

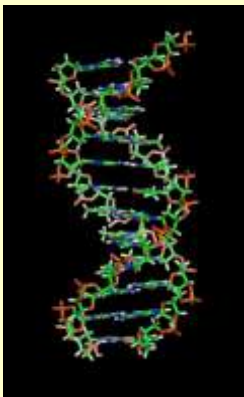


GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

High density LD map of chromosome 14 in German Holstein cattle

S.Qanbari, E.Pimentel, H. Simianer

Institute of Animal Breeding & Genetics
Georg-August-University, Goettingen



GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN



Introduction

- Linkage disequilibrium (LD) is the tendency for alleles to be inherited together more often than would be expected under random segregation.
- LD can be affected by many factors related to the evolutionary history of the population, including random genetic drift, founder effects, mutation, selection, population admixture and stratification.
- linkage disequilibrium (LD) has received considerable attention among livestock geneticists primarily to determine the actual genes responsible for variation of economically important traits and also because of the increasing use of haplotype data for historical or evolutionary inference of animal populations.
- Recently, rapid increase in newly identified SNP markers and continuous decline in genotyping cost have made it more realistic to develop high density maps of animal populations.



Aims:

1. In this study we aimed to develop a dense chromosome-wide linkage disequilibrium map of Chr14 in German Holstein cattle.
1. To measure the extent of haplotypic patterns and blocking structures.
1. To Analyze DGAT1 region to find any signs of selection signature

Material & Methods

Genetic DATA

Genotypic data from 95 unpedigreed (assumed as unrelated) Doutche Holstein cows obtained from 1722 SNP markers on bovine chromosome *14* used as a genetic material.

STATISTICAL METHODS & SOFTWARES

We used:

- R. for Data Processing
- PHASE (Stephens, et al; 2001) for haplotype reconstruction of data.
- Phasing done for whole chromosome with in 14 separate segments with 10 overlapping markers. Phasing process done with 300 iteration, 6 thinning interval and burning value of 30.
- GOLD (Abecasis and Cookson, 2000) for LD statistics (D , D' and r^2) as well as graphical summary of LD.
- HAPLOVIEW v4.1 (Barrett *et al.* 2005) for haplotype frequencies and plotting the LD blocks and Tagging SNPs.
- SWEEP (Sabeti, 2002) for mesearing positive selection based on Extended Haplotype Homozygosity (EHH) aproach.

RESULTS

- **Marker Statistics and Data Filtering**

Total Number of in SNPs on Illumina Chip=54001

Total Number of Chip-SNPs on BTA14= 1722

Total length of Chr14= 81,323,942 bp \approx 81cM

Mean SNP Dist on Chr14=47.22 (kp)

Filtering:

Number of Monomorphic Markers= 148

Number of Markers with MAF < 0.05 = 236

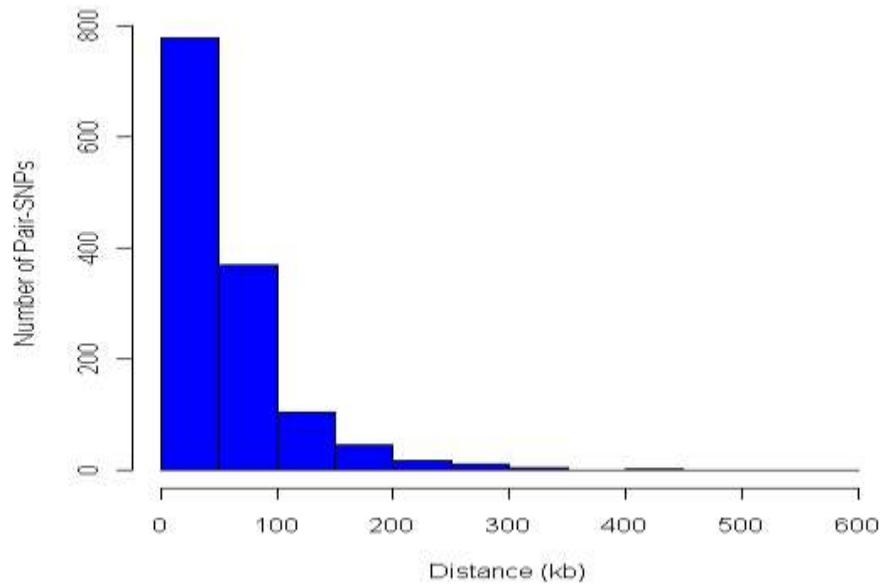
Number of Markers Entered for Analysis= 1338

Mean ObsHet & ExpHet=0.39

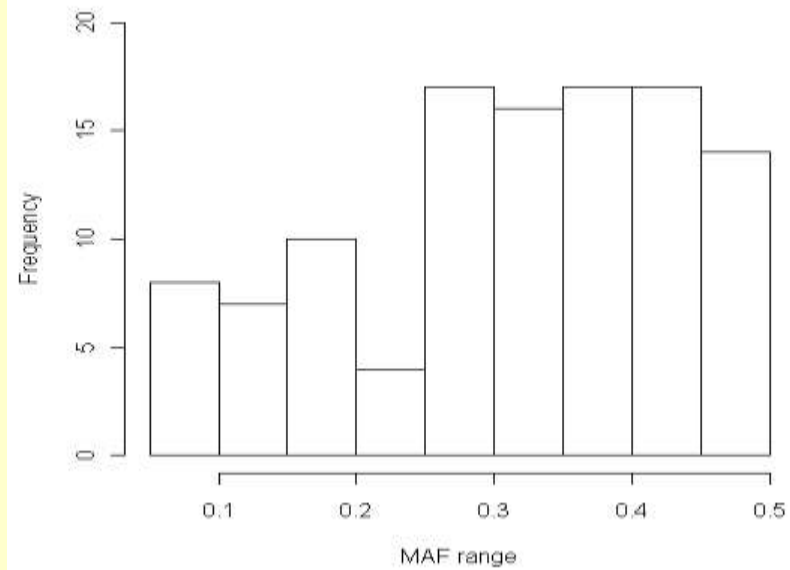
Mean MAF=0.30

Mean SNP Dist after filtering=60.82 (kb)

