

ORIGINAL ARTICLE

Expression quantitative trait loci analysis of genes in porcine muscle by quantitative real-time RT-PCR compared to microarray data

S Ponsuksili¹, E Murani², C Phatsara³, M Schwerin¹, K Schellander³ and K Wimmers²

¹Research Group 'Functional Genomics', Research Institute for the Biology of Farm Animals (FBN), Dummerstorf, Germany; ²Research Unit 'Molecular Biology', Research Institute for the Biology of Farm Animals (FBN), Dummerstorf, Germany and ³Animal Breeding and Husbandry Group, Institute of Animal Science, University of Bonn, Bonn, Germany

Genetic analysis of transcriptional profiling is a promising approach for identifying biological pathways and dissecting the genetics of complex traits. Here, we report on expression quantitative trait loci (eQTL) that were estimated from the quantitative real-time RT-PCR data of 276 F₂ animals and compared with eQTL identified using 74 microarrays. In total, 13 genes were selected that showed trait-dependent expression in microarray experiments and exhibited 21 eQTL. Real-time RT-PCR and microarray data revealed seven *cis* eQTL in total, of which one was only detected by real-time RT-PCR, one was only detected by microarray analysis, three were consistently found in overlapping intervals and two were in neighbouring intervals on the same chromosome; whereas no *trans* eQTL was confirmed. We demonstrate that *cis* regulation is a stable characteristic of

individual transcripts. Consequently, a global microarray eQTL analysis of a limited number of samples can be used for exploring functional and regulatory gene networks and scanning for *cis* eQTL, whereas the subsequent analysis of a subset of likely *cis*-regulated genes by real-time RT-PCR in a larger number of samples is relevant to narrow down a QTL region by targeting these positional candidate genes. In fact, when modelling SNPs of six genes as fixed effects in the eQTL analysis, eQTL peaks were shifted downwards, experimentally confirming the impact of the respective polymorphic genes, although these SNPs were not located in the regulatory sequence and these shifts occur as a result of linkage disequilibrium in the F₂ population. *Heredity* (2010) **105**, 309–317; doi:10.1038/hdy.2010.5; published online 10 February 2010

Keywords: eQTL; *cis* eQTL; *trans* eQTL; microarrays; real-time RT-PCR; pig

Introduction

Quantitative differences in gene expression provide a significant source of variation in natural populations, thereby representing an important substrate for evolution and accounting for a considerable fraction of phenotypic differences. The abundance of a gene transcript could be directly affected by polymorphism in regulatory elements. Therefore, the analysis of DNA sequence variations, related to variations in gene expression, results in a better understanding of transcriptional regulation. Variants that affect gene expression (expression quantitative trait loci, eQTL) may have a substantial impact on quantitative traits, whereas qualitative traits, including hereditary disease, may more often depend on structural variation (Farrall, 2004; Yan and Zhou, 2004). A literature survey revealed that of 107 genes with experimentally verified functional *cis* regulatory polymorphisms 82% had proven effects on phenotypes (Rockman and Wray, 2002). In human, more than 100 *cis* regulatory polymorphisms were associated with

phenotypic variation (Wray, 2007). Moreover, in livestock current findings of quantitative trait nucleotides suggest high importance of *cis* regulatory polymorphisms. Transcript levels can be thought of as a quantitative phenotype that can be subjected to statistical genetic analyses in an effort to identify genomic regions that are likely to contain at least one causal gene with regulatory effects on the expression level, termed expression QTL or eQTL (Jansen and Nap, 2001; Schadt *et al.*, 2003; Yvert *et al.*, 2003). Using microarray technology, the abundance of many transcripts can be assessed simultaneously. However, the high cost of microarrays limits the sample size for global genetic mapping of transcript abundance. To overcome this limitation, several methods based on the use of selective phenotyping/profiling strategies, maximum recombination or other selection strategies (Jin *et al.*, 2004; Nettleton and Wang, 2006; Cardoso *et al.*, 2008) have been proposed. Most of these studies were performed using statistical simulations. Many eQTL studies in human or mice were limited in sample size (57–195 individuals) (Schadt *et al.*, 2003, 2005; Morley *et al.*, 2004; Cheung *et al.*, 2005; Stranger *et al.*, 2005), including our previous eQTL study using the microarray expression profiles of 74 F₂ animals. We propose combining eQTL analyses based on microarrays and on real-time RT-PCR in a two-step procedure with increasing numbers of samples to deal with the limitations of microarrays (high costs and limited sensitivity) in eQTL

