Linkage Disequilibrium Estimation of Chinese Beef Simmental Cattle Using High-density SNP Panels

M. Zhu[1], B. Zhu[1,2], Y. H. Wang[1], Y. Wu[1], L. Xu[1], L. P. Guo[1], Z. R. Yuan[1], L. P. Zhang[1],
X. Gao[1], H. J. Gao[1], S. Z. Xu[1] and J. Y. Li[1,*]

[1] Laboratory of Molecular Biology and Bovine Breeding, Institute of Animal Science,
Chinese Academy of Agricultural Sciences, Beijing, China

ABSTRACT: Linkage disequilibrium (LD) plays an important role in genomic selection and mapping quantitative trait loci (QTL). In this study, the pattern of LD and effective population size (N_e) were investigated in Chinese beef Simmental cattle. A total of 640 bulls were genotyped with IlluminaBovinSNP50BeadChip and IlluminaBovinHDBeadChip. We estimated LD for each autosomal chromosome at the distance between two random SNPs of <0 to 25 kb, 25 to 50 kb, 50 to 100 kb, 100 to 500 kb, 0.5 to 1 Mb, 1 to 5 Mb and 5 to 10 Mb. The mean values of r^2 were 0.30, 0.16 and 0.08, when the separation between SNPs ranged from 0 to 25 kb to 50 to 100 kb and then to 0.5 to 1 Mb, respectively. The LD estimates decreased as the distance increased in SNP pairs, and increased with the increase of minor allelic frequency (MAF) and with the decrease of sample sizes. Estimates of effective population size for Chinese beef Simmental cattle decreased in the past generations and N_e was 73 at five generations ago. (**Key Words:** Chinese Beef Simmental Cattle, Bovine Genome, Linkage Disequilibrium, Single Nucleotide Polymorphisms, Effective Population Size)

INTRODUCTION

Linkage disequilibrium (LD) denotes non-random association between alleles at different loci. LD is the theoretical basis of genomic selection (GS) and genome-wide association study (GWAS), that is also important in gene mapping, estimates for effective population size, population structure and so on (Nachman, 2002). Molecular markers such as single nucleotide polymorphisms (SNPs) and microsatellites were widely used to estimate the extent of LD. The level of LD is usually influenced by non-genetic factors and genetic factors containing genetic linkage, selection, the rate of recombination, the rate of mutation, genetic drift, non-random mating and population structure.

The effective population size (N_e) is defined as the number of individuals in an ideal population that would show the same amount of dispersion of allele frequencies under random genetic drift or the amount of inbreeding as in the population under consideration, and is usually less than the absolute population size (Wright, 1938). N_e is an important parameter, as it can help to explain how cattle populations evolve and expand, and by definition describe the rate of inbreeding accumulation and loss of genetic variation. Estimates for N_e can be obtained from heterozygote excess or LD. Presently, estimates for N_e based LD data are are more frequently used than heterozygote excess, and therefore complement evolutionary studies of cattle populations (Hayes et al., 2003).

Recently the discovery of large numbers of SNP through sequencing of the cattle genome has generated extensive research in quantifying LD characteristics (Farnir et al., 2000; Odani et al., 2006; McKay et al., 2007; Sargolzaei et al., 2008; Kim and Kirkpatrick, 2009; Qanbari et al., 2010). A recent report showed that high density markers were used to study the extent of LD in Angus, Charolais and Crossbred beef cattle (Lu et al., 2012). However, similar studies were not reported in Simmental cattle which are an important economic breed of beef cattle. LD indicates population characteristics and has different pattern on each chromosome. Hence, it is necessary to study the extent of LD, and then to estimate effective population size in Simmental cattle.

China's role in international beef markets has grown

* Corresponding Author: J. Y. Li. Tel: +86-10-62818176, Fax: +86-10-62817806, E-mail: JL1@iascaas.net.cn
[2] Laboratory of Animal Genetic and Breeding, Agricultural University of Hebei, Baoding, China.

significantly in the past years, and domestic production is projected to continue to increase (Longworth et al., 2001). However, China does not have special-purpose beef cattle. To increase beef production, American, Canadian and Australian Simmental cattle have been introduced into China and crossbred with native dual-purpose Simmental cattle, which are named Chinese beef Simmental cattle. In current research, high density SNPs data from Chinese beef Simmental cattle were used to analyze the pattern of LD, and to infer the effective population size up to 2000 generations ago. Meanwhile, we evaluated the effects of minor allelic frequency (MAF) and sample size on LD estimations.

