ISSN NO. 2320-5407

Journal homepage: http://www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH

RESEARCH ARTICLE

Comparative analysis of genome wide difference in Red Sindhi and Holstein cattle breeds using dense SNP marker

Hamid Mustafa[1], Huson J. Heather[3], Kim EuiSoo[4], Nisar Ahmad[1], Afzal Ali[1], Waqas Ahmad Khan[2], Talat Naseer Pasha[5], Muhammad Zahid Farooq[6], Khalid Javed[1], Adeela Ajmal[1] and Tad S. Sonstegard[8]

1. Department of Livestock Production, University of Veterinary and Animal Sciences, Lahore-Pakistan.
2. Department of Biotechnology, University of Sargodha, Pakistan.
3. Department of Animal Nutrition, University of Veterinary and Animal Sciences, Lahore-Pakistan.
4. Department of Animal Sciences, Cornell University, Ithaca, NY, USA.
5. Department of Animal Sciences, Iowa State University, USA.
6. College of Veterinary and Animal Sciences, Jhang.
7. Department of Livestock Management, University of Agriculture Faisalabad.
8. Bovine Functional genomics Laboratory, USDA/ARS, Beltsville, USA

Manuscript Info

Manuscript History:

Received: 11 February 2014
Final Accepted: 25 March 2014
Published Online: April 2014

Key words:
Red Sindhi, Holstein, Polymorphism, SNP frequency

Corresponding Author

Hamid Mustafa

Abstract

Genomic selection programs for the improvement of dairy and beef cattle breed has become a reality due to the availability of detailed genomic information. In this study, we compared the whole genome of Red Sindhi (**Bos** indicus) and Holstein (**Bos** taurus) cattle breeds using the Illumina Bovine HD SNP BeadChip for better understanding of genomic selection and utilization for improved breeding programme in Pakistan. One hundred and five samples consisting of 25 Red Sindhi and 70 Holstein were genotyped. Analysis of both breeds showed different spectra of single nucleotide polymorphisms (SNP) frequencies. Significant ($p< 0.001$) difference between these two cattle breeds in minor allele frequency (MAF) was observed. The average minor allele frequencies (MAF) were 0.18 ± 0.15 and 0.22 ± 0.16 for Red Sindhi and Holstein, respectively. Total of 777, 962 SNPs were identified successfully. About 71% and 79% were polymorphic in Red Sindhi and Holstein, respectively. Polymorphic and fixed SNPs were not distributed uniformly across the chromosomes between these two breeds. Fixed SNPs numbers were higher on all chromosomes in Red Sindhi cattle, which show genetic uniqueness of Red Sindhi breed. The results of this study suggest that the rate of polymorphisms can be effectively used for breed identification, evolutionary studies and whole genome association studies..

Introduction

Study of genetic variation at molecular level has evolved very rapidly since the late 20th century. Molecular data provides reliable genetic variation due to the detection at DNA level and different techniques are available for identifying genetic variation between individuals, breeds, and species (Avise, 2004). Protein polymorphisms were the first genetic markers for the studies of livestock species in the 1970s and replaced by the restriction enzymes analysis in 1980s. Genetic markers are useful in both basic (phylogenetic analysis and gene selection) and applied (marker assisted selection and paternity testing) research. In recent years, polymerase chain reaction based genotyping methods has provided the rapid and easy assay for genetic data analysis. Single nucleotide polymorphism (SNPs) are the most common form of polymorphism and believed to contain valuable information, widely used as genetic markers in population genetics and molecular ecology studies. Btau_4.2 and UMD 3.1 are two different bovine genome assemblies, which are available (Bovine genome sequencing and analysis consortium, 2009; Bovine HapMap consortium, 2009; Zimin et al., 2009 and Hailu et al., 2011) and led to the development of Illumina Bovine HD SNP (more than 777,000 SNPS) BeadChip genotyping array (Van Tassel et al., 2008;

