

## A GENOME-WIDE ASSOCIATION STUDY (GWAS) FOR CARCASS TRAITS IN HANWOO CATTLE USING IMPUTED WHOLE GENOME SEQUENCE DATA

H. Kim, S. de las Heras-Saldana, N. Moghaddar and J.H.J van der Werf

School of Environmental & Rural science, University of New England, Armidale, NSW, 2351 Australia

### SUMMARY

The identification of genomic region that are associated with phenotypic traits differences is important for improving genomic prediction accuracy. In this study, we aimed to find significant genomic regions for carcass traits in Hanwoo cattle using imputed whole genome sequence data on 13,715 animals. For carcass weight we found 285 SNPs in 7 QTL regions in which 54 candidate genes were identified on BTA4, BTA6 and BTA14. For back fat thickness we found 249 SNPs in 2 QTL regions containing 27 candidate genes on BTA17 and BTA19. The candidate genes from the top 5 significant SNPs were *ZFAT*, *TG* and *TOX* for carcass weight and *NOG* for back fat thickness. No significant SNPs for eye muscle area and marbling score were observed.

### INTRODUCTION

The fast development of the genomic technology enables the use of genomic information to improve the selection of animals in breeding programs. A genome-wide association study (GWAS) identifies associations between genetic variants along the genome and variation in phenotypes. These associations have been used to identify quantitative trait loci (QTL) and candidate genes for complex traits for humans and for diseases and production traits in livestock. QTL information can be used to prioritize genetic markers in order to improve the accuracy of genomic prediction of breeding value. Hanwoo is a Korean native beef cattle breed with the characteristic of high quality meat, mainly caused by high levels of intra muscular fat, also known as marbling. For finding significant QTLs for carcass traits, many GWAS studies have been reported on Hanwoo, However, many of those studies are limited due to their small sample size and low density of genetic markers. The objective of this study is to identify candidate genes for carcass traits using a larger number of samples with imputed sequence data.

### MATERIALS AND METHODS

In total, 13,715 animals with genotypes and phenotypes for carcass traits were used in this study. The four carcass and meat quality traits recorded were carcass weight (CWT), back fat thickness (BFT), eye muscle area (EMA) and marbling score (MS).

**Table 1. Summary statistics for carcass traits**

Traits	Mean	Standard deviation	Min	Max	Coefficient of variation
CWT (kg)	425.48	59.84	152	692	0.14
BFT (mm)	13.42	5.23	1	57	0.39
EMA (cm <sup>2</sup> )	92.61	12.56	22	156	0.14
MS (1-9)	5.68	1.98	1	9	0.35

The phenotypic data were adjusted for fixed effects using a linear mixed model in ASReml v.4.1. (Gilmour *et al.* 2014):

$$y = CG + \text{Sex} + \text{age} + e$$

Where  $y$  is the observation vector, CG is the fixed effect of contemporary group, defined by birth season (4 season per year) and farm, Sex and age are covariates and  $e$  is the vector with residuals, which will be used as adjusted phenotypes in our GWAS.

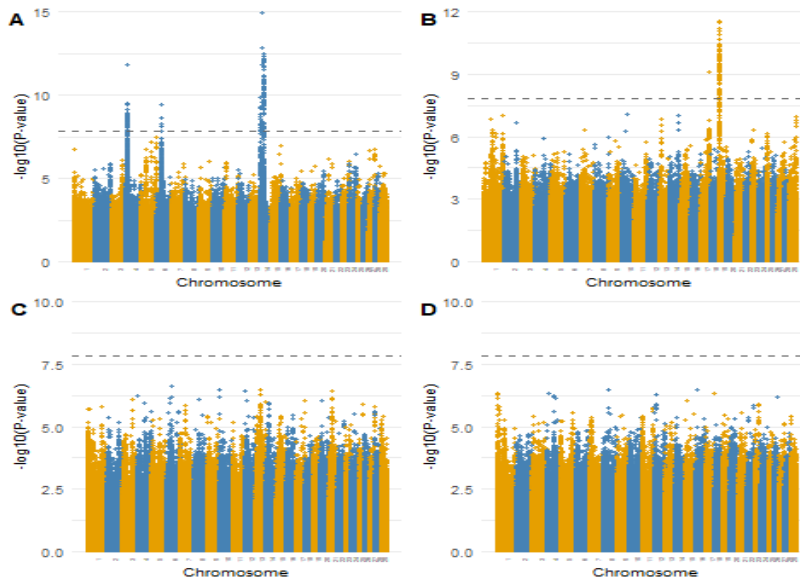
Animals were genotype with the Illumina Bovine SNP 50K Bead Chip. After quality control, only 14K SNPs were remaining and these were used to first impute from 14K to 50k, then to high density, and finally to sequence with 203 reference animals using Beagle V5.1. The imputed SNPs with an accuracy of imputation ( $R^2$ ) lower than 0.4 were removed. Finally, 17,549,506 SNPs and 13,715 animals were used in this study. A single SNP regression, GWAS was performed under a mixed linear model in (MLMA) in GCTA v.1.93 (Yang *et al.* 2011):

