

Whole genome study of linkage disequilibrium in Sahiwal cattle

H. Mustafa^{1,2#}, N. Ahmad¹, H. J. Heather², K. Eui-soo², W. A. Khan³, A. Ajmal¹, K. Javed¹, T. N. Pasha⁴, A. Ali¹, J. J. Kim⁵ & T. S. Sonstegard²

¹ Department of Livestock Production, University of Veterinary and Animal Sciences, Lahore 54000, Pakistan

² United States Department of Agricultural, Maryland 27030, USA

³ Department of Biotechnology, University of Sargodha, Pakistan

⁴ Department of Animal Nutrition, University of Veterinary and Animal Sciences, Lahore 54000, Pakistan

⁵ Department of Biotechnology, Yeungnam University Gyeongsan, Gyeongbuk 712749, Republic of Korea

(Received 26 September 2017; Accepted 27 December 2017; First published online 30 December 2017)

Copyright resides with the authors in terms of the Creative Commons Attribution 4.0 South African Licence.

See: <http://creativecommons.org/licenses/by/4.0/za>

Condition of use: The user may copy, distribute, transmit and adapt the work, but must recognise the authors and the South African Journal of Animal Science.

Abstract

The linkage disequilibrium (LD) is an important tool to study quantitative trait locus (QTL) mapping and genetic selection. In this study, we identified the extent of linkage disequilibrium (LD) in Sahiwal ($n = 14$) cattle using the bovine high density single-nucleotide polymorphisms (SNPs) BeadChip. After data filtering, 500,968 SNPs comprising 2518.1 Mb of the genome, were used for the LD estimation. The minor allele frequency (MAF) was 0.21 in a substantial proportion of SNPs and mean distance between adjacent markers was 4.77 ± 2.83 kb. The overall mean LD between adjacent markers was 0.18 (r^2) and 0.55 ($|D'|$), respectively. The LD (r^2) values reduced with the increase in distance between adjacent markers from 1 kb (0.35) to 100 kb (0.12) and $|D'|$ specified distinct decay of the LD. Chromosomes 1, 27, 28 and 29 presented the LD at some distance between markers. The extent of LD was higher, except these four chromosomes, for markers separated by 20 kb. At < 3 kb distance, the upper value of the linkage disequilibrium (LD) was observed at 0.30. High level of the linkage disequilibrium (LD) between markers was observed at high minor allele frequency (MAF) threshold (0.15), at the short distance between markers. The results of this study revealed that the Bovine high density SNPs BeadChip will be informative for the estimation of breeding value in Sahiwal cattle.

Keywords: Breeding value, linkage disequilibrium, minor allele frequency, molecular marker

Corresponding author: hamidmustafapasha@gmail.com

Introduction

Sahiwal is an important dairy breed of Pakistan, belonging to *Bos indicus*. This breed is well known in the world for dairy traits and is present in twenty-nine countries (FAO, 2007). The population of Sahiwal cattle is under threat, due to crossbreeding. There are about 2, 753 thousand heads of this breed in Pakistan (GOP, 2006). Over the last two decades, the interest in the genetic evaluation of indigenous cattle breeds has increased. Several genetic evaluation schemes for the genetic improvement of Sahiwal cattle are under progress. The main objective of these programs is the enhancement of production and reproduction performance of this breed. Sahiwal is an internationally important dairy breed and is used in various synthetic breed's development (Khan *et al.*, 2008).

Animal breeding value can be estimated from whole genome data by using marker assisted selection, also known as genetic selection (Bennewitz *et al.*, 2009). Genetic selection identifies the linkage disequilibrium (LD) between adjacent markers and reveal that chromosome segments will affect the whole population and genes responsible for traits expression (QTL) since the markers are in LD. The LD is a non-random allelic association with different loci in a population. The LD maps are significant approaches for determining the genetic basis of economically key traits in cattle. Similarly, LD maps comparison permits to institute the diversity between cattle breeds with diverse biological characteristics and to classify genome regions that were open to dissimilar selection pressures (McKay *et al.*, 2007).

The two methods most usually used to calculate LD between markers (biallelic) are r^2 and $|D'|$ and these parameters can vary from 0 to 1 (Hill & Robertson, 1968; Bohmanova *et al.*, 2010). When $|D'| < 1$, this