Correspondence: Dr K Wimmers, Research Unit 'Molecular Biology', Research Institute for the Biology of Farm Animals (FBN), FBN-Dummerstorf, Wilhelm-Stahl-Allee 2, 18196 Dummerstorf, Germany.
E-mail: wimmers@fhn-dummerstorf.de
Received 18 June 2009; revised 23 December 2009; accepted 2 January 2010; published online 10 February 2010

studies. We report experimental data from eQTL analyses obtained by quantitative real-time RT-PCR on 276 F₂ animals and compare this to our previous eQTL study based on microarrays. The relatively large sample size of this study provides the means to assess the relationship between genetic variants and gene expression with improved statistical power. The candidate genes were selected to include genes with *cis*- and *trans*-acting eQTL (according to the microarray expression profiling study) and genes that were differentially expressed in *M. longissimus dorsi* samples that differed in meat waterholding capacity (measured as so-called drip loss) and in pH (Ponsuksili *et al.*, 2008a). We demonstrate that a global microarray eQTL analysis of a limited number of samples can be used for exploring functional and regulatory gene networks and scanning for *cis* eQTL. Subsequent analysis of a subset of likely *cis*-regulated genes by real-time RT-PCR in a larger number of samples serves for verification and narrowing down the QTL regions by targeting these positional candidate genes. Additionally, polymorphisms of *cis*-acting genes (SNPs) were identified and, when used as fixed effects for eQTL analysis, confirmed the impact of the respective polymorphic genes.

Materials and methods

Animals and tissue collection

In this study, a total of 276 animals obtained from the F₂ generation of an experimental cross between the commercial breeds Duroc and Pietrain (DuPi population) were used (Liu *et al.*, 2007, 2008). The grandparental purebred F₀ animals were reciprocally mated, and 32 F₁ animals were used to produce the F₂ generation. Here, the F₂ progeny that were used formed 20 full-sib families.

Microarray analysis

The expression patterns of *M. longissimus dorsi* samples taken between the 13th and 14th ribs of 74 F₂ animals were assessed using Affymetrix GeneChip Porcine Genome Arrays according to the Affymetrix protocols as previously described (Ponsuksili *et al.*, 2008a,b). In brief, 5 µg of total RNA was used to prepare antisense biotinylated RNA target products. Hybridization, washing and scanning were performed using Affymetrix equipment (Hybridization Oven 645, Fluidics Station 450, Scanner 3000 7G; Affymetrix UK Ltd, High Wycombe, UK). The quality of hybridization was assessed in all samples following the manufacturer's recommendations. Data were analysed with Affymetrix GCOS 1.1.1 software using global scaling to a target signal of 500 and were processed with MAS5.0 to generate cell intensity files (present or absent). Quantitative expression levels for the present transcripts were estimated using PLIER (probe logarithmic intensity error) for normalization purposes, as implemented in the Array Assist software (Stratagene Europe, Amsterdam, the Netherlands). The microarray data related to all samples were deposited in the Gene Expression Omnibus (GEO (<http://www.ncbi.nlm.nih.gov/geo/>)) public repository (GEO accession number: GSE10204).

