Meta-analysis of Genome Wide Association Studies for Pork Quality Traits

Y.L. Bernal Rubio[*†], J.L. Gualdrón Duarte[†], R.O. Bates[*], C.W. Ernst[*], D. Nonneman[‡], G.A. Rohrer[‡], A. King[‡], S.D. Shackelford[‡], T.L. Wheeler[‡], R.J.C. Cantet[†] and J.P. Steibel[*]

[*]Michigan State University, East Lansing, MI, USA, [†]Department of Animal Science, University of Buenos Aires, Buenos Aires, Argentina, [‡]USDA/ARS, Clay Center, NE, USA

ABSTRACT: Genome-wide association studies (GWA) were performed for eight meat quality traits and also, a meta-analysis (MA) of GWA was implemented combining independent results from pig populations. Data from three pig datasets (USMARC, Commercial and MSUPRP) were used. MA was implemented by combining z-scores derived for each SNP in every population, and then, weighting them using the inverse of estimated variance of SNP effects. In population specific GWA, several regions were identified as significantly associated with most of the traits. MA-GWA permitted detection of significant peaks that were not reported within population-GWA. Thus, MA-GWA methodology is an attractive alternative to integrate results for economically relevant traits from multiple GWA, when several populations are available.

Keywords: genome-wide association; meta-analysis; pork quality traits

Introduction

Genomic selection allows estimation of genomic breeding values or GEBV for economically relevant traits (Hayes et al. (2009)). A common practice after a successful genomic evaluation consists of performing genome-wide association analyses (GWA; Wang et al. (2012)). Among the various statistical methodologies allowing GEBV prediction and simultaneous GWA (Minozzi et al. (2012)), the most commonly used is GBLUP (Wang et al. (2012)). We have recently modified this method to efficiently compute GWA p-values (Gualdrón-Duarte et al. (2014)).

A limitation of many GWA applications in animal breeding is that due to the effect sizes of GWA tests, large sample sizes are required to guarantee acceptable power (Minozzi et al. (2012)). However, increasing sample size may not be possible in animal production settings (Houlston et al. (2010)). Thus, one alternative is to perform a meta-analysis of GWA by quantitatively combining GWA information from independent GWA (Evangelou and Ioannidis (2013)). Meta-analysis allows more accurate estimates of SNP effects to be obtained than those derived from population-specific GEBV, improving power and accuracy to detect genomic associations that are consistent across populations (Minozzi et al. (2012)). Therefore, a combined analysis or meta-analysis of several populations using GBLUP is an attractive approach to increase power to identify variants with small effects. The goal of this research is to develop and implement a meta-analysis of GWA for pork quality traits, combining results from multiple GBLUP evaluations, assuming independence between evaluations.

Materials and Methods

Data. Records from purge loss % (PRL), cooking loss % (CKL), ultimate pH 24 hours post-slaughter (pHu), shear force (SF), intramuscular fat % (IMF), CIE L*, CIE a* and CIE b*, were obtained from three pig populations. 1) F_2 generation (n=1259) of the Michigan State University Pig Resource Population (MSUPRP) from which records were analyzed and details were reported in Edwards et al. (2008) and Gualdrón-Duarte et al. (2013), 2) Meat Animal Research Center Population (MARC, n=1237) that has been described by Nonneman et al. (2013), and 3) a commercial population (N=2001). All three populations were genotyped using the PorcineSNP60 BeadChip (Illumina, Inc. Ramos et al., 2009) or the Neogen Porcine GGP LD (version 1) chip (Badke et al. (2012)) and imputed with high accuracy (Badke et al. (2012); Gualdrón-Duarte et al. (2013)).