indicates the existence of recombination between two loci, whereas there is lack of recombination between two loci at $|D'| = 1$. The r^2 parameter denotes the correlation between two loci and is used in association studies. Nevertheless, inverse relationships exist between r^2 and the sample size required to detect the same power. The LD is required to study the association between a marker and a QTL (Pritchard & Przeworski, 2001). Marques *et al.* (2008) studied the chromosome 14 with 505 SNPs in Holstein cattle and found moderate LD level (0.2), separated by 100 kb. McKay *et al.* (2007) found similar results for 2, 670 markers in eight different cattle breeds. In a comprehensive study of 19 taurine and indicine breeds with 32, 826 SNPs dataset, a higher proportion of low allele frequencies and low LD level was observed in indicine breeds (Villa-Angulo *et al.*, 2009). Silva *et al.* (2010) genotyped Gyr breed with 54,000 SNPs and found LD (0.21) between adjacent markers. The authors of different studies identified that *Bos indicus* breeds have a higher percentage of low alleles frequency and a lower level of LD than *Bos taurus* breeds (McKay *et al.*, 2007; Sargolzaei *et al.*, 2008; Kim & Brian *et al.*, 2009; Espigolan *et al.*, 2013). However, there was no single report of LD estimation in Sahiwal cattle breed, which is an internationally important dairy breed. The LD identifies population features and different chromosomes pattern in a population. Thus, this present study aims to identify the LD in Sahiwal cattle breed using the bovine high density SNPs BeadChip.

Materials and Methods

In this present study, fourteen Sahiwal individuals from two different cattle farms were used. Sample details are described in a study by Mustafa *et al.* (2017). There was no need to take approval from animal's ethical committee for the present study, because existing data set were used. Genotyping was completed by using the bovine HD SNPs (700 K) BeadChip (Illumina HD assay^(R)). The bovine HD SNPs BeadChip contained 777, 962 SNP spread across the genome with mean 3.43 kb distance between markers. The Genome Studio^(R) software (Illumina) was used to analyze genotypes and total 1467 markers were excluded from the data set due to their unknown position in the genome. Monomorphic markers were found 15, 186. Only autosomal markers were included in LD analysis with MAF higher than 0.05, 0.10 and 0.15. Additionally, only markers with call rate >0.90 and heterozygote (<0.30) were considered. The LD was evaluated using r^2 and $|D'|$ between two SNPs. The r^2 was measured as follows:

$$r^2 = \frac{(freq. XY * freq. xy - freq. xy * freq. xY^*)}{(freq. X * freq. x * freq. Y * freq. y)}$$

$$= \frac{(D)^2}{(freq. X * freq. x * freq. Y * freq. y)}$$

Where,

$$D = freq. XY - freq. x * freq. Y$$

and

$$|D'| = \begin{cases} \frac{D}{(freq. X * freq. x * freq. Y * freq. y)} & \text{if } D > 0 \\ \frac{D}{(freq. X * freq. x * freq. Y * freq. y)} & \text{if } D < 0 \end{cases}$$

Where, frequency X, frequency x, frequency Y and frequency y are alleles frequencies of alleles X, x, Y and y, respectively.

Frequency XY, frequency xy, frequency xY and frequency Xy are the haplotypes frequencies of XY, xy, xY and Xy, respectively. If the two loci are independent, the expected frequency of haplotype XY (freq. XY) is calculated as the product between freq. X and freq. Y. X freq. XY higher or lower than the expected value indicates that these two loci in particular tend to segregate together and are in LD. Higher or lower frequency of XY indicates that two loci separate together and in LD. The linkage disequilibrium (r^2 and $|D'|$) were measured for all markers pairs at each chromosomes using SnppldHD software (Espigolan *et al.*, 2013). The present study measured LD (r^2 and $|D'|$) estimation using maternal haplotypes as commonly used in LD estimation studies, where, half sib families exist (Reich & Lander, 2001).

Results and Discussion

The descriptive statistic results for each autosome of LD between adjacent markers and the SNP markers are shown in Table 1. A total of 500, 968 (64.4 %) markers met the filtering criteria and were comprised in the final analysis. This subset of markers included 2, 518.1 Mb of the genome, with a mean

distance between markers of 4.77 ± 2.89 kb. The density of markers was similar across all the chromosomes ranging from 5.1 to 5.4 kb.

Table 1 SNP summary of analyzed markers and average linkage disequilibrium (r^2 and $|D'|$) between synthetic adjacent markers

