Received: 16 February 2025 | Accepted: 19 June 2025

DOI: 10.1111/age.70027

SHORT COMMUNICATION

ANIMAL GENETICS WILEY

Fine mapping highlights *ITGAL* and *MUS81* loss-of-function mutations modulating recessive impacts in dairy cattle

Abstract

We recently described several major-effect recessive loci impacting anatomical and lactation traits in dairy cattle. Two of these loci in particular presented multiple candidate causative variants, comprising tightly linked coding variants that could not be easily differentiated on a statistical or functional basis. Here, we re-examine the candidacy of these variants by leveraging a dataset of 1 million genotyped animals. Assessing lactation and bodyweight effects in conjunction with rare, recombined genotypes for the *IL4R*, *KIAA0556*, *ITGAL*, *DPF2*, and *MUS81* candidates, we highlight *ITGAL* and *MUS81* as the most likely causative genes for the two QTL. Recombinant homozygotes for these genes present larger, more significant effects than other candidates at the same loci, with both representing premature stop mutations anticipated to inactivate *ITGAL* and *MUS81*. We further examined homozygotes for the *ITGAL* mutation to better understand the range of phenotypes impacted. While outwardly normal, *ITGAL* mutants showed significant differences in the number and composition of circulating leukocytes, consistent with the role of *ITGAL* as a key mediator of leukocyte signalling, adhesion, and migration. These results demonstrate how near-perfectly linked candidate mutations can be differentiated given population-scale data, and highlight the *ITGAL* and *MUS81* mutations as diagnostic targets to help manage the frequency of these variants.

We recently reported two large sequence-resolution screens of cattle (Reynolds et al., 2021, 2022) with the aim of discovering non-additive quantitative trait loci (QTL) for bodyweight, lactation traits, and other phenotypes. These studies highlighted multiple recessive loci, some with large effects probably reflective of undiagnosed genetic disorders (Reynolds et al., 2021). While some of the QTL presented single compelling candidate mutations that could be assumed to underlie these effects (e.g. see Dittmer et al., 2022), two of the QTL harboured multiple mutations as plausible candidates. These included a *DPF2* missense and *MUS81* premature stop mutation for a recessive QTL on chromosome 29 (Reynolds et al., 2021), and missense variants (*IL4R*, *KIAA0556*) and a premature stop (*ITGAL*) for a QTL on chromosome 25 (Reynolds et al., 2022). These variants were in strong to very strong LD with each other, presenting r^2 values of 0.74–0.90 for chromosome 25, and $r^2 \sim 0.98$ for the chromosome 29 locus. These LD relationships prevented easy differentiation between causative and non-causative variants using these datasets. Hence, we reasoned that very large sample sizes might provide added power. We therefore added all five candidate variants to two iterations of a low-density Illumina Infinium XT SNP-chip. These panels included an additional 492 variants used for calculation of genotype-based principal components in the current study, representing parentage markers or other custom content common to the two panels. At time of writing, >2 million animals had been physically genotyped on these platforms, with approximately 1 million of these animals also measured for lactation traits, and approximately 160 000 measured for juvenile bodyweight. Table S1 and Figures S1–S3 show summary statistics and trait distributions for genotypes and phenotypes of interest, and the Appendix S2 describe filtering and modelling of these data prior to analysis.

Since differentiation of candidate mutation effects would require observation of rare haplotypes separating the alleles of interest, we first assessed the relationships between these variants in the new datasets and found modest numbers of animals that are uniquely homozygous for each of the candidate variants of interest. For example, amongst 941 388 genotyped cows measured for the lactation traits, there were 22 animals uniquely homozygous for the *IL4R* chr25:24904939C>T variant, four animals for the *KIAA0556* chr25:25161613G>A variant, and 13 animals for the *ITGAL* chr25:26689392G>A variant. For the chromosome 29 locus, 28 animals were uniquely homozygous for the *DPF2* chr29:43546966A>G variant and 17 animals for the *MUS81* chr29:43979681G>T variant for the lactation traits. For bodyweight, six and four animals, respectively,

were uniquely homozygous. To assess whether rare haplotypes of these alleles might represent spurious genotype calls, we manually visualised genotype clusters from all uniquely homozygous animals using GenomeStudio software (v2.0; Illumina). This analysis supported the validity of these genotypes, with animals of interest presenting clean calls that were well delineated within clusters.