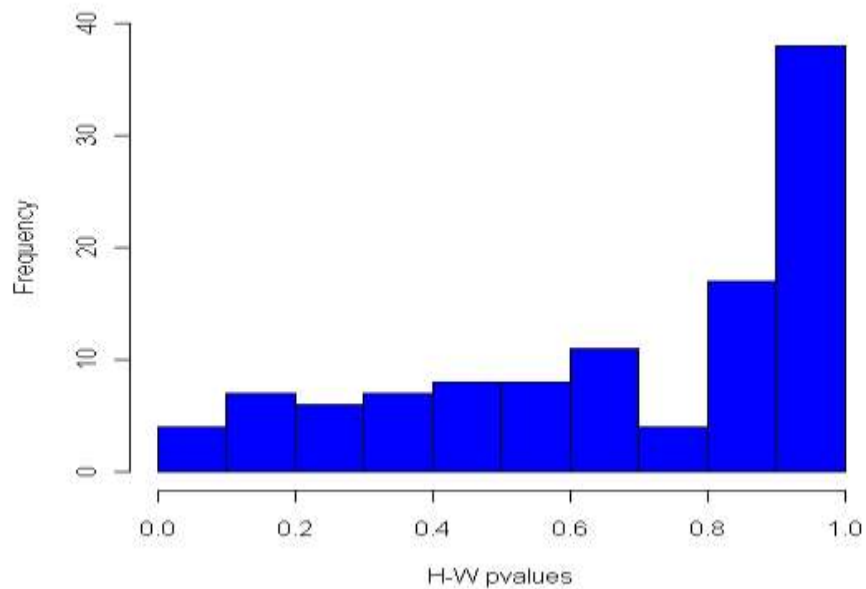
Distribution of distances between neighboring SNPs



