ARTICLE
Genomics & Informatics
Vol. 9(2) 69-73, June 2011

DOI: 10.5808/GI.2011.9.2.69
pISSN 1598-866X eISSN 2234-0742

Identification of Genomic Differences between Hanwoo and Holstein Breeds Using the Illumina Bovine SNP50 BeadChip

Hailu Dadi Melka[1¶], Eun-Kyeong Jeon[1¶], Sang-Wook Kim[1], James-Bond Han[1], Duhak Yoon[2] and Kwan-Suk Kim[1]*

[1]Department of Animal Science, Chungbuk National University, Cheongju 361-763, Korea, [2]Department of Animal Science, Kyungbook National University, Sangju 742-711, Korea

Abstract

The use of genomic information in genomic selection programs for dairy and beef cattle breeds has become a reality in recent years. In this investigation, we analyzed single-nucleotide polymorphisms (SNPs) for Hanwoo (n=50) and Holstein (n=50) breeds using the Illumina Bovine SNP50 BeadChip to facilitate genomic selection and utilization of the Hanwoo breed in Korea. Analysis of the entire genomes showed different spectra of SNP frequencies for Hanwoo and Holstein cattle. The study revealed a highly significant ($p < 0.001$) difference between Hanwoo and Holstein cattle in minor allele frequency (MAF). The average MAFs were 0.19 ± 0.16 and 0.22 ± 0.16 for Hanwoo and Holstein, respectively. From the total of 52,337 SNPs that were successfully identified, about 72% and 79% were polymorphic in Hanwoos and Holsteins, respectively. Polymorphic and fixed SNPs were not distributed uniformly across the chromosomes within breeds or between the two breeds. The number of fixed SNPs on all chromosomes was higher in Hanwoo cattle, reflecting the genetic uniqueness of the Hanwoo breed. In general, the rate of polymorphisms detected in these two breeds suggests that the SNPs can be used for different applications, such as whole-genome association and comparative genetic studies, and are a helpful tool in developing breed identification genetic markers.

Keywords: Hanwoo, Holstein, genome, SNP frequency, rate of polymorphism

Introduction

The application of molecular markers to study genetic variation has evolved very rapidly since the mid-1960s (Avise, 2004). The dominance of protein electrophoretic methods to study population genetics occurred in the 1970s, which was replaced by the use of restriction enzyme analysis (animal mtDNA, DNA fingerprinting, single-copy nuclear DNA) in the 1980s. In recent decades, the introduction of polymerase chain reaction (e.g., STRs, AFLP, SNPs)-based DNA genotyping approaches has provided the first rapid and easy assay for genetic data analysis.

Single nucleotide polymorphisms (SNPs) are believed to contain applicable genomic information and have been used widely as genetic markers in population genetics and molecular ecology studies. To date, two full genome assemblies for bovines are available, Btau_4.2 and UMD3.1 (Bovine Genome Sequencing and Analysis Consortium, 2009; Zimin *et al.*, 2009). The availability of genome assemblies led to the development of the Illumina Bovine SNP50 BeadChip genotyping array (Matukumalli *et al.*, 2009; Van Tassell *et al.*, 2009). The Illumina Bovine SNP50 BeadChip (Illumina Inc.) comprises 54,001 evenly spaced SNP probes. The chip presents an average SNP spacing of 51.5 Kb across the entire genome, offering more-than-sufficient SNP density for different applications, including whole-genome association studies, genomic selection, evaluation of genetic merit, and comparative genetic studies (Alam *et al.*, 2011; Kijas *et al.*, 2009; Lee *et al.*, 2010; Matukumalli *et al.*, 2009; Michelizzi *et al.*, 2011; Uemoto *et al.*, 2010; Wade *et al.*, 2009).

Evaluations of genetic merit based on SNPs have become an important area, leading to enhanced genetic improvements in dairy and beef breeds (Hayes *et al.*, 2009). For instance, the application of SNPs for quantitative trait loci detection and genomic selection practices in dairy cattle has been reported (Cole *et al.*, 2009; Hayes *et al.*, 2009). Genomic selection requires a careful understanding of the level of SNPs in each cattle breed. The SNPs in public databases are not validated, and the level of polymorphisms is unknown for many cattle breeds. Therefore, in this study, we evaluated SNP frequency spectra and the rate of polymorphisms in the Hanwoo and Holstein breeds using the Bovine SNP50 BeadChip genotyping array to facilitate genomic evaluation and selection of the Hanwoo breed.