Chr.	Size (Mb)	SNP (n)	Mean distance (SD)	Mean r^2 (SD)	Median r^2	Mean $ D' $ (SD)	Median $ D' $	Mean MAF (SD)
1	158.49	28,649	5.1 ± 2.9	0.17 ± 0.22	0.015	0.38 ± 0.34	0.24	0.22 ± 0.13
2	136.7	23,788	5.1 ± 2.9	0.19 ± 0.25	0.077	0.57 ± 0.32	0.7	0.21 ± 0.13
3	121.5	22,981	5.2 ± 2.9	0.19 ± 0.25	0.081	0.57 ± 0.32	0.7	0.23 ± 0.13
4	120.5	20,896	5.1 ± 2.9	0.18 ± 0.24	0.071	0.56 ± 0.32	0.58	0.23 ± 0.13
5	121.2	20,089	5.1 ± 2.9	0.20 ± 0.25	0.085	0.58 ± 0.32	0.61	0.22 ± 0.13
6	119.3	23,601	5.2 ± 2.9	0.19 ± 0.25	0.086	0.57 ± 0.32	0.63	0.21 ± 0.13
7	112.2	20,193	5.1 ± 2.9	0.20 ± 0.25	0.082	0.56 ± 0.32	0.64	0.23 ± 0.13
8	113.4	21,721	5.1 ± 2.8	0.20 ± 0.25	0.097	0.58 ± 0.32	0.63	0.22 ± 0.13
9	106.5	20,988	5.2 ± 2.9	0.19 ± 0.25	0.078	0.56 ± 0.32	0.56	0.22 ± 0.13
10	104.4	17,241	5.1 ± 2.9	0.18 ± 0.25	0.064	0.57 ± 0.32	0.53	0.21 ± 0.13
11	107.1	18,341	5.1 ± 2.9	0.19 ± 0.25	0.07	0.56 ± 0.32	0.58	0.21 ± 0.13
12	91.3	15,571	5.1 ± 2.9	0.18 ± 0.25	0.07	0.55 ± 0.32	0.51	0.21 ± 0.13
13	84.1	14,152	5.1 ± 2.9	0.21 ± 0.26	0.094	0.59 ± 0.32	0.67	0.21 ± 0.13
14	83.7	16,917	5.2 ± 2.8	0.18 ± 0.25	0.079	0.56 ± 0.32	0.55	0.25 ± 0.13
15	85.3	14,914	5.1 ± 2.9	0.18 ± 0.25	0.064	0.54 ± 0.32	0.52	0.23 ± 0.13
16	81.9	14,731	5.2 ± 2.9	0.18 ± 0.25	0.072	0.56 ± 0.32	0.51	0.21 ± 0.13
17	75.3	14,191	5.1 ± 2.9	0.19 ± 0.26	0.068	0.55 ± 0.32	0.59	0.22 ± 0.13
18	66.8	11,491	5.1 ± 2.9	0.15 ± 0.22	0.051	0.52 ± 0.32	0.55	0.22 ± 0.13
19	63.7	9,971	5.1 ± 2.9	0.20 ± 0.26	0.088	0.59 ± 0.32	0.66	0.21 ± 0.13
20	71.7	12,831	5.1 ± 2.9	0.18 ± 0.24	0.067	0.54 ± 0.32	0.53	0.21 ± 0.13
21	71.5	12,981	5.1 ± 2.9	0.17 ± 0.42	0.064	0.55 ± 0.32	0.52	0.23 ± 0.13
22	61.4	10,531	5.1 ± 2.9	0.16 ± 0.23	0.061	0.54 ± 0.32	0.51	0.21 ± 0.13
23	52.4	9,541	5.1 ± 2.9	0.17 ± 0.24	0.063	0.55 ± 0.32	0.56	0.21 ± 0.13
24	62.2	11,201	5.1 ± 2.9	0.17 ± 0.23	0.063	0.54 ± 0.32	0.52	0.22 ± 0.13
25	42.91	7,613	5.1 ± 2.9	0.16 ± 0.22	0.058	0.55 ± 0.32	0.52	0.21 ± 0.13
26	51.8	9,514	5.1 ± 2.9	0.17 ± 0.23	0.062	0.54 ± 0.31	0.52	0.22 ± 0.13
27	45.5	7,981	5.1 ± 2.9	0.09 ± 0.18	0.02	0.42 ± 0.31	0.31	0.21 ± 0.13
28	46.3	7,941	4.9 ± 2.9	0.02 ± 0.07	0.008	0.25 ± 0.25	0.15	0.21 ± 0.13
29	51.3	8,321	4.9 ± 2.9	0.002 ± 0.01	0.001	0.12 ± 0.14	0.09	0.21 ± 0.13

Chromosome 1 was found longer and chromosome 25 was shortest with 158.49 Mb and 42.91 Mb, respectively. The SNPs proportion < 0.20 MAF was observed (Figure 1), which is similar to the previous studies in different indicine breeds (McKay *et al.*, 2007; Silva *et al.*, 2010; Espigolan *et al.*, 2013). Khatkar *et al.* (2008) described that the MAF threshold affects the extent of LD and distribution. Although, the mean MAF 0.21 was found in this study, which is considerably high than that reported by Matukumalli *et al.* (2009) and HapMap consortium (2009) in Nellore cattle (0.19 to 0.20). However, the mean MAF was similar to that reported by Espigolan *et al.* (2013) in the study conducted on Nellore cattle. In this study, Chromosomes 2, 4, 7, 15, 17, 25 and 26 showed a higher minor allele frequency proportion (< 0.10), while chromosomes 6, 8, 16, 22 and 23 showed a lower proportion of minor allele frequency (< 0.10). All possible SNP pairs on the same chromosome separated by ≤ 100 kb formed 9,162,314 combinations of SNP pairs to estimate linkage

disequilibrium across the 29 autosomes. The overall mean linkage disequilibrium (r^2 and $|D'|$) between marker pairs measured was 0.18 and 0.55, respectively. Espigolan *et al.* (2013) genotyped Nellore cattle using 444, 986 SNPs and found mean linkage disequilibrium (r^2 and $|D'|$) between adjacent markers of 0.17 and 0.52, respectively. The present study results and those reported in other studies endorse that $|D'|$ parameter overestimates linkage disequilibrium exclusively in case of low MAF.