MATERIALS AND METHODS

Animals

Experimental animals consisted of 640 young Simmental bulls, born in 2008 to 2010, originated from Ulgai, located at Xilingol league, Inner Mongolia, China. DNA was extracted from blood of the bulls using the routine procedures. The IlluminaBovineHD chip was used to genotype 504 young bulls and their autosomal chromosomes contained total of 735,293 SNPs. Additionally, 136 young bulls were genotyped with IlluminaBovineSNP50, and 51,582 SNPs were detected on their autosomal chromosomes. There were 46,000 SNPs in common between two chips. In the present study, quality control standards for SNPs data were Hardy-Weinberg equilibrium (p>10^{-3}), MAF>0.05, SNP call rate >0.95 and Mendel error rate <0.05. 35,079 common SNPs survived after being filtered on quality control standards, which were used to analyze the extent of LD.

LD estimation

Several statistics parameters were proposed to measure the extent of LD. D′ (Lewontin, 1964) and r^2 (Hill, 1974) were widely used in practice, but their functions are different. r^2 was considered to be a better descriptor of LD as it is more robust and not sensitive to changing gene frequency and effective population size (Terwilliger et al., 2002; Zhao et al., 2007).

Assume two loci A and B, each locus has two alleles (denoted A$_1$, A$_2$ and B$_1$, B$_2$, respectively). P$_{A1}$, P$_{A2}$, P$_{B1}$ and P$_{B2}$ are the frequency of each of the alleles. P$_{11}$, P$_{12}$, P$_{21}$ and P$_{22}$ show the frequency of haplotypes A$_1$B$_1$, A$_1$B$_2$, A$_2$B$_1$ and A$_2$B$_2$. Thus, r^2 can be expressed as:

$$r^2 = \frac{(P_{11}P_{12} - P_{12}P_{21})^2}{P_{A1}P_{A2}P_{B1}P_{B2}}$$
(Equation 1)

PLINK (Purcell et al., 2007) includes a set of options to calculate pair-wise linkage disequilibrium between SNPs,

and to present or process this information in various ways. In this study, we used the command *plink -cow -bfile filename -ld-window-r2 0 -out outname*. To display the decay of LD, distances of pair-wise SNPs were binned into seven types of intervals (0 to 25 kb, 25 to 50 kb, 50 to 100 kb, 100 to 500 kb, 0.5 to 1 Mb, 1 to 5 Mb and 5 to 10 Mb) along the first 10 Mb of each chromosome, and mean r^2 was computed for each interval. Table 2 shows information for all the SNP pair groups.

Three factors, chromosomes, MAF and sample sizes, affecting LD estimation were studied based on r^2 data computed above.

Genetic distance

In the high-density SNP chip, genetic distance for SNP pairs could not be obtained. Therefore, physical distance was used to replace genetic distance for the estimation of effective population size in the current study. l00 kb of physical distance in genetic distance is approximate equivalent to 0.1 cM. SNP physical position from the UMD 3.1 bovine assembly (http://www.ncbi.nlm.nih.gov/ assembly/313678/) was used in this study.

Effective population size estimation

LD data make it feasible to estimate N_e. Sved (1971) has proposed the relationship formula for LD and N_e as follows:

$$r^2 = \frac{1}{kN_ec+\alpha} + \frac{1}{n}$$
(Equation 2)

$$N_e = \frac{1}{(r^2 - \frac{1}{n})kc} - \frac{2}{kc}$$
(Equation 3)

Where N_e is the effective population size, r^2 is the mean value of LD estimation value of SNP pairs. k is 4 or 2, which denotes autosomal chromosomes or sex chromosomes respectively. c represents the genetic distance in Morgan (M). Generation is calculated with 1/2c. n is the chromosome experimental sample size. $\alpha = 1$ (Sved, 1971) in the absence of mutation, otherwise $\alpha = 2$ (Lewontin, 1964; Weir and Hill, 1980; McVean, 2002). In the majority of studies, the formula for the absence of mutation is chosen to estimate N_e. Hence, in our study, $k = 4$ and $c = 1$ were chosen. Seven types of SNP pairs with the physical distances 25 kb, 50 kb, 100 kb, 500 kb, 1 Mb, 5 Mb and 10 Mb were respectively chosen to estimate the N_e of Chinese beef Simmental cattle from 5 generations ago.