Statistical analyses. For all traits, a GBLUP model (Aguilar et al. (2010); Wang et al. (2012)) was fitted, incorporating fixed effects of age at slaughter, contemporary group and sex. SNP effects $\boldsymbol{g}$ were obtained from a linear transformation of estimated breeding values $\hat{\boldsymbol{a}}$ for genotyped individuals (Wang et al. (2012)). SNP effects variance (Gualdrón-Duarte et al. (2014)) were given by

$$\mathrm{Var}(\hat{\boldsymbol{g}}) = \boldsymbol{Z}'\boldsymbol{G}^{-1}\boldsymbol{Z}\sigma^2_{\mathrm{A}} - \boldsymbol{Z}'\boldsymbol{G}^{-1}\boldsymbol{C}^{\boldsymbol{aa}}\boldsymbol{G}^{-1}\boldsymbol{Z} \quad [1]$$

Also, *P*-values for significance of SNP effects (Gualdrón-Duarte et al. (2014)) were computed using

$$p\text{-value}_{ij} = 2\left[1 - \Phi\left(\left|\frac{\hat{\boldsymbol{g}}_{ij}}{\sqrt{\mathrm{Var}(\hat{\boldsymbol{g}}_{ij})}}\right|\right)\right] \quad [2]$$

The Z-scores from independent GWA were combined into a single meta-analysis *Z-score*,

$$Z_{overall} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{M} Z_{ij}\sqrt{w_{ij}}}{\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{M} w_{ij}}} \quad [3]$$

where Z_{ij} is the *Z*-score for the *j*th SNP in *i*th population, with *j*=1,… *M*, *k* is the number of populations, with *k*=1,… *n*, and w_{ij} is the weight for $\hat{\boldsymbol{g}}_{ij}$ calculated as $w_{ij} = 1/\mathrm{Var}(\hat{\boldsymbol{g}}_{ij})$. *P*-values were obtained using [2].

Results and Discussion

Table 1 contains SNP associations with pork quality traits based on applying population-specific GWA. For PRL in the MSUPRP, we found a significant SNP on SSC4 located 6Mb from the position of a QTL previously reported for this trait by Ma et al. (2013). Also in two

populations (Commercial and MSUPRP) we found significant overlapping genomic regions on SSC15 associated with pHu. This region of SSC15 coincides with the location of the gene *PRKAG3* (protein kinase, AMP-activated, gamma 3 non-catalytic subunit), which has been confirmed to affect glycogen content in the muscle and resulting meat quality (Ciobanu et al. (2001)).

Table 1. SNP associations for pork quality traits across populations.

Tr[1]	Pop[2]	SNP	SSC	Peak[3], Mb	Reg[4], Mb	P-val
PRL	MS	ALGA-0087273	4	75.0	75	3.22 e-8
		MARC-0093624	15	135.5	133.1-145.6	4.44 e-15
pHu	C	ASGA-0070646	15	133.9	133.6-134.1	1.19 e-10
	MS	H3GA-0052416	15	135.2	122.9-138.7	1.21 e-11
SF	C	DRGA-0003281	2	109.3	109.3	2.56 e-7
	MA	H3GA-0055977	2	5.4	5.4	2.15 e-7
	MS	M1GA-0002229	2	2.9	2.9-5.4	2.89 e-8
CKL		H3GA-0052416	15	135.2	133.1-135.5	1.48 e-8
	C	M1GA-0020450	15	133.9	133.6-134	3.61 e-9
	MS	MARC-0047188	15	135.2	133.1-137.5	2.67 e-15
CIE L*	C	M1GA-0020450	15	133.9	133.9	1.66 e-8

For SF, we identified three significant SNP on SSC2 (one in each population) and one on SSC15 (MSUPRP). The SNP detected on SSC2 in the MARC population was 500kb from a significant SNP reported by Nonneman et al. (2013) using the same dataset. Also, an additional significant peak on SSC2 identified in the Commercial population, was located 17Mb from a significant SNP reported by Nonneman et al. (2011) for the MARC and other pig populations. In this case, the nearest candidate gene is *CAPN1*, which is located 700kb from the peak reported in this study. Additionally, two peaks (one on SSC2 and one on SSC15) were detected in the MSUPRP. For the same dataset and trait, Choi et al. (2011) detected a significant QTL on SSC15, flanked by two microsatellite markers located between 127.8Mb and 135.3Mb. Thus, our interval was narrower (2.4Mb) than the previous report from our group.