$$y^* = \mu + Xb + g + e$$

Where  $y^*$  is a vector with adjusted phenotypes one for each of the four traits,  $\mu$  is the overall mean,  $b$  is the allele substitution effects and  $X$  is the vector of genotype codes for SNP fitted.  $g$  is a vector of additive genetic effects with  $g \sim N(0, G\sigma_g^2)$ , where  $G$  is the genomic relationship matrix (GRM) calculate from 17,549,506 SNPs in PLINK v.1.9 and  $e$  is the residual effect. Manhattan plots were produced using ggplot2 packages in R. To reduce type-1 errors, the significance threshold was set at ( $P < 1.54 \cdot E08$ ), derived as 0.05 divided by the number of independent variants, which in turn was calculated after not counting SNPs in linkage disequilibrium  $> 0.5$  with other SNPs (Sham and Purcell 2014; Bedhane *et al.* 2019). We used the Ensamble database for *Bos taurus* UMD3.1 to identify the candidate genes that were located within 1Mb of the significant SNPs.

## RESULTS AND DISCUSSION

We identified 9 QTL regions for carcass traits in Hanwoo.



**Figure 1. Manhattan plot for carcass weight (A), back fat thickness (B), eye muscle area (C) and marbling (D)**

The Manhattan plot from the results of GWAS are shown in Figure 1 for all traits. In total, 285 SNPs in 7 QTL regions were detected for CWT on *Bos Taurus* autosome 4 (BTA4), BTA6 and BTA14, these chromosomes include 80, 8 and 197 SNPs respectively. The most significant QTLs for CWT were located on BTA14 which contained 32 candidate genes. The 22 candidate genes were

found on BTA4 and nine candidate genes were found on BTA6 (Table 2). Previous GWAS also reported QTL for CWT on BTA4, BTA6 and BTA14 in Hanwoo (Lee *et al.* 2012; Srikanth *et al.* 2020) and BTA6 and BTA14 contains QTL regions in Chinese Simmental and multiple beef cattle breeds (Lu *et al.* 2013; Chang *et al.* 2018; Wang *et al.* 2020).

Significant SNPs associated with BFT were located on BTA17 and seven genes were found close to significant SNPs on BTA19. Significant QTL regions for BFT in Hanwoo were detected on BTA13 and BTA16 (Lee *et al.* 2012) and BTA19 (Srikanth *et al.* 2020) similar to the results obtained in this study.

**Table 2. QTL regions and candidate genes associates with carcass traits**

Traits	Chr	Position (Mb)	Candidate genes
Carcass weight	4	7.82 - 9.07	FZD1
	4	9.35 - 12.4	KRIT1, ANKLB1, TMBIM7, GATAD1, PEX1, RBM48, CDK6, SAMD9, CALCR, TFPI2, GNGT1, GNG11, BET1, VPS50, HEPACAM2, COLA2, CASD1, SGCE, PEG10, PPP1R9A, PON1
	6	38.52 - 39.52	LAP3, MED28, FAM184B, LCORL, DCAF16, NCAPG
	6	40.4 - 42.11	SLIT2, PACRGL, KCNIP4
	14	4.91 - 6.33	COL22A1, FAM135B
	14	6.58 - 10.89	KDHRBS3, ZFAT, ST3GAL1, NDRG1, WISP1, TG, SLA, PHF20L1, TMEM71, LRRRC6, KCNQ3
	14	23.99 - 27.65	RP1, KR4, TMEM68, TGS1, LYN, RPS20, MOS, PLAG1, SDR16C5, SDR16C6, PENK, IMPAD1, FAM110B, UBXXN2B, SDCBP, CYP7A1, NSMAF, TOX, CA8
Back fat thickness	17	64.84 - 65.84	SIRT4, MS11, SRSF9, GATC, TRIAP1, COX6A1, COQ5, DYNLL1, RNF10, POP5, CABP1, MLEC, UNC119B, ACADS, SPPL3, HNF1A, OASL, C17H12orf43, ANKRD13A, GIT2, TCHP, GLTP, TRPV4
	19	7.02 - 8.15	FAM222A, ANKFN1, NOG, C19H17orf67, DGKE, TRIP25, COIL, SCPEP1

No significant SNPs were detected for EMA and MS. In another recent study, no significant QTLs were detected for the MS in Hanwoo (Srikanth *et al.* 2020). MS appears to be mainly affected by many genes, each with a small effect.

