Genome-wide linkage analysis of gene expression of loin muscle tissue identifies candidate genes in pigs

J.P. Steibel[1,2], R.O. Bates[1], G.J.M. Rosa[3], R.J. Tempelman[1], V.D. Rilington[1], A. Ragavendran[1], N.E. Raney[1], A.M. Ramos[1], F.F. Cardoso[1,4], D.B. Edwards[1] and C.W. Ernst[1]

[1]Department of Animal Science, and [2]Department of Fisheries & Wildlife, Michigan State University, East Lansing, MI, USA; [3] Department of Dairy Science University of Wisconsin, Madison, USA; [4]Embrapa Pecuária Sul, Bagé, RS, Brazil.

Introduction

Integration of transcriptional profiling with genotyping data in segregating populations allows mapping of expression quantitative trait loci or eQTL (Jansen and Nap, 2001, Schadt, et al., 2003). In the past decade, such studies have been conducted in human cell lines and model organisms (de Koning and Haley, 2005) , but in livestock global eQTL analyses are still sorely lacking and most of the literature have concentrated on experimental design and modeling issues (Bueno Filho, et al., 2006, Cardoso, et al., 2008, de Koning, et al., 2007) or on the analysis of selected transcripts (Ponsuksili, et al., 2008). The implementation of eQTL mapping has the potential to uncover gene networks and the genetic control of gene activity, as well as shed light on the genetic architecture of phenotypic variation, through integration with phenotypic QTL (pQTL) results (Kadarmideen, et al., 2006). In this paper we present the first genomewide linkage analysis of global gene expression in a livestock species, using an F2 intercross of two pig breeds.

Material and methods

Genotype data. 510 F2 Duroc X Pietrain pigs together with their parents and grandparents from a resource population at Michigan State University were genotyped for 124 microsatellite markers as described previously (Edwards, et al., 2008b).

Gene expression data. Transcriptional profiling was performed on a subset of 176 F2 pigs using a 70-mer long oligonucleotide microarray containing 20,400 oligos (Pigoligoarray; www.pigoligoarray.org). RNA processing and labeling, and microarray procedures were performed as previously reported (Steibel, et al., 2009). Pigs were selected for transcriptional profiling using a selective phenotyping strategy which consisted in choosing the two extreme males and females for a trait of interest within each litter (Cardoso, et al., 2008).

Model and significance test. The following linear mixed model was fit to normalized log-intensity data on an oligo-by-oligo basis:

$$Y = \mu + Dye + Array + Sex + Litter + growth\_group + c_a a + e$$

where Y is the normalized log-intensity, Dye, Sex and growth_group are fixed effects accounting for systematic variation and Array and Litter are random effects. The additive

QTL coefficient c_a was derived assuming that the parental breeds were fixed for alternative QTL alleles (Haley, et al., 1994). A t-test for the additive QTL effect *a* was performed at each putative QTL position for each expression trait and the p-values were corrected for multiple testing (q-value) across all traits and positions. A preliminary analysis considered p<0.0001 (FDR=0.56) as significance threshold. Candidate eQTL analysis on individual genes used p<0.000005 (FDR=0.1).

Physical localization of oligonucleotides. All oligonucleotides were aligned against the pig genome (Build 9; www.ensembl.org) using the BLAT (Kent, 2002) sequence alignment tool. Up to 3 mismatches were allowed and multiple alignments of the same oligonucleotide sequence were discarded.

Gene networks subject to genetic control. Oligonucleotide annotation was retrieved from our previous work (Steibel, et al., 2009). The corresponding HGNC name and its associated QTL p-value were input into the Ingenuity Pathways Analysis software (Ingenuity Systems, Redwood City, CA, USA) to test for enrichment of functional categories.

Co-localization of eQTL and pQTL. A set of 67 phenotypic traits has been previously analyzed for QTL (pQTL) in this cross (Edwards, et al., 2008a, Edwards, et al., 2008b). Given a particular eQTL region, delimited by a 10 cM interval to each side of the peak, all overlapping pQTL regions were selected (West, et al., 2007).

