Genome-wide Association Study for Loin Marbling Score in Canadian Duroc pigs

A. Neustaeter[1], D. A. Grossi[1], M. Jafarikia[1,2], M. Sargolzaei[1,3] and F. Schenkel[1]
[1]Centre for Genetic Improvement of Livestock, University of Guelph, Guelph, Ontario, Canada,
[2]Canadian Centre for Swine Improvement Inc., Ottawa, Ontario, Canada, [3]The Semex Alliance, Guelph, Ontario, Canada

ABSTRACT: The objective of this investigation was to perform a genome-wide association study to identify chromosomal regions associated with loin marbling score (LMS) in Duroc pigs. A Generalized quasi-likelihood scoring test and a 10% false discovery rate at chromosome-wise level were used to identify the SNPs associated with LMS estimated breeding values. A total of 35 SNPs located on chromosomes 3, 5, 9, 10 and 18 were significantly associated with LMS and 47 genes were mapped either harboring or being next to these SNPs. Some of the significant SNPs were located close to or within genes (APOB, SORL1, SC5DL, FGF23, GRIK4, CLCN1) that play role in biological processes and pathways related to lipid metabolism and fat deposition, which were significantly enriched in functional analyses. The identified genes are potential candidate genes for the underlying genetic variation in marbling score in Duroc pigs.
Keywords:
candidate genes
marbling
Duroc

Introduction

With the continuous advancements in the global agricultural markets, pork quality is becoming more important to both producers and consumers. Intramuscular fat (IMF) is a notable trait in pork that is related to increased palatability for the consumer (Albrecht et al. (1996); Ellis et al. (1996); Gandemer, (2002)). The Duroc breed is noted to have good on-farm performance traits, carcass characteristics, and high economic merit for producers due to superior growth, efficient food utilization, and higher IMF levels than traditional white breeds (Sutherland et al. (1984); Staene (1986); Wood et al. (1987)). However, heavy selection on leanness in the Duroc breed starting in the mid 1980's has resulted in decreased sensorial characteristics of the meat, such as taste, aroma, tenderness, juiciness, and visual marbling scores (Oksbjerg et al. (2000); Ruiz-Carrascal et al. (2000); Schwab et al. (2006)). Obtaining and validating area(s) in the genome associated to IMF and marbling with genomics tools like the porcine 60k single nucleotide polymorphism (SNP) panel can assist with the accuracy of genetic evaluations for IMF and marbling using genetic evaluations enhanced by genetic markers. This will aid in allowing producers to select pigs with increased IMF levels.

The objective of this investigation was to perform a genome-wide association study to identify chromosomal regions and candidate genes associated with loin marbling score in Duroc pigs, contributing to a better understanding of the genetic control of this trait and providing new insights for genetic selection enhanced by genetic markers in Duroc pigs.

Materials and Methods

Data. Estimated breeding values (EBVs) and corresponding reliabilities for loin marbling score (LMS) were from various breeding stock suppliers across Canada and were provided by the Canadian Centre for Swine Improvement Inc. (CCSI). The model for estimation of marbling breeding values included carcass weight as a covariate, sex and the interaction of plant by slaughter date by technician as fixed effects. Litter was included as a random effect. A minimum reliability threshold of 10% was applied to keep an EBV in the analysis. All the pedigree information available was used to build the relationship matrix.

Animals were genotyped for the PorcineSNP60 panel (Illumina Inc., San Diego, CA, USA) and SNPs were mapped to their corresponding genomic locations based on the current pig genome annotation, Sus Scrofa Build 10.2. A total of 37,284 SNPs were mapped to autosomal chromosomes. After exclusion of possible misplaced SNPs, the genotype file was filtered according to four criteria: SNP call rate ≥ 90%, minor allele frequency ≥ 0.05, p-value of $\chi 2$ test for Hardy-Weinberg equilibrium ≥ 10^{-6}, and animal call rate ≥ 90%. The analyzed data included 426 animals with EBVs, with a mean of reliability equal to 46.11 and standard deviation equal to 20.50, and 35,429 SNPs after editing

Genome-wide association analysis. A Generalized quasi-likelihood scoring (GQLS) test (Feng et al. (2011)) was used to evaluate the association of the SNPs with LMS EBVs. In this methodology the relationship matrix is accounted for and a logistic regression model is used to link the phenotypic distribution of the trait (EBVs) to the distribution of allelic frequencies. A 10% false discovery rate (Benjamini and Hochberg (1995)) at chromosome-wise level (FDR_CW) was used to account for multiple tests.

Gene Mapping and *in silico* functional analysis. The significant SNPs were mapped to their corresponding genes or surrounding genes within a distance of 100kb, using the NGS-SNP scripts (Grant et al. (2011)) and Sus Scrofa 10.2 genome assembly. *In silico* functional investigation and enrichment analysis included: genes harboring SNPs and genes within a distance of 100kb from the significant SNPs. The list of genes was submitted to

Database for Annotation, Visualization, and Integrated Discovery (DAVID) 6.7 software (Huang et al. (2009)) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al. (2014)) for functional and enrichment analysis, using human orthologs when possible.