With respect to CKL, population-level GWA identified a significant region on SSC15 in the Commercial and MSUPRP populations, which was similar to a QTL reported for this trait by Nonneman et al. (2013). The region observed in the MSUPRP was wider than the one reported for the Commercial population, principally due to the high linkage disequilibrium (LD) observed in the F_2 dataset.

A significant SNP on SSC15 in the Commercial population was associated with CIE L*. This SNP was located in a similar region of a QTL reported as significant for this trait in the MSUPRP by Choi et al. (2011). Nonetheless, we failed to replicate these results in the MSUPRP, which were only significant in the Commercial population.

Table 2. SNP associations for pork quality traits using Meta-analysis GWA.

Tr[1]	SNP	SSC	Peak[3], Mb	Region[4], Mb	P-val
PRL	ASGA-0070779	15	137.5	133.2 - 145.6	3.83 e-9
pHu	DIAS-0000678	15	134.5	133.2 – 135-8	4.7 e-12
SF	ALGA-0011623	2	6.2	5.8-6.2	1.33 e-10
	DRGA-0015526	15	136.5	136.5	1.03 e-6
CKL	MARC-0036560	5	68.3	68.3	4.59 e-7
	H3GA-0052416	15	135.2	133.1-135.2	8.88 e-16
CIEa*	INRA-0021243	6	21.1	21.1-49.8	7.05 e-7
	ASGA-0057129	13	35.4	35.4	8.54 e-8
CIEb*	DRGA-0006183	5	95.1	95.1	1.84 e-12

[1]Traits:PRL=Purge;pHu=UltimatepH;SF=Shear force;CKL=Cook loss.
[2]Population analyzed: Commercial (C), MSUPRP (MS), MARC (MA).
[3]Position of peak in Megabases (Sus scrofa genome build 10.2).
[4]Significant region in Megabases.

All significant SNP found on SSC15 were located near the region containing the gene *PRKAG3*, which has been shown to affect meat quality. However, this gene is not represented in our SNP set. Therefore, further research is needed in order to test the role of *PRKAG3* in the pork quality traits assessed for the populations in this study.

Comparing results from MA to those obtained in individual GWA, significant SNP for PRL, pHu, SF and CKL on SSC15 were confirmed. MA did not detect significant associations for PRL on SSC4 or for CIE L* on SSC15.

Interestingly, MA allowed identification of four additional genomic regions that were not detected in population-specific GWA (Table 2). These included associations with CKL on SSC5, CIE a* on SSC6 and SSC13, and CIE b* on SSC5. A significant SNP on SSC5 was associated with CKL. An association for this trait had previously been reported on SSC5 by Rohrer et al. (2005), but in a different chromosomal region, flanked by two microsatellite markers located between 7.9Mb and 19.6Mb. We searched for potential candidate genes for this CKL QTL by querying the respective region of the pig genome (www.ensembl.org, accessed on 02/19/2014). Specifically,

the search was done looking in a 2Mb window around the peak detected for CKL (SSC5, 68.3Mb). We also identified genes in this region that exhibited cis-acting expression QTL in a subset of the MSUPRP (Steibel et al. (2014)). Among the 59 annotated genes in the region, several genes function as ion channels which are potential candidate genes for CKL since effects on membrane polarization could impact moisture loss during cooking. From the ion channel genes, the gene *KCNA5* (potassium voltage-gated channel, shaker-related subfamily, member 5) emerges as a strong candidate because it showed allelic specific expression in a subset of the MSUPRP population (Steibel et al. (2014)).

We also observed a significant peak on SSC6 for CIE a*, located 1.4Mb from a peak reported by Ma et al. (2013) for the same trait. However, we found a wider significance region (28Mb in length) than reported by Ma et al. (2013). This can be explained by the high LD observed on this chromosome in our populations. In addition, it should be noted that for this color score across each individual population, peaks were observed on SSC6 but none of them reached the genome-wide significance threshold. Thus, MA increased power of detection by combining results obtained from individual GWA, detecting association between SNP on SSC6 with CIE a*.

