

Potential influence of κ -casein and β -lactoglobulin genes in genetic association studies of milk quality traits

José Luis Zepeda-Batista¹, Luis Antonio Saavedra-Jiménez¹, Agustín Ruiz-Flores^{1,*}, Rafael Núñez-Domínguez¹, and Rodolfo Ramírez-Valverde¹

* **Corresponding Author:** Agustín Ruiz-Flores
Tel: +52-5959521621, **Fax:** +52-5959521621,
E-mail: arf@correo.chapingo.mx

¹Posgrado en Producción Animal, Universidad Autónoma Chapingo, Chapingo, Estado de México 56230, México

Submitted Jun 27, 2016; Revised Sept 30, 2016;
Accepted May 22, 2017

Objective: From a review of published information on genetic association studies, a meta-analysis was conducted to determine the influence of the genes κ -casein (*CSN3*) and β -lactoglobulin (*LGB*) on milk yield traits in Holstein, Jersey, Brown Swiss, and Fleckvieh.

Methods: The GLIMMIX procedure was used to analyze milk production and percentage of protein and fat in milk. Models included the main effects and all their possible two-way interactions; not estimable effects and non-significant ($p > 0.05$) two-way interactions were dropped from the models. The three traits analyzed used Poisson distribution and a log link function and were determined with the Interactive Data Analysis of SAS software. Least square means and multiple mean comparisons were obtained and performed for significant main effects and their interactions ($p < 0.0255$).

Results: Interaction of breed by gene showed that Holstein and Fleckvieh were the breeds on which *CSN3* ($6.01\% \pm 0.19\%$ and $5.98\% \pm 0.22\%$), and *LGB* ($6.02\% \pm 0.19\%$ and $5.70\% \pm 0.22\%$) have the greatest influence. Interaction of breed by genotype nested in the analyzed gene indicated that Holstein and Jersey showed greater influence of the *CSN3* AA genotype, $6.04\% \pm 0.22\%$ and $5.59\% \pm 0.31\%$ than the other genotypes, while *LGB* AA genotype had the largest influence on the traits analyzed, $6.05\% \pm 0.20\%$ and $5.60\% \pm 0.19\%$, respectively. Furthermore, interaction of type of statistical model by genotype nested in the analyzed gene indicated that *CSN3* and *LGB* genes had similar behavior, maintaining a difference of more than 7% across analyzed genotypes. These results could indicate that both Holstein and Jersey have had lower substitution allele effect in selection programs that include *CSN3* and *LGB* genes than Brown Swiss and Fleckvieh.

Conclusion: Breed determined which genotypes had the greatest association with analyzed traits. The mixed model based in Bayesian or Ridge Regression was the best alternative to analyze *CSN3* and *LGB* gene effects on milk yield and protein and fat percentages.

Keywords: Dairy Cattle; Genetic Improvement; Polymorphism

INTRODUCTION

Genetic association studies have been increasingly used in cattle breeding programs. However, the results have been inconsistent for milk protein genes. Positive, negative, or absence of association with similar genotypes have been reported [1]. Examples of these are the conclusions of Duifhuis-Rivera et al [2] and Dogru [3] who reported that different κ -casein (*CSN3*) and β -lactoglobulin (*LGB*) genotypes were not associated with milk yield in a Mexican herd of Holstein and in a Turkish herd of Brown Swiss cattle, respectively. On the contrary, Gustavsson et al [4] reported that the composite genotype BB/A1A2/AB of *CSN1S1/CSN2/CSN3* has positive effects on cheese yield and percentage of protein and fat in milk.

The inconsistencies have been attributed to various issues affecting the production and composition of milk. Bernabucci et al [5] concluded that the sampling station affects total protein

concentration and acidity in milk, ranging from 3.2% to 6.0% and from 2.0% to 5.6%, respectively. Moreover, Streit et al [6] documented the importance of determining the alleles in the diacylglycerol O-acyltransferase 1 genotype in German Holstein sires to avoid differences that can be as high as 20% in the association of these alleles with milk production values as a result of the allele substitution effect.