RESULTS

SNP statistics

SNPs information for every autosomal chromosome is

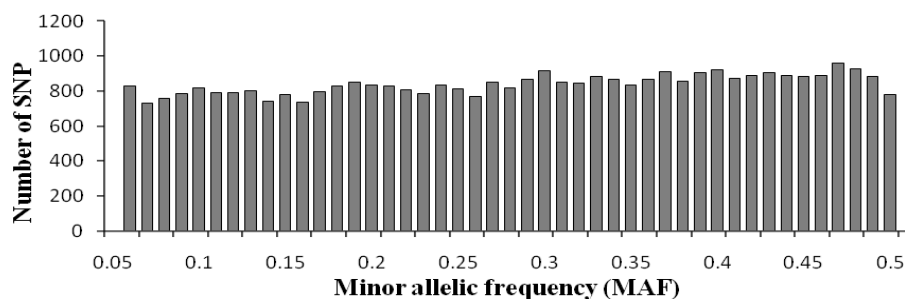**Table 1.** Statistical information for analyzed SNP

Chromosome	Length (Mb)	Number of SNP	Average SNP Interval (Mb)	Longest interval (Mb)	Shortest interval (kb)
BTA1	158.14	2,290	0.05	1.13	0.13
BTA2	136.52	1,815	0.06	1.47	0.08
BTA3	121.37	1,726	0.05	1.32	0.11
BTA4	120.61	1,731	0.05	0.59	4.90
BTA5	119.73	1,454	0.06	0.61	0.15
BTA6	119.21	1,769	0.05	1.60	1.36
BTA7	112.61	1,543	0.05	1.49	0.00
BTA8	113.36	1,629	0.05	0.52	1.80
BTA9	105.46	1,381	0.06	0.93	0.45
BTA10	104.21	1,491	0.05	2.41	0.28
BTA11	107.04	1,514	0.05	1.10	0.88
BTA12	91.09	1,128	0.06	3.34	0.24
BTA13	84.15	1,242	0.05	1.95	0.38
BTA14	84.61	1,205	0.05	3.62	0.11
BTA15	85.05	1,141	0.05	0.86	2.90
BTA16	80.92	1,071	0.05	2.39	0.17
BTA17	74.96	1,080	0.05	0.81	4.80
BTA18	65.97	928	0.05	0.97	5.60
BTA19	64.01	967	0.05	0.54	1.40
BTA20	71.79	1,096	0.05	0.51	3.50
BTA21	70.61	921	0.05	1.22	0.90
BTA22	60.54	863	0.05	0.46	0.20
BTA23	52.13	749	0.05	1.14	4.20
BTA24	62.64	910	0.05	0.47	0.10
BTA25	42.80	687	0.04	0.49	1.30
BTA26	51.68	732	0.05	0.73	6.60
BTA27	45.36	671	0.05	1.28	11.00
BTA28	46.19	657	0.05	2.14	0.02
BTA29	50.97	688	0.05	0.91	1.80

given in Table 1. The total autosomal chromosome length of Chinese beef Simmental cattle was 2,541.30 Mb. The longest *Bos taurus* autosomal chromosome is BTA1 (length = 158.14 Mb), and the shortest is BTA25 (length = 42.80 Mb). 35,079 common SNPs between two chips covered the whole genome in this study. Average adjacent SNPs spacing was 54.17±61.44 kb, and the largest spacing situated on BTA14 was 3620 kb (between ARS-BFGL-NGS-37733 and Hapmap42739-BTA-95927). The mean MAF of the genome

was 0.28±0.13, and followed an almost uniform distribution, as can be seen in Figure 1.

Extent of LD across the genome

The mean values of r^2 for each autosomal chromosome for distance bins of 0 to 25 kb, 25 to 50 kb, 50 to 100 kb, 100 to 500 kb, 0.5 to 1 Mb, 1 to 5 Mb and 5 to 10 Mb were calculated. Table 2 shows that the average r^2 is 0.30, 0.23, 0.16, 0.08, 0.05, 0.04 and 0.03 at different distance bins for

Figure 1. Minor allele frequency (MAF) distribution for total SNPs.
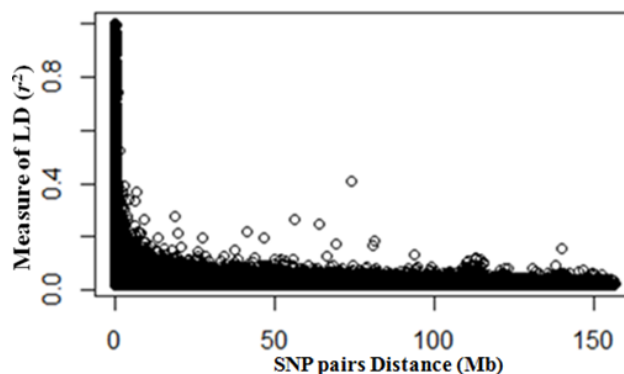
Table 2. Statistical information for LD over various distances

Distance	Number of SNP pairs	Average r^2
0-25 kb	4,100	0.30
25-50 kb	12,412	0.23
50-100 kb	18,855	0.16
100-500 kb	92,378	0.08
0.5-1 Mb	179,984	0.05
1-5 Mb	368,467	0.04
5-10 Mb	654,755	0.03

r^2: denotes the extent of LD.