We next performed two different association analyses to compare the candidacy of these variants. For the first analysis, we restricted the genotype dataset to exclude animals that were homozygous across multiple candidates for each locus (i.e., only the informative animals uniquely homozygous for the derived alleles highlighted above were retained, see Table 1; Appendix S2). We then fitted linear models with recessive genotype encoding for each phenotype using R software (v4.2.1, R Core Team, 2021), incorporating genotype-based principal components to address confounding due to population stratification (see Appendix S2). Tables 1 and 2 show the effect estimates for the non-reference allele and *p*-value ranges from these models (details about genotype coding can be found in the Appendix S2).

For the chromosome 25 locus, only the *ITGAL* chr25:26689392G>A SNP showed significant negative effects for homozygous animals (smallest $p = 6.9 \times 10^{-10}$ for fat yield). The same was true for the *MUS81* chr29:43979681G>T SNP at the chromosome 29 locus with negative effects on all traits (smallest $p = 8 \times 10^{-8}$ for milk volume) in contrast to the *DPF2* SNP. Together, these data highlighted the *ITGAL* and *MUS81* variants as the better candidates for the two QTL. For the second analysis, we performed similar association testing yet used all the genotype data, including animals that were multiply homozygous for the different candidate variants (see Appendix S2). Comparing the effect sizes and *P*-values between variants yielded a similar result and implication to the primary analysis, with *ITGAL* and *MUS81* mutations presenting the largest and most significant effects for all traits (Tables S2, S3). Notably, differential effects were more clearly resolved for the chromosome 25 locus for these 'all data' models, since the lower LD between these candidates enabled more effective fine mapping of the QTL.

The *ITGAL* gene (aka *CD11a*) encodes the integrin α L chain, a protein that is expressed on all white blood cells as part of the integrin lymphocyte function-associated antigen-1 (LFA-1; Nueda et al., 1993). LFA1 plays multiple roles in cellular adhesion and signalling in the immune system, and a loss-of-function variant in *ITGAL* has previously been associated with compromised immunity in mice (Zhang et al., 2019). Human *ITGAL* variants have also been implicated in various autoimmune phenotypes and disorders (de Lange et al., 2017; Knight et al., 2014; Park et al., 2014). Having demonstrated the *ITGAL*

TABLE 1 Effect size estimates for the mutant allele when only using the homozygous individuals that are unique to each chromosome 25 candidate variant (retained diplotypes are shown in column one) and excluding animals that are homozygous across more than one of these variants. All other animals are included without filter.

Chromosome 25 locus					Fat yield		Protein yield		Milk volume	
IL4R	***KIAA0556***	***ITGAL***	**Count**	**#**	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**
1|1	1|0	1|0	21	22	−7.4	6.2	−3.8	4.9	−36.4	133.3
1|1	1|0	0|0	1							
0|1	**1|1**	0|1	4	4	−27.6	14.6	−13.9	11.6	−256.0	312.6
0|1	0|1	**1|1**	10	13	−50.1***	8.1	−39.1***	6.4	−931.0***	173.4
0|0	0|0	**1|1**	3							
Observations					941 327					

Note: The traits under investigation for this locus were milk fat yield (kg/lactation), milk protein yield (kg/lactation) and milk volume (L/lactation). Focal diplotypes are highlighted in bold. Count = number of times this specific diplotype occurred. # = total count of uniquely homozygous animals for each variant. Observations = total number of animals included in model. SE = standard error. *p*-values: *$p<0.05$; **$p<0.01$; ***$p<0.001$.

TABLE 2 Effect size estimates for the mutant allele when only using the homozygous individuals that are unique to each chromosome 29 candidate variant (retained diplotypes are shown in column one) and excluding animals that are homozygous across more than one of these variants. All other animals are included without filter.