The mean LD between SNPs ranged from 0.002 to 0.19 for r^2 and from 0.13 to 0.58 for $|D'|$ across all autosomes (Table 1). This is similar to that previously reported in Nellore cattle (Espigolan *et al.*, 2013). In this study low levels of r^2 were estimated at chromosomes 1, 27, 28 and 29, which were comparatively lower than that reported in zebu cattle (McKay *et al.*, 2007; Silva *et al.*, 2010), but Espigolan *et al.* (2013) reported a similar value of these chromosomes in Nellore cattle. Diversity in LD pattern of the different genomic region is due to variation in autosomal recombination rate (Arias *et al.*, 2009). Therefore, the results of this study revealed that chromosomes 1, 27, 28 and 29 might be attributed to variation in sampling since markers number, density, MAF mean or MAF proportion did not differentiate from the other previous autosomes studies.

To investigate the LD decline, synthetic SNPs pairs were categorized into bins (intervals) according to the distance between markers. Figures 2 and 3 depict the LD decline and mean r^2 and $|D'|$ value for each interval per chromosome and for whole genome and LD decreased with increasing distance between markers (Table 2). Comparatively, $|D'|$ change showed a less marked LD decrease. At distance < 30 Kb moderate r^2 were observed (0.19 to 0.33). The mean r^2 decreased in 0.19 to 0.10 were observed with increase in the distance between markers from 30 to 100 kb. At 10 kb marker distances, high r^2 variation was estimated. Average 38.81 and 41.74 kb spacing was observed, when LD was higher than 0.30 and 0.15 and there was not all markers with 40 to 50 kb spacing and have higher r^2 value than 0.3. The r^2 (> 0.15 and > 0.30) for markers proportion ranging from 33 to 54 percent and from 22 to 41 percent, respectively, which was similarly reported on a distance between fewer than 40 kb as reported in Nellore cattle (Espigolan *et al.*, 2013). However, this proportion of markers for r^2 was lower than reported by Sargolzaei *et al.* (2008) in 821 Holstein sires using 5, 564 SNPs at 0.3 threshold for LD (r^2) and 68.34 % markers spacing found from 0 to 0.1 Mb. Qanbari *et al.* (2010) reported the proportion of 29 % markers with $r^2 > 0.25$ separated by less than 100 kb.

Table 2 Linkage disequilibrium (r^2 and $|D'|$) between pairs (n) of synthetic SNPs separated by different distances

Distance (kb)	N	r^2 (SD)	r^2 (median)	$ D' $ (SD)	$ D' $ (median)	$r^2 > 0.3$	$r^2 > 0.15$	$ D' > 0.8$
0-1	54304	0.34 ± 0.33	0.21	0.71 ± 0.32	0.89	41	56	43
1-2	99588	0.32 ± 0.33	0.19	0.69 ± 0.33	0.87	39	54	41
2-3	104665	0.31 ± 0.32	0.18	0.69 ± 0.33	0.85	37	53	40
3-4	112749	0.30 ± 0.32	0.17	0.67 ± 0.33	0.83	37	52	41
4-5	108789	0.29 ± 0.31	0.16	0.66 ± 0.33	0.81	35	50	38
5-10	514150	0.27 ± 0.30	0.14	0.63 ± 0.33	0.76	32	47	39
10-20	977976	0.23 ± 0.28	0.1	0.61 ± 0.33	0.67	29	42	35
20-30	946585	0.20 ± 0.26	0.08	0.56 ± 0.33	0.6	23	37	31
30-40	927758	0.18 ± 0.24	0.07	0.53 ± 0.33	0.56	22	34	29
40-50	9185546	0.16 ± 0.23	0.06	0.51 ± 0.33	0.52	18	31	24
50-60	908620	0.15 ± 0.22	0.05	0.49 ± 0.32	0.51	16	28	26
60-70	901968	0.14 ± 0.21	0.05	0.49 ± 0.32	0.47	14	27	27
70-80	895954	0.13 ± 0.20	0.04	0.47 ± 0.32	0.46	13	25	24
80-90	893182	0.12 ± 0.19	0.04	0.46 ± 0.31	0.44	12	23	23
90-100	889674	0.11 ± 0.18	0.04	0.46 ± 0.31	0.43	11	22	20