An additional significant SNP for CIE a* was detected on SSC13. Ma et al. (2009) had previously reported an association for this trait on SSC13 but in a different region, flanked by two microsatellite markers located between 132.4Mb and 142.8Mb. Finally, a significant SNP for CIE b* was identified on SSC5. For the same trait and chromosome, an associated QTL was reported by Li et al. (2010). In that case, flanking markers were located between 78.5Mb and 83.9Mb, that is, approximately 11.2Mb from the region detected in the present study.

Conclusion

Implementation of population specific GWA identified and confirmed significant regions associated with pork quality traits. In many cases, narrower intervals were found in this study and refined QTL association regions reported previously. Combining these datasets for a meta-analysis GWA allowed identification of additional significant regions for CKL and CIE a* and also, novel QTL for CIE a* (35.4Mb) and for CIE b* (95.1Mb). MA detected associations that did not reach genome-wide significance threshold in population-specific GWA, providing an increase in power of GWA as a result of the MA. Thus, MA-GWA methodology is an attractive alternative to synthesize information from multiple genomic evaluations of independent populations into a single GWA scan.

Acknowledgments

This work was supported by Agriculture and Food Research Initiative Competitive Grant no. 2010-65205-20342 from the USDA National Institute of Food and Agriculture, and by funding from the National Pork Board Grant no. 11–042. Partial funding was also provided by the US Pig Genome Coordinator and Animal Agriculture Initiative of Michigan. Computer resources were provided by the Michigan State University High Performance Computing Center (HPCC). USDA is an equal opportunity provider and employer.

Literature Cited

Aguilar, I., Misztal, I., Johnson, D.L. et al. (2010). *J. Dairy Sci.* 93 :743–752.

Badke, Y.M., Bates, R.O., Ernst, C.W. et al. (2012). *BMC Genomics.*13:24.

Ciobanu, D., Bastiaansen, J., Malek, M. et al. (2001). *Genetics* 159: 1151–1162

Choi, I, Steibel, J.P., Bates, R.O. et al. (2011). *Frontiers in Genetics*. 2,18.

Edwards, D.B., Ernst, C. W., Raney, N.E. et al. (2008). *J Anim Sci*. 86:254-266

Evangelou, E. and Ioannidis, J.P.A. (2013). *Nat Rev. Genet.* 14(6), 379-389.

Gualdrón-Duarte,J.L., Bates,R.O., Ernst,C.W. et al. (2013). *BMC Genetic*s. 14:38

Gualdrón-Duarte,J.L, Cantet,R.J.C., Bates, R.O. et al. (2014). *BMC Bioinformatics* 2014. (*Submitted*)

Hayes, B.J., Bowman, P.J., Chamberlain, A.C. et al. (2009). *Genet. Sel. Evol*. 41:51.

Houlston, R.S., Cheadle, J., Dobbins, S.E. et al. (2010). *Nat Genet* . 42: 973–7.

Li, H.D., Lund, M.S., Christensen, O.F. et al. (2010). *J Anim Sci*. 88:2904-2912.

Ma, J., Ren, J., Guo, Y. et al. (2009). *Animal Genetics*. 40:637-647.

Ma, J., Yang, J., Zhou, L. et al. (2013). *PloS one.* 8 (5): e64047.

Minozzi, G., Williams, J.L., Stella, A. et al. (2012). *PLoS One.* 7(3):e32578.

Nonneman, D., Lindholm-Perry, A.K., Shackelford, S.D. et al. (2011). *J. Anim. Sci*. 89(9):2663-72

Nonneman, D.J., Shackelford, S.D., King, D.A. et al. (2013). *J Anim Sci*. 91 (9): 4043-50.

Ramos, A.M., Crooijmans, R.P., Affara, N.A. et al. (2009). *PLoS ONE*.4, e6524.

Rohrer, G.A, Thallman, R.M. Shackelford, S. et al. (2005). *Animal Genetics*, 37, 17–27.

Steibel, J.P., Reeb, P., Ernst, C.W. et al. (2014). Proc 10th WCGALP. *Submitted.*

Wang, H., Misztal, I., Aguilar, I. et al. (2012). *Genet. Res., Camb*. 94: 73–83.