Quantitative real-time RT-PCR

For real-time RT-PCR analysis, 13 genes were selected to represent candidate genes to be further analysed for possible effects on meat quality (Srikanchai *et al.*, 2009). The criteria used to select the genes included their differential gene expression depending on meat quality traits (Ponsuksili *et al.*, 2008b) and the involvement of both *cis*- and *trans*-acting eQTL (Ponsuksili *et al.*, 2008a). Four of the selected genes (VTN, unknown transcript probe set no. SSC.25503.1.S1.at, CYP3A and LYZ) were differentially expressed between high and low drip loss in our previous study (Ponsuksili *et al.*, 2008b). Two genes showed a differential expression between high and low pH at 24 h post mortem (LAMA4, KERA). The chromosomal positions of the genes are shown in Table 1, which were used to distinguish *cis* or *trans*-acting eQTL. The mapping information in the linkage group was accessed through PigQTLdb (<http://www.animalgenome.org/QTLdb/pig.html>). The rest of the genes (AHNAK, SLC3A2, MAP4K4, COQ9, ANKRD13, ZDHHC5, CS) exhibited trait correlated expression and possessed *cis*- or *trans*-acting eQTL according to microarray expression profiling (Ponsuksili *et al.*, 2008a). To normalize the variation in RT-PCR efficiency and initial RNA input, RPL32 and HPRT were used as internal standards. Both of the reference genes were first checked for lack of variation and eQTL, on the basis of microarray data. All 15 transcripts were quantified by real-time RT-PCR using the iCycler apparatus (Bio-Rad Laboratories GmbH, Munich, Germany) and iQ SYBR Green Supermix (Bio-Rad). Real-time RT-PCR was performed in duplicate using 276 animals from full-sib families including 56 individual animals that were previously used for microarray analysis. RNA was isolated as described above. In all, 2 µg of RNA was reverse transcribed to cDNA using SuperScript III MMLV reverse transcriptase (Invitrogen, Karlsruhe, Germany) in a reaction containing 500 ng oligo (dT)₁₁VN primer and 500 ng random hexamer primer according to the manufacturer's protocol. Templates were amplified, using the gene-specific primers (Supplementary Table 1), by 45 cycles of 95 °C for 15 s (denaturation), 60 °C for 30 s (annealing), and 72 °C for 30 s (extension) preceded by an initial denaturation step of 95 °C for 10 min as a universal thermal cycling parameter. On the basis of the analysis of PCR product melting curves, a high-temperature fluorescence acquisition point was estimated and included in the amplification cycle programme. For all assays, a standard curve was generated by amplifying serial dilutions of specific PCR products. After completing the qPCR melting curve analysis, agarose gel electrophoresis was performed to confirm the specificity of the amplification. Normalization of variation in RT-PCR efficiency and initial RNA input was performed using the RPL32 and HPRT genes as internal standards by dividing the calculated mRNA copy numbers by a mean normalization factor derived from the expression of the reference genes (relative abundance).

Linkage analysis

The eQTL estimation was performed using the F2 option of QTLexpress (Seaton *et al.*, 2002). The model included full-sib family and sex as fixed effects and carcass weight (kg) and age at slaughter (days) as covariates. The genotypes of 106 microsatellite markers were analysed.

Table 1 Genes used for RT-PCR and eQTL analysis

Probe set ID	Gene	Chromosome location	Gene name	Selection criteria from microarray data
SSC.16645.1.S1.at	AHNAK	SSC2 25–30 cM	AHNAK nucleoprotein	<i>Cis</i> eQTL, gene located in drip QTL SSC2 ^a
SSC.26213.1.S1.at	ANKRD13	SSC14 30–35 cM	Ankyrin repeat domain 13A	<i>Trans</i> eQTL ^a
SSC.5334.1.S1.at	COQ9	SSC6 45–50 cM	Coenzyme Q9 homolog	<i>Cis</i> eQTL, gene located in drip QTL SSC 6 ^a
SSC.16315.1.S2.at	CS	SSC5 70–75 cM	Citrate synthase	<i>Trans</i> eQTL, gene located in drip QTL SSC5 ^a
SSC.30334.1.A1.at	CYP3A4	SSC 9 80–90 cM	Cytochrome P450, family 3, subfamily A, polypeptide 4	Downregulation in high vs low drip loss ^b
SSC.12936.1.A1.at	KERA	SSC5 130–135 cM	Keratocan	Upregulation in high vs low pH ^b
SSC.24909.1.S1.at	LAMA4	SSC1 40–45 cM	Laminin, alpha 4	Downregulation in high vs low pH ^b
SSC.670.1.S1.at	LYZ	SSC5 75–80 cM	Lysozyme	Upregulation in high vs low drip loss ^b
SSC.3574.1.A1.at	MAP4K4	SSC3 15–45 cM	Mitogen-activated protein kinase kinase kinase 4	<i>Cis</i> eQTL, gene located in drip QTL SSC3 ^a
SSC.2330.1.S1.at	SLC3A2	SSC2 20–25 cM	Solute carrier family 3 member 2	<i>Cis</i> eQTL, gene located in drip QTL SSC2 ^a
SSC.25503.1.S1.at	Unknown	SSC7 110–115 cM	Unknown	<i>Cis</i> eQTL; downregulation in high vs low drip loss ^{a,b}
SSC.309.1.S1.at	VTN	SSC12 90–95 cM	Vitronectin	Downregulation in high vs low drip loss ^b
SSC.3084.1.S1.a.at	ZDHHC5	SSC2 40–60 cM	Zinc-finger, DHHC-type containing 5	<i>Cis</i> eQTL, gene located in drip QTL SSC2 ^a

Abbreviation: eQTL, expression quantitative trait loci.