Figure 2. LD decay for Chinese beef Simmental cattle in whole autosomal chromosome.

Simmental cattle, respectively. Figure 2 shows the LD decay over varying distances of the genome. The measured LD was high for pairs of SNPs within close proximity. However, there is a strong LD in the long distance SNP pairs.

The extent of LD was significantly different among chromosomes. The average r^2 for SNPs separated by intervals 0 to 25 kb, 25 to 50 kb, 50 to 100 kb, 100 to 500 kb, 0.5 to 1 Mb, 1 to 5 Mb and 5 to 10 Mb in each autosomal chromosome are presented in Table 3. The mean value of r^2 for distances less than 25 kb was 0.30, but higher for BTA9 and BTA21 (0.363 and 0.364, respectively), and lower for BTA27 (0.209). The average of r^2 was 0.30 in

Table 3. Statistical information for average r^2 as distance between pairs of SNP up to 10Mb for the genome

CHR	SNP pairs Distance						
	0-25 Kb	25-50 Kb	50-100 Kb	100-500 Kb	0.5-1 Mb	1-5 Mb	5-10 Mb
BTA1	0.322±0.298	0.255±0.257	0.186±0.201	0.083±0.103	0.049±0.034	0.039±0.022	0.034±0.015
BTA2	0.335±0.317	0.251±0.252	0.165±0.179	0.085±0.097	0.051±0.037	0.040±0.023	0.034±0.015
BTA3	0.302±0.291	0.229±0.235	0.166±0.188	0.078±0.089	0.048±0.032	0.038±0.023	0.033±0.016
BTA4	0.277±0.267	0.242±0.245	0.173±0.186	0.089±0.108	0.053±0.042	0.040±0.023	0.033±0.015
BTA5	0.282±0.282	0.250±0.255	0.169±0.197	0.085±0.109	0.055±0.045	0.044±0.029	0.035±0.018
BTA6	0.322±0.318	0.249±0.257	0.174±0.195	0.089±0.103	0.058±0.047	0.044±0.029	0.035±0.017
BTA7	0.318±0.296	0.253±0.251	0.179±0.189	0.085±0.106	0.051±0.041	0.041±0.029	0.034±0.016
BTA8	0.291±0.287	0.246±0.256	0.158±0.177	0.079±0.096	0.049±0.032	0.039±0.022	0.034±0.015
BTA9	0.363±0.321	0.234±0.250	0.166±0.188	0.073±0.083	0.048±0.032	0.039±0.023	0.034±0.015
BTA10	0.286±0.306	0.231±0.241	0.164±0.187	0.072±0.082	0.045±0.031	0.037±0.018	0.033±0.014
BTA11	0.318±0.291	0.251±0.257	0.169±0.195	0.079±0.092	0.051±0.040	0.041±0.026	0.034±0.016
BTA12	0.257±0.258	0.207±0.231	0.160±0.188	0.074±0.086	0.050±0.036	0.040±0.024	0.034±0.016
BTA13	0.291±0.306	0.184±0.196	0.137±0.160	0.069±0.080	0.047±0.034	0.037±0.020	0.032±0.014
BTA14	0.290±0.299	0.233±0.230	0.162±0.181	0.081±0.099	0.049±0.035	0.038±0.021	0.033±0.014
BTA15	0.285±0.293	0.213±0.229	0.164±0.184	0.076±0.085	0.049±0.037	0.039±0.023	0.033±0.017
BTA16	0.330±0.304	0.248±0.238	0.165±0.190	0.075±0.087	0.048±0.034	0.037±0.019	0.033±0.014
BTA17	0.289±0.288	0.227±0.241	0.151±0.167	0.075±0.085	0.048±0.033	0.038±0.021	0.033±0.014
BTA18	0.343±0.318	0.231±0.237	0.140±0.169	0.073±0.077	0.048±0.030	0.039±0.021	0.033±0.013
BTA19	0.260±0.245	0.201±0.222	0.148±0.183	0.069±0.079	0.046±0.030	0.037±0.020	0.032±0.017
BTA20	0.234±0.241	0.215±0.226	0.152±0.175	0.074±0.081	0.049±0.034	0.039±0.023	0.034±0.018
BTA21	0.364±0.308	0.248±0.231	0.178±0.192	0.079±0.090	0.050±0.036	0.040±0.025	0.033±0.022
BTA22	0.337±0.324	0.228±0.227	0.143±0.169	0.074±0.081	0.05±0.036	0.04±0.022	0.033±0.014
BTA23	0.277±0.272	0.196±0.200	0.142±0.167	0.077±0.104	0.052±0.046	0.038±0.019	0.032±0.013
BTA24	0.329±0.286	0.259±0.251	0.163±0.191	0.076±0.080	0.049±0.031	0.04±0.023	0.034±0.015
BTA25	0.265±0.250	0.19±0.220	0.129±0.148	0.064±0.068	0.045±0.026	0.036±0.018	0.031±0.012
BTA26	0.291±0.279	0.231±0.239	0.154±0.183	0.070±0.071	0.049±0.035	0.039±0.021	0.033±0.017
BTA27	0.209±0.220	0.193±0.221	0.132±0.163	0.064±0.070	0.048±0.033	0.037±0.019	0.032±0.013
BTA28	0.225±0.253	0.198±0.225	0.13±0.145	0.061±0.058	0.044±0.028	0.036±0.019	0.032±0.012
BTA29	0.296±0.278	0.206±0.233	0.130±0.152	0.064±0.072	0.046±0.030	0.038±0.023	0.032±0.013