[¶]Authors are equally contributed.
*Corresponding author: E-mail Kwanskim@chungbuk.ac.kr
Tel +82-43-261-2547, Fax +82-43-273-2240
Accepted 26 April 2011

Methods

Molecular data

Blood samples were collected from genetically unrelated Hanwoo (n=50) and Holstein (n=50) cattle of the Nonghyup Animal Genetic Improvement Programs (http://www.limc.co.kr; http://www.dcic.co.kr/). Genomic DNA was extracted based on laboratory procedures. Approximately, 200 ng of genomic DNA was used to commercially genotype each sample on the Illumina Bovine SNP50 BeadChip (SNP Genetics, Seoul). Samples were processed according to the Illumina Infinium-II assay manual. Each sample was whole-genome-amplified, fragmented, precipitated, and resuspended in an appropriate hybridization buffer. Denatured samples were hybridized on the prepared Bovine SNP50 BeadChip for a minimum of 16 h at 48°C. Following the hybridization, the BeadChips were processed for the single-base extension reaction, stained, and imaged on an Illumina Bead Array Reader. Normalized bead intensity data for each sample were loaded into the Beadstudio 3.0 software (Illumina), which converted fluorescent intensities into SNP genotypes. SNP clusters for genotype calling were examined for all SNPs using Beadstudio 3.0 software.

Data analysis

The SNP frequencies were estimated from the genotypic data. The calculated SNP frequencies were used to construct figures and examine rates of polymorphisms and variable frequency spectra in the two breeds using Microsoft EXCEL (2007) and SPSS package, version 17.0. The significance differences in SNP distribution between Hanwoo and Holstein for minor allele frequencies were tested using SPSS.

Results

Variable SNP frequency spectra

The variable SNP frequency spectra detected in Hanwoo and Holstein breeds are summarized in Table 1. A total of 52,337 SNPs were detected in both Hanwoo and Holstein genomes. Approximately, 72% of these SNPs fell within the range of 0.1 to 0.9 for Hanwoo. Similarly, Holstein displayed about 78% of these SNPs, which are considered suitable markers for genomic evaluation. The smallest SNP frequency that could be observed in this study was 0.010 for Hanwoos and 0.0004 for Holstein.

Averaged across the two breeds, 75% of SNPs displayed polymorphisms, which indicates that the majority of the identified SNPs can represent genetic characteristics of the two breeds. On the other hand, about 27% (Hanwoo) and 20% (Holstein) of SNPs were monomorphic in the current analysis (Fig. 1).

The SNP data were also used to compare minor allele frequency (MAF) patterns between the Hanwoo and Holstein breeds. As indicated in Fig. 2, the analysis revealed a highly significant (Pearson's χ^2 test=865, $p < 0.001$) difference between Hanwoo and Holstein for MAF spectra. The average MAFs were 0.19 ± 0.16 and 0.22 ± 0.16 for Hanwoo and Holstein breeds, respectively. The intermediate MAF spectrum was similar for both breeds, but the distribution was different for monomorphic or fixed alleles. Of the total SNPs, 47% of the Hanwoo comprised less common variants ($\geq 0.05 - \leq 0.5$).

Further breakdown of the distribution of MAFs revealed that 25% and 28% highly polymorphic SNPs ($\geq 0.3 - \leq 0.5$) were observed in the Hanwoo and Holstein genomes, respectively. Hanwoo displayed a higher level of monomorphic SNPs (9%), which was significantly ($p < 0.01$) higher than in Holstein (Fig. 1). The frequency of rare alleles ($> 0 - < 0.05$) was about 3,767 (7.2%)

Table 1. Variable frequency spectra of 52,337 SNPs observed in Hanwoo and Holstein breeds

Breed	Variable SNPs				
	< 0.1 or > 0.9	0.1 ~ 0.9	0.2 ~ 0.8	0.3 ~ 0.7	0.2 ~ 0.6
Hanwoo	14,397	37,940	28,491	20,359	12,297
Holstein	11,420	40,917	32,403	23,485	14,275