However, there is no accurate information related to the potential influence of *CSN3* and *LGB* genes that could be used in genetic association studies to analyze samples obtained under different racial conditions and sampling methodologies. Therefore, in this study we conducted a meta-analysis of genetic association studies to determine the behavior of genes *CSN3* (Genbank Accession No. AC_000163.1) and *LGB* (Genbank Accession No. AC_000168.1) on milk yield and percentage of protein and fat under different conditions and methods of analysis.

MATERIALS AND METHODS

Paper selection criteria

Scientific journals were searched for published papers on genetic association studies. Initially, one hundred and forty-seven papers were chosen. The used criteria aimed to eliminate their heterogeneity, to look for representability of the results, and to ensure their replicability. In a first step, the papers published from 2003 to 2016 were selected. Subsequently, we considered only studies dealing with the most studied milk protein genes, *CSN3* and *LGB*. The remaining papers described studies on milk production and protein and fat percentage in milk. The final sample included 26 papers dealing with Holstein, Jersey, Brown Swiss, and Fleckvieh dairy breeds.

Meta-analysis

Traits analyzed were milk production and percentage of milk protein and fat. The random and fixed effects included in the final model were determined by first establishing a complete model by trait. This model included the main effects and all their possible two-way interactions; some of these were not estimable and were dropped from the model. Additionally, non-significant ($p > 0.05$) two-way interactions were deleted as well. The best fit model and link function were determined with the modules Distribution (Y) and Fit (Y, X), respectively; both are modules from the Interactive Data Analysis of SAS software [7]. The three traits analyzed had the best fit with a final model that used Poisson distribution and a log link function. The final models for milk production, protein and fat percentages were as follows:

$$MP = A_i + B_z + M_y + T_x + AG_w + GN_v + SS_u + BAG_t + BGN_s + MGN_r + E$$

Where MP = milk production; A_i = random effect of i -th paper included in the study; B_z = z -th effect of the breed; M_y = y -th effect of the used model in the original study; T_x = x -th effect of the

used test to recover the original data; AG_w = w -th effect of the analyzed gene; GN_v = v -th effect of the genotype nested in the analyzed gene; SS_u = u -th effect of the sample size used; BAG_t = t -th effect of the interaction of breed by analyzed gene; BGN_s = s -th effect of the interaction of breed by genotype nested in the analyzed gene; MGN_r = r -th effect of the interaction model by genotype nested in the analyzed gene; and E is the residual random effect.

$$P = A_i + B_z + M_y + T_x + AG_w + GN_v + SS_u + E$$

Where P = protein percentage; A_i = random effect of i -th paper included in the study; B_z = z -th effect of the breed; M_y = y -th effect of the used model in the original study; T_x = x -th effect of the used test to recover the original data; AG_w = w -th effect of the analyzed gene; GN_v = v -th effect of the genotype nested in the analyzed gene; SS_u = u -th effect of the sample size used; and E is the residual random effect.

$$F = A_i + B_z + M_y + T_x + AG_w + GN_v + SS_u + E$$

Where F = fat percentage; A_i = random effect of i -th paper included in the study; B_z = z -th effect of the breed; M_y = y -th effect of the used model in the original study; T_x = x -th effect of the used test to recover the original data; AG_w = w -th effect of the analyzed gene; GN_v = v -th effect of the genotype nested in the analyzed gene; SS_u = u -th effect of the sample size used; and E is the residual random effect.

Once all the information from the papers of the sample had been gathered, an adjustment of the information expressed as deviations from their mean was conducted. The population mean reported in each article was used as a reference mean for the analyzed traits. The adjusted values were included in a final database and the statistical analysis was conducted. The GLIMMIX procedure [7] was used to analyze the information. For the main effects and their interactions that were significant ($p < 0.0255$), least square means were obtained and multiple mean comparisons were performed.

