

## Genome-wide association study for birth weight in Irish sheep breeds

F. McGovern<sup>1</sup>, A. O'Brien<sup>2</sup>, D. P. Berry<sup>2</sup>, N. McHugh<sup>2</sup> and D. C. Purfield<sup>2</sup>

<sup>1</sup>Teagasc, Athenry, Co. Galway, Ireland

[Fiona.mcgovern@teagasc.ie](mailto:Fiona.mcgovern@teagasc.ie) (Corresponding Author)

<sup>2</sup>Teagasc, Moorepark, Fermoy, Co. Cork, Ireland

### Summary

Lamb birth weight is a quantitative trait which impacts animal productivity, and in particular perinatal mortality. This is reflected by its inclusion as a correlated trait to lambing difficulty, lamb dystocia and perinatal mortality in Irish national genetic evaluations. Research has focused on the genetic variance of lamb survival traits but less emphasis has been given to lamb birth weight. The objective of the present study was to detect regions of the ovine genome associated with lamb birth weight. Genotypes from 40,500 single nucleotide polymorphisms (SNPs) from 9,909 animals of five main breeds namely Belclare (n=1,236), Charollais (n=2,895), Suffolk (n=1,884), Texel (n=3,148) and Vendeen (n=746) were available. Genomic regions of significance were detected in four of the five breeds analysed. In conclusion, the genomic architecture of lamb birth weight is complex with breed specific associations identified in the current study.

*Keywords: sheep, lamb birth weight, genomic*

### Introduction

Productivity and profitability on sheep farms are directly affected by the level of assistance required at lambing and ultimately lamb survival (Morris *et al.*, 2000). Lamb survival rates globally have been documented to vary between 70% and 92% (Nowak & Poindron, 2006) with the highest rates of pre-weaning mortality recorded within the first three days of life (Matheson *et al.*, 2012). Lamb birth weight is one of the primary risk factors associated with lamb mortality (McHugh *et al.*, 2016). Optimum lamb birth weights of between 3.0kg and 5.5kg are targeted in order to enhance lamb viability while minimising the incidence of dystocia at parturition (Nowak & Poindron, 2006). Terminal sire breeds are used extensively in Ireland due to their superior lamb growth rates and carcass conformation. However, increased selection for growth traits may indirectly contribute to heavier lamb birth weights and thus greater dystocia, as evidenced by a strong genetic correlation between birth weight and dystocia (Everett-Hincks *et al.*, 2014). Although lamb birth weight is not included as a goal trait in the national Irish sheep breeding programme, it is included as a correlated trait, highlighting its importance in the overall production system.

Accurate genomic predictions depend on the genetic architecture of the trait, in particular by the number of loci that affect that trait and the distribution of the size of their effects (Hayes *et al.*, 2010). Furthermore, accurate across-breed genomic predictions are reliant on the presence of common genomic variants with similar allele substitution effects in each breed. In comparison to their small ruminant counterparts, numerous studies have investigated quantitative trait loci (QTL) associated with birth weight in both dairy and beef cattle. The objective of the present study therefore was to perform a genome-wide association study (GWAS) to identify regions of the genome associated with lamb birth weight in Irish sheep breeds.

## Materials and Methods

### Genotypic data

Single nucleotide polymorphism (SNPs) genotypes were available from animals genotyped on one of four genotyping panels namely; the Illumina OvineSNP50 Beadchip (54,241 SNPs; 50K panel; 3,081 animals), a custom Illumina Infinium panel (15,000 SNPs; 15K panel; 7,340 animals), a custom AgResearch Ovine HD Beadchip (606,006 SNPs; HD panel; 93 animals) and a custom Affymetrix platform (51,121 SNPs; Affy panel; 81 animals) for a population of Irish purebred sheep breeds including Belclare, Charollais, Suffolk, Texel and Vendéen animals. Only SNPs common between each panel and the Illumina OvineSNP50 panel were retained for analysis resulting in the retention of 11,302, 41,634, and 50,773 SNPs on the 15K, HD, and Affy panels, respectively.

All animals had a genotype call rate >85%, but individual SNPs with a call rate <90% (n=717) were discarded. In addition, only autosomal SNPs and SNPs of known genomic position were retained. Single nucleotide polymorphism with an Illumina GenCall (GC) score <0.55 or with Mendelian inconsistency >0.02% were discarded. Minor allele frequency (MAF) was calculated within each breed and SNPs with a MAF <2% were not considered for further analysis. Finally, SNPs that deviated from Hardy-Weinberg equilibrium ( $P < 1 \times 10^{-9}$ ) within each individual breed were discarded. After all SNP edits, 40,500 SNPs, 35,797 SNPs, 40,494 SNPs, and 9,389 SNPs remained on the 50k, HD, Affy and LD panels, respectively. All genotypes were imputed to the 40,500 SNPs on the 50K panel using the FImpute version 2.2 (Sargolzaei *et al.*, 2014). After imputation, 40,269, 40,420, 40,058, 40,410 and 39,889 non-monomorphic autosomal SNPs remained for the analysis of the Belclare, Charollais, Suffolk, Texel and Vendéen, respectively.