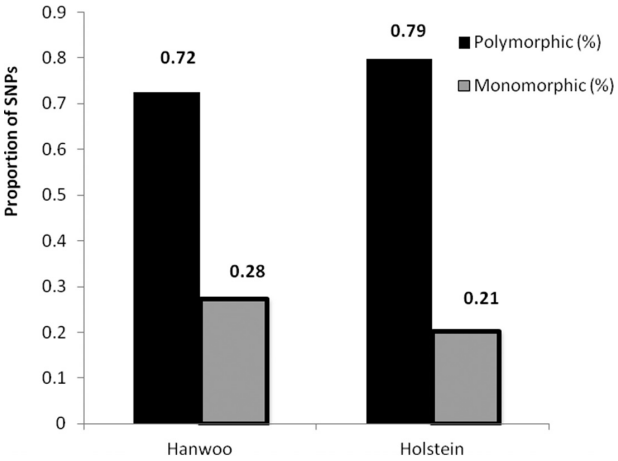

Fig. 1. Polymorphic and monomorphic SNP dynamics in Hanwoo and Holstein cattle; polymorphic ($\geq 0.05 - \leq 0.95$), monomoprphic (< 0.05 or > 0.95).

and 4,861 (9.2%) in Hanwoo and Holstein, respectively. The intermediate allele frequency (≥0.05－<0.10) of Hanwoo and Holstein accounted for 5% of the total genomic SNPs.

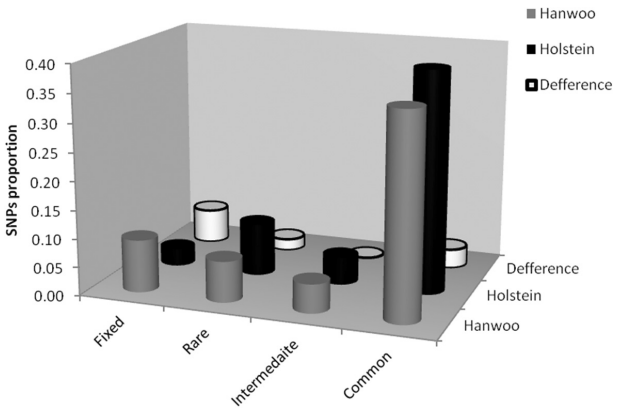

Fig. 2. Minor allele frequency spectra in the genomes of Hanwoo and Holstein cattle, (monomorphic, frequency of 0; rare, >0－<0.05; intermediate, ≥0.05－<0.10; common, ≥0.1－≤0.50).

Chromosomal level variations in SNPs

We further examined the distribution of SNPs at the chromosomal level (Fig. 3a, b). SNPs were not uniformly distributed over the entire genome of all chromosomes in Hanwoo or Holstein. Each chromosome consisted of variable number of polymorphic and fixed SNPs (Figure 3a, b). The two breeds generally displayed a similar pattern of SNP distribution over autosomes and X chromosomes, but Holstein had a relatively higher level of polymorphisms on all chromosomes.

The number of available SNPs on a chromosome depends on its length. From this investigation, it was observed that the total number of SNPs, as well as genetic variation, was the lowest for the X chromosome. The proportion of SNPs on X chromosomes was about 0.9% (Hanwoo) and 1% (Holstein). There was a slight difference between the number of SNPs reported on the BeadChip and the present study. We deleted about 1,679 SNPs from unknown chromosomes from the data

