat and milk production. The discoveries of many thousands of SNPs (DNA markers) lead to the accurate computation of genomic breeding values for genomic selection in the livestock industry (Meuwissen et al., 2001; VanRaden et al., 2009). Several SNPs were genotyped by the Illumina Bovine SNP50K chip in New Zealand, United States and Australian for breeds improvement (Hayes et al., 2009). The efforts of these countries can offer important lessons, and holds great promise for improving production and saving money and

time in livestock industry.

In this study, we observed differences between the cattle breeds for MAFs. The difference in allele frequencies among the breeds may attributable to recent genetic drift and those arising from more ancient divergence of the breeds. A higher proportion of the SNPs that were polymorphic in Hanwoo breed had lower allele frequencies. Significantly, lowest average MAF was noted within Hanwoo compared to other breeds. The implication is that the utility of the Bovine SNP50K assay will have reduced power for genome-wide association studies within Hanwoo breed. Despite that the Bovine SNP50K assay will be useful for implementing genome wide association studies to identify QTL regions and improve production traits in all breeds studied. In another study, relatively higher average MAFs for Holstein, Limousine, Angus, Simmental and Charolais were reported by Matukumalli et al. (2009). Major allele frequencies are also essentially equivalent to MAFs in information content for different applications. However, markers with allele frequency >0.8 have reduced information content (Kruglyak, 1997).

Genetic variation within or among breeds is usually explained in terms of allele frequencies. Most SNP markers have alleles which were shared among the breeds, or have breed specific alleles at low frequencies as we have seen in this study. With regard to Hanwoo specific SNPs, seven SNPs were found in Hanwoo, not detected in other breeds. A number of other alleles were found only in Hanwoo in comparison to each breed. Markers private to Hanwoo most likely arose following breeds divergence and that all SNPs common to all breeds arose before the divergence of the breeds. The influence of rare alleles on most quantitative traits has not been systematically evaluated in livestock species. Even so, the significant influences of rare alleles on economically important traits have been well reported in livestock species (Grobet et al., 1997; McPherron and Lee, 1997; Freking et al., 2002; Smit et al., 2003). In the case of human disease, rare alleles carrying significant disease risks have been also reported. For example, there is evidence that

**Table 2.** The number of private SNPs found in Hanwoo compared to each breed genotyped by the Bovine SNP50K assay

Breed	Number of SNP	Alleles frequency in Hanwoo		
		Minimum	Maximum	Mean
Angus	91	0.01	0.280	0.05
Limousine	39	0.01	0.278	0.06
Charolais	207	0.01	0.406	0.047
Simmental	15	0.01	0.213	0.05
Holstein	178	0.01	0.277	0.042

rare alleles contribute significantly to low plasma levels of HDL-C, a common quantitative trait in human (Cohen et al., 2004). To summarize, the Bovine SNP50K described in this study offer a vital genotyping platform for mapping QTL in cattle breeds. The SNPs data represent a vast and largely untapped resource to assist the investigation of several complex QTLs in cattle, and also useful for cattle genetic improvement programs. The pattern of allele frequency variability observed among the breeds signal the genetic imprint of past and presumably ongoing episodes of natural and artificial selection.

### ACKNOWLEDGEMENT

This research was supported by grants titled “Development of a DNA kit for selection on growth and meat quality at early stages in Hanwoo” and “Development of production technologies for high quality & nutritional values of beef in Hanwoo” from the Technology Development Program for Agriculture and Forestry, Ministry for Food, Agriculture, Forestry and Fisheries, Republic of Korea.

