Polymorphisms in the growth hormone and insulin-like growth factor 1 genes are associated with milk production, somatic cell count, and survival in Holstein-Friesian dairy cattle

M.P. Mullen[1], S.M. Waters[2], D.J. Howard[1], C.O. Lynch[3], M.G. Diskin[3], D.A. Magee[4], D.E. MacHugh[4,5], L Giblin[6], D.P. Berry[6]
[1]Teagasc Animal Bioscience Centre, Mellows Campus, Athenry, Co. Galway, Ireland, [2]Teagasc Animal Bioscience Centre, Grange, Co. Meath, Ireland, [3]Teagasc Animal Production Research Centre, Mellows Campus, Athenry, Co. Galway, Ireland, [4]Animal Genomics Laboratory, UCD School of Agriculture, Food Science and Veterinary Medicine, University College Dublin, Belfield, Dublin 4, Ireland, [5]Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Belfield, Dublin 4, Ireland, [6]Teagasc Moorepark Dairy Production Research Centre, Fermoy, Co. Cork, Ireland *Email: michael.mullen@teagasc.ie*

Introduction Milk production, udder health, fertility and survival represent a large proportion of the emphasis within international dairy cattle breeding objectives. In this study, we used a candidate gene approach to detect novel single nucleotide polymorphisms (SNPs) associated with performance. The growth hormone (*GH*) and insulin like growth factor 1 (*IGF1*) genes were chosen because of their known role in milk production and reproduction (Velazquez, 2008). The objective of this study was to discover novel SNPs in these genes and to quantify their association with production, somatic cell count, calving interval and survival in dairy cattle.

Materials and methods A panel of 22 cattle (four Belgian Blue crossbreds, four Charolais, four Simmental, four Aberdeen Angus crossbreds and six Holstein-Friesians) was selected for SNP discovery. Regions of the *GH* and *IGF-1* genes encompassing both promoter and regulatory flanking sequences were PCR-amplified and sequenced. Sequence validation and *de novo* polymorphism detection was carried out using a combination of software packages including BLAST, ClustalW and Chromas. Discovered SNPs, as well as previously published SNPs, were then genotyped using iPLEX-Mass technologies (Sequenom Inc.) across 848 HF sires with progeny in Ireland. The association between each SNP and performance was quantified using weighted mixed models in ASREML (Gilmour *et al*., 2009) with genotyped individual included as a random effect, and average expected relationships among individuals accounted for. Year of birth (divided into five yearly intervals) and percent Holstein of the individual sire were included as fixed effects in the model. In all instances the dependent variable was daughter yield deviation for milk yield, fat yield and percent, protein yield and percent, and somatic cell score (SCS) and de-regressed PTA for calving interval and functional survival, weighted by their respective reliability less the parental contribution. Genotype was included in the model as a continuous variable. A multiple regression model was developed for each gene separately by backward elimination of non-significant ($P > 0.05$) SNPs.

Results Sequence analysis of ~ 13 kb across these two genes identified; 44 *de novo* SNPs in the *GH* gene and nine *de novo* in the *IGF1* gene. Significant associations were evident between both novel and previously published SNPs in both genes with performance (Table 1). For example: in the *GH* gene, a C to T substitution at *GHr17*, was associated with increased milk fat and protein composition of 0.03% and 0.02%, respectively; and in the *IGF1* gene for example, a substitution of the C allele with a T allele at *IGF1r10* was associated with increased functional survival of 0.79 %. The *IGF1* SNP *rs2901285* has previously being associated with carcass traits in cattle; however, this study has also shown its association to milk fat composition, additionally four novel SNPs were identified across both genes with associations to milk production traits.

Table 1 Associations (standard errors in parenthesis) between SNPs in the *GH* and *IGF-1* genes and performance

Gene	SNP[a]	dbSNP	Allele	Fat yield (kg)	Protein yield (kg)	Fat % (*100)	Protein % (*100)	SCS (log_e*100)	Survival (%*100)
GH	*GHi36*	*rs41923523*	C → T					2.03 (0.62)	-29.54
	GHi24	*de novo*	C → T					1.54 (0.64)	
	GHe5	*rs41923484*	C → G		-0.85 (0.42)				
	GHr17	*rs41923483*	C → T			3.35	1.54		
	GHr19	*rs41923481*	C → T	-0.98 (0.45)					
IGF1	*Igf1i3*	*de novo*	A → G	-2.17 (0.78)	-1.31 (0.61)			-2.90 (1.30)	
	Igf1i4	*rs29012855*	A → G			-3.07			
	Igf1i6	*de novo*	A → G	-2.16					
	Igf1r10	*de novo*	C → T						78.56

[a] :*Igf1rX* = SNP X located downstream i.e. 3' of the *IGF-1* gene, *GHi/eY* = SNP Y located within an intron/exon of the *GH* gene.

Conclusion SNPs were identified in both genes that were associated with traits of economic importance substantiating previous research that reported a role of GH and IGF-1 in milk production. Further sequencing would be required to identify the any additional SNPs located in both the *IGF1* and *GH* genes, and functional genomic studies undertaken to determine causation.

References Gilmour, A. R., Cullis, B. R., Welham, S. J. and Thompson, R. 2009. ASREML Reference Manual 2nd edition. Velazquez, M. 2008. Annual Review Biomedical Science 101, 36-62.