Chromosome 29 locus			Fat yield		Protein yield		Milk volume		Bodyweight		
DPF2	***MUS81***	**#**	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**	**#**	**Estimate**	**SE**
1|1	0|1	28	−1.7	5.5	0.9	4.3	37.6	118	6	−2.4	3.4
0|1	**1|1**	17	−32.8***	7.1	−26.8***	5.6	−812.9***	151.4	4	−8.4*	4.2
Observations		948 377							161 191		

Note: The traits under investigation for this locus were milk fat yield (kg/lactation), milk protein yield (kg/lactation), milk volume (L/lactation) and bodyweight (kg). Focal diplotypes are highlighted in bold. Count = number of times this specific diplotype occurred. # = total diplotype count of uniquely homozygous animals for each variant. Observations = total number of animals included in model. SE = standard error. *p*-values: *$p<0.05$; **$p<0.01$; ***$p<0.001$.

chr25:26689392G>A variant as the likely causative mutation for the chromosome 25 QTL, we wondered whether direct immunological effects might also be apparent in animals homozygous for this variant. As a simple way to investigate these effects, we weighed and sampled blood from seven homozygous cows, and seven broadly matched control animals (Appendix S2). Analysis of blood metabolites and haematology found significant differences in white blood cell count ($5.1 \pm 1.0 \times 10^9$/L; $p = 0.002$), absolute lymphocytes ($3.0 \pm 0.6 \times 10^9$/L; $p = 0.001$), and absolute number and proportion of eosinophils ($-4.9 \pm 1.5 \times 10^9$/L; $p = 0.02$ and $-0.3 \pm 0.1\%$; $p = 0.04$; Table S4, Figure S4). Differences were also apparent in the amount of albumin (-2.9 ± 0.7 g/L; $p = 0.01$), bicarbonate (2.6 ± 1.0 mmol/L; $p = 0.03$), haemoglobin (-8.0 ± 3.1 g/L; $p = 0.04$), and haematocrit/packed cell volume ratio (-0.02 ± 0.01 L/L; $p = 0.02$). Although most cows were within the reference range for these measures, four (of five) homozygous cows had bicarbonate ranges above the reference range (>30 mmol/L), although the significance of this is uncertain.

The main takeaways from this study are several-fold. First, we have validated the chromosome 25 and 29 QTL effects first reported by Reynolds et al., 2021, 2022, highlighting major recessive impacts on lactation and bodyweight traits in a much larger dataset. These effects are substantial and include per-season reductions in milk volume by ~750 L and ~620 L for the chromosome 25 and chromosome 29 loci respectively. Second, we highlight *ITGAL* and *MUS81* premature stop mutations as probably responsible for these effects. In the Reynolds et al. (2021) and Reynolds et al. (2022) papers, candidates at these QTL each proved difficult to differentiate—despite sample sizes of ~100 k animals. While this was an issue of statistical power, the genetic properties of the two QTL were also subtly different, with the chromosome 25 effect presenting low minor allele frequency (MAF) mutations in moderately strong LD (MAF ~ 0.01; $r^2 \sim 0.74$–0.90) and chromosome 29 presenting higher MAF candidates that were near perfectly linked (MAF ~ 0.04; $r^2 \sim 0.98$). In the current study, we used physical genotype data in ~1 million animals to overcome these challenges, enabling identification of sufficient numbers of diplotypes representing the very rare (chromosome 25) and very tightly linked (chromosome 29) candidate variants. Third, these results should enhance biological understanding of these QTL and support future diagnostic testing to manage the frequencies of the variants within breeding schemes. Diagnostics will be of particular relevance in dairy animals of New Zealand ancestry, and other international breeds if these variants are found to segregate in those populations. Analysis of leukocytes in animals homozygous for the *ITGAL* p.Trp731* premature stop mutation showed changes consistent with altered immune function in these cows. While this was a relatively superficial analysis, we can speculate that the lactation effects attributed to the variant are secondary to some form of immune dysfunction, that may extend to other immune phenotypes such as those observed in humans and mice (de Lange et al., 2017; Knight et al., 2014; Park et al., 2014; Zhang et al., 2019). While we did not perform functional studies of *MUS81* mutant homozygotes, it is intriguing to consider how these effects may manifest—given that *Mus81* knockout mice show differences in DNA replication, repair, and structure (Fu et al., 2015). Whether the increased number of chromosomal abnormalities observed in *Mus81* null cells (Dendouga et al., 2005) is relevant to the bodyweight and lactation effects demonstrated in cows is unknown, although future analyses could attempt to quantify DNA replication and/or integrity effects to assess this possibility.