### REFERENCES

- Bovine Genome Sequencing and Analysis Consortium, C. G. Elsik, R. L. Tellam and K. C. Worley. 2009. The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science* 324:522-528.
- Cohen, J. C., R. S. Kiss, A. Pertsemlidis, Y. L. Marcel, R. McPherson and H. H. Hobbs. 2004. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science* 305:869-872.
- Decker, J. E., J. C. Pires, G. C. Conant, S. D. McKay, M. P. Heaton, K. Chen, A. Cooper, J. Vilkki, C. M. Seabury, A. R. Caetano, G. S. Johnson, R. A. Breneman, O. Hanotte, L. S. Eggert, P. Wiener, J. J. Kim, K. S. Kim, T. S. Sonstegard, C. P. Van Tassell, H. L. Neibergs, J. C. McEwan, R. Brauning, L. L. Coutinho, M. E. Babar, G. A. Wilson, M. C. McClure, M. M. Rolf, J. Kim, R. D. Schnabel and J. F. Taylor. 2009. Resolving the evolution of extant and extinct ruminants with high-throughput phylogenomics. *Proc. Natl. Acad. Sci. USA*. 106: 18644-18649.
- Falconer, D. S. and T. F. C. Mackay. 1997. Introduction to quantitative genetics. 4<sup>th</sup> Edition. Longman limited.
- Freking, B. A., S. K. Murphy, A. A. Wylie, S. J. Rhodes, J. W. Keele, K. A. Leymaster, R. L. Jirtle and T. P. Smith. 2002. Identification of the single base change causing the callipyge muscle hypertrophy phenotype, the only known example of polar overdominance in mammals. *Genome Res.* 12:1496-1506.
- Goddard, M. E. 2008. Genomic selection: prediction of accuracy and maximization of long term response. *Genetica* 136:245-257.
- Graur, D. and Wen Hsiun Li. 2000. Fundamentals of molecular evolution. 2<sup>nd</sup> ED. Li. Sinauer Association inc., Sunderland, MA.
- Grobet, L., L. J. Martin, D. Poncelet, D. Pirottin, B. Brouwers, J. Riquet, A. Schoeberlein, S. Dunner, F. Méniéssier, J. Massabanda, R. Fries, R. Hanset and M. Georges. 1997. A deletion in the bovine myostatin gene causes the double-muscling phenotype in cattle. *Nat. Genet.* 17:71-74.
- Gupta, P. K., J. K. Roy and M. Prasad. 2001. Single nucleotide polymorphisms: a new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. *Curr. Sci.* 80:524-535.
- Hayes, B. J., P. J. Bowman, A. J. Chamberlain and M. E. Goddard. 2009. Invited review: Genomic selection in dairy cattle: progress and challenges. *J. Dairy Sci.* 92:433-443.
- Kruglyak, L. 1997. The use of a genetic map of biallelic markers in linkage studies. *Nat. Genet.* 17:21-24.
- Matukumalli, L. K., C. T. Lawley, R. D. Schnabel, J. F. Taylor, M. F. Allan, M. P. Heaton, J. O'Connell, S. S. Moore, T. P. Smith, T. S. Sonstegard and C. P. Van Tassell. 2009. Development and characterization of a high density SNP genotyping assay for cattle. *PLoS One.* 4(4):e5350.
- McPherron, A. C. and S. J. Lee. 1997. Double muscling in cattle due to mutations in the myostatin gene. *Proc. Natl. Acad. Sci. USA.* 94:12457-12461.
- Melka, H. D., E. Y. Jeon, S. W. Kim, J. B. Han, D. Yoon and S. K. Kim. 2011. Identification of genomic difference between Hanwoo and Holstein breeds using the Illumina Bovine SNP50K BeadChip. *Genomics & Informatics* 9:69-73.
- Meuwissen, T. H. E., B. J. Hayes and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819-1829.
- Micklos, D. A., G. A. Freyer and D. A. Crotty. 2003. DNA science, A first course. 2<sup>nd</sup> Edition. Cold Spring Harbor Laboratory Press.
- Smit, M., K. Segers, L. G. Carrascosa, T. Shay, F. Baraldi, G. Gyapay, G. Snowden, M. Georges, N. Cockett and C. Charlier. 2003. Mosaicism of Solid Gold supports the causality of a noncoding A-to-G transition in the determinism of the callipyge phenotype. *Genetics* 163:453-456.
- VanRaden, P. M., C. P. Van Tassell, G. R. Wiggans, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor and F. S. Schenkel. 2009. Invited review: Reliability of genomic predictions for North American Holstein bulls. *J. Dairy Sci.* 92:16-24.
- Zimin, A. V., A. L. Delcher, L. Florea, D. R. Kelley, M. C. Schatz, D. Puiu, F. Hanrahan, G. Perlea, C. P. Van Tassell, T. S. Sonstegard, G. Marçais, M. Roberts, P. Subramanian, J. A. Yorke and S. L. Salzberg. 2009. A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome Biol.* 10:R42.