

## **Extreme antagonistic pleiotropy effects of *DGAT1* on fat, milk and protein yields**

*J. Jiang<sup>1</sup>, L. Ma<sup>1</sup>, D. Prakapenka<sup>2</sup>, M.E. Tooker<sup>3</sup>, P.M VanRaden<sup>3</sup>, J.B. Cole<sup>3</sup>, Y. Da<sup>2</sup>*

<sup>1</sup>*Department of Animal and Avian Sciences, University of Maryland, USA*

<sup>2</sup>*Department of Animal Science, University of Minnesota, USA*

[yda@umn.edu](mailto:yda@umn.edu) (Corresponding Author)

<sup>3</sup>*Animal Genomics and Improvement Laboratory, ARS-USDA, USA*

### **Summary**

An integrated analysis of large-scale genome-wide association study using 294,079 first lactation Holstein cows and selection signature by direct comparison between contemporary Holsteins and a Holstein line unselected since 1964 was conducted to assess why *DGAT1* had the most significant effects on milk production. The ‘G’ allele had the largest positive effect on fat yield and the largest negative effects on milk and protein yields. These largest effects of the ‘G’ allele were the reasons for the statistical significance of *DGAT1*, and constituted extreme antagonistic pleiotropy between positive fat effect and negative milk and protein effects. The antagonism extended to all 68 SNPs in the 2.08 Mb region of 1.38-3.46 Mb but became negligible for SNPs downstream of *PLEC*. Most of these alleles with positive fat effects had higher allele frequencies in contemporary Holsteins than in the 1964 Holstein line. The ‘A’ allele of *DGAT1* had positive effects on milk and protein yields and had negative effect on fat yield but the antagonism of the ‘A’ allele was much weaker than that of the ‘G’ allele. The largest positive effects on milk yield were in the *LY6E-ARC* region and in *EEF1D*, and these alleles with positive milk effects had higher frequencies in the contemporary Holsteins than in the unselected Holsteins. The extreme antagonism between fat yield and milk and protein yields likely was the reason for the lack of genetic selection for *DGAT1* since the mid-1980’s.

*Keywords: DGAT1, antagonistic pleiotropy, GWAS, selection signature, milk production*

### **Introduction**

The *DGAT1* gene at 1.80 Mb of Chr14 has been widely confirmed to be the statistically most significant gene for milk, fat and protein yields (Grisart et al., 2002, Spelman et al., 2002, Schennink et al., 2007, Cole et al., 2011, Jiang et al., 2017). However, the comparison of allele frequencies between a Holstein line unselected since 1964 and contemporary Holsteins showed that the *DGAT1* allele with positive effect on milk yield had a lower frequency in the contemporary Holsteins than in the 1964 line of Holsteins as shown by our data for selection signature analysis. This raised question why *DGAT1* was highly significant but was unselected for its positive effect on milk yield. To answer this question, we conducted an integrated analysis of selection signature and large-scale genome-wide association study.

## **Material and methods**

The GWAS analysis used 294,079 first lactation Holstein cows with 60,671 genotyped and imputed SNPs available from the USDA genomic evaluation. The phenotypic observations of fat, milk and protein yields were yield deviations corrected for polygenic effects as observed breeding values, and the GWAS analysis was implemented using the EPINPmpi1 program (Ma et al., 2008, Weeks et al., 2016). The phenotypic observations for haplotype analysis were yield deviations without correction for breeding values. The Holstein line unselected since 1964 had 301 cattle representing the Holstein genomes of the 1950's and these unselected Holsteins were genotyped with the bovine 50K SNP chip.