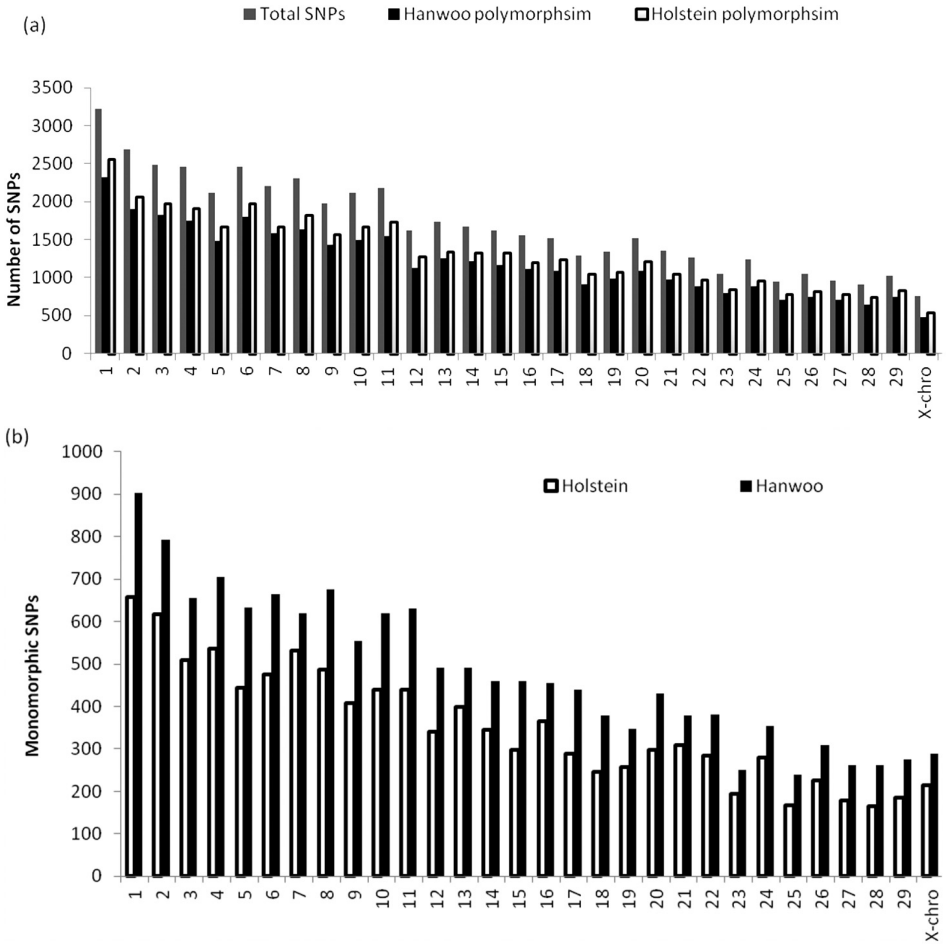

Fig. 3. (a) Distribution of polymorphic and (b) monomorphic SNPs on chromosomes of Hanwoo and Holstein breeds.

during the analysis; this could be the reason for the difference.

The distribution and level of fixed SNPs across the genome of these animals were also evaluated. SNPs were fixed across the chromosomes of Hanwoo and Holstein (Fig. 3b). It was also observed that the number of fixed SNPs varied across chromosomes and between breeds. The number of fixed SNPs was generally higher in Hanwoo as compared to Holstein (Fig. 3b).

Discussion

In the human genome, the most common form of genetic variation is a simple change of a base pair, a single nucleotide polymorphism (Broeckel and Hessner, 2006). Population history, selection pressure, and other population genetic factors determine the frequency and patterns of SNPs in the genome. Hence, in this study, we were interested in examining the frequency spectra of the entire genome, comprising 52,337 SNPs in Hanwoo and Holstein breeds. The SNPs detected using the Illumina Bovine SNP50 Chip were informative. SNPs that are significantly associated with one population but not the other can be used as informative SNPs for that particular breed. The chi-square test was performed to investigate allele frequency spectra differences between Hanwoo and Holstein. The test revealed a significant variation in MAF proportion between the two breeds. This genetic difference, detected from allele frequency spectra, can be explained by the fact that Hanwoo and Holstein have followed unique evolutionary pathways in different geographic locations.

The majority of SNPs in public databases has not been validated in several cattle breeds, and the level of polymorphisms is unknown for many cattle breeds in the world. This may limit the application of SNPs, such as for population structure or whole-genome association studies. In this preliminary study, about 37,971 and 41,724 polymorphic SNPs were detected in Hanwoos and Holsteins, respectively, with an average polymorphism rate of 75%. In future genome-wide association and population genetics studies, the selection of highly polymorphic SNPs is paramount to guarantee the most optimal result. In another study, Matukumalli *et al.* (2009) used the Bovine SNP50 BeadChip and found that the number of SNPs with minor allele frequencies of $\geq$ 0.05 ranged from 31,633 to 42,711 among 14 taurine breeds, from 28,823 to 35,425 between two African breeds, and from 23,284 to 30,139 among three indicine breeds. About 30% of SNPs had MAFs $>$0.3 within the taurine breeds, and only about 19% had MAFs $>$0.3 within the indicine breeds (Bovine HapMap Consortium, 2009). In sheep, averaged across the breeds, 81% of SNPs displayed polymorphisms, which indicates that the majority of identified SNPs predates the radiation of the domestic breeds that were sampled (Kijas *et al.*, 2009).