Results and discussion

Physical location of oligonucleotides. Sequence alignment of 20,400 oligonucleotides allowed determination of positions for 13,611 oligos. The number of oligonucleotide per chromosome varied between 286 (SSC16) and 1,399 (SSC1). Chromosome 12 presented the highest density of oligos per megabase whereas the Chromosomes 11, 16 and X had the lowest densities. Although the physical distribution of oligonucleotides was not uniform, the whole genome was satisfactorily covered by the gene sequences spotted on the microarray.

Significance tests and putative eQTL. Testing for 1,279 putative QTL positions in 20,400 expression traits produced over 26 million p-values that required multiple test correction. A first inspection of the quantile-quantile plots of p-values (Fig. 1a) reveals an excess of smaller p-values compared to what is expected under the null hypothesis. Using a p-value cutoff of P<0.0001, a total of 397 putative eQTL peaks were detected and 253 of those were associated with oligonucleotides with known physical position. Notably, cis-acting eQTL had in general smaller p-values compared to trans-acting eQTL. The global pattern is represented in Figure 1b.

For individual gene analysis, a more stringent significance threshold (p<0.000005) was used. This produced a total of 978 significant tests, corresponding to 62 unique linkage peaks comprising 59 genes with comparative human annotation. The positional analysis of these oligonucleotides indicated that 40 of these 59 eQTL were located in close proximity of the physical location of the oligonucleotide (cis-acting).

Gene networks subject to genetic control. Three gene networks were enriched for differentially expressed genes between alleles of alternative breed origin (Table 1). This suggests that significant eQTL genes exist for loin muscle tissue function in common metabolic pathways. These results suggest that these genes and networks are involved in growth and meat quality in pigs, and that they contribute to the phenotypic variation observed in this population.

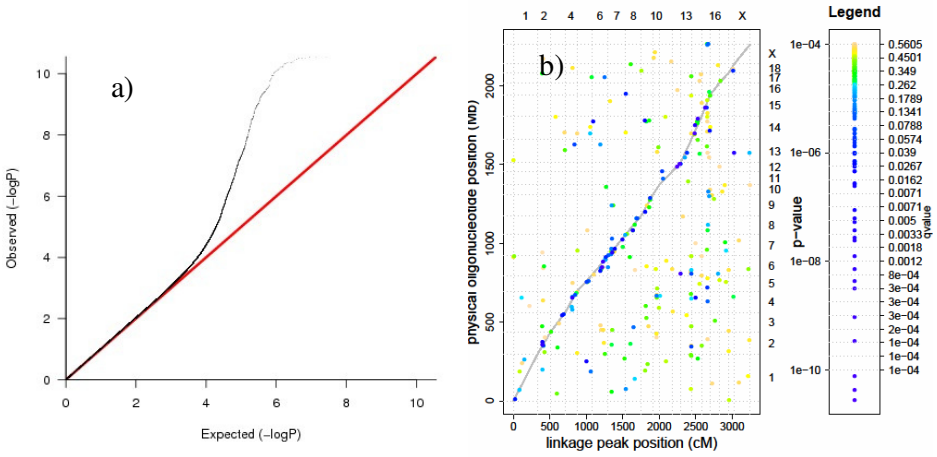

Figure 1: a) Quantile-quantile plot of p-values. There is an excess of small p-values compared to the null model represented by the red line. b) Global plot of physical versus linkage position of eQTL across the pig genome. Points along the gray curve represent likely cis-acting eQTL, while points off the line represent trans-acting eQTL.

Table 1: description of three networks enriched for eQTL

Associated terms	Gene Symbols[a]
Lipid Metabolism, Small Molecule Biochemistry, Post-Translational Modification	RAP2C, AKR7A2, CASP7, DUSP12, CYP4F2, ZFYVE20, CYP4F3, NPO, DYNLT1, RNF167, S100A1, TXNDC12, LSM3, MRPS14, TIMM44, ACTN1, CASP2, CASP4, CCNE1, CTSD, DNAJA3, DYNLL2, HNF4A, PAWR, PSMA1, SART3, SLC2A4, SOCS3, UBE2M
Cell Cycle, Drug Metabolism, Lipid Metabolism	COIL, GSTM4, ZNF24 (includes EG:7572), LRBA, CASP7, ERH, FOLH1, ZBTB5, GSTM1, GPX8, GCSH, AIFM2, NID1, GSTM5, BYSL, CBL, CCDC130, CCNE1, CEP70, IGSF21, MAD1L1,MYO5B, MYO6, PINK1, PSME3, PTPRK, TGFB1, TP53, TSPYL2, ZBTB16
DNA Replication, Recombination, and Repair, Cell Cycle, Cell Death	OAS2, CASP7, EXOSC6, WRN, SSX2IP, RFXANK, LIMCH1, CKMT2, SUFU, MRPL14, KCNS3, CDK2, RNF17, METAP1, BCL2L12, CASP3, CENPC1, CTNNB1, CXCL12, DFFB, E2F4, GAS2, HSH2D, MYC