Results and Discussion

Genome-wide association analysis. Results from the GQLS analyses are shown in Figure 1. A total of 35 SNPs located on chromosomes 3 (6 SNPs), 5 (2 SNPs), 9 (18 SNPs), 10 (8 SNPs) and 18 (1 SNP) were associated (FDR_CW<0.10) with LMS and 47 genes were mapped either harboring SNPs (7 genes) or being next to these SNPs (40 genes).

Functional analyses. The functional analyses indicated four relevant biological processes (steroid metabolic process, sterol metabolic process, transmission of nerve impulse, and neurological system process) and one significant pathway (regulation of actin cytoskeleton). The genes playing role in the biological processes and pathways, significantly ($p<0.10$) enriched in functional analyses, are described in Table 1 and the SNPs located within these genes are showed in Table 2.

Some of the mapped genes, such as SORL1, APOB, SC5DL, and FGF23 play role in the steroid or sterol metabolic processes that are lipid metabolic processes and directly related to marbling adipose tissue, also known as interfascicular or intramuscular adipose tissue (Hausman et al. (2009)). The genes GRIK4, CLCN1, and CACNB2 are involved in the transmission of nerve impulse that is a cell communication process and is important for the communication between adipocytes and myogenic cells. The interactions between myogenic cells and adipocytes play a significant role in growth and development, including the rate and extent of myogenesis, muscle growth, adipogenesis, lipogenesis/lipolysis, and in the utilization of energy substrates (Kokta, et al. (2004)). Hormones and growth factors involved in the regulation of these processes have important implications in their influence on relative fat and lean deposition and the efficiency of energy utilization in growth and development (Kokta, et al. (2004)).

A minimum reliability threshold of 10% is a useful threshold for screening possible candidate genes but allows for greater error rates in the input EBVs, so a validation study on an independent group of animals is needed to confirm results obtained from this study.

Conclusion

Some of the significant SNPs identified were located close to or within genes that are involved in lipid metabolism. Thus, these identified genes are potential candidate genes for the underlying genetic variation in marbling score in the Duroc pigs. Further research around these significant regions might also help to identify more SNPs or genes that can contribute to marbling score variation. An additional study with more genotyped animals with reliable EBVs is suggested to validate these results.

Acknowledgement

The authors would like to express their sincere gratitude to breeders who actively participated in the project and contributed valuable time, pigs, and support towards the project. This project was made possible thanks to funding from many different public and private sources including Agriculture Adaptation Councils in Québec, New Brunswick, Nova Scotia, Manitoba, and Ontario, the *Ministère de l'Agriculture, des Pêcheries et de l'Alimentation du Québec* (MAPAQ) and Agriculture and Agri-Food Canada through Swine Innovation Porc and the Canadian Agricultural Adaptation Program. Funding contributions from CCSI's regional centres, PigGen Canada, *la Fédération des producteurs de porcs du Québec* (FPPQ) and Canadian swine breeders is also gratefully acknowledged.

Literature Cited

Albrecht, E., Wegner, J., and Ender, K. (1996). Fleischwirtschaft. 76:1145-1148.

Benjamini, Y., and Hochberg, Y. (1995). J. Royal Stat. Soci. Series B, 57:289-300.

Ellis, M., Webb, A. J., Avery, P. J. et al. (1996). J. Anim. Sci. 74:521-530

Feng, Z., Wong, W., Gao, X. et al. (2011). Ann. Appl. Stat. 5:2109-2130.

Gandemer, G. (2002). Meat Sci. 62:309-321.

Grant J. R., Arantes, A. S., Liao, X. et al. (2011). Bioinformatics, 27:2300-2301.

Huang, D. W., Sherman, B. T. and Lempicki, R. A. (2009). Nat. Protoc. 4: 44-57.

Hausman, G. J., Dodson, M. V., Ajuwon, K. et al. (2009). J. Anim. Sci. 87:1218–1246

Kanehisa, M., Goto, S., Sato, Y. et al. (2014). Nucleic Acids Res. 42: 199-205.

Kokta, T. A., Dodson, M. V., Gertler, A. et al. (2004). Domest. Anim. Endocrin., 27:303–331.

Oksbjerg, N., Peterson, J. S., Sorensen, I. L. et al. (2000). Anim. Sci. 71:81-92.

Ruiz-Carrascal, J., Ventanas, J., Cava, R. et al. (2000). Food Res. Int. 33: 91-95.

Schwab, C. R., Baas, T. J., Stalder, K. J. et al. (2006). J. Anim. Sci. 84: 1577-1583.

Steane, D. E. (1986). Anim. Sci. 3: 153-157.

Sutherland, R. A., Webb, A.I., and King, J. W. B. (1984). J. Agr. Sci. 103: 561-570.