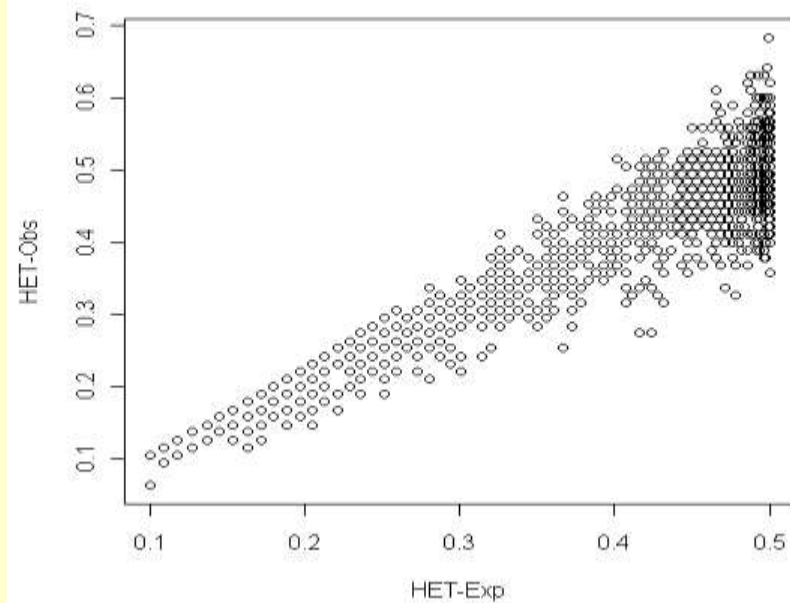
MAF Frequency



Distribution of H-W Proportions

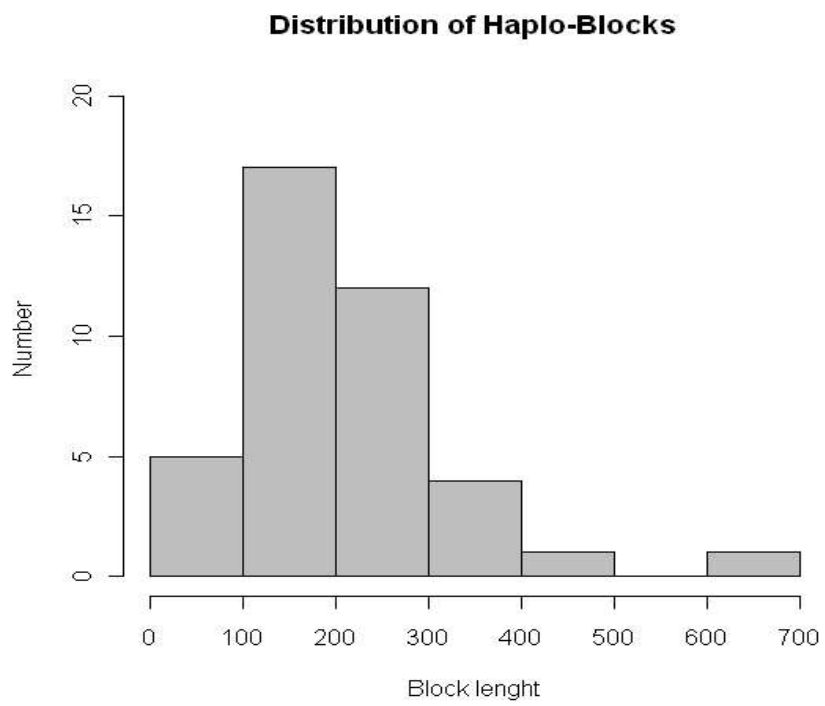


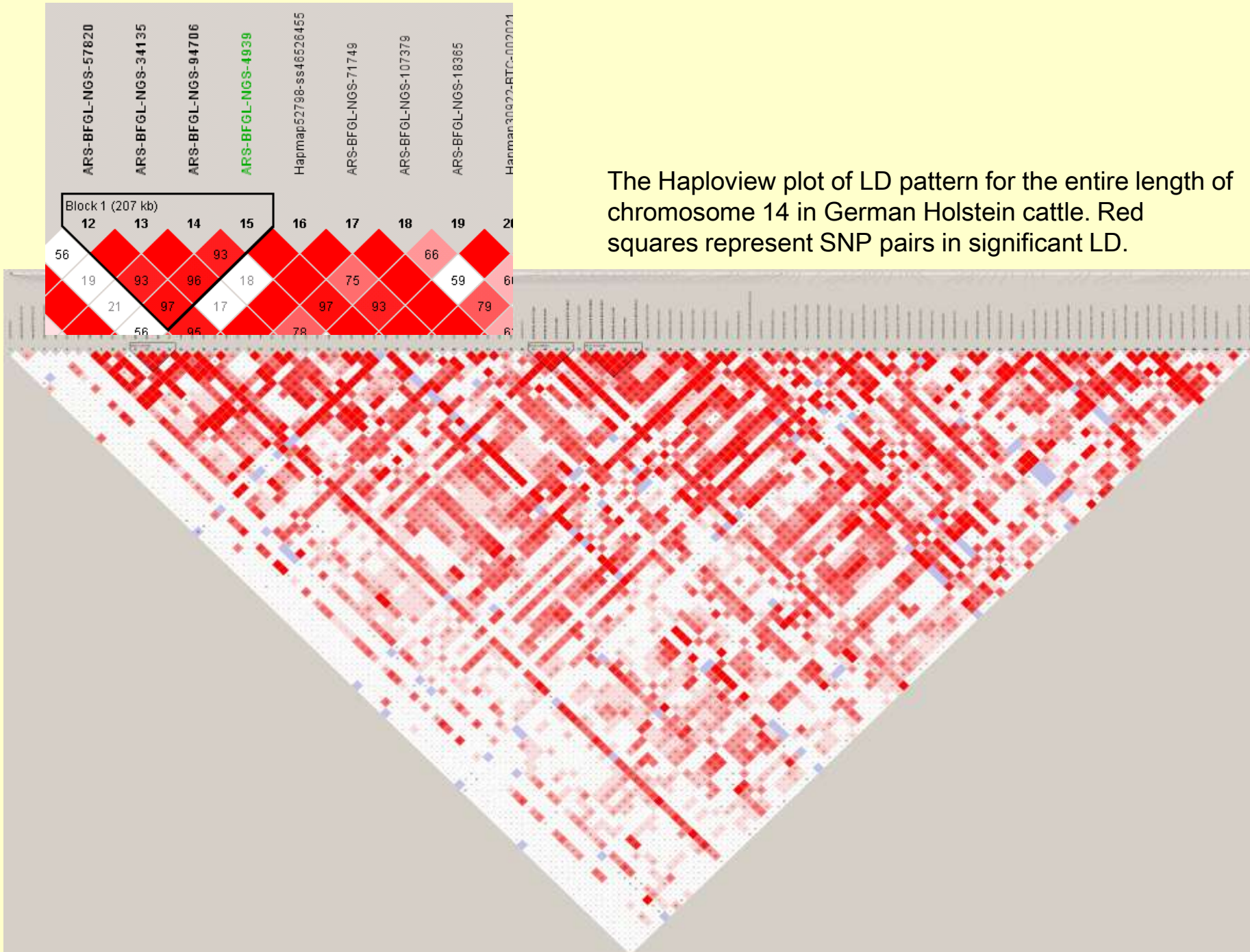
Obs-Hetrezygosity vs. Exp-Heterozygosity



DGAT1

- A total of 40 Haplo-Blocks were detected
- Totally 203 SNPs formed Block with 2-10 SNPs per tract
- The haplotype tracts are spanned 8165 kb (10 %) of the Chr14
- Mean Block length= 204.12 kb with (Min=19 & Max=610 kb)
- Using the tagger option incorporated in HAPLOVIEW, 1120 SNP markers were tagged in Holstein population.