Matukumalli et al., 2009). The Bovine HD BeadChip is the most comprehensive genome-wide SNPs genotyping array with superior power to identify genetic variation in many breeds of cattle (Illumina Inc.). The Bovine HD BeadChip comprises more than 777,000 SNPs, which are more than enough SNP density for broad range applications including genome wide association studies, identification of quantitative trait loci (QTL), genetic evaluation of genetic merit, linkage disequilibrium studies (LD) studies, evolutionary studies and breed characterization (Wade et al., 2009; Matukumalli et al., 2009; Kijas et al ., 2009; Uemoto et al., 2010; Lee et al., 2010; Alama et al., 2011; Michelizzi et al., 2011; Hailu et al., 2011). Public database of SNPs information is not a valid source because of unknown level of polymorphisms. However, genomic selection for dairy or beef cattle breeds require a level of understanding of SNPs information (Cole et al., 2009; Hayes et al., 2009). Thus, the main objective of this study was to evaluate the performance of SNP frequency and their rate of polymorphism in Red Sindhi and Holstein breeds using the Bovine HD SNP BeadChip genotyping array to explore the potential of Red Sindhi breed for genomic selection programmes.

Materials and Methods

Animals sampling and genomic DNA extraction:

Blood samples were obtained from 25 Red Sindhi animals from Livestock Experimentation Station, Rak Mani, Bhakar. A 10 ml blood was collected using 50 ml EDTA containing falcon tubes. Genomic DNA was extracted by modified salt method (Sambrook and Russel, 2001). DNA samples were quantified using Nano drop Spectrophotometer (Nano-drop ND 1000). Holstein (n=70) DNA samples were obtained from the BFGL / USDA DNA library.

Genotyping:

Genotyping was conducted at the Bovine Functional Genomics laboratory (BFGL), Beltsville, MD, USA, platform with Illumina Bovine high density SNPs BeadChip v2 features more than 777,000 SNPs distributed across the whole genome of bovine. Approximately, 200 ng DNA was used to genotyped each sample. Samples were preceded following the defined procedure and protocol of Illumina infinium-II assay (Illumina, Inc, San Diego, CA, USA). Each animal sample was amplified for 20 hours at 37^0C. The samples were hybridized on the prepared bovine high density SNP BeadChip for 20 hours at 48^0C. The samples were stained and imaged on an Illumina iScan Reader.

Data analysis

Genotype data was generated from the iScan system loaded into Illumina genome Studio version 1.9.0 software to perform primary data analysis, clustering and genotyping calling. The SNP frequencies were estimated from the genotypic data. Rate of polymorphisms and variable frequency spectra was calculated using Microsoft Excel (MS excel 2010) and SAS (version 9.1). The significance difference in SNP distribution between Red Sindhi and Holstein for minor allele frequencies (MAF) were tested using SAS (ver. 9.1)

Results

SNP frequency:

SNP frequency variation of Red Sindhi and Holstein cattle breeds are presented in Table 1. A total of 547,615 SNPs were detected in whole genome of Red Sindhi and Holstein. Red Sindhi cattle breed shows 71% of the total SNPs fell within the range of 0.1-0.9, whereas 78% was detected in Holstein breed. This considerable difference of SNP variations are suitable genetic markers for genomic evaluation in both breeds. Average SNPs across the two breeds was displayed 75% polymorphism. This shows that the majority of the SNPs can represent genetic characteristics of two breeds. Average monomorphic SNPs were identified in Red Sindhi breed 28% and 21% in Holstein (Fig. 1). Pattern of minor allele frequency (MAF) was also compared between these two breeds. Analysis revealed a highly significant ($p<0.001$) difference between Red Sindhi and Holstein for minor allele frequency (MAF). The average minor allele frequency (MAF) were 0.18±0.15 and 0.22±0.16 for Red Sindhi and Holstein breeds, respectively. The intermediated MAF was similar for both breeds, but distribution was different for monomorphic or fixed alleles. Common variants ($\geq 0.05- \leq 0.5$) SNPs was comprised 47% of the total SNPs in Red Sindhi breed. Highly polymorphic SNPs ($\geq 0.3 - \leq 0.5$) were 26% and 29% in the Red Sindhi and Holstein, respectively. High level of monomorphic SNPs (11%) was observed in Red Sindhi breed, which was significantly ($p< 0.01$) higher than in Holstein breed (Fig 1). Rare alleles ($> 0 - < 0.05$) was about 9734 (9%) and 13,412 (11%) in Red Sindhi and Holstein, respectively. The intermediate allele frequency (($\geq 0.05 - < 0.10$) of Red Sindhi and Holstein was 5% of the total genomic SNPs.