Espigolan *et al.* (2013) reported LD (r^2) from 0.11 to 0.05 at 100 kb to 1, 000 kb. The results of this study revealed the same values at this distance but McKay *et al.* (2007) estimated mean r^2 between adjacent markers range from 0.15 to 0.20 at distance of 100 kb. In this study, a linear relationship was

observed between LD (r^2) and chromosomes length, as the length of chromosome increases due to the decrease in recombination rate (Yan *et al.*, 2009). Bohmanova *et al.* (2010) did not find any association with LD level and chromosome size. The low allele frequency of SNPs pairs underestimates LD (r^2). Estimation of LD (r^2) is considered less biased when allele frequency polymorphism is high in data set (Reich and Lander, 2001). Therefore, in this study, we examined the MAF effects on the estimation of LD (r^2) and $|D'|$ (Figure 4 and 5). At high MAF threshold (0.15), the distance between markers was short. The LD was higher between markers, similarly reported by Yan *et al.* (2009) in maize. The $|D'|$ was observed unchanged at different MAF threshold for the adjacent marker (< 10 kb) and at the threshold of MAF increased, the $|D'|$ was lower as reported in the study of different cattle breeds (Bohmanova *et al.*, 2010). SNPs pair with low frequencies of allele, $|D'|$ divided by a small number and resulting in $|D'|$ large value indicated in the formula, where the denominator is the allele frequencies product (Reich & Lander, 2001). The results of this study revealed a significant variation in the r^2 magnitude and pattern of the Sahiwal genome. The cause of this variation might be recombination rate, genetic drift, heterozygosity and selection (Reich & Lander, 2001).

The r^2 level between adjacent markers (30 - 40 kb) detected in this study was similar as reported in *Bos indicus* but lower than *Bos taurus* cattle (McKay *et al.*, 2007). The breed difference between indicus and taurus separated markers by 80 to 100 kb, because of population's recent history. Generally, LD comparison is difficult to obtain in different studies due to the difference in sample size, LD measures, markers type and density of markers.

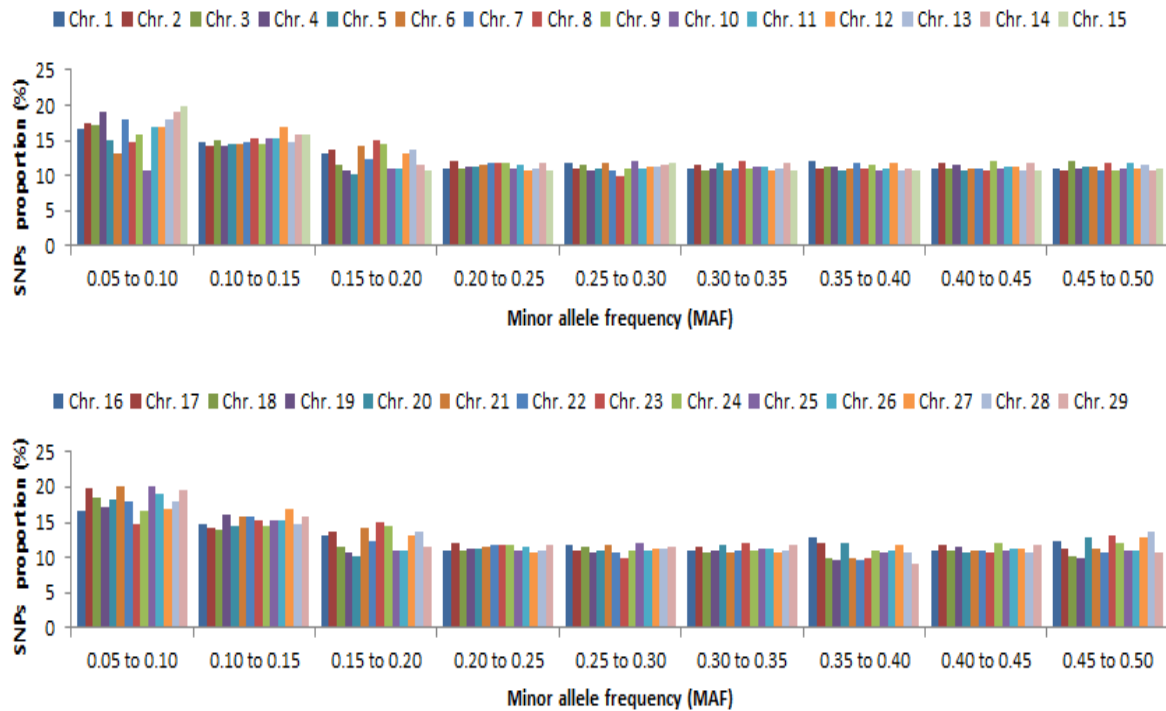


Figure 1 SNPs proportion for each chromosome for various minor allele frequencies