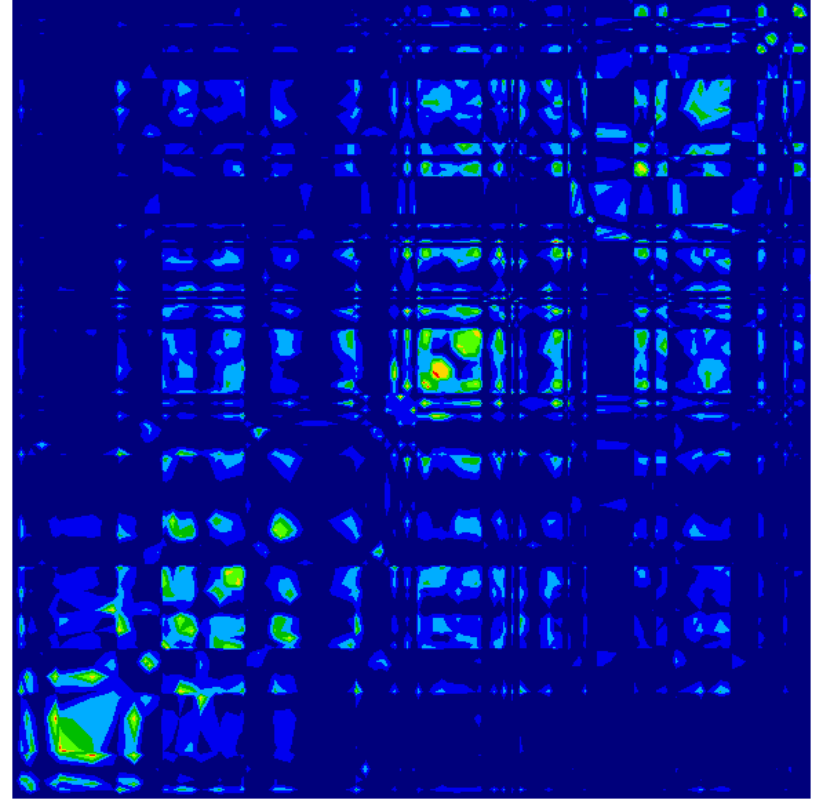
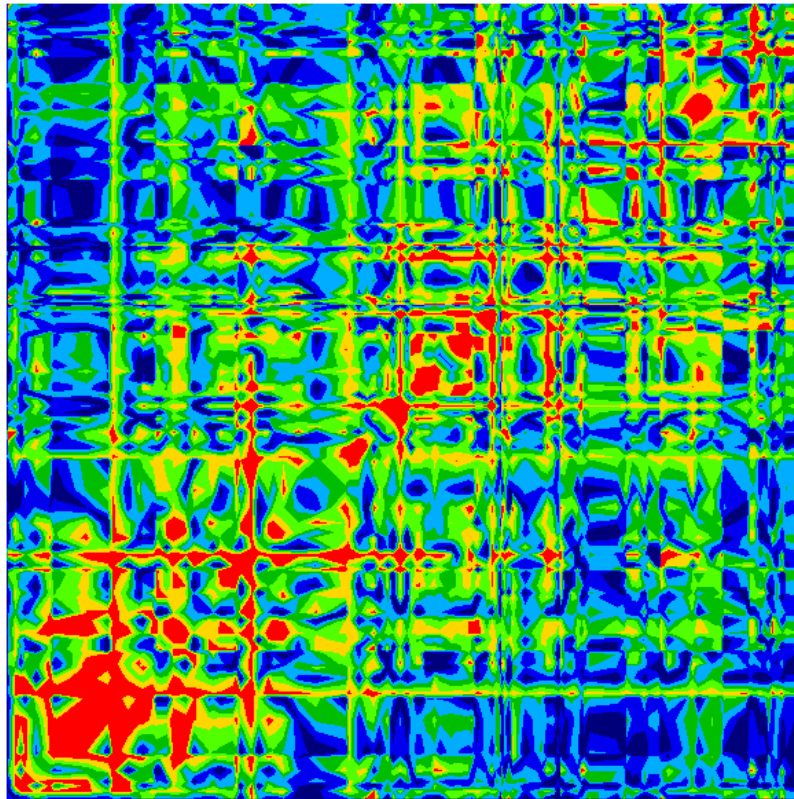


The Haploview plot of LD pattern for the entire length of chromosome 14 in German Holstein cattle. Red squares represent SNP pairs in significant LD.

D'-GOLD

Delta-Gold

- C-056740
- GS-88263
- AC-24804
- AC-24839
- S-110563
- VGS-3571
- VGS-6240
- S-109902
- S-110993
- C-053907
- S-102710
- C-005059
- C-064942
- C-064822
- GS-10284
- C-002159
- GS-41248
- S-103064
- GS-26520
- C-001997
- GS-71749
- GS-94706
- ASA-5815