### Phenotypic Data

In Ireland, lamb birth weight is recorded by producers participating in the national breeding programme, using a weighing scale, within 24 hours of birth. Direct estimated breeding values (EBVs) for lamb birth weight were available from the national evaluation. Only animals with a lamb birth weight EBV and a genotype were retained for analysis. A total of 9,909 animal EBVs were used comprising of 1,236 Belclare, 2,895 Charollais, 1,884 Suffolk, 3,148 Texel and 746 Vendéen animals. In addition raw phenotypic weight measurement data was available for all animals and was used to estimate the proportion of phenotypic explained by all SNPs.

### Whole genome association

The proportion of phenotypic variance accounted for by all SNPs (SNPs-based heritability/genomic heritability) was estimated using the genomic restricted maximum likelihood approach in GCTA across all breeds (Yang *et al.*, 2010). Whole genome association analysis was also performed in GCTA (Yang *et al.*, 2011) using a mixed linear model association analysis based on the leave-one-chromosome-out method (Yang *et al.*, 2014), within each breed. This approach accounts for population substructure and relatedness through the construction of genomic relationship matrixes. The following model was used for analysis;

(1)

where  $y$  is the dependent variable,  $\mu$  is the overall mean,  $a$  is the additive effect of the candidate

SNP tested for association, is the SNP genotype indicator variable coded as 0,1, or 2, is the accumulated polygenic effect of all SNPs except those on the chromosome where the candidate SNP is located, and is the residual. False discovery rate (FDR) control was performed using the Benjamini-Hochberg method using a FDR of 0.05.

To identify genomic regions associated with birth weight across each breed a meta-analysis was performed using the weighted Z-score method within the METAL software (Willer et al., 2010). The combined weighted Z-score was calculated as:

$$(2)$$

where the weight  $w_i$  was the square root of the sample size of breed  $i$  and  $Z_i = \Phi^{-1}(1 - \frac{p_i}{2})$  ( $\pm 1$ ) for the sign of the direction of the effect for each breed  $i$  where  $\Phi$  is the standard normal cumulative distribution function and  $p_i$  is the p-value for that SNP in breed  $i$ . Weighted Z-scores were then transformed into their corresponding p-values. Gene search was completed using Ensembl (<http://ensembl.org/>) and NCBI map viewer (<http://www.ncbi.nlm.nih.gov/mapview/>) on the Ovine 3.1 genome build.

## Results & Discussion

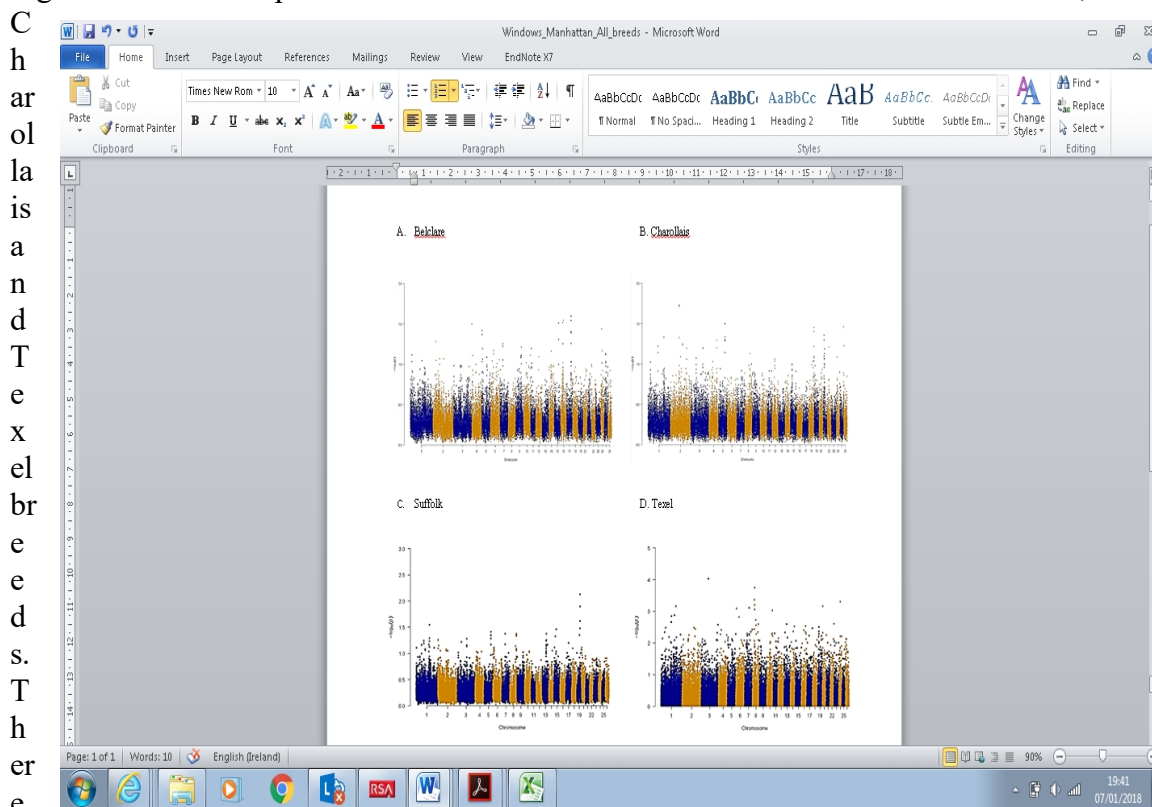
The direct heritability of lamb birth weight calculated in the present study was  $0.17 \pm 0.03$ , which is similar to the heritability estimated for lamb birth weight by Everett-Hincks *et al.* (2014) in a crossbred population.