^aPonsuksili *et al.*, 2008a.

^bPonsuksili *et al.*, 2008b.

The linkage maps, according to Liu *et al.* (2007), were constructed using CRIMAP version 2.4 (Green *et al.*, 1990). The single QTL model analysis with additive and dominance effects was fitted at 1-cM intervals along every single chromosome. Chromosome-wide and genome-wide significance levels were estimated by permutation tests using 5000 permutations (Churchill and Doerge, 1994). The 5% chromosome-wide threshold corresponds approximately to the suggestive linkage threshold proposed by Lander and Kruglyak (1995). Average significance thresholds were 4.94 and 5.20 for relative transcript abundance and microarray expression levels, respectively, corresponding to LOD 2.0 and 1.9. The fraction of the phenotypic variance indicated by the QTL was defined as the percentage reduction of the residual variance resulting from inclusion of the QTL in the model (Knott *et al.*, 1996).

Genotyping and association study

Polymorphisms were identified in the genes AHNAK, SLC3A2, ZDHHC5, MAP4K4, COQ9 and CS and in the unknown transcript SSC.25503.1.S1.at by comparative sequencing of PCR fragments in an SNP discovery panel of commercial breed animals (German Landrace, German Large White, Pietrain, Duroc). The primer sequences and the method of genotyping are shown in Supplementary Table 2. The associations between transcript abundance and their polymorphisms were calculated using family, sex and genotype as fixed effects and carcass weight (kg) and age at slaughter (d) as covariates. Furthermore, the SNPs of these genes were fitted as fixed effects in the eQTL analysis.

Results

eQTL analysis based on quantitative RT-PCR

Expression levels for the 13 genes were measured by quantitative RT-PCR in 276 animals from the F₂ generation of a cross between the commercial breeds Duroc and Pietrain (Liu *et al.*, 2007, 2008). For the quantification of transcripts by real-time RT-PCR, the selected genes were functional candidate genes for meat or carcass quality traits that showed trait-associated expression. Compared to the microarray-based eQTL analysis, a greater sample size was used for real-time RT-PCR to assess the relationship between genetic variation and gene expression with improved statistical power.

Using the relative abundance of transcripts, a total of 16 eQTL for the expression level of 11 genes were mapped with LOD scores of at least 2, that is, the 5% chromosome-wide significance level. There were no eQTL for LYZ and KERA. Plots of F-values obtained for the relative abundance of each transcript across all chromosomes are shown in Figure 1. Five eQTL were mapped to the map position of the corresponding gene (*cis*-acting) with relatively high LOD scores (2.4–15.5) (Table 2), whereas one eQTL was mapped on the same chromosome as the gene itself, although in a more distant position (COQ9). The expression level of ZDHHC5 was mapped with the highest LOD score, having a value of 15.5. The eQTL for the expression level of AHNAK was mapped (LOD score 3.26) in the same vicinity. The eQTL for ZDHHC5 and AHNAK were assigned as *cis*-acting on SSC2. These eQTL explained 26.05 and 6.16% of the expression variation for ZDHHC5