500 Kb

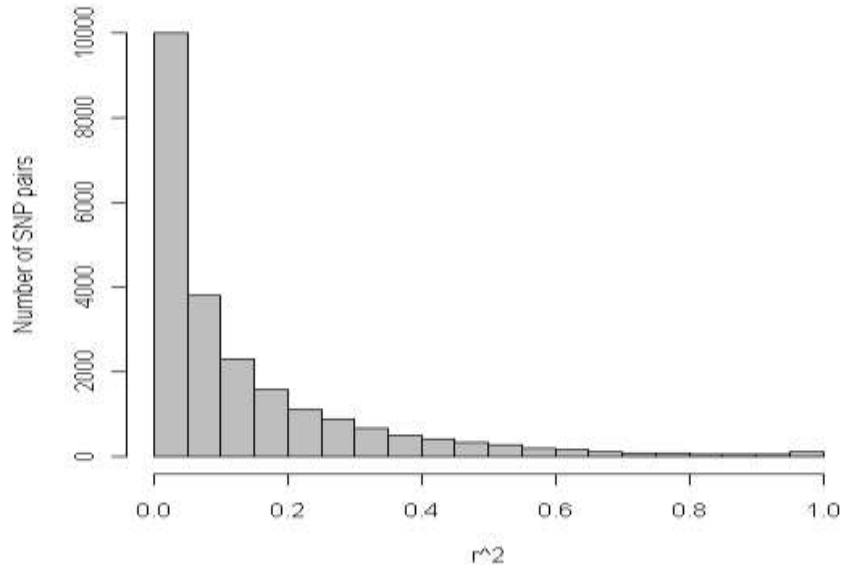
2500 Kb

500 Kb

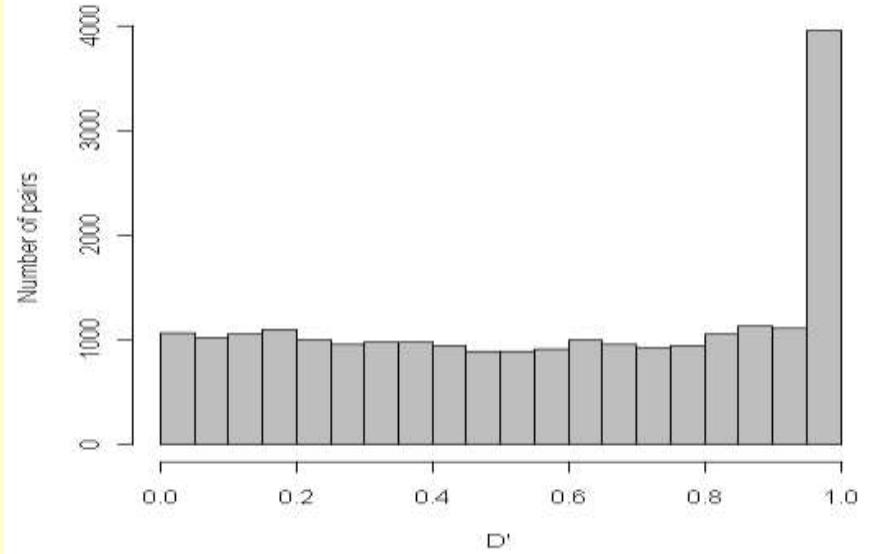
2500 Kb

Properties of D , D' and r^2

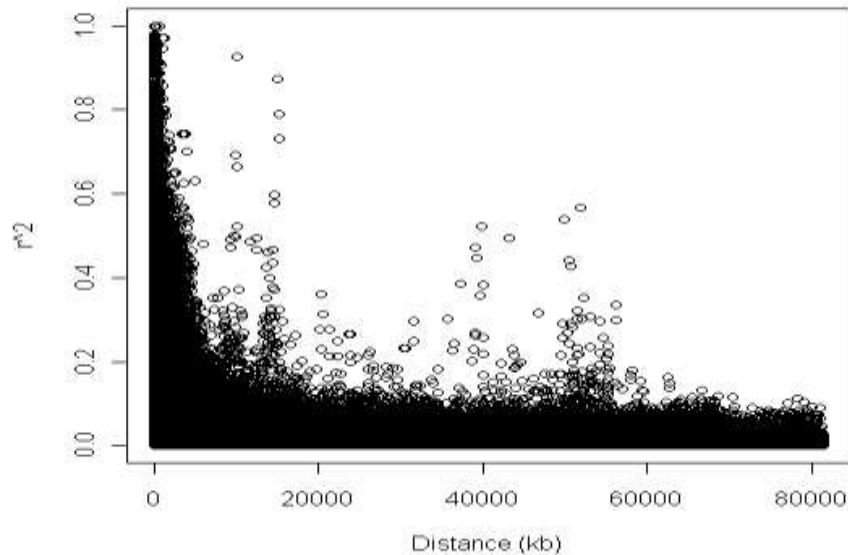
Distribution of Pair-wise r^2



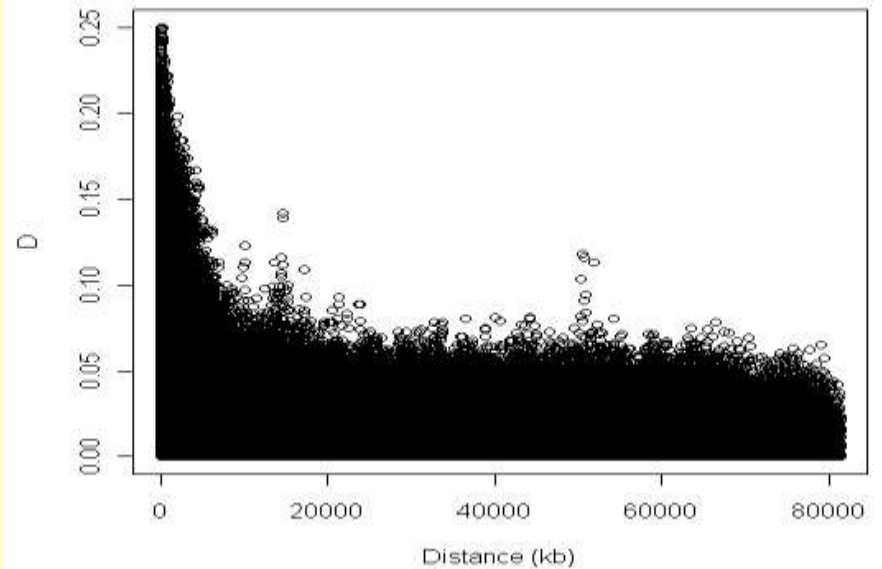
Distribution of Pair-wise D'



Linkage Disequilibrium(r^2) vs. Marker Distance



Linkage Disequilibrium(D) vs. Marker Distance



Signatures of Selection-Theory

- Under neutral evolution, new variants require a long time to reach high frequency in the population, and LD around the variants will decay substantially during this period owing to recombination. As a result, common alleles will typically be old and will have only short-range LD. Rare alleles may be either young or old and thus may have long- or short-range LD.
- The key characteristic of positive selection, however, is that it causes an unusually rapid rise in allele frequency, occurring over a short enough time that recombination does not substantially break down the haplotype on which the selected mutation occurs. A signature of positive natural selection is thus an allele having unusually long-range LD given its population frequency.

