

Association of UDP-galactose-4-epimerase (*GALE*) with milk protein concentration in Chinese Holstein population

Cong Li¹, Wentao Cai², Shuli Liu², Chenghao Zhou², Mingyue Cao², Hongwei Yin²,
Dongxiao Sun², Shengli Zhang^{2*}, Juan J. Loo^{3*}

¹Shaanxi Key Laboratory of Molecular Biology for Agriculture, College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi 712100, China

²College of Animal Science and Technology, Key Laboratory of Animal Genetics and Breeding of Ministry of Agriculture, National Engineering Laboratory for Animal Breeding, China Agricultural University, Beijing 100193, China

³Mammalian NutriPhysioGenomics, Department of Animal Sciences and Division of Nutritional Sciences, University of Illinois, Urbana, Illinois 61801, United States

Supplementary Files

Table S1 PCR primers information of *GALE* gene

Primer	Location	Primer sequences (5'-3')	Product size (bp)	Annealing Temp (°C)
<i>GALE</i> -1	5' flanking region	F-AAGGAGGTGGACAGGAGTTG	517**	59
		R-AAGCCTCCGTTTTCTCTCA		
<i>GALE</i> -2	5' flanking region	F-AGAGGAAAACGGAGGCTTCA	639	59
		R-ACACGACTGAGCCACTAACA		
<i>GALE</i> -3	5' flanking region	F-AGTCGTGTCCAGTCGTGTC	523	59
		R-TGTTAGTAAGCCTGGCGTGA		
<i>GALE</i> -4	Exon-1	F-GACCCACTGTTTTCCATCGG	533	59
		R-CCCTCACCAGCTTCCCAA		
<i>GALE</i> -5	Exon-2	F-ATCTCGTCTGACCCGTACAC	516	59

		R-CAAAATCAGAGGCGATGGGG		
		F-TCCCCATCGCCTCTGATTTT		
GALE-6	Exon-3/4	R-CAGAGGAGAGGTGTCAGTGG	564	59
		F-ACACACTGGACACGTCTTCA		
GALE-7	Exon-5	R-CCAAAAGGCCTGAAGTCACC	444	59
		F-TCAGGAAGCAGTGGTACTT		
GALE-8	Exon-6/7	R-AGAGGAGGAAGTGGTTGCAG	510	59
		F-AGGGAGAAGGGAAAGGAAGC		
GALE-9	Exon-7/8	R-CCAGGTTGTAGATCTGCCCA	442	59
		F-AGGGGAGAAGGGAAAAGTGG		
GALE-10	Exon-9/10	R-CACATCCTGTCCAGCCCTAG	515**	59
		F-ATTTCTGGAGGGCTGGTCAG		
GALE-11	Exon-10	R-TCTCTCCCTTTTCTGGTCCG	456	59
		F-CCTGGGTGCTTACCTGACC		
GALE-12	Exon-11	R-TCTGGTCGTGTCCTGGATTC	633	59

Note: **SNP are detected.

Table S2. Basic data quality checks for the markers

#	Name	Position	ObsHET	PredHET	HWpval	%Geno	FamTrio	MendErr	MAF	Alleles	Rating
1	ss1996900612	129705167	0.455	0.455	1.0	99.4	0	0	0.351	G:A	<input checked="" type="checkbox"/>
2	ss1996900613	129705244	0.482	0.482	1.0	100	0	0	0.405	A:G	<input checked="" type="checkbox"/>
3	rs211659075	129711117	0.325	0.325	0.662	99.3	0	0	0.199	G:C	<input checked="" type="checkbox"/>

Note: # is the marker number. **Name** is the marker ID specified. **Position** is the marker position specified. **ObsHET** is the marker's observed heterozygosity. **PredHET** is the marker's predicted heterozygosity (i.e. $2*MAF*(1-MAF)$). **HWpval** is the Hardy-Weinberg equilibrium p value, which is the probability that its deviation from H-W equilibrium could be explained by chance. **%Geno** is the percentage of non-missing genotypes for this marker. **FamTrio** is the number of fully genotyped family trios for this marker (0 for datasets with unrelated individuals). **MendErr** is the number of observed Mendelian inheritance errors (0 for datasets with unrelated individuals). **MAF** is the minor allele frequency (using founders only) for this marker. **Alleles** are the major and minor alleles for this marker. **Rating** is checked if the marker passes all the tests and unchecked if it fails one or more tests (highlighted in red).