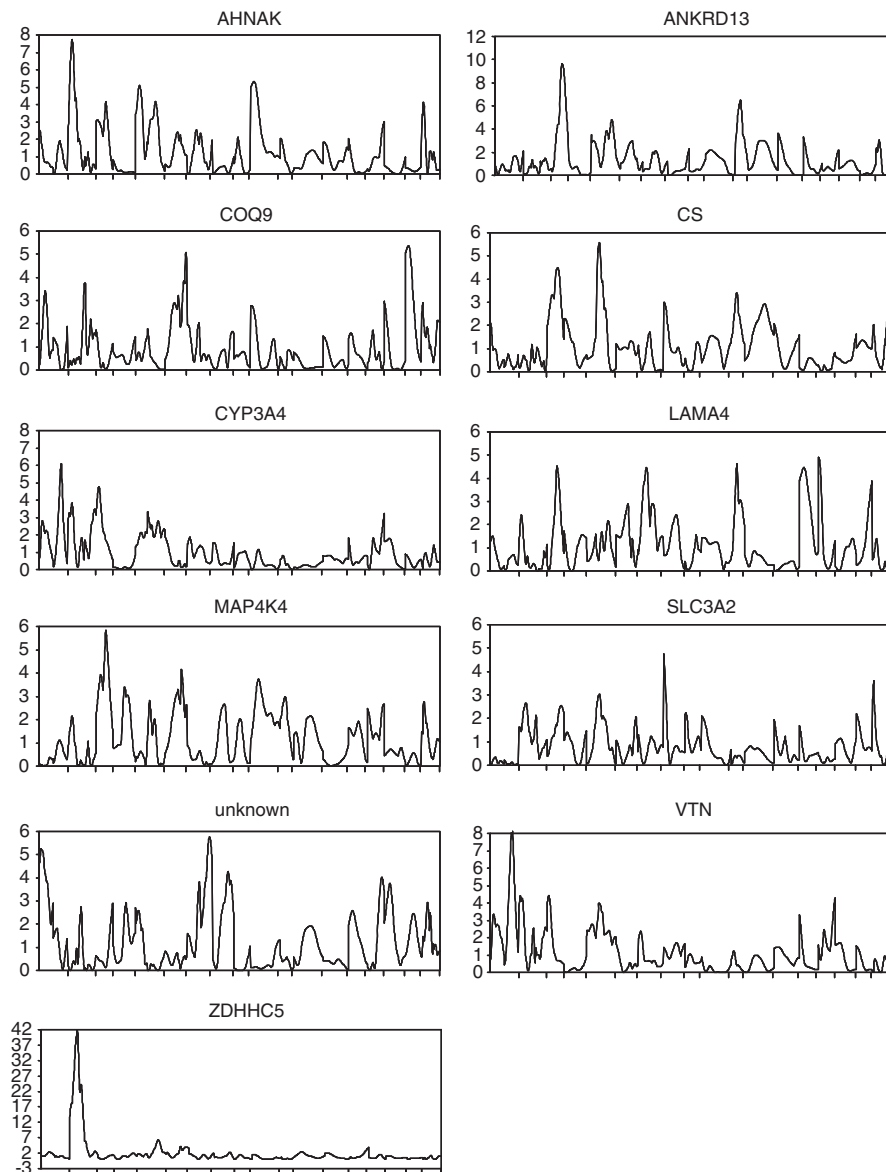


Figure 1 F-ratio obtained for the relative abundance of each transcript, plotted at 1 cM distances along all chromosomes. *x* axis—boundaries of chromosome, SSC1-18; *y* axis—F-ratio from interval mapping.

Table 2 Details of eQTL detected based on relative transcript abundance of 11 genes obtained by real-time RT-PCR ($n = 276$)

Gene	Peak of eQTL on SSC (cM)	Interval (cM)	F-value	LOD score	Variance explained (%)
AHNAK	2 (22)	11–31	7.75	3.26	6.16
	10 (18)	0–32	5.31	2.26	4.31
ANKRD13A	3 (54)	37–79	9.65	4.03	7.48
	11 (28)	12–35	6.50	2.75	5.16
COQ9	6 (106)	103–109	5.06	2.15	4.14
	17 (14)	0–31	5.37	2.28	4.39
CS	5 (73)	65–77	5.67	2.41	4.53
CYP3A4	1 (128)	112–141	6.33	2.68	5.03
	3 (15)	11–13	4.75	2.02	3.82
LAMA4	15 (0)	0–7	4.91	2.09	4.01
MAP4K4	3 (50)	44–61	5.84	2.48	4.66
SLC3A2	8 (0)	0–1	4.75	2.02	3.85
Unknown	7 (113)	104–125	5.77	2.45	4.61
VTN	1 (127)	122–132	9.19	3.84	7.14
ZDHHC5	2 (39)	0–82	41.7	15.5	26.1
	5 (116)	110–123	6.30	2.67	5.05

Abbreviation: eQTL, expression quantitative trait loci.

and AHNAK, respectively. The eQTL for the transcript abundance of MAP4K4 was assigned to its own mapping interval on SSC3 (LOD score 2.48). On SSC5 a *cis*-acting eQTL of CS (LOD score 2.41) was found. The transcript SSC.25503.1.S1.at had a *cis*-acting eQTL located on SSC7 (LOD score 2.45). The eQTL for the expression level of COQ9 was mapped distal on SSC6, whereas COQ9 mapped in the neighbouring marker interval proximal on SSC6.