Extended Haplotype Homozygosity (EHH)

The extended haplotype homozygosity (EHH) is defined as the probability that two randomly chosen chromosomes carrying the core haplotype of interest are identical by descent (as assayed by homozygosity at all SNPs) for the entire interval from the core region to a distance x (Sabeti, Reich et al. 2002). EHH thus detects the transmission of an extended haplotype without recombination. The EHH of a tested core haplotype t is mathematically calculated as:

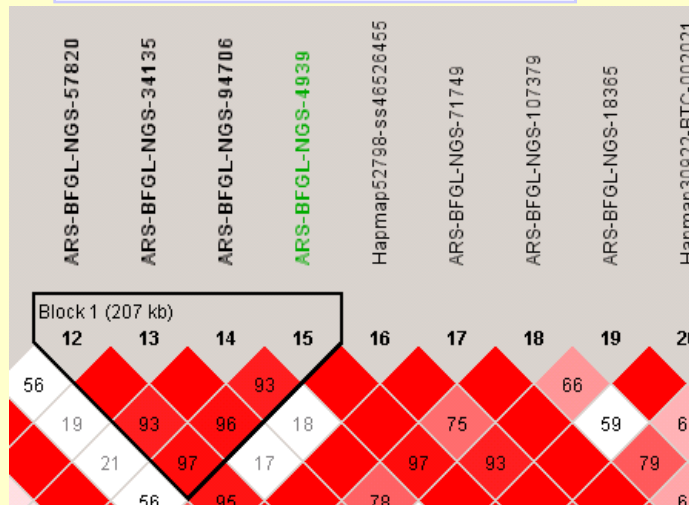
$$EHH_t = \frac{\sum_{i=1}^s \binom{e_{ti}}{2}}{\binom{c_t}{2}}$$

where c is the number of samples of a particular core haplotype, e is the number of samples of a particular extended haplotype, and s is the number of unique extended haplotypes.

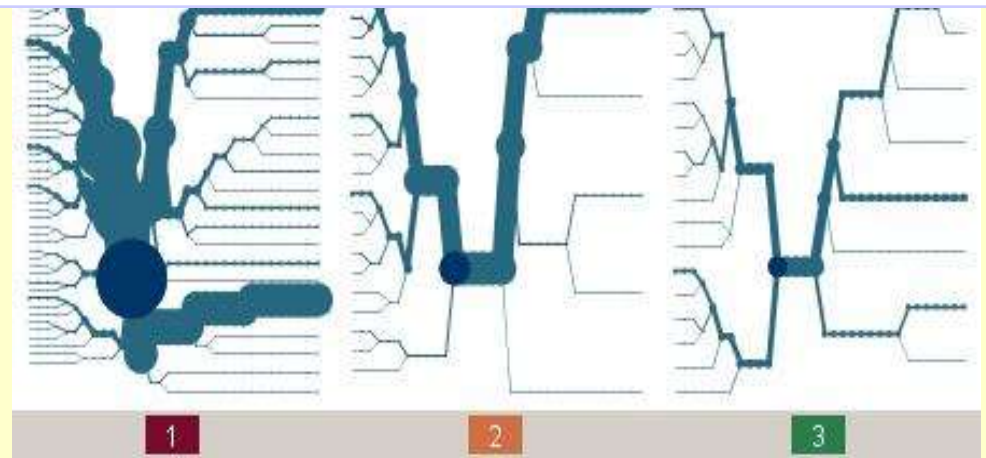
Let's See does DGAT1 region shows any signs of positive selection?

Method: We applied the EHH approach proposed by Sabeti et al. (2002) to compare the rates of allelic specific LD decay.

DGAT1 as Core Region



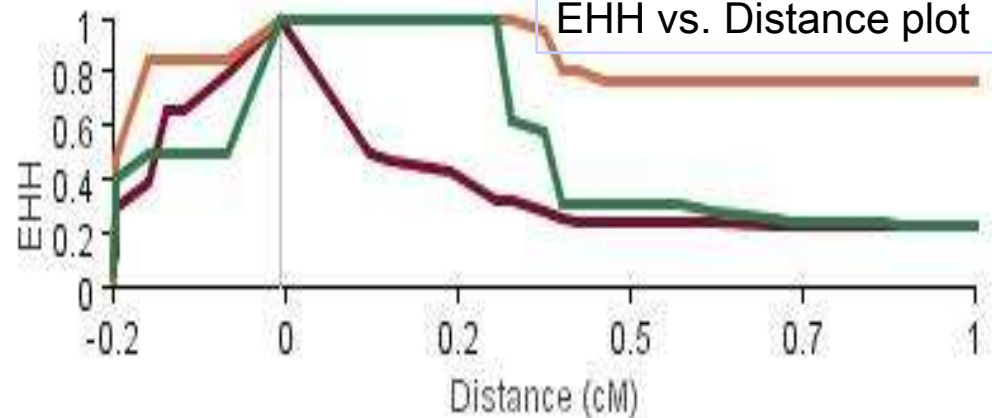
Haplotype bifurcation plots of the 3 core haplotypes