Wood, J. D., Kempster, A. J., David, P. J. et al. (1987). Anim. Prod. Sci. 44: 488.

Figure 1: Plots of the genome-wide association analysis for loin marbling score in Duroc pigs

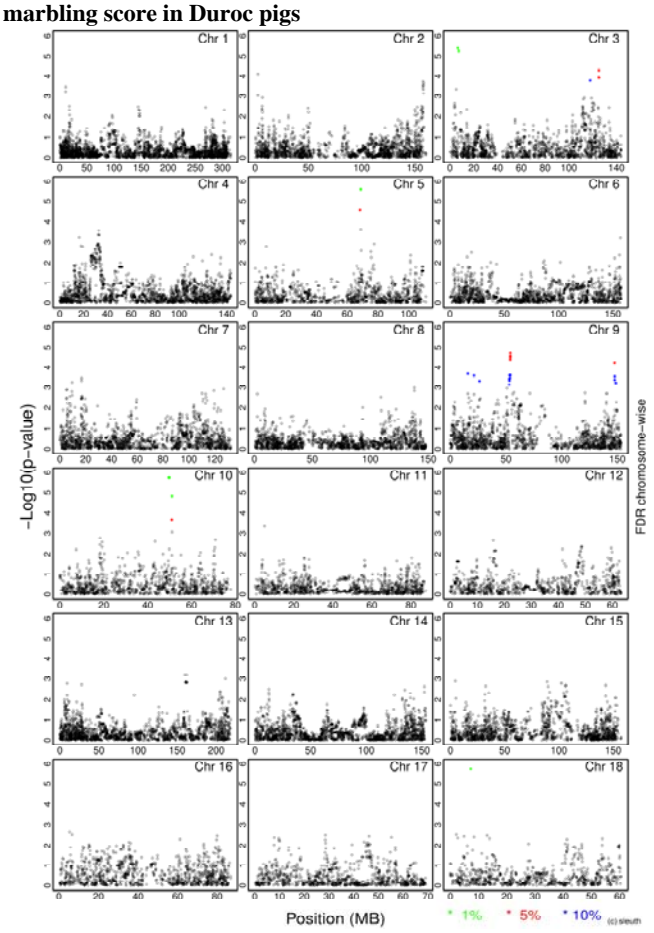

Table 1. Relevant biological process (BP) terms and Kyoto encyclopedia of genes and genomes (KEGG) pathways significantly enriched in functional analyses of genes associated with loin marbling score in Duroc pigs

BP or KEGG Pathway	p-values*	Gene Symbol [SNP]
BP: steroid metabolic process	0.003	SORL1[4a2i] APOB[5a] SC5DL[1a] FGF23[1a]
BP: sterol metabolic process	0.010	SORL1[6a] APOB[5a] SC5DL[1a]
BP: transmission of nerve impulse	0.093	GRIK4[3a] CLCN1[1a] CACNB2[1a1i]
BP: neurological system process	0.097	GRIK4[3a] ITGA8[2i] TAS2R40[1a] CLCN1[1a] CACNB2[2a]
Pathway: Regulation of actin cytoskeleton	0.064	ARPC1B[1a] FGF23[1a] ITGA8[2i]

*modified Fisher exact p-value from DAVID gene-enrichment analysis; [a]SNP located up to 100kb of the gene; [i]intron variant. The numbers in superscript refers to the number of SNPs.

Table 2. Description of the SNP located close to or within the genes identified by functional analyses.

SNP Name	Chr	Position (bp)	MAF	Genes
ASGA0096844***	3	6,587,683	0.10	ARPC1B[a]
ASGA0016343**	3	125,197,050	0.10	APOB[a]
INRA0011611**	3	125,293,055	0.10	APOB[a]
ASGA0016339**	3	125,339,052	0.10	APOB[a]
H3GA0016584**	5	68,341,184	0.50	FGF23[a]
M1GA0013010*	9	52,973,032	0.24	GRIK4[a]
ALGA0052977*	9	53,063,345	0.22	GRIK4[a]
H3GA0055629*	9	53,317,006	0.23	GRIK4[a]
H3GA0027334*	9	53,604,565	0.24	SORL1[a]
ASGA0043045**	9	53,789,765	0.13	SORL1[i]
ALGA0052992*	9	53,807,005	0.14	SORL1[i]
ASGA0043057**	9	53,879,442	0.14	SORL1[a]
ASGA0043056**	9	53,892,198	0.15	SORL1[a]
ALGA0052993**	9	53,906,236	0.14	SORL1[a]
ASGA0089200***	10	49,588,426	0.09	CACNB2[i]
H3GA0030196**	10	50,993,659	0.11	ITGA8[i]
H3GA0030197**	10	51,028,164	0.11	ITGA8[a]
ASGA0104235***	18	7,219,093	0.13	CLCN1[a]

***1% FDR_CW; **5% FDR_CW; *10% FDR_CW; MAF: minor allele frequency; [a]SNP located up to 100kb of the gene; [i]intron variant