In summary, we report fine mapping of two major recessive QTL impacting lactation and bodyweight traits in dairy cattle, highlighting *ITGAL* and *MUS81* premature stop mutations as probably responsible. These findings demonstrate how population-scale data can be used to differentiate rare and tightly linked candidate variants, and present a starting point to better understand the biology and practical management of these effects.

KEYWORDS

dairy cattle, genomic variation, genotyping, lactation traits, linkage, non-additive association, recessive variants

AUTHOR CONTRIBUTIONS

Liam Williams: Resources. **Mathew D. Littlejohn:** Conceptualization; funding acquisition; writing – review and editing; supervision; project administration. **Edwardo G. M. Reynolds:** Resources. **Swati Jivanji:** Validation; resources; visualization. **Thomas Lopdell:** Resources. **Laura Duntsch:** Conceptualization; investigation; writing – original draft; methodology; validation; visualization; formal analysis; data curation; writing – review and editing.

ACKNOWLEDGMENTS

We are grateful for the use of New Zealand eScience Infrastructure (NeSI) high performance computing facilities as part of this research. The authors would like to thank Kathryn Tiplady, Ahmed Ismael and Mohammad Nilforooshan for their feedback on statistical modelling, Ric Sherlock and Jasper Friedeman for their explanation of cow phenotypes and Keren Dittmer, Hayley Hunt and Gemma Worth for assisting with the *ITGAL* variant trial and ethics approvals. We would also like to thank the reviewers for their helpful comments and suggestions. Open access publishing facilitated by Massey University, as part of the Wiley - Massey University agreement via the Council of Australian University Librarians.

CONFLICT OF INTEREST STATEMENT

L.D., S.J., T.L., L.W. and M.D.L. are paid employees of the Livestock Improvement Corporation, a breeding company and supplier of bovine germplasm. The

Livestock Improvement Corporation is also the applicant for several patent applications related to some of the mutations detailed in this article, with E.G.M.R. and M.D.L. named inventors on these applications. Specifically, these patents relate to genetic testing applications of mutations impacting the *DPF2* (751917), *MUS81* (768801), *IL4R* (768803), *KIAA0556* (768804) and *ITGAL* (790674) genes. All other authors declare no competing interests.

FUNDING INFORMATION

The research was financially supported by the Ministry of Business, Innovation and Employment (MBIE; Wellington, New Zealand) and the Ministry of Primary Industries (MPI; Wellington, New Zealand), who independently co-funded the work through the Endeavour Fund (LICX1802) and (now historical) Primary Growth Partnership research programmes, respectively. Funders had no role in the analysis or interpretation of the data, or in writing the manuscript.

ETHICS STATEMENT

All animal experiments were conducted in strict accordance with the rules and guidelines outlined in the New Zealand Animal Welfare Act 1999. The majority of genotype and phenotype data were generated as part of routine commercial activities outside the scope of that requiring formal committee assessment (as defined by the above guidelines). Approval was sought for additional ear tissue sampling, liveweight measurements, photographing, and blood sampling and subsequently approved by the AgResearch Animal Ethics Committee, Hamilton, New Zealand (approval AEC 1867).

DATA AVAILABILITY STATEMENT

Genotypic and phenotypic data has been deposited on Dryad (DOI: 10.5061/dryad.x3ffbg7wp).