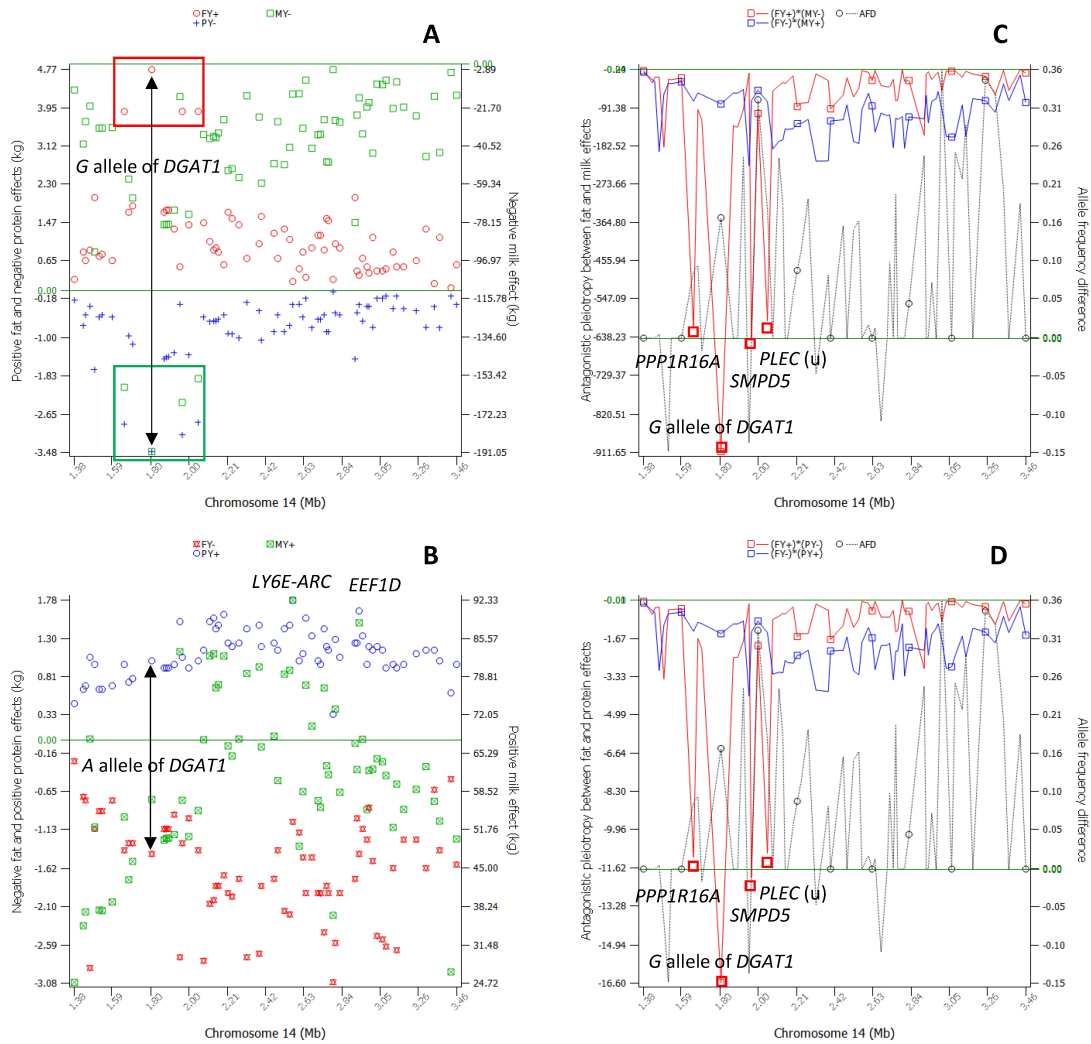
## **Results and discussion**

### **DGAT1 significance was due to extreme effects of the 'G' allele**

The GWAS results showed that the DGAT1 SNP (rs109421300) was the most significant SNP for milk, fat and protein yields, consistent with many previous reports. However, the analysis of effect size and direction revealed that the 'G' allele of the DGAT1 SNP had the largest positive effect on fat yield and the largest negative effects on milk and protein yields, i.e., #1 for positive fat yield, and #1 for negative milk and protein yields, indicating extreme antagonistic pleiotropy effects of the 'G' allele. (Figure 1A). The 'A' allele of the DGAT1 SNP had positive effects on milk and protein yields and had negative effect on fat yields (Figure 1B), but none of the 'A' allele effects was extremely positive or negative, i.e., top #139<sup>th</sup> for milk yield, top 332<sup>nd</sup> for protein yield, and bottom 46<sup>th</sup> for fat yield in terms of allelic means. Large positive milk and protein effects were located in LY6E-ARC region and in EEF1D (Figure 1B). The highly significant DGAT1 effects on milk production were due to the extreme antagonistic pleiotropy effects of the 'G' allele. Specifically, the DGAT1 significance for milk production was due to the extremely large positive effect for fat yield and the extremely large negative effects for milk and protein yields of the 'G' allele.