Consistency in eQTL between microarray data and quantitative RT-PCR

To examine whether the evidence for eQTL was consistent between the different methods, eQTL data from microarray experiments and quantitative RT-PCR were compared. In total, an analysis of the microarray data revealed 21 eQTL from 12 genes with LOD scores that were higher than 1.9, corresponding to 5% chromosome-wide significance. There was no eQTL for LYZ. The percent variance of the expression level given by eQTL analysis when using microarrays was higher than the corresponding quantitative RT-PCR value.

Every *trans*-acting eQTL that was found using the 74 porcine Affymetrix chips was different from the *trans*-acting eQTL found from the 276 quantitative RT-PCR samples (Tables 2, 3). Linkage analysis using either microarray data or real-time RT-PCR data from the 13 genes revealed a total of seven eQTL mapping in close vicinity to the corresponding gene, that is, *cis*-acting eQTL. For AHNAK, ZDHHC5 and SSC.25503.1.S1.at the eQTL F-values reached chromosome-wide significance within overlapping intervals, whereas MAP4K4 and COQ9 consistently showed eQTL on the same chromosome, although in more distant positions. The eQTL for SLC3A2 only reached significance with the analysis using microarray data, yet the shapes of the F-value plots derived from real-time RT-PCR were similar, with a peak at the same position on SSC2. Only real-time RT-PCR

revealed chromosome-wide significant eQTL for CS at its own mapping position.

Effects of SNPs at *cis*-acting eQTL and association with their transcript abundance

All genes with *cis*-acting eQTL or eQTL that were mapped on their own chromosome according to real-time RT-PCR analysis were used to identify SNPs. Owing to a lack of sequence information for the promoter regions, the 5'- and 3'-UTRs, and the coding regions of the genes were used for SNP identification and for further genotyping. The detected SNPs were subsequently used as fixed effects in the eQTL analysis. In total, five SNPs from AHNAK and one SNP each for ZDHHC5, CS, MAP4K4, SSC.25503.1.S1.at, and COQ9 were identified. The eQTL effect for AHNAK transcript abundance remained significant when three of the five SNPs were included, whereas the other two SNPs had a significant influence on the QTL effect. No eQTL for the AHNAK transcript level could be detected if either of these two SNPs or the haplotype of all five SNPs were used as fixed effects in the model to estimate eQTL. The expression level of ZDHHC5, which was mapped to SSC2, was maintained after its own SNP was used as a fixed effect. The SNPs of COQ9, CS and MAP4K4 cause a reduction in the F-values at the eQTL peaks (Figure 2). An association between the haplotype or genotype of AHNAK, COQ9 and CS and their transcript abundances was found with $P < 0.001$, 0.01 and 0.05, respectively.

Discussion

Microarray technology that allows the simultaneous assessment of the expression of all known transcripts throughout the genome has recently become available. However, the high cost of microarrays limits the sample size for global genetic mapping of transcript abundance. Many investigations have reported the

Table 3 Details of eQTL detected based on expression levels of 12 genes obtained by microarray analysis ($n = 74$)

Prob set ID	Gene	Peak of eQTL on SSC (cM)	Interval (cM)	F-value	LOD score	Variance explained (%)
SSC.16645.1.S1.at	AHNAK	2 (18)	0–69	37.6	9.44	63.6
		8 (0)	0–18	7.95	2.94	26.9
SSC.26213.1.S1.at	ANKRD13	2 (3)	0–21	7.05	2.65	24.7
		12 (151)	150–151	5.08	1.98	19.1
SSC.5334.1.S1.at	COQ9	2 (103)	86–107	8.76	3.19	28.9
		4 (78)	50–104	9.32	3.36	30.2
		6 (43)	23–67	7.51	2.80	25.9
		16 (70)	61–75	5.61	2.17	20.7
SSC.16315.1.S2.at	CS	18 (5)	0–9	7.52	2.80	25.9
SSC.30334.1.A1.at	CYP3A4	2 (53)	52–54	6.38	2.43	22.9
SSC.12936.1.A1.at	KERA	17 (31)	12–48	6.96	2.62	24.4
SSC.24909.1.S1.at	LAMA4	3 (0)	0–12	6.77	2.56	23.9
		7 (119)	119–121	5.54	2.14	20.5
SSC.3574.1.A1.at	MAP4K4	3 (0)	0–14	8.34	3.06	27.9
SSC.2330.1.S1.at	SLC3A2	2 (31)	30–32	6.21	2.37	22.4
		14 (22)	5–49	5.75	2.21	21.1
SSC.25503.1.S1.at	Unknown	1 (145)	132–161	8.59	3.14	28.6
		7 (109)	109–112	5.10	1.96	14.2
SSC.309.1.S1.at	VTN	2 (54)	48–58	8.00	2.95	27.1
		11 (29)	21–37	7.24	2.71	25.2
SSC.3084.1.S1.a.at	ZDHHC5	2 (71)	61–95	13.4	4.52	38.4