Laura Duntsch[1]
Swati Jivanji[1]
Thomas Lopdell[1]
Edwardo G. M. Reynolds[2]
Liam Williams[1]
Mathew D. Littlejohn[1,3]

[1]*Livestock Improvement Corporation, Hamilton, New Zealand*
[2]*Department of Cardiology, Angiology and Intensive Care Medicine, Deutsches Herzzentrum der Charite (DHZC), Campus Virchow Klinikum, Berlin, Germany*
[3]*AL Rae Centre for Genetics & Breeding, School of Agriculture & Environment, Massey University, Hamilton, New Zealand*

Correspondence
Laura Duntsch, Livestock Improvement Corporation, Hamilton, New Zealand.
Email: laura.duntsch@lic.co.nz

ORCID

Laura Duntsch https://orcid.org/0000-0002-4347-456X
Thomas Lopdell https://orcid.org/0000-0002-7684-4870

REFERENCES

de Lange, K.M., Moutsianas, L., Lee, J.C., Lamb, C.A., Luo, Y., Kennedy, N.A. et al. (2017) Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nature Genetics*, 49(2), 256–261. Available from: https://doi.org/10.1038/ng.3760

Dendouga, N., Gao, H., Moechars, D., Janicot, M., Vialard, J. & McGowan, C.H. (2005) Disruption of murine Mus81 increases genomic instability and DNA damage sensitivity but does not promote tumorigenesis. *Molecular and Cellular Biology*, 25(17), 7569–7579. Available from: https://doi.org/10.1128/MCB.25.17.7569-7579.2005

Dittmer, K. E., Neeley, C., Perrott, M. R., Reynolds, E., Garrick, D. J., & Littlejohn, M. D. (2022) Pathology of the peripheral neuropathy Charcot-Marie-Tooth disease type 4H in Holstein Friesian cattle with a splice site mutation in FGD4. *Veterinary Pathology*, 59(3), 442–450. Available from: https://doi.org/10.1177/03009858221083041

Fu, H., Martin, M.M., Regairaz, M., Huang, L., You, Y., Lin, C.-M. et al. (2015) The DNA repair endonuclease Mus81 facilitates fast DNA replication in the absence of exogenous damage. *Nature Communications*, 6(1), 6746. Available from: https://doi.org/10.1038/ncomms7746

Knight, J.M., Lee, S.-H., Roberts, L., Smith, C.W., Weiss, S.T., Kheradmand, F. et al. (2014) CD11a polymorphisms regulate TH2 cell homing and TH2-related disease. *Journal of Allergy and Clinical Immunology*, 133(1), 189–197. Available from: https://doi.org/10.1016/j.jaci.2013.03.049

Nueda, A., López-Cabrera, M., Vara, A. & Corbí, A.L. (1993) Characterization of the CD11a (alpha L, LFA-1 alpha) integrin gene promoter. *Journal of Biological Chemistry*, 268(26), 19305–19311. Available from: https://doi.org/10.1016/S0021-9258(19)36514-7

Park, S.R., Park, K.S., Park, Y.J., Bang, D. & Lee, E.-S. (2014) CD11a, CD11c, and CD18 gene polymorphisms and susceptibility to Behçet's disease in Koreans. *Tissue Antigens*, 84(4), 398–404. Available from: https://doi.org/10.1111/tan.12420

R Core Team. (2021) *R: A language and environment for statistical computing [computer software]*. Vienna, Austria: R Foundation for Statistical Computing. Available from: https://www.R-project.org/

Reynolds, E.G.M., Lopdell, T., Wang, Y., Tiplady, K.M., Harland, C.S., Johnson, T.J.J. et al. (2022) Non-additive QTL mapping of lactation traits in 124000 cattle reveals novel recessive loci. *Genetics Selection Evolution*, 54(1), 5. Available from: https://doi.org/10.1186/s12711-021-00694-3

Reynolds, E.G.M., Neeley, C., Lopdell, T.J., Keehan, M., Dittmer, K., Harland, C.S. et al. (2021) Non-additive association analysis using proxy phenotypes identifies novel cattle syndromes. *Nature Genetics*, 53(7), 949–954. Available from: https://doi.org/10.1038/s41588-021-00872-5

Zhang, J., Teh, M., Kim, J., Eva, M.M., Cayrol, R., Meade, R. et al. (2019) A loss-of-function mutation in the integrin alpha L (Itgal) gene contributes to susceptibility to *Salmonella enterica* Serovar Typhimurium infection in collaborative cross strain CC042. *Infection and Immunity*, 88(1), e00656-19. Available from: https://doi.org/10.1128/iai.00656-19

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.