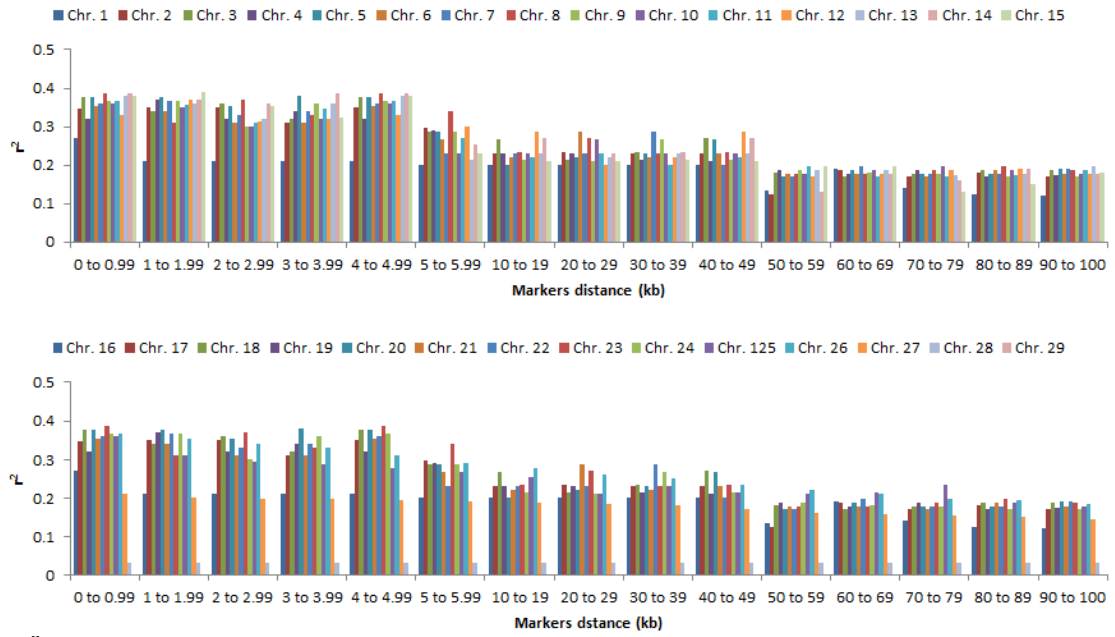


Figure 2 r^2 value for distance between markers per chromosomes

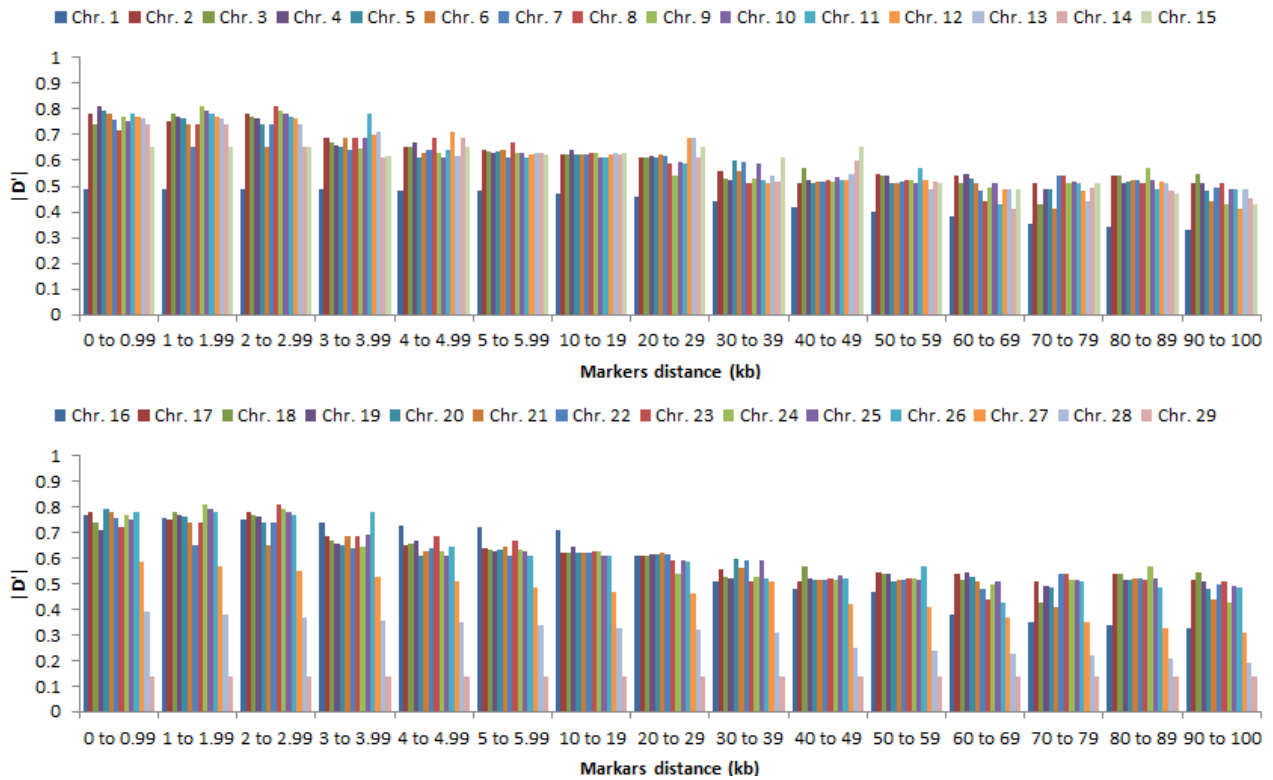


Figure 3 $|D'|$ value for markers distance per chromosome

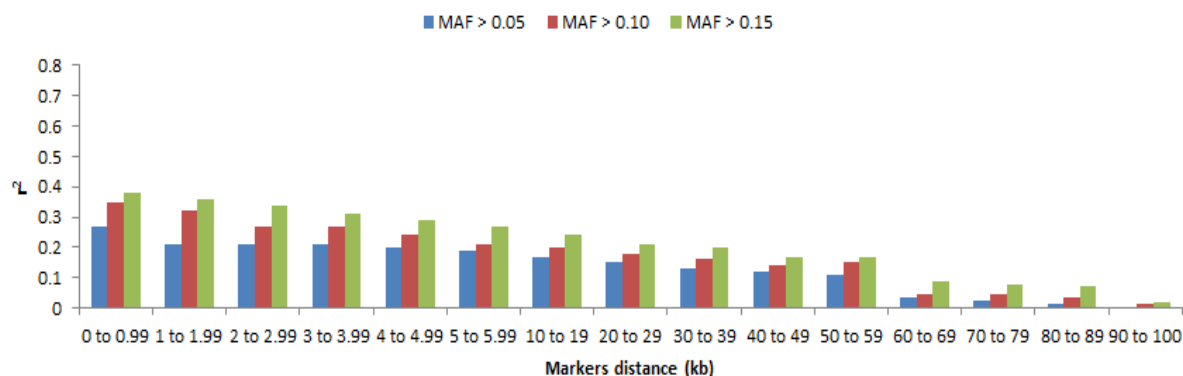


Figure 4 r^2 value for different thresholds of MAF (>0.05, >0.10, >0.15)

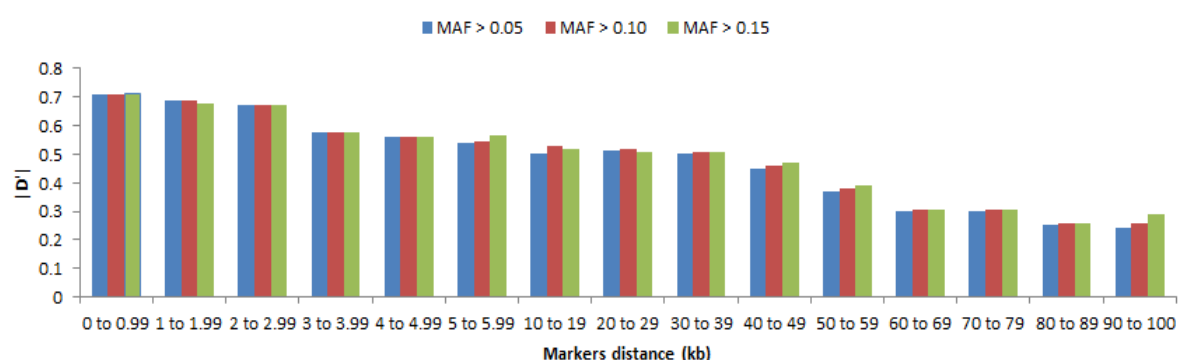


Figure 5 $|D'|$ value for different thresholds of MAF (>0.05, >0.10, >0.15)

However, it is difficult to compare LD level found in other studies due to change in sample size, LD measures, markers type, markers density, and recent population history (Pritchard & Przeworski, 2001). Nevertheless, *Bos taurus* and *Bos indicus* difference emerged during the domestication process and selection, and consequently effective population size seem to elucidate the inconsistency in linkage disequilibrium at short distance between markers (Tenesa *et al.*, 2007). Another reason is the fact that indicine populations present higher proportion of low allele frequency in the bovine high density SNPs BeadChip than the taurine, which in turn affects LD estimations (Weiss & Clark, 2002; McKay *et al.*, 2007).

Conclusions

The LD level assessed for markers segregated by less than 30 kb designates that the high density bovine SNPs BeadChip will probably be an appropriate tool for the predication of genomic breeding values in Sahiwal cattle. Additional studies considering the linkage disequilibrium in larger samples of Sahiwal population are required to ratify the linkage disequilibrium estimates found in this study.

Acknowledgements

The authors would like to acknowledge cooperation and support of the animal breeder and research institutions for animal's blood provision. ARS/USDA platform is acknowledged for laboratory facilities. Financial support from Higher Education Commission (HEC) of Pakistan is also appreciated.

Authors' Contributions

HM and TSS designed the experiment, HM, HJH, KE carried out the analysis. HM, AA drafted the manuscript. JJK assisted with statistical analysis. HM, AA and TSS structured scientific content. All authors provided editorial suggestions and approved the final manuscript draft.

Conflict of Interest Declaration

The authors of this manuscript declare that they have no conflict of interests.