CHR denotes chromosome. r^2: Means±SE.
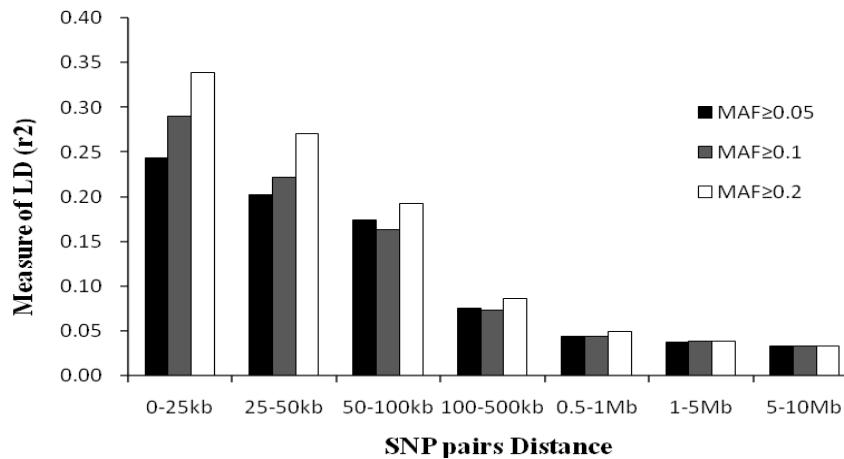
Figure 3. Average r^2 estimates at different physical distances for three different minor allelic frequency (MAF) thresholds. Mean LD estimates are pooled over all autosomal chromosomes, and three different minimum threshold cut off levels for minimum allele frequency are shown.

SNP pairs with physical distances of <25 kb and decreased to 0.16 at distances of 50 to 100 kb, this result was similar to that previously reported (Qanbari et al., 2010; Lu et al., 2012). A similar study found the extent of LD ($r^2 = 0.59$) in approximately 50 kb on north American Holstein cattle, which was much larger than that found in our study (Sargolzaei et al., 2008).

MAF and LD

In this study, three different minimum allelic frequency (MAF) thresholds (0.05, 0.1 and 0.2) were used to study the effects of MAF on the extent of LD. Figure 3 shows that MAF has a significantly effect on the mean value of r^2, especially over short distances (0 to 25 kb). The mean values of r^2 increase significantly with an increasing MAF. For example, from 0 to 25 kb, the mean value of r^2 for MAF≥0.05 was 0.24, however, with MAF≥0.1 and 0.2, the

mean value of r^2 increased to 0.29 and 0.34, respectively.

Sample size and LD estimates

As can be seen in Figure 4, sample sizes affect the LD estimation value. In this paper, five different sample sizes of n = 25, n = 50, n = 100, n = 200 and n = 400 were randomly selected from the total set to study the effect of sample size on estimates of the level of LD. The mean r^2 were greater when sample size is smaller, and this phenomenon is more noticeable for LD estimation across a SNP interval more than 500 kb. There were no significant differences for LD estimates when sample sizes were greater than 400 and SNP distances less than 50 kb.