### **The 'DGAT1 region' had long-range antagonistic pleiotropy effects**

We refer to the 2.08 Mb region of 1.38-3.46 Mb of chromosome 14 with 68 SNPs including the DGAT1 SNP as the 'DGAT1 region', because all 68 SNPs in this region had antagonistic pleiotropy effects between fat yield and milk and protein yields in the same directions as the DGAT1 alleles (red and blue lines in Figures 1C and 1D). A SNP with a '+' or '-' effect on fat yield necessarily had effects on milk and protein yields with the opposite sign. The 'G' allele of DGAT1 SNP had the most extreme antagonistic pleiotropy effects between positive fat yield and negative milk and protein yields, followed by three other SNPs in PPP1R16A, SMPD5 and upstream of PLEC (red line in Figure 1d). It is interesting that the pleiotropy effects between negative fat effects and positive milk and protein effects of these four SNPs were among the weakest pleiotropy effects in this region (blue lines in Figures 1C and 1D). Other than the four SNPs in DGAT1, PPP1R16A, SMPD5 and upstream of PLEC, the pleiotropy effects of the remaining 64 SNPs were weak (red and blue lines in Figures 1C and 1D).



**Figure 1.** Antagonistic pleiotropic effects of milk production traits in the 2.08 Mb region of 1.38-3.46 Mb of chromosome 14. **A.** Positive fat effects and negative milk and protein effects. **B.** Negative fat effects and positive milk and protein effects. **C.** Antagonistic pleiotropy effects between fat and milk effects, and allele frequency differences between the GWAS and unselected Holsteins for alleles with positive fat effects. **D.** Antagonistic pleiotropy effects between fat and milk effects, and allele frequency differences between the GWAS and unselected Holsteins for alleles with positive fat effects

### Alleles with positive fat effects in the ‘DGAT1 region’ was subjected to genetic selection

The GWAS population of 294,079 cows with 98% of cows born in 2006-2015 had 42 overlapping SNPs with the 317 Holsteins unselected since 1964. The allele frequency differences (AFD) between these populations revealed that the 294,079 cows had higher allele frequencies for the alleles with positive effects on fat yield and negative effects on milk and protein yields for 38 of the 42 overlapping SNPs (black line in Figures 1C and 1D). This result showed that the selection in the DGAT1 region was for high fat yield rather than high milk or protein yield. It is interesting to note this 2.08 Mb region containing DGAT1 has been relatively unchanged since the mid 1980’s although this region is still the statistically most significant region affecting milk production today. The extreme DGAT1 antagonism between fat yield and milk and protein yields likely was the reason for the lack of genetic selection for DGAT1 since

the mid 1980's.

### Haplotype analysis of the DGAT1 surrounding region

Haplotype analysis of 27 SNPs in the 1.43-2.24 Mb region among 398,845 Holstein cows showed that 23 haplotypes with high imputing confidence accounted for 88% of the haplotypes in this region (Figure 2). Of these 23 haplotypes, the four haplotypes containing the ‘G’ allele of DGAT1 (SNP 15 in red in Figure 2) had high fat yield and low milk yield. However, these four haplotypes had variations in milk and fat yields, and two other haplotypes also had large effects, i.e., haplotype 3 had high fat yield and haplotype 11 had high milk yield. These results indicated that DGAT1 was not the only gene responsible for the significant effects on milk production in the 2.08Mb ‘DGAT1 region’.