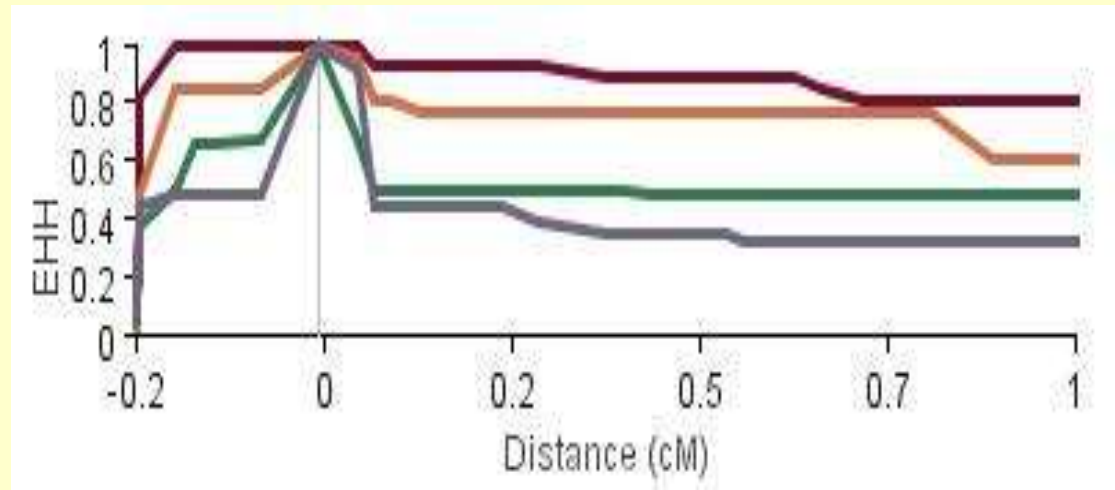
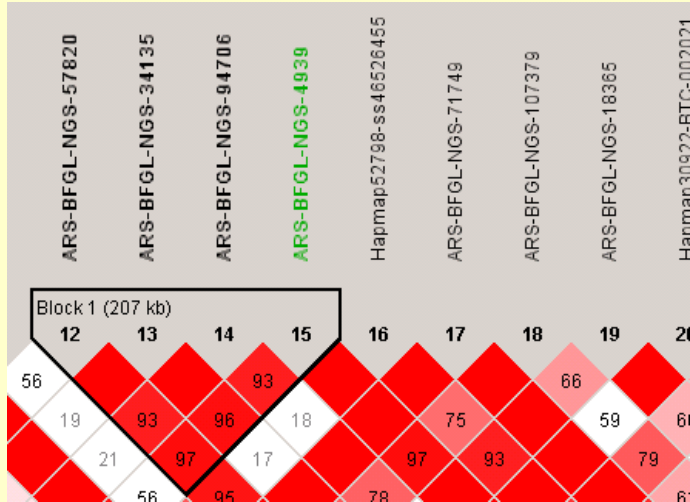


Core Haplotypes

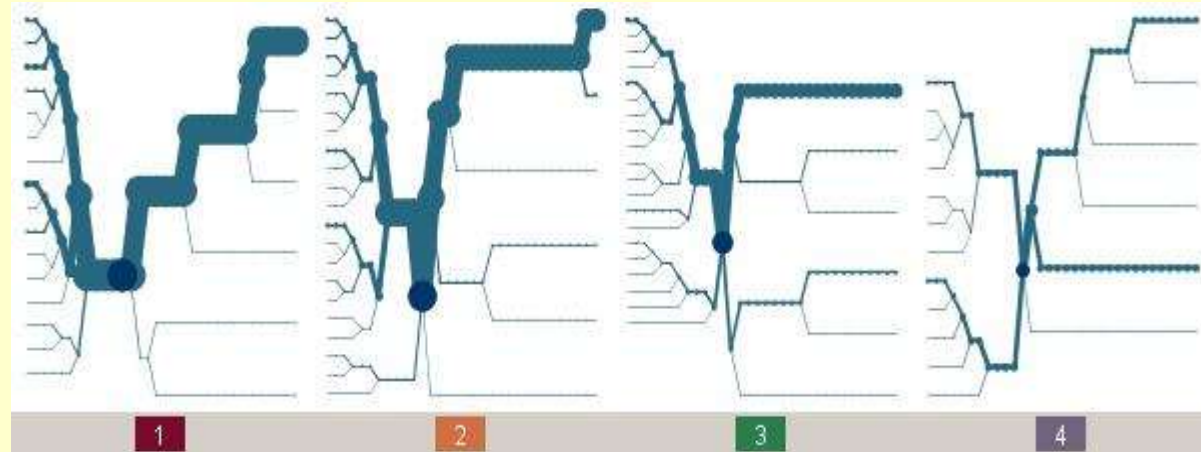
<input type="checkbox"/>	1	58%	AAAA
<input checked="" type="checkbox"/>	2	25%	GGCG
<input checked="" type="checkbox"/>	3	15%	AGCA
<input type="checkbox"/>	4	1%	AAAAG
<input type="checkbox"/>	5	1%	GGAA
<input type="checkbox"/>	6	1%	GGAG