[a]Gene symbols in red denote genes with significant eQTL and gene symbols in black denote genes in the network with no eQTL detected. Networks and associated terms were determined using the Igenuity Pathways Analysis software (Ingenuity Systems, Redwood City, CA, USA).

Co-localization analysis. Joint analysis of pQTL and eQTL traits revealed 12 common genomic regions (Table 2). Thus, these loci are strong cis-acting candidate genes for the pQTL. This analysis has revealed new gene targets for further validation. For example OCA2 (oculocutaneous albinism II) has been studied extensively for its role in the mammalian pigmentary system (Sturm, 2009). Allelic variants of OCA2 define human blue-

brown eye color, it functions in melanin synthesis within melanocytes, and aberrant OCA2 alleles cause type 2 oculocutaneous albinism in humans. Our results identify a significant OCA2 eQTL coincident with a pQTL for both objective and subjective meat color phenotypes on pig chromosome 15. No previous studies have reported an association of this gene with muscle color. Additional research will be needed to confirm this association and to determine if the cellular mechanism involves tyrosine transport similar to the mechanism in melanocytes.

Table 2: overlapping eQTL and pQTL regions

SSC	cM	eQTL (gene symbols)	pQTL
1	12-21	DYNLT1	Loin Muscle Area
2	89-95	RFXANK	Meat Off Flavor
6	110-119	ETV2	A*
6	142-163	AKR7A2, ZNF24,SSX2IP	Marbling, Fat, Backfat
6	185-194	TXNDC12	Protein, Loin Muscle Area, Chop weight
6	220-221	TMEM	Ultimate meat temperature
7	90-101	MRPL14	Carcass length, Loin Muscle Area
8	87-106	LIMCH1	B*, Moisture, number of ribs
8	152-162	METAP1	Back fat, L*
12	54-69	PNPO, COIL	Fat, Belly, Moisture
15	49-60	WRN	Loin Muscle Area, Tenderness
15	60-80	OCA2, CDK2	L*, A*, Meat color, Protein

Conclusion

To the best of our knowledge, this is the first global linkage analysis of genome-wide gene expression in a livestock species. Our results show that mapping eQTL in this pig resource population can bring further insight on the molecular mechanisms affecting trait variation and generate further candidate genes or genomic regions for fine mapping studies.

References

Bueno Filho, J. S., Gilmour, S. G., Rosa, G. J. (2006). *Genetics*, 174 (2): 945.
Cardoso, F. F. *et al.* (2008). *Genetics*, 180 (3): 1679.
de Koning, D. J., Cabrera, C. P., Haley, C. S. (2007). *Poult Sci*, 86 (7): 1501.
de Koning, D. J., Haley, C. S. (2005). *Trends Genet*, 21 (7): 377.
Edwards, D. B. *et al.* (2008a). *J Anim Sci*, 86 (2): 254.
Edwards, D. B. *et al.* (2008b). *J Anim Sci*, 86 (2): 241.
Haley, C. S., Knott, S. A., Elsen, J. M. (1994). *Genetics*, 136 (3): 1195.
Jansen, R. C., Nap, J. P. (2001). *Trends Genet*, 17 (7): 388.
Kadarmideen, H. N., von Rohr, P., Janss, L. L. (2006). *Mamm Genome*, 17 (6): 548.
Kent, W. J. (2002). *Genome Res*, 12 (4): 656.
Ponsuksili, S. *et al.* (2008). *Bmc Genomics*, 9
Schadt, E. E. *et al.* (2003). *Nature*, 422 (6929): 297.
Steibel, J. P. *et al.* (2009). *Anim Gen*, 40 (6): 883.
Sturm, R. A. (2009). *Human Molecular Genetics*, 18 R9.
West, M. A. L. *et al.* (2007). *Genetics*, 175 (3): 1441.