Effective population size

The extent of LD for different chromosome fragment length could reflect the effective population size of different
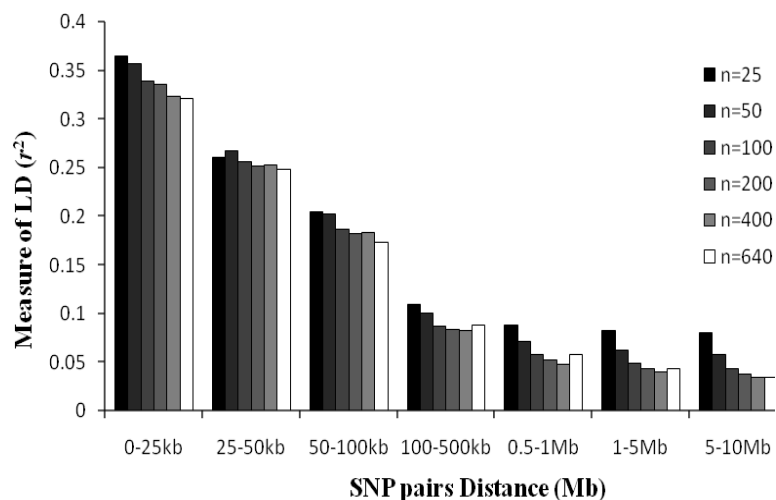
Figure 4. Average r^2 estimates at different physical distances for six different sample sizes. Mean LD estimates are pooled over all chromosomes, and six different sample sizes are shown.

Table 4. Statistical information for effective population sizes of Simmental cattle

	SNP pairs distance						
	25 kb	50 kb	100 kb	500 kb	1 Mb	5 Mb	1 Mb
Genetic distance	0.00025	0.0005	0.001	0.005	0.01	0.05	0.1
Generations ago	2,000	1,000	500	100	50	10	5
N_e	2377	1697	1344	611	484	123	73

Genetic distance: Morgan. N_e = Effective population size.

past generations. Table 4 shows N_e of Simmental cattle in past generations. Estimates of N_e for 2,000 generations ago was approximately 2,377 and down to 73 at 5 generations ago. Estimates of N_e for Chinese beef Simmental cattle show an increasing trend when plotted against increasing past generations (Figure 5).

DISCUSSION

Recent developments in high-throughput SNP panels have generated enthusiasm and interesting in GS and GWAS on cattle. Linkage disequilibrium maps can increase power and precision in association mapping. Qanbari et al. (2010) reported an average level LD of 0.30 over pair wise distances less than 25 kb based on 40,854 SNPs in 810 German Holstein cattle. Kim and Kirkpatrick (2009) reported LD of >0.80 over genomic regions of approximately 50 kb using 7119 SNPs in North America Holstein cattle. Lu et al. (2012) reported the extent of LD in Angus, Charolais and crossbred beef cattle based on Illumina BovineSNP50_v2 Beadchip and Illumina BovineSNP50_v1 Beadchip, with the level of LD being 0.29, 0.22 and 0.15 when the distance range between markers is 0 to 30 kb, respectively. This could be attributed in part to the difference in populations between the current study and previously reported research. Furthermore, in the current study, we used 35,079 SNPs distributed across the entire bovine autosomal chromosome for the analysis of LD

in Chinese beef Simmental cattle. The r^2 statistic denotes the extent of LD. The extent of LD showed a decreasing tendency with increasing distances of the genome. The mean r^2 was much higher between close loci, and the result was the same as previously reported estimates (Farnir et al., 2000; Smith et al., 2006; Kim and Kirkpatrick, 2009; Qanbari et al., 2010; Lu et al., 2012). However, a low level of LD can exist between two SNPs that are closely adjacent, while markers that are more distant can show a higher than expected level of LD. This situation also appeared in linkage disequilibrium studies on human and model animals (Reich et al., 2001). It could be caused by selection, the rate of recombination, mutation and genetic drift (Nachman, 2002).

The mean r^2 values were different for the same fragment length on different autosomal chromosomes. Higher LD was found for BTA21. This may reflect selection for traits that are strongly influenced by QTL on this chromosome in this breed. Chinese beef Simmental cattle are a popular breed in Chinese beef production and genetic trends suggest a strong selection for growth and meat traits. A majority of studies have shown highly significant evidence for the presence of QTLs affecting meat traits (McClure et al., 2010) on BTA21. In addition, when selection operates at a locus, the neighboring loci in close linkage with the locus under selection will have an enhanced extend of LD. When selection occurs at multiple loci in epistasis, LD between loci under epistatic selection and their tightly linked loci
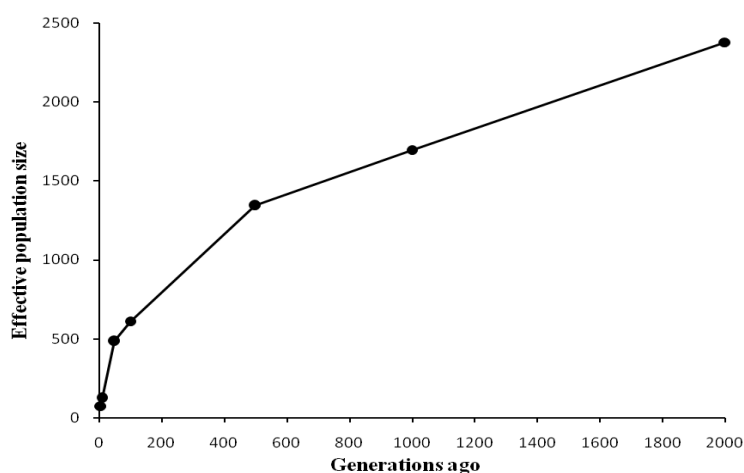
Figure 5. Estimated N_e for Chinese beef Simmental cattle over time from linkage disequilibrium data.