References

- Arias, J.A., Keehan, M., Fisher, P., Coppieters, W. & Spelman, R., 2009. A high density linkage map of the bovine genome. *BMC Genet.* 10, 18.
- Bennewitz, J., Solberg, T. & Meuwissen, T., 2010. Genomic breeding value estimation using nonparametric additive regression models. *Genet. Sel. Evo.* 41, 20.
- Bohmanova, J., Sargolzaei, M. & Schenkel, F., 2010. Characteristics of linkage disequilibrium in North American Holsteins. *BMC Genom.* 11, 421.
- Espigolan, R., Fernando, B., Arione, A.B., Fabio, R.P.S., Daniel, G.M.G., Rafael, L.T., Diércles, F.C., Henrique, N.O., Humberto, T. & Mehdi, S., 2013. Study of whole genome linkage disequilibrium in Nellore cattle. *BMC Genom.* 14, 305.
- FAO., 2007. The state of the world's animal genetic resources for food and agriculture. Food and Agriculture Organization; Rome, Italy.
- Farnir, F., Coppieters, W. & Arranz, J.J., 2000. Extensive genome-wide linkage disequilibrium in cattle. *Genom. Res.* 10, 220-227.
- 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., 2009. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science* 324, 528-532.
- Government of Pakistan. 2006. Livestock Census 2006. Agricultural Census Organization, Statistics Division; Government of Pakistan.
- Hill, W.G. & Robertson, A., 1968. Linkage disequilibrium in finite populations. *Theo. App. Genet.* 38, 226-231.
- Khan, M.S., Zia, U.R., Muqarrab, A.K. & Sohail, A., 2008. Genetic resources and diversity in Pakistani cattle. *Pak. Vet. J.* 28, 95-102.
- Khatkar, M.S., Nicholas, F.W., Collins, A.R., Zenger, K.R., Cavanagh, J.A., Barris, W., Schnabel, R.D., Taylor, J.F. & Raadsma, H.W., 2008. Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high density SNP panel. *BMC Genom.* 9, 187.
- Kim, E.S. & Brian, W.K., 2009. Linkage disequilibrium in the North American Holstein population. *Anim. Genet.* 40, 279-288.
- Marques, E., Schnabel, R., Stothard, P., Kolbehdari, D., Wang, Z., Taylor, J.F. & Moore, S.S., 2008. High density linkage disequilibrium maps of chromosome 14 in Holstein and Angus cattle. *BMC Genet.* 9, 45.
- 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., Tad, S.S. & Van Tassell, C.P., 2009. Development and characterization of a high density SNP genotyping assay for cattle. *PLoS One.* 4, e5350.
- McKay, S.D., Schnabel, R.D., Murdoch, B.M., Matukumalli, L.K., Aerts, J., Wouter C. W., Crews, D., Dias Neto, E. & Gill, C.A., 2007. Whole genome linkage disequilibrium maps in cattle. *BMC Genet.* 8, 74.
- Mustafa, H., Kim, E., Huson, J.H., Adeela, A., David, R., Talat, N.P., Afzal, A., Khalid, J. & Tad, S.S., 2017. Genome-wide SNPs analysis of indigenous Zebu breeds in Pakistan. *Bio. Anim. Hus.* 33, 3-25.
- Pritchard, J.K. & Przeworski M., 2001. Linkage disequilibrium in humans: models and data. *Am. J. Hum. Genet.* 69, 1-14.
- Qanbari, S., Pimentel, E.C.G., Tetens, J., Thaller, G., Lichtner, P., Sharifi, A.R. & Simianer, H., 2010. The pattern of linkage disequilibrium in German Holstein cattle. *Anim. Genet.* 41, 346-356.
- Reich, D.E. & Lander, E.S., 2001. On the allelic spectrum of human disease. *Trend. Genet.* 17, 502-10.
- Sargolzaei, M., Schenkel, F.S., Jansen, G.B. & Schaeffer, L.R., 2008. Extent of linkage disequilibrium in Holstein cattle in North America. *J. Dair. Sci.* 91, 2106-2117.
- Silva, C.R., Neves, H.H.R., Queiroz, S.A., Sena, J.A.D. & Pimentel, E.C.G., 2010. Extent of linkage disequilibrium in Brazilian Gyr dairy cattle based on genotypes of AI sires for dense SNP markers. In: *Proceedings of the 9th World Congress on Genetics Applied to Livestock Production*. Leipzig, Germany: WGAPL.
- Tenesa, A., Navarro, P., Hayes, B.J., Duffy, D.L., Clarke, G.M., Goddard, M.E. & Visscher, P.M., 2007. Recent human effective population size estimated from linkage disequilibrium. *Genom. Res.* 17, 520-526.
- Villa-Angulo, R., Matukumalli, L.K., Gill, C.A., Choi, J., Van Tassell, C.P., John, J. & Grefenstette, J., 2009. High-resolution haplotype block structure in the cattle genome. *BMC Genet.* 10, 19.
- Weiss, K.M. & Clark, A.G., 2002. Linkage disequilibrium and the mapping of complex human traits. *Trend. Genet.* 18, 19-24.
- Yan, J., Shah, T., Warburton, M.L., Buckler, E.S., McMullen, M.D. & Crouch, J., 2009. Genetic characterization and linkage disequilibrium estimation of a global maize collection using SNP markers. *PLoS One.* 4, e845.