The number of polymorphic and monomorphic SNPs was not uniformly distributed across the chromosomes between the two breeds or within a breed. The number of polymorphic SNPs was higher in Holstein on all chromosomes. In contrast, several alleles were monomorphic in Hanwoo, meaning that only a single fixed gene was present. As alleles became fixed in Hanwoo, there was an overall decline in heterozygosity as the breed became homozygous for one allele or the other. In the past, if the Hanwoo population remained small and isolated for several generations, there was a possibility of random fixation and the loss of some alleles by genetic drift. We also hypothesize that high levels of allelic fixation may reflect a certain level of inbreeding. Previously, Lee *et al.* (2010) reported that average inbreeding varied from 0.00% to 0.97% and that the average annual rate of increase in inbreeding was 0.14% in Hanwoo. Therefore, the level of inbreeding in Hanwoo is still low and may not be a major concern. Generally, genetic variability in Hanwoo should provide balance in a gene pool and also offer genetic variants for natural and artificial selection. Furthermore, the current breeding strategy implemented for Hanwoo cattle in Korea should also focus on minimizing genetic homogeneity and maintaining genetic uniqueness of the breed while improving beef production.

In conclusion, Hanwoo and Holstein cattle have a significant variation of genomic SNP (alleles) frequency spectra, implying that the breeds follow unique evolutionary paths that lead to genetic differences between the two breeds. The rate of SNP polymorphisms detected in these two breeds suggests that the SNPs could potentially be used in genomic evaluations and as a helpful tool in developing breed identification genetic markers.

Acknowledgments

This research work was supported by a research grant of Chungbuk National University in 2009.

References

Alama, M., Lee, Y.M., Park, B. L., Kim, J.H., Lee, S.S., Shin, H.D., Kim, K.S., Kim, N.S., and Kim, J.J. (2011). A Whole Genome Association Study to Detect Single Nucleotide Polymorphisms for Body Conformation Traits in a Hanwoo Population. *Asian-Aust. J. Anim. Sci.* 24, 322-329.

Avise, J.C. (2004). *Molecular markers, natural history, and evolution*, (Sunderland, MA: Sinauer Associates, Inc.).

Bovine Genome Sequencing and Analysis Consortium, Elsik, C.G., Tellam, R.L., and Worley, K.C. (2009). The Genome Sequence of Taurine Cattle: a window to ruminant biology and evolution. *Science* 324, 522-528.

Bovine HapMap Consortium, Gibbs, R.A., Taylor, J.F., Van Tassell, C.P., Barendse, W., Eversole, K.A., Gill, C.A., Green, R.D., Hamernik, D.L., Kappes, S.M., Lien, S., Matukumalli, L.K., McEwan, J.C., Nazareth, L.V., Schnabel, R.D., Weinstock, G.M., Wheeler, D.A., Ajmone-Marsan, P., Boettcher, P.J., Caetano, A.R., Garcia, J.F., Hanotte, O., Mariani, P., Skow, L.C., Sonstegard, T.S., Williams, J.L., Diallo, B., Hailemariam, L., Martinez, M.L., Morris, C.A., Silva, L.O., Spelman, R.J., Mulatu, W., Zhao, K., Abbey, C.A., Agaba, M., Araujo, F.R., Bunch, R.J., Burton, J., Gorni, C., Olivier, H., Harrison, B.E., Luff, B., Machado, M.A., Mwakaya, J., Plastow, G., Sim, W., Smith, T., Thomas, M.B., Valentini, A., Williams, P., Womack, J., Woolliams, J.A., Liu, Y., Qin, X., Worley, K.C., Gao, C., Jiang, H., Moore, S.S., Ren, Y., Song, X.Z., Bustamante, C.D., Hernandez, R.D., Muzny, D.M., Patil, S., San Lucas, A., Fu, Q., Kent, M.P., Vega, R., Matukumalli, A., McWilliam, S., Sclep, G., Bryc, K., Choi, J., Gao, H., Grefenstette, J.J., Murdoch, B., Stella, A., Villa-Angulo, R., Wright, M., Aerts, J., Jann, O., Negrini, R., Goddard, M.E., Hayes, B.J., Bradley, D.G., Barbosa da Silva, M., Lau, L.P., Liu, G.E., Lynn, D.J., Panzitta, F., and Dodds, K.G. (2009). Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science* 324, 528-532.

Broeckel, U., and Hessner, M.J. (2006). Single nucleotide polymorphism. Testing DNA variation for disease association. In *Molecular diagnostics: For the Clinical Laboratorian*, W.B. Coleman and G.J. Tsongalis, ed. (New York, USA: Humana Press), pp.111-120.

Cole, J.B., Van Raden, P.M.O., Connell J.R., Van Tassell C.P., Sonstegard, T.S., Schnabel, R.D., Taylor, J.F., and Wiggans, G.R. (2009). Distribution and location of genetic effects for dairy traits. *J. Dairy Sci.* 92, 2931-2946.