**Table 3. Top 5 significant SNPs and candidate genes associated with carcass traits**

Traits	Chromosome	Position	P-value	Candidate genes
Carcass weight	14	8,160,456	1.20E-15	ZFAT
	14	9,518,339	1.41E-13	TG
	14	26,619,895	3.60E-13	TOX
	14	26,621,673	3.60E-13	TOX
	14	26,622,060	3.60E-13	TOX
Back fat thickness	19	7,645,081	3.02E-12	NOG
	19	7,620,479	3.31E-12	NOG
	19	7,646,102	6.87E-12	NOG
	19	7,620,249	8.93E-12	NOG
	19	7,618,889	1.21E-11	NOG

The *ZFAT* (zinc finger and AT-hook domain containing) gene which is located near to the most

significant SNP for CWT has the potential for semi-lethality in Aberdeen Angus (Jenko *et al.* 2019) but was also associated with growth in humans and horses (Lango Allen *et al.* 2010; Makvandi-Nejad *et al.* 2012). The *TG* (thyroglobulin) gene plays a role in metabolism and has been associated with carcass and growth traits in cattle (Zhang *et al.* 2015). The *TOX* (thymocyte selection associated high mobility group box) gene has been associated with reproductive traits (de Camargo *et al.* 2015). All top five significant SNPs for BFT were located on BTA 19 were close to the *NOG* (Noggin) gene. *NOG* plays a role in inducing adipogenesis (Sawant *et al.* 2012) and was previously associated with BFT in Hanwoo (Srikanth *et al.* 2020).

## CONCLUSIONS

This study shows identified QTL regions and candidate genes associated with carcass traits in Hanwoo. Seven QTL regions with 63 candidate genes were found for carcass weight and two QTL regions with 31 candidate genes for back fat thickness. There were no significant genomic regions for eye muscle area and marbling score. This result can be helpful as genomic information to improve the accuracy of genomic prediction in Hanwoo breeding.

## REFERENCES

- Bedhane M., van der Werf J.H.J., Gondro C., Duijvesteijn N., Lim D., Park B.H., Park M.N., Hee R.S and, Clark, S. (2019) *Front Genet.* **10**: 1235.
- Chang C., Chow C., Tellier L.C., Vattikuti S., Purcell S.M. and Lee J.J. (2015) *Gigascience.* **4**: 7.
- Chang T., Xia J., Xu L., Wang X., Zhu B., Zhang L., Gao X., Chen Y., Li J and Gao H. (2018) *Anim. Genet.* **49**(4): 312.
- de Camargo G.M., Costa R.B., de Albuquerque L.G., Regitano L.C., Baldi F. and Tonhati H. (2015) *Reprod. Fertil. Dev.* **27**(3): 523.
- Gilmour A.R., Gogel B.J., Cullis B.R., Welham S.J., Thompson R., Butler D., Cherry M., Collins D., Dutkowsky G. and Harding SA. (2014) VSN international Ltd
- Jenko J., McClure M.C., Matthews D., McClure J., Johnsson M., Gorjanc G. and Hickey, J. M. (2019) *Genet. Sel. Evol.* **51**(1): 9.
- Lango Allen H., Estrada K., Lettre G., Berndt S.I., Weedon M.N., Rivadeneira F., . . . Hirschhorn, J.N. (2010) *Nature* **467**(7317): 832.
- Lee S.H., van der Werf J., Lee S.H., Lim D.J., Park E.W., Gondro C., Yoon D., Oh S.J., Kim O.H., Gibson J. and Thompson J. (2012) *Genes & Genomics.* **34**(1): 43.
- Lu D., Sargolzaei M., Kelly M., Vander Voort G., Wang Z., Mandell I., Moore S., Plastow G. and Miller P. (2013) *BMC Genet.* **14**: 80.
- Makvandi-Nejad S., Hoffman G.E., Allen J.J., Chu E., Gu E., Chandler A.M., Loreda A.I., Bellone R.R., Mezey J.G., Brooks S.A and Sutter N.B. (2012). *PLoS One* **7**(7): e39929.
- McClure M.C., Morsci N.S., Schnabel R.D., Kim J.W., Yao P., Rolf M.M., McKay S.D., Gregg S.J., Chapple R.H., Northcutt S.L. and Taylor, J. F. (2010) *Anim Genet.* **41**(6): 597.
- Sawant A., Chand D., Isayeva T., Tsuladze G., Garvey W.T., and Ponnazhagan S. (2012). *J Biol Chem.* **287**(15): 12241.
- Sham P.C. and Purcell S.M. (2014) *Nat. Rev. Genet.* **15**(5): 335.
- Srikanth K., Lee S.H., Chung K.Y., Park J.E., Jang G.W., Park M.R., Kim N.Y., Kim T.H., Chai H.H., Park W.C. and Lim D. (2020) *Genes (Basel)* **11**(3).
- Wang Y., Zhang F., Mukiiibi R., Chen L., Vinsky M., Plastow G., Basarab J., Stothard P. and Li, C. (2020) *BMC Genomics* **21**(1): 38.
- Yang J., Lee S.H., Goddard M.E. and Visscher P.M. (2011) *Am. J Hum. Genet.* **88**(1): 76.
- Zhang L.P., Gan Q.F., Hou G.Y., Gao H.J., Li J.Y. and Xu, S. Z. (2015) *Genet. Mol. Res.* **14**(2): 5320.