will be created and enhanced (Du et al., 2007).

Estimates of LD across the whole genome could be affected by many factors. In this study, removing SNPs with very low MAFs also lead to lower numbers of SNPs available for study, which can also lead to bias of LD estimates. There are several published papers observing a similar phenomenon in other species (Khatkar et al., 2008; Yan et al., 2009; Qanbari et al., 2010). LD estimation is greater with MAF increasing at a short SNP pairs distance, but the phenomenon is not sensitive when SNP pairs achieve a distance of 1 Mb. Sample size is another factor that affects estimation of the extent of LD (Khatkar et al., 2008; Yan et al., 2009). A small sample size (n = 25) can also lead to the biased estimates for LD. However, there are no significant differences for the mean r^2 when sample sizes exceed 100, especially when the given extent interval of LD is less than 100 kb. In addition, previous research on Holstein cattle demonstrated that a sample of 400 or more was required for reliable estimation of LD (Khatkar et al., 2008). A similar study in humans that found sample sizes would be even higher, which may be due to humans having a larger effective population size (Chen et al., 2006).

Hill (1974) proposed a method for estimating effective population sizes. In this method, estimates of N_e depend on the number of animals alive at any time and the variance of progeny number per sire. In addition, previous research showed that the latter played a key role in the decrease of the population size (Mukai et al., 1989; Nomura et al., 2001). To maximize the net response in economic merit for dairy cattle, FAO (1998) reported an effective population size of 50 per generation was required to maintain the fitness in a breed. Goddard and Smith (1990) suggested a minimum effective number of 10 bull sires per generation, equivalent to 40 individuals per generation. McParland et al. (2007) used this traditional method to estimate N_e for 550,591 Ireland Simmental cattle, the result showed that N_e was 127 at the current generation. However, pedigree information was often missing or error that caused the decline of N_e estimated accuracy. In our study, the estimate for N_e of Chinese beef Simmental cattle was approximately 73 for 5 generations ago, well above the reported numbers. This could be attributed to a sufficiently large number of sires being used to produce animals in the current dataset, and thus a small variance of family size was generated. The slope of the N_e suggests that the population sizes were decreasing consistently fast, possibly due to the use of artificial selection, and therefore actions is required to maintain a larger N_e.

ACKNOWLEDGEMENTS

REFERENCES

Chen, Y., C. H. Lin and C. Sabatti. 2006. Volume measures for linkage disequilibrium. BMC Genet. 7:54.

Du, F. X., A. C. Clutter and M. M. Lohuis. 2007. Characterizing linkage disequilibrium in pig populations. Int. J. Biol. Sci. 3:166-178.

FAO. 1998. Secondary Guidelines for Development of National Farm Animal Genetic Resources Management Plans: Managementof Small Populations at Risk. FAO, Rome, Italy.

Farnir, F., W. Coppieters and J. J. Arranz. 2000. Extensive genome-wide linkage disequilibrium in cattle. Genome Res. 10:220-227.

Goddard, M. G. and C. Smith. 1990. Optimum number of bull sires in dairy cattle breeding. J. Dairy Sci. 73:1113-1122.

Hayes, B. J., P. M. Visscher, H. C. McPartlan and M. E. Goddard. 2003. Novel multilocus measure of linkage disequilibrium to estimate past effective population size. Genome Res. 13:635-643.

Hill, W. G. 1974. Estimation of linkage disequilibrium in randomly mating populations. Heredity 33:229-239.

Khatkar, M., F. Nicholas, A. Collins, K. Zenger, J. Cavanagh, W. Barris, R. Schnabel, J. Taylor and H. Raadsma. 2008. Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. BMC Genomics 9:187.