Distribution of SNPs at chromosomal level:

Distribution of SNPs variation at chromosomal level was also examined. Single nucleotide polymorphisms (SNPs) were not uniformly distributed over the all chromosomes in Red Sindhi and Holstein genome. Variable numbers of

polymorphic and fixed SNPs were found on each chromosome of both breeds. Generally, both breeds displayed a similar pattern of SNP distribution over autosomes and X chromosomes. Relatively, higher level of polymorphisms was observed on all chromosomes in Holstein breed. Available number of SNPs on a chromosome depends on its length. In this study, the total number of SNPs and genetic variation was lowest for X chromosome. The distribution and level of fixed SNPs across the genome of both breeds were also evaluated. It was observed that the fixed SNPs number varied across chromosomes and between breeds. Fixed SNPs number was observed generally higher in the Red Sindhi than Holstein.

Table 1. Observed frequency variation of 547,615 SNPs in Red Sindhi and Holstein cattle breeds

Breed	SNPs variation				
	< 0.1 or > 0.9	0.1 ~0.9	0.2~0.8	0.3~0.7	0.2~0.6
Red Sindhi	232,128	314,281	212,871	211,173	151,932
Holstein	214,253	452,871	319,891	261,762	191,321

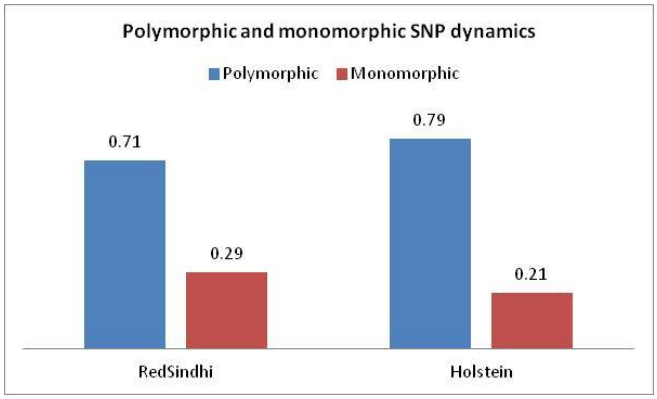

Fig.1 Polymorphic (≥0.05−≤0.95) and monomorphic (＜0.05 or＞0.95) SNP in Red Sindhi and Holstein cattle breed