Table S3. Pairwise measures of LD

L1	L2	D'	LOD	r ²	CI _{low}	CI _{hi}	Dist	T-int
ss1996900612	ss1996900613	1	299.68	0.792	0.99	1	77	427.55
ss1996900612	rs211659075	0.963	127.87	0.432	0.93	0.99	5950	-
ss1996900613	rs211659075	1	118.77	0.367	0.98	1	5873	246.64

Note: **L1** and **L2** are the two loci, referenced by their name or name. **D'** is the value of D prime between the two loci. **LOD** is the log of the likelihood odds ratio, a measure of confidence in the value of D'. **r²** is the correlation coefficient between the two loci. **CI_{low}** is 95% confidence lower bound on D'. **CI_{hi}** is the 95% confidence upper bound on D'. **Dist** is the distance (in bases) between the loci, and is only displayed if a marker info file has been loaded. **T-int** is a statistic used by the HapMap Project to measure the completeness of information represented by a set of markers in a region.

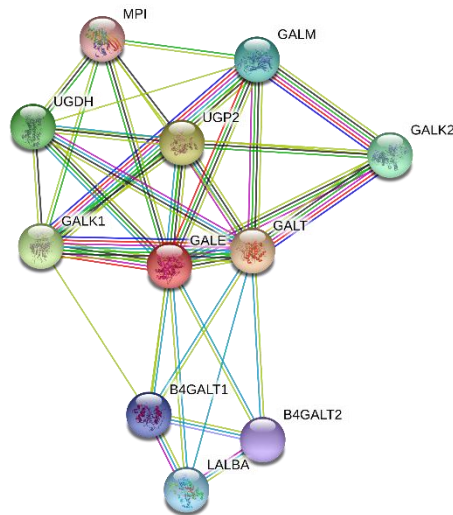


Figure S1. The interaction of GALE with LALBA, B4GALT1 and UGDH

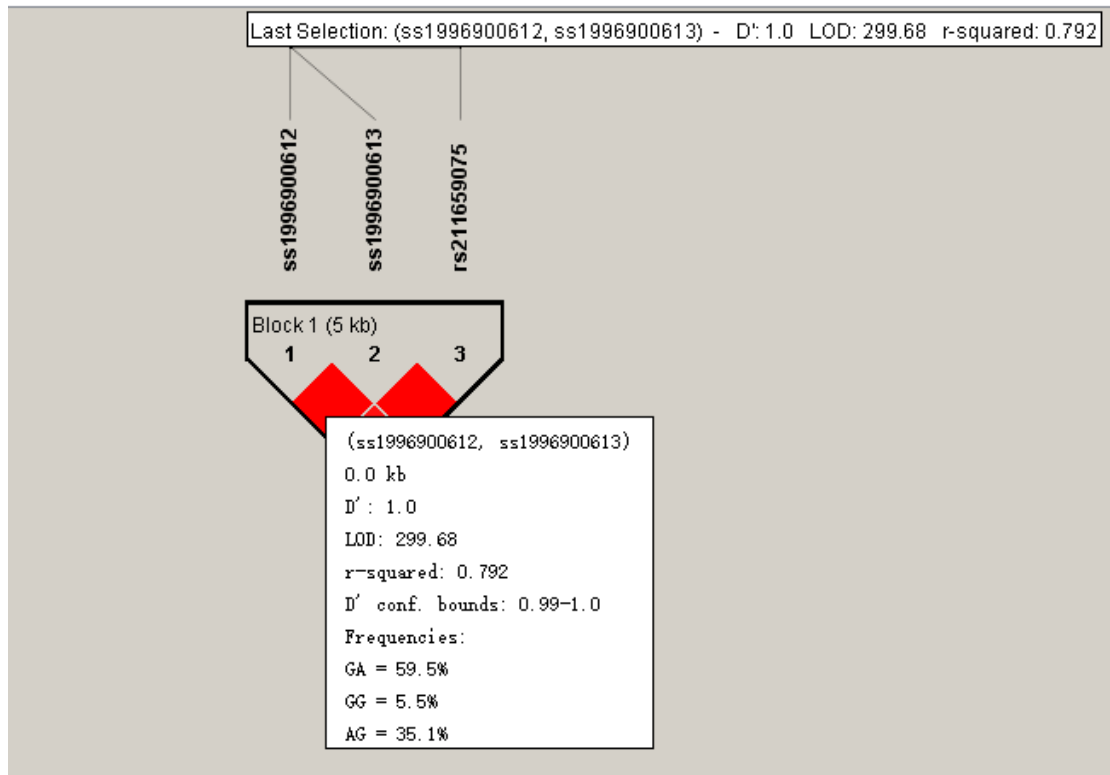


Figure S2. Block 1 comprises marker ss1996900612 and marker ss1996900613

Legend: The white box shows the correlation between marker ss1996900612 and marker ss1996900613 of haplotype block, $D' = 1$, $r^2 = 0.792$.