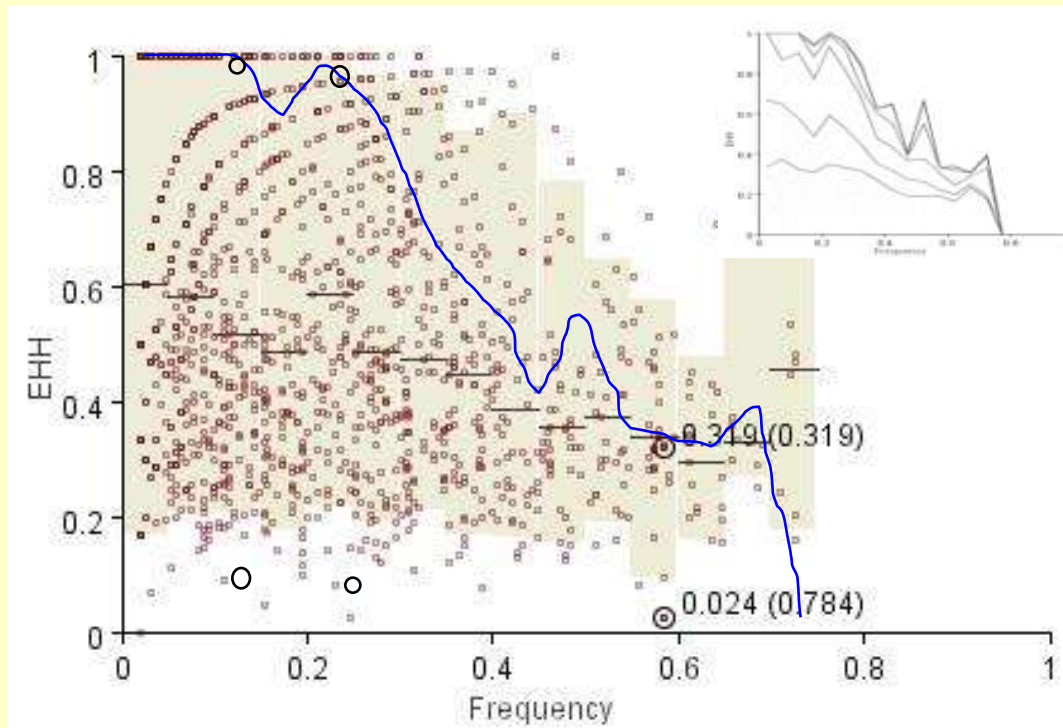
EHH vs. Distance plot





<input type="checkbox"/>			
<input checked="" type="checkbox"/>	1	26%	AAAAGGAGG
<input checked="" type="checkbox"/>	2	25%	GGCGGAGGG
<input checked="" type="checkbox"/>	3	18%	AAAAAAAAGG
<input checked="" type="checkbox"/>	4	12%	AGCAAAAAAA
<input type="checkbox"/>	5	9%	AAAAAAA
<input type="checkbox"/>	6	4%	AGCAAAAAG
<input type="checkbox"/>	7	3%	AAAAAAGAA
<input type="checkbox"/>	8	1%	AAAAAAG
<input type="checkbox"/>	9	1%	AAAAAAGGG
<input type="checkbox"/>	10	1%	AAAAAAGGG
<input type="checkbox"/>	11	1%	AAAGGAAGG
<input type="checkbox"/>	12	1%	GGAAAAAAGG
<input type="checkbox"/>	13	1%	GGAGGAGGG





EHH vs. Frequency Distribution of Holstein Chromosome 14. EHH was calculated at 1 cM distances on both sides for all possible core haplotypes from Chromosome 14 and represented with brown dots and given with 95th percentile. EHH values of DGAT1-H1 were represented with circled dots. Small figure on the corner shows the 50th, 75th, 90th, 95th, 99th and 99.99th percentiles.

Discussion

Population Diversity

The average heterozygosity observed in the elite Holstein sample (39%) was close to the average heterozygosity of 38.9% reported for SNPs loci typed in scanning of bovine chromosome 6 in Australian Holstein bulls (Khatkar et al; 2006).

Blocking Structure

It seems in the case of experimental Holstein population observed pattern of continental regions could be as a result of

1. the breeding strategy of this breed comprising more than one hundred years of intense selection on regions harboring QTLs,
2. the small effective size of population leading to the strong effect of random drift and
3. especially since the wide use of AI bull as founder animals.

Block Size

Comparing to the results of human studies, on average block sizes observed in the present study (=204kb) are 50- 100 times larger than haplotype blocks described in the human genome where typically it extends for only a few kilobases (Hinds *et al.* 2005).

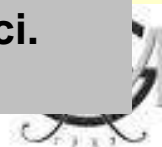
However, it must be noted that there is a large difference between the marker density used in this study and those reported on human genome. Hence, some of the long blocks observed in the present study may be broken down to smaller tracts by the increased availability of SNP density in the adjacent region and may not represent the actual boundary of the block.

Selection Signature

Although there is a long-range LD for DGAT1 region in Holstein cattle, the approach did not show an unusual signature of selection for any haplotypes of this region.

This observation could be as a result of the continues selection for the both major variants of DGAT1-correlated with fat and high milk yield performance- that have fixed them in moderate frequencies in Holstein population. Furthermore there are enough published reports emphasizing the multi-source nature of DGAT1 region of milk production traits.

1. Kuehn, C., G. Thaller, A. Winter, O. R. Bininda-Emonds, B. Kaupe, G. Erhardt, J. Bennewitz, M. Schwerin, and R. Fries. 2004. Evidence for multiple alleles at the DGAT1 locus better explains a quantitative trait locus with major effect on milk fat content in cattle. *Genetics* 167:1873–1881.
1. Bennewitz, J., N. Reinsch, S. Paul, C. Looft, B. Kaupe, C. Weimann, G. Erhardt, G. Thaller, C. Kuhn, M. Schwerin, H. Thomsen, F. Reinhardt, R. Reents, and E. Kalm. 2004. The DGAT1 K232A mutation is not solely responsible for the milk production quantitative trait locus on the bovine chromosome 14. *J. Dairy Sci.* 87:431–442.



CONCLUSION

- The results of this study indicate that despite of intensive selective breeding over successive past generations the degree of genetic variation observed in this purebred Holstein population is still high, which allows selection for traits of economic importance.
- The larger block size in Holstein cattle observed in this study indicates substantially greater LD in cattle than in humans.
- This analysis identified a number of regions on the bovine chromosome 14 where there is very limited or no evidence of historical recombination in this population. These regions may be indicators of tracts harboring candidate genes for interested traits in dairy industry.
- Since the tract coverage tends to be co-localized in all populations, the results of this study may also provide an opportunity to identify common conserved haplotypes along the chromosome among Holstein populations world wide.



Thank You For Attention

The End