Abbreviation: eQTL, expression quantitative trait loci.

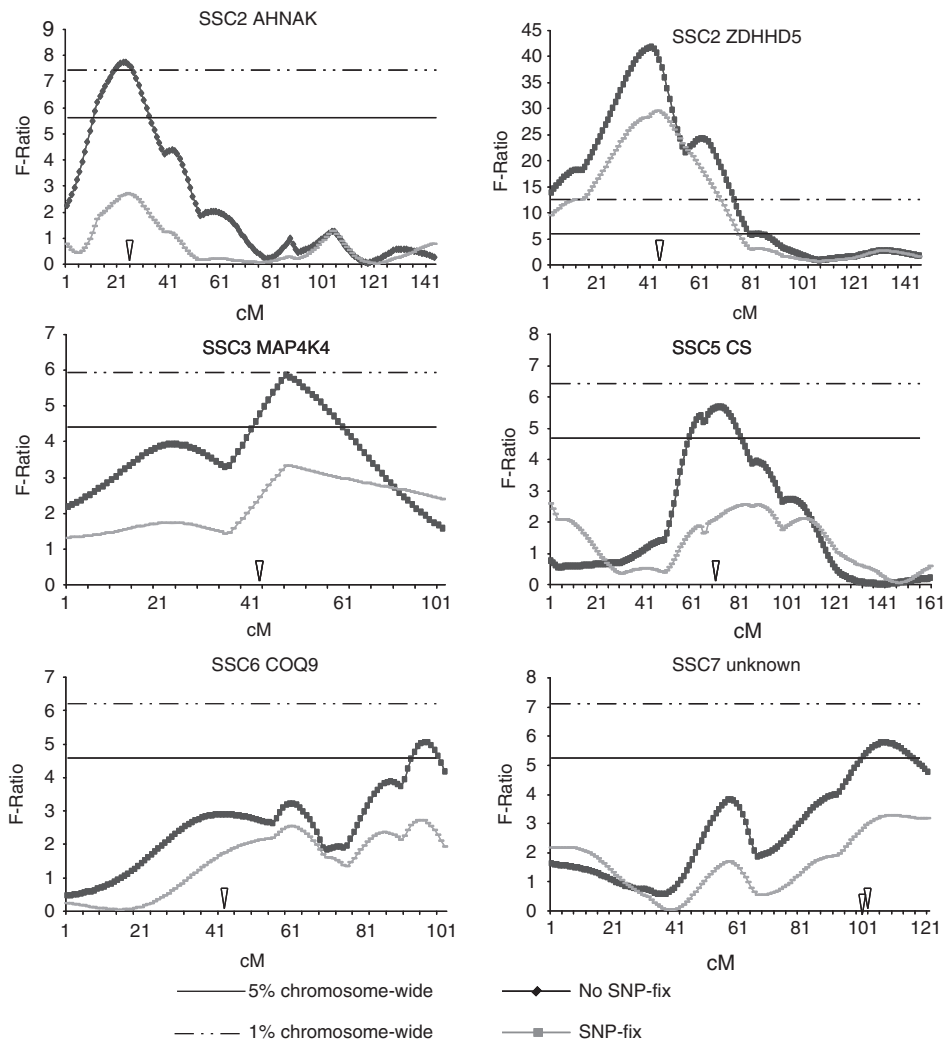


Figure 2 F-ratio plots of eQTL analyses