hap	SNP																											milk	fat	frequency
	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29			
23	2	0	0	2	2	2	0	2	0	0	0	0	0	0	2	2	0	2	2	0	0	0	0	0	2	0	2	-1109.0	14.9	0.0051
8	0	0	2	2	2	2	0	2	0	0	0	0	0	0	0	2	2	0	2	2	0	0	0	0	2	0	2	-954.1	26.8	0.0347
1	2	0	2	2	2	2	0	2	0	0	0	0	0	0	2	2	0	2	2	0	0	0	0	0	2	0	2	-837.1	33.5	0.1533
6	0	0	0	2	2	2	0	2	0	0	0	0	0	0	2	2	0	2	2	0	0	0	0	0	2	0	2	-824.7	20.1	0.0247
16	2	0	2	2	2	0	0	2	0	2	0	0	2	2	0	0	2	2	0	2	2	2	2	2	2	0	2	-152.5	-13.4	0.0104
13	0	2	0	0	2	0	2	2	2	2	2	2	2	2	0	0	2	2	0	2	2	0	2	0	2	0	2	-33.9	19.6	0.0162
17	2	0	2	2	2	0	0	2	0	2	0	0	2	2	0	0	2	2	0	2	2	2	2	2	2	0	2	4.7	-0.8	0.0118
12	2	0	2	2	0	0	2	0	2	2	2	2	2	2	2	0	0	2	2	0	2	2	2	2	2	0	2	89.7	-4.0	0.0131
21	2	2	2	2	2	2	0	2	0	2	2	2	2	2	2	0	0	2	2	0	2	2	2	2	2	0	2	92.9	-10.0	0.0096
22	0	0	0	0	0	0	2	0	2	2	2	2	2	2	0	2	2	0	0	2	2	2	2	0	2	0	2	101.2	-1.5	0.0165
10	2	2	2	2	2	2	0	2	0	2	2	2	2	2	0	0	2	2	0	2	2	0	0	2	2	2	103.3	-9.2	0.0163	
5	0	0	2	2	2	0	0	2	0	2	0	0	2	2	0	0	2	2	0	2	2	2	2	2	2	0	2	149.0	-5.9	0.0466
3	2	2	0	0	2	0	2	0	2	2	2	2	2	2	2	0	0	2	2	2	2	0	0	0	2	0	2	165.3	28.1	0.1777
7	2	2	0	2	2	2	0	2	0	2	2	2	2	2	2	0	0	2	2	0	2	2	2	2	0	0	2	173.7	3.8	0.0352
24	2	0	2	2	0	0	2	0	2	2	2	2	2	2	2	0	0	2	2	0	2	2	2	2	0	0	2	216.7	-7.6	0.0071
2	2	0	0	0	0	0	2	0	2	2	2	2	2	2	0	2	2	0	0	0	0	2	2	0	0	2	227.6	-5.7	0.1296	
18	2	0	0	0	0	2	0	2	2	2	2	2	2	2	0	0	0	2	2	2	2	0	0	0	2	0	2	239.1	-5.5	0.0090
15	2	2	2	2	2	2	0	2	0	2	2	2	2	2	2	0	0	2	2	0	2	2	2	2	2	0	2	244.2	-11.7	0.0159
9	0	0	2	2	2	0	0	2	0	2	2	0	0	2	2	0	0	2	2	0	2	2	0	0	2	2	0	305.2	5.3	0.0181
14	2	2	0	0	2	0	2	0	2	2	2	2	2	2	2	0	0	2	2	0	2	2	2	0	2	2	0	351.3	3.4	0.0222
4	2	0	2	2	2	0	0	2	0	2	0	0	2	2	0	0	2	2	0	2	2	2	0	2	2	0	2	357.6	4.8	0.0732
19	0	0	0	0	0	0	2	0	2	2	2	2	2	2	2	0	0	2	2	2	2	0	0	0	2	0	2	419.0	-3.7	0.0071
11	2	0	0	2	2	2	0	2	0	0	0	2	2	2	2	0	0	2	2	0	2	2	0	2	2	2	0	876.0	16.5	0.0208

Figure 2. Haplotype association analysis of 27 SNPs among 398,845 Holstein cows.

### Conclusion

The ‘G’ allele of the DGAT1 SNP had extreme antagonistic pleiotropy effects between positive fat effect and negative milk and protein effects, and the extremely large positive fat effect and negative milk and protein effects were responsible for the statistical significance of DGAT1.

### List of References

Cole, J. B. et al, 2011. BMC Genomics 12(1):408.  
 Grisart, B. et al, 2002. Genome research 12(2):222-231.  
 Jiang, J. et al , 2017. BMC genomics 18(1):425.  
 Ma, L. et al, 2008. BMC bioinformatics 9(1):315.  
 Schennink, A. et al, 2007. Anim. Genet. 38(5):467-473.

Spelman, R. et al, 2002. *J. Dairy Sci.* 85(12):3514-3517.

Weeks, N. T. et al, 2016. *Int. J. High Perform. Comp. Appl.*:1094342016658110.