The main effects considered in the analysis were the following: article (A), each article included in the study; breed (B), Holstein, Jersey, Brown Swiss, and Fleckvieh; model (M), each of the statistical analysis approaches used in the papers (least square means and mixed models based on Bayesian or Ridge Regression: MMBRR); test (T), the test to recover the data used in the papers (305-day and one-day test); gene (AG), *CSN3* and *LGB*; genotype (GN), AA, AB, AE, BB, BE, and BE for *CSN3*, and for *LGB* were AA, AB, and BB; sample size (SS), size of the population analyzed in the paper, i) 1 to 500 animals was considered small, and ii) from 501 to 2,000 was regarded as medium. The final model for milk production included some two-way interactions: breed by gene; breed by genotype nested in analyzed gene; and model by genotype nested in analyzed gene. The effect of paper was

regarded as random, the rest of them were considered fixed.

RESULTS AND DISCUSSION

Table 1 presents the level of significance for the effects considered in the models to analyze milk production (MP) and protein (PP), and fat percent (FP) in milk. The type of test used was highly significant ($p < 0.0001$) for milk production, probably because the data used in the meta-analysis included one-day tests as well as 305-day tests. On this regard, Gustavsson et al [4] and Poulsen et al [8] concluded that the biases associated with these tests could be controlled by grouping measurements obtained with the same test.

The sample size was significant ($p < 0.0001$) for milk production. On this regard, Vidović et al [9] concluded that bias caused by the sample size is due to the difficulty of differentiating the effects of each gene in the analysis of polygenic traits. On the other hand, the effect of the interaction model by genotype nested in the analyzed gene was significant ($p \leq 0.0255$) probably because of the variance explained by the random effects considering the mixed model [9,10].

The interaction of breed by analyzed gene, was highly significant ($p < 0.0001$). The effect of the gene analyzed had different behavior across breeds and between genes within breed (Table 2). The largest difference in CSN3 and LGB was between the Holstein and Jersey breeds, with 7.65% and 7.64% higher in Holstein than in Jersey, respectively. In Fleckvieh and Brown Swiss breeds, both genes are similarly associated with milk production and total solids, showing greater influence of both genes in Fleckvieh, up to 7.65% and 7.64% for CSN3 and LGB, respectively, compared with Brown Swiss.

Deb et al [11], mentioned that breed has a marked effect on

Table 2. Least square means and multiple comparison for milk production for the interaction breed by gene

Breed	Gene	
	CSN3	LGB
Brown Swiss	5.60 ± 0.46 ^{ab}	5.64 ± 0.46 ^{ab}
Holstein	6.01 ± 0.19 ^a	6.02 ± 0.19 ^a
Jersey	5.55 ± 0.25 ^b	5.56 ± 0.25 ^b
Fleckvieh	5.98 ± 0.22 ^{ab}	5.70 ± 0.22 ^{ab}

CSN3, κ-casein; LGB, β-lactoglobulin.
^{a,b} Means with different literal in the same row or column are different ($p < 0.001$).

milk yield but somewhat less of an effect on milk composition. The heavier breeds tend to produce more milk. Therefore, the breeds with high milk production (Holstein and Fleckvieh) showed the greatest response. On this regard, since the improvement programs in cattle are based on the genetic potential that the individuals could show. These results may be due to selection for traits such as milk production and total solids that are associated with CSN3 and LGB in both breeds [12,13]. Additionally, genetic make-up of dairy animals plays a great role in the variation of milk yield and composition. Raven et al [14] and Ramayo-Caldas et al [15] mentioned the importance of determine the proportion of genetic markers shared between breeds to study and compare them in multi-breed multi-trait association studies in commercial herds. On this regard, Ramayo-Caldas et al [15], reported around 206 genes with the same effect in three French breeds. Meanwhile, Raven et al [14], determined that despite the different linkage disequilibrium patterns, Holstein and Jersey share between 8% to more than 38% genetic markers with similar effects on economic important traits in dairy cattle.

The interaction of breed by genotype nested in the analyzed gene (Table 3) was highly significant ($p < 0.0001$). Holstein and

Table 1. Level of significance for the effects included in the models to analyze milk production (MP), and protein (PP) and fat (FP) percentage in milk

Variable	B _z	M _y	T _x	AG _w	GN _v	SS _u	BAG _t	BGN _s	MGN _i
MP	0.1528	0.2411	<0.0001	0.1833	0.0974	<0.0001	<0.0001	<0.0001	0.0255
PP	0.7529	0.7358	0.6484	0.9966	1.000	0.9090	-	-	-
FP	0.1512	0.4161	0.8409	0.8604	1.000	0.8199	-	-	-

B_z, breed; M_y, model used in the original study; T_x, test used to recover original data; AG_w, analyzed gene; GN_v, genotype nested in analyzed gene; SS_u, sample size; BAG_t, interaction of breed by analyzed gene; BGN_s, interaction of breed by genotype nested in analyzed gene; MGN_i, interaction of model by genotype nested in analyzed gene.