Kim, E. S and B. W. Kirkpatrick. 2009. Linkage disequilibrium in the North American Holstein population. Anim. Genet. 40:279-288.

Lewontin, R. C. 1964. The interaction of selection and linkage .i. general considerations; heterotic models. Genetics 49:49-67.

Longworth, J. W., C. G. Brown and S. A. Waldron. 2001. Beef in China: agribusiness opportunities and challenges. University of Queensland Press.

Lu, D., M. Sargolzaei, M. Kelly, C. Li, V. V. Gordon, Z. Wang, G. Plastow, S. Moore and S. P. Miller. 2012. Linkage disequilibrium in Angus, Charolais, and Crossbred beef cattle. Front Genet. 3:152-161.

McParland, S., J. F. Kearney and M. Rath. 2007. Inbreeding trends and pedigree analysis of Irish dairy and beef cattle populations. J. Anim. Sci. 85:322-331.

McKay, S., R. Schnabel, B. Murdoch, L. K. Matukumalli, J. Aerts, W. Coppieters, D. Crews, E. DiasNeto, C. A. Gill, C. Gao, H. Mannen and P. Stothard. 2007. Whole genome linkage

disequilibrium maps in cattle. BMC Genet. 8:74.

McVean, G. A. T. 2002. A genealogical interpretation of linkage disequilibrium. Genetics 162:987-991.

McClure, M. C., N. S. Morsci, R. D. Schnabel, J. W. Kim, P. Yao, M. M. Rolf, S. D. Mckay, S. J. Greqq and J. F. Taylor. 2010. A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. Anim. Genet. 41:597-607.

Mukai, F., S. Tsuj, K. Fukazawa, S. Ohtagaki and Y. Nambu. 1989. History and population structure of a closed strain of Japanese black cattle. J. Anim. Breed. Genet. 106:254-264.

Nachman, M. W. 2002. Variation in recombination rate across the genome: evidence and implications. Curr. Opin. Genet. Dev. 12:657-663.

Nomura, T., T. Honda and F. Mukai. 2001. Inbreeding and effective population size of Japanese Black cattle. J. Anim. Sci. 79:366-370.

Odani, M., A. Narita, T. Watanabe, K. Yokouchi, Y. Sugimoto, T. Fujita, T. Oguni, M. Matsumoto and Y. Sasaki. 2006. Genome-wide linkage disequilibrium in two Japanese beef cattle breeds. Anim. Genet. 37:139-144.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. Bakker, M. J. Daly and P. C. Sham. 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81:559-575.

Qanbari, S., E. C. G. Pimentel, J. Tetens, G. Thaller, P. Lichtner, A. R. Sharifi and H. Simianer. 2010. The pattern of linkage disequilibrium in German Holstein cattle. Anim. Genet. 41:346-356.

Reich, D. E., M. Cargill, S. Bolk, J. Ireland, P. C. Sabeti, D. J. Richter, T. Lavery, R. Kouyoumjian, S. F. Farhadian, R. Ward and E. S. Lander. 2001. Linkage disequilibrium in the human genome. Nature 411:199-204.

Sargolzaei, M., F. S. Schenkel, G. B. Jansen and L. R. Schaeffer. 2008. Extent of linkage disequilibrium in Holstein cattle in North America. J. Dairy Sci. 91:2106-2117.

Smith, E. M., X. Wang, J. Littrell, R. Eckert, R. Cole, A. H. Kissebah and M. Olivier. 2006. Comparison of linkage disequilibrium patterns between the HapMap CEPH samples and a family-based cohort of Northern European descent. Genomics 88:407-414.

Sved, J. A. 1971. Linkage disequilibrium and homozygosity of chromosome segments in finite populations. Theor. Popul. Biol. 2:125-141.

Terwilliger, J. D., F. Haghighi, T. S. Hiekkalinna and H. H. H. Goring. 2002. A biased assessment of the use of SNPs in human complex traits. Curr. Opin. Genet. Dev. 12:726-734.

Wright, S. 1938. Size of population and breeding structure in relation to evolution. Science 87:430-431.

Weir, B. S. and W. G. Hill. 1980. Effect of mating structure on variation in linkage disequilibrium. Genetics 95:477-488.

Yan, J., T. Shah, M. L. Warburton, E. S. Buckler, M. D. McMullen and J. Crouch. 2009. Genetic characterization and linkage disequilibrium estimation of a global maize collection using SNP markers. PLoS ONE 4:e8451.

Zhao, H., D. Nettleton and J. C. M. Dekkers. 2007. Evaluation of linkage disequilibrium measures between multi-allelic markers as predictors of linkage disequilibrium between single nucleotide polymorphisms. Genet. Res. 89:1-6.