Hayes, B.J., Bowman, P.J., Chamberlain, A.J., and Goddard, M.E. (2009). Invited review: Genomic selection in dairy cattle: progress and challenges. *J. Dairy Sci.* 92, 433-443.

Kijas, J.W., Townley, D., Dalrymple, B.P., Heaton, M.P., Maddox, J.F., McGrath, A., Wilson, P., Ingersoll, R.G., McCulloch, R., McWilliam, S., Tang, D., McEwan, J., Cockett, N., Oddy, V.H., Nicholas, F.W., and Raadsma, H. (2009). A Genome Wide Survey of SNP Variation Reveals the Genetic Structure of Sheep Breeds. *PLoS,* 4, e4668.

Lee, J., Dang, C., and Kim, N. (2010). Inbreeding Trends and Pedigree Structure of Hanwoo (Korean Cattle). Proceedings Vol. II.

Lee, Y.M., Han, C.M., Li, Y., Lee, J.J., Kim, L.H., Kim, J.H., Kim, D.I., Lee, S.S., Park, B.L., Shin, H.D., Kim, K.S., Kim, N.S., and Kim, J.J. (2010). A whole genome association study to detect single nucleotide polymorphisms for carcass traits in Hanwoo populations. *Asian-Aust. J. Anim. Sci.* 23, 417-424.

Matukumalli, L.K., Lawley, C.T., Schnabel, R.D., Taylor, J.F., Allan, M.F., Heaton, M.P., O'Connell, J., Moore, S.S., Smith, T.P., Sonstegard, T.S., and Van Tassell, C.P. (2009). Development and characterization of a high density SNP genotyping assay for cattle. *PLoS,* 4, e5350.

Michelizzi1, V.N., Wu, X., Dodson, M.V., Michal, J.J., Varon, J.Z., McLean, D.J., and Jiang, Z. (2011). A Global View of 54,001 Single Nucleotide Polymorphisms (SNPs) on the Illumina BovineSNP50 BeadChip and Their Transferability to Water Buffalo. *Int. J. Biol. Sci.* 7, 18-27.

SPSS Statistics 17.0 Brief Guide. SPSS Inc., 233 South Wacker Drive, 11 Floor, Chicago.

Uemoto, Y., Abe, T., Tameoka, N., Hasebe, H., Inou, K., Nakajima, H., Shoji, N., Kobayashi, M., and Kobayashi, E. (2010). Whole-geneome association study for fatty acid composition of oleic acid in Japanese Black cattle. *Anim. Genet.* 1-8.

Van Tassell, C.P., Smith, T.P., Matukumalli, L.K., Taylor, J.F., Schnabel, R.D., Lawley, C.T., Haudenschild, C.D., Moore, S.S., Warren, W.C., and Sonstegard, T.S. (2008). SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. *Nat. Methods* 5, 247-252.

Wade, C.M., Giulotto, E., Sigurdsson, S., Zoli, M. Gnerre, S., Imsland, F., Lear, T.L., Adelson, D.L., Bailey, E., Bellone, R.R., Blöcker, H., Distl, O., Edgar, R.C., Garber, M., Leeb, T., Mauceli, E., MacLeod, J.N., Penedo, M.C., Raison, J.M., Sharpe, T., Vogel, J., Andersson, L., Antczak, D.F., Biagi, T., Binns, M.M., Chowdhary, B.P., Coleman, S.J., Della Valle, G., Fryc, S., Guérin, G., Hasegawa, T, Hill, E.W., Jurka, J., Kiialainen, A., Lindgren, G., Liu, J., Magnani, E., Mickelson, J.R., Murray, J., Nergadze, S.G., Onofrio, R., Pedroni, S., Piras, M.F., Raudsepp, T., Rocchi, M., Røed, K.H., Ryder, O.A., Searle, S., Skow, L., Swinburne, J.E., Syvänen, A.C., Tozaki, T., Valberg, S.J., Vaudin, M., White, J.R., Zody, M.C. Broad Institute Genome Sequencing Platform, Broad Institute Whole Genome Assembly Team, Lander, E.S., and Lindblad-Toh, K. (2009). Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science* 326, 865-867.

Zimin, A.V., Delcher, A.L., Florea, L., Kelley, D.R., Schatz, M.C., Puiu, D., Hanrahan, F., Pertea, G., Van Tassell, C.P., Sonstegard, T.S., Marçais, G., Roberts, M., Subramanian, P., Yorke, J.A., and Salzberg, S.L. (2009). A whole-genome assembly of the domestic cow, Bos taurus. *Genome Biol.* 10, R42.