Previous studies have identified QTLs associated with lamb birth weight on sheep (*Ovis aries*; OAR) chromosome 6 (OAR6) and OAR16 (Gholizadeh *et al.*, 2015). In the present study, only a few SNPs were identified as significant (Adjusted  $P < 0.05$ ) in Belclare, Charollais, Suffolk and Texel animals. There were no significant SNPs identified within the cohort of Vendéen animals examined. The gene RAB11FIP is a protein coding gene which has been associated with adipocyte transport in humans; and is located within the same region of the genome as SNP rs37855315 (OAR22) which was found to have a significant impact on lamb birth weight in Suffolk animals. Biologically relevant genes located in the regions surrounding the significant SNP rs31649270 and SNP rs45777821 include RIPOR2 and SLC35B. SLC35B belongs to a family of nucleotide sugar transporters and is a protein coding gene which is involved in cell function and is linked to glucose transporting pathways in the human.

Table 1. Chromosome location, breed and candidate genes associated with significant SNPs associated with lamb birth weight.

SNP	Chromosome	p-value	Adj. p-value (q value)	Breed	Candidate Gene
rs103152324	OAR1	$2.89 \times 10^{-6}$	0.01	Belclare	AQP10
rs7856616	OAR14	$1.40 \times 10^{-6}$	0.03	Suffolk	PLCG2
rs31649270	OAR20	$1.31 \times 10^{-6}$	0.04	Charollais	RIPOR2
rs45777821	OAR20	$2.82 \times 10^{-7}$	0.01	Texel	SLC35B3
rs37855315	OAR22	$9.7 \times 10^{-7}$	0.03	Suffolk	RAB11FIP

Results from this study indicate that genetic variation exists for lamb birth weight. Significant breed specific SNP associations were observed for the Belclare, Suffolk,



were some potential QTLs identified however a larger cohort of animals is required in order to investigate genomic regions associated with lamb birth weight in greater detail. Further investigation is warranted to determine whether the proportion of genomic variation that is attributed to goal traits and correlated traits in the Irish national evaluation is due to pleiotropic effects within individual breeds. In addition, if genomic evaluations are implemented, the presence of putative QTLs across breeds should be investigated as a possible approach to increasing the accuracy of evaluations.

**Figure 1.** Manhattan plots for single SNP regression for (a) Belclare, (b) Charollais, (c) Suffolk and (d) Texel animals analysis for lamb birth weight.

## List of References

- Everett-Hincks, J.M., H. C. Mathias-Davis, G. J. Greer, B. A. Auvray & K. G. Dodds, 2014. Genetic parameters for lamb birth weight, survival and death risk traits1. *J. Anim. Sci.*, 92(7): 2885-2895.
- Gholizadeh, M., G. Rahimi-Mianji & A. Nejati-Javaremi, 2015. Genomewide association study of body weight traits in Baluchi sheep. *J Genet.*, 94(1): 143-146.
- Hayes BJ, J. Pryce, A.J. Chamberlain, P.J Bowman & M.E Goddard, 2010. Genetic architecture of complex traits and accuracy of genomic prediction: Coat colour, milk-fat percentage, and type in Holstein cattle as contrasting model traits. *PLoS Genet.* 6: 1001139.
- Matheson, S.M., L. Bünger & C.M. Dwyer, 2012. Genetic parameters for fitness and neonatal behavior traits in sheep. *Behavior Gen.*, 42(6): 899-911.
- McHugh, N., D.P. Berry & T. Pabiou, 2016. Risk factors associated with lambing traits. *Anim.* 10(1): 89-95.
- Morris, C.A., S.M. Hickey & J.N. Clarke, 2000. Genetic and environmental factors affecting

- lamb survival at birth and through to weaning. *NZ J. Agr. Res.*, 43(4): 515-524.
- Nowak, R. & P. Poindron, 2006. From birth to colostrum: early steps leading to lamb survival. *Rep. Nut. Dev.* 46(4): 431-446.
- Purfield, D.C., D.G. Bradley, J.F. Kearney & D.P. Berry, 2014. Genome-wide association study for calving traits in Holstein–Friesian dairy cattle. *Anim.* 8(2): 224-235.
- Purfield, D.C., D.G. Bradley, R.D. Evans, F.J. Kearney & D.P. Berry, 2015. Genome-wide association study for calving performance using high-density genotypes in dairy and beef cattle. *Genet. Sel. Evol.*, 47(1): 1-13.