Table 3. Least square means, standard error, and multiple comparison for milk production for the subclasses of the interaction breed by genotype in the loci of the CSN3 gene and LGB gene

Breed	Genotype of the CSN3 gene			Genotype of the LGB gene		
	AA	AB	BB	AA	AB	BB
Brown Swiss	5.54 ± 0.47 ^a	5.62 ± 0.47 ^a	5.64 ± 0.47 ^{ab}	5.54 ± 0.47 ^a	5.62 ± 0.47 ^{ab}	5.64 ± 0.47 ^{ab}
Holstein	6.04 ± 0.22 ^a	6.01 ± 0.22 ^a	6.00 ± 0.22 ^{ab}	6.05 ± 0.20 ^a	6.01 ± 0.20 ^a	6.01 ± 0.20 ^a
Jersey	5.59 ± 0.31 ^a	5.57 ± 0.31 ^a	5.50 ± 0.31 ^b	5.60 ± 0.19 ^a	5.54 ± 0.19 ^b	5.57 ± 0.19 ^b
Fleckvieh	5.80 ± 0.24 ^a	5.86 ± 0.24 ^a	6.30 ± 0.22 ^a	5.71 ± 0.22 ^a	5.73 ± 0.22 ^{ab}	5.67 ± 0.22 ^{ab}

CSN3, κ-casein; LGB, β-lactoglobulin.
^{a,b} Means with different letter in the same row or column are different ($p < 0.001$).

Table 4. Least square means, standard error, and multiple comparison for milk production for the subclasses of the interaction model by genotype in the loci of the *CSN3* gene and *LGB* gene

Model ¹⁾	Genotype of the <i>CSN3</i> gene			Genotype of the <i>LGB</i> gene		
	AA	AB	BB	AA	AB	BB
LS	5.96 ± 0.18 ^a	5.98 ± 0.18 ^a	6.09 ± 0.18 ^a	5.95 ± 0.25 ^a	5.98 ± 0.25 ^a	5.94 ± 0.25 ^a
MM	5.53 ± 0.19 ^b	5.54 ± 0.19 ^b	5.62 ± 0.19 ^b	5.54 ± 0.19 ^a	5.50 ± 0.19 ^b	5.50 ± 0.19 ^a

CSN3, κ -casein; *LGB*, β -lactoglobulin.

¹⁾ LS, least square model; MM, mixed model based on Bayesian and ridge regression.

^{a,b} Means with different letter in the same row or column are different ($p < 0.05$).

Jersey had a greater influence of genotype AA in *CSN3* gene on milk yield, up 1.61% from other genotypes studied; Miciński et al [12] reported a difference of 0.92% to 11.94% for Polish Jersey.

Estimates for Brown Swiss and Fleckvieh were different from those for Holstein and Jersey; the BB genotype was the most closely associated with milk production, 7.94% higher than the other studied genotypes. Similarly, Chrenek et al [16] found that, for Fleckvieh, the genotype most closely associated with milk yield is BB, 7.42%. In contrast, Matějíček et al [17] concluded that genotype AB was the most closely associated with that trait.

Similar behavior shown in the study by Holstein and Jersey and Brown Swiss and Fleckvieh may also attend mainly the common geographic origin of the analyzed breeds. In this regard, Negrini et al [18] calculated genetic distances through genetic fingerprinting of 51 cattle breeds. They placed the Holstein and Jersey breeds in the Nordic genetic type group, while the Brown Swiss and Fleckvieh were grouped with genetic types from the Alpine region or Central-France.