Discussion

Single Nucleotide polymorphisms (SNP) are the most common form of polymorphism and genetic variation among individuals and arising approximately every 200 base pairs in livestock (Broeckel and Hessner, 2006). Population history selection pressure and other population genetic factor determine the frequency and patterns of SNPs in the genome (Hailu et al., 2011). Therefore, the objective of this study was to identify the frequency variation of whole genome of Red Sindhi and Holstein cattle breeds. A total of 547,615 SNPs were detected in whole genome of these breeds. The SNPs detected using the Illumina Bovine HD BeadChip was informative. SNPs that were significant associated with one population but not other can be used as informative SNPs for that particular breed. Investigation of allele frequency between two breeds was performed using chi-square test and the results of this test revealed a significant variation in minor allele frequency (MAF) proportion between these two breeds. This genetic difference of allele frequency clearly explained the unique evolutionary pathways in different geographic locations as described by Loftus et al., 1994 and MacHugh et al., 1997, they investigate genetic variation at mtDNA sequence and microsatellite and suggest that the ancestors of Bos indicus and Bos taraus diverged some hundreds of thousands of years ago and must therefore be the results of at least two biologically independent domestication events. This study is also very informative due to the high level of polymorphisms which is obviously lacks in majority of SNPs available in public databases and has not been validated in several cattle breeds in the world. Low level of SNPs polymorphism may limits the application for population structure or whole genome association studies. This is a first study of its kind to compare a taurine cattle breed with indicine at high magnitude of SNPs panel. About 500,939 and 594, 290 polymorphic SNPs were detected in Red Sindhi and Holstein cattle breeds, respectively, with average polymorphism rate of 75%. In future, high polymorphic SNPs are dominant to assure the most optimal results for genome-wide association and population genetics studies (Gupta et al., 2001). In a study, Hailu et al. (2011) used the Bovine SNP 50 BeadChip and found the SNPs with minor allele frequency of ＜0.1 or ＞0.9 ranged

from 37,971 and 41,724 polymorphic SNPs were detected in Hanwoos and Holsteins cattle breed, respectively, with an average polymorphism rate of 75%. In another study, Matukumalli et al. (2009) used also Bovine SNP50 and found SNPs number with minor allele frequencies of 0.05 and estimate the range from 31,633 to 42,711polymorphic SNPs among 14 taurine breeds. This range between two African cattle breeds from 28,823 to 35,425 between two African breeds, while, this range from 23,284 to 30,139 polymorphic SNPs 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. In this study, the number of polymorphic SNPs was higher in Holstein cattle on all chromosomes, in contrast several alleles were found monomorphic in Red Sindhi , which shows that only a single fixed gene was present in this cattle population. As alleles became fixed in Red Sindhi , there was an overall decline in heterozygosity as the breed became homozygous for one allele or the other. If the Red Sindhi cattle population remained small and isolated for several generations, there was a possibility of random fixation and loss of some allele by genetic drift. This high level of allelic fixation may reflect a certain level of inbreeding.

Genetic variation in Red Sindhi cattle breed should provide balance in a gene pool and also offer genetic variants for natural and artificial selection. Therefore, the breeding strategy for Red Sindhi cattle breed in Pakistan should focus on minimizing genetic homogeneity and maintaining genetic uniqueness for the improvement of milk and beef production.

In conclusion, Red Sindhi and Holstein cattle have a significant variation of SNPs frequency implying that the breeds follow unique evolutionary pathways that lead to genetic difference between two breeds. The rate of SNP polymorphisms detected in both breeds suggests that SNPs could potentially used in evolutionary studies and also helpful in developing breed identification genetic markers.

Acknowledgments

The authors of this study would like to thank the Bovine Functional Genomics Laboratory, USDA/ARS, MD, USA, for laboratory resources and genotyping facility. Financial support from the Higher Education Commission (HEC), Pakistan is greatly appreciated. Provision of Blood Samples by Government Livestock Experimental Stations is also acknowledged.

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.

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.

Hailu Dadi Melka, Eun-Kyeong Jeon, Sang-Wook Kim, James-Bond Han, Duhak Yoon and Kwan-Suk Kim (2011). Identification of Genomic Differences between Hanwoo and Holstein Breeds Using the Illumina Bovine SNP50 BeadChip. Gen. Inf. 9(2): 69-73.

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, 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.

Loftus RT, MacHugh DE, Bradley DG, Sharp PM, Cunningham EP. 1994. Evidence for two independent domestications of cattle. Proc. National Acad. Sci. USA. 91: 2757-2761.

MacHugh DE, Shriver MD, Loftus RT, Cunningham P, Bradley DG. 1997. Microsatellite DNA variation and the evolution, domesticationand phylogeography of taurine and zebu cattle (Bos Taurus and Bos inducus). Genet. 146: 1071-1086.

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.

Michelizzi, 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.

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.