Holstein and Jersey showed similar behavior for AA genotype of the *LGB* gene (Table 3). This estimate had the largest influence on milk yield, up to 0.66% and 1.07% above the other genotypes. This result is similar to that found for the *CSN3* gene in this study since, according to Bonfatti et al [19], the *LGB* gene is also associated with milk production and total solids. Fleckvieh and Brown Swiss are influenced in a major way by genotypes BB and AB, showing up to 1.77% and 1.05% difference, relative to other genotypes. On this regard, Gustavsson et al [4] determined the effect of breed and genotype using composite genotypes of κ -casein and β -lactoglobulin genotypes to determine their genetic association with production traits. These authors concluded that the differences between the genetic association values, ranging between 0.5% and 30%, were mainly influenced by genotype. Moreover, Oltenacu and Broom [20], report rising of inbreeding rates of 0.2% per year in Holstein and Jersey which would cause a decreasing in the response to improvement and selection program and the genetic association values due to loss of the genetic variation in the population.

On the other hand, the interaction of type of statistical model used by genotype nested in the analyzed gene was highly significant ($p < 0.0001$), and mean estimates are shown in Table 4.

The differences between the models used for each of the studied

genes may be explained with the conclusions of Miciński et al [12] and Monir and Zhu [21], who demonstrated that the inclusion or omission of the effect of each particular gene and polygenic effects, influence the results of production traits and only the models that include random effects are able to differentiate those changes.

The results of comparing models of *LGB* and *CSN3* genes (Table 4) were similar, maintaining a difference of just over 7% across all genotypes. Here, Comin et al [10] and Vidović et al [9] concluded that the MMBRR explained the effect of milk protein genotypes on performance and composition of different breeds of dairy cattle, nearing a 9% difference, relative to studies using only models with fixed effects based on least squares.

Oleński et al [22] and Pärna et al [23] used models similar to ours when attempted to explain the variability present in genetic association studies. However, the general linear mixed models were not able to differentiate changes between AB and BB genotypes of the *LGB* gene, while for the *CSN3* MMBRR had the best fit. Here, Kučerová et al [24] concluded that the mixed model is the best suited for studies of genetic association with genes of casein; while such models shown less power to determine associations for genes of other milk-whey proteins, including β -lactoglobulin.

CONCLUSION

Differences in the magnitude of the influence of *CSN3* and *LGB* genotypes depending on the breed could change according on the shared genetics. The mixed model based in Bayesian or Ridge Regression was the best alternative for analysis in genetic association studies involving the *CSN3* and *LGB* genes. Due to their higher substitution allele effect and the minor inbreeding level, Brown Swiss and Fleckvieh could show more progress in selection programs that include the *CSN3* and *LGB* genes, relative to Holstein or Jersey.

CONFLICT OF INTEREST

We certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

ACKNOWLEDGMENTS

The authors thank Mexican Council of Science and Technology (CONACYT) for the financial support to the first author during his Doctoral studies. Authors thank Dr. Gaspar Manuel Parra-Bracamonte for the revision of the manuscript.

REFERENCES

1. Heck JML, Schennink A, Van Valenberg HJF, et al. Effects of milk protein variants on the protein composition of bovine milk. *J Dairy Sci* 2009;92:1192-202.
2. Duifhuis-Rivera T, Lemus-Flores C, Ayala-Valdovinos MA, et al. Polymorphisms in beta and kappa casein are not associated with milk production in two highly technified populations of Holstein cattle in México. *J Anim Plant Sci* 2014;24:1316-21.
3. Dogru U. β -Lactoglobulin genetic variants in Brown Swiss dairy cattle and their association with milk yield and quality traits. *J Anim Plant Sci* 2015;25:595-8.
4. Gustavsson F, Buitenhuis AJ, Johansson M, et al. Effects of breed and casein genetic variants on protein profile in milk from Swedish Red, Danish Holstein, and Danish Jersey cows. *J Dairy Sci* 2014;97:3866-77.
5. Bernabucci U, Basiricò L, Morera P, et al. Effect of summer season on milk protein fractions in Holstein cows. *J Dairy Sci* 2015;98:1815-27.
6. Streit M, Neugebauer N, Meuwissen THE, Bennewitz J. Evidence for a major gene by polygene interaction for milk production traits in German Holstein dairy cattle. *J Dairy Sci* 2011;94:1597-600.
7. SAS. SAS 9.3 User's guide. Cary, NC: SAS Institute Inc; 2011.
8. Poulsen NA, Buitenhuis AJ, Larsen LB. Phenotypic and genetic associations of milk traits with milk coagulation properties. *J Dairy Sci* 2015;98:2079-87.
9. Vidović V, Lukać D, Nemes Ž, Trivunović S. β -Lactoglobulin genetic variants in Serbian Holstein Friesian dairy cattle and their association with yield and quality of milk. *Anim Sci Pap Rep* 2014;32:179-82.
10. Comin A, Cassandro M, Chessa S, et al. Effects of composite β - and κ -casein genotypes on milk coagulation, quality, and yield traits in Italian Holstein cows. *J Dairy Sci* 2008;91:4022-7.
11. Deb R, Singh U, Kumar S, et al. Genetic polymorphism and association of kappa-casein gene with milk production traits among Frieswal (HF×Sahiwal) cross breed of Indian origin. *Iran J Vet Res* 2014;15:406-8.
12. Miciński J, Klupczyński J, Mordas W, Zablotna R. Yield and composition of milk from Jersey cows as dependent on the genetic variants of milk proteins. *Pol J Food Nutr Sci* 2007;57:95-9.
13. Sitkowska B, Neja W, Milczewska A, Mroczkowski S, Markowska A. Milk protein polymorphisms and effect of herds on cow milk composition. *J Central European Agric* 2013;14:78-90.
14. Raven L, Cocks BG, Hayes BJ. Multibreed genome wide association can improve precision of mapping causative variants underlying milk production in dairy cattle. *BMC Genomics* 2014;15:62.
15. Ramayo-Caldas Y, Renand G, Ballester M, Saintilan R, Rocha D. Multi-breed and multi-trait co-association analysis of meat tenderness and other meat quality traits in three French beef cattle breeds. *Genet Sel Evol* 2016;48:37.
16. Chrenek P, Huba J, Vasicek D, Peskovicová D, Bulla J. The relation between genetic polymorphism markers and milk yield in Brown Swiss cattle imported to Slovakia. *Asian-Australas J Anim Sci* 2003;16:1397-401.
17. Matějčiček A, Matějčičková J, Němcová E, et al. Joint effects of CSN3 and LGB genotypes and their relation to breeding values of milk production parameters in Czech Fleckvieh. *Czech J Anim Sci* 2007;52:83-7.
18. Negrini R, Nijman IJ, Milanese E, et al. Differentiation of European cattle by AFLP fingerprinting. *Anim Genet* 2007;38:60-6.
19. Bonfatti V, Di Martino G, Cecchinato A, Vicario A, Carnier P. Effects of β - κ -casein (CSN2-CSN3) haplotypes and β -lactoglobulin (BLG) genotypes on milk composition traits and detailed protein composition of individual milk Simmental cows. *J Dairy Sci* 2010;93:3797-808.
20. Oltenacu PA, Broom DM. The impact of genetic selection for increased milk yield on the welfare of dairy cows. *Anim Welf* 2010;19:39-49.
21. Monir M, Zhu J. Comparing GWAS results of complex traits using full genetic model and additive models for revealing genetic architecture. *Sci Rep* 2017;7:38600.
22. Oleński K, Cieślińska A, Suchocki T, Szyda J, Kamiński S. Polymorphism in coding and regulatory sequences of beta casein gene is associated with milk production traits in Holstein-Friesian cattle. *Anim Sci Pap Rep* 2012;30:5-12.
23. Pärna E, Kaart T, Kiiman H, Bulitko T, Viinalass H. Milk protein genotype association with milk coagulation and quality traits. In: Chaiyabutr N, editor. Milk production-advanced genetic traits, cellular mechanism, animal management and health. Rijeka, Croatia: InTech; 2012. p. 155-72.
24. Kučerová J, Matějčiček A, Jandurová OM, et al. Milk protein genes CSN1S1, CSN2, CSN3, LGB and their relation to genetic values of milk production parameters in Czech Fleckvieh. *Czech J Anim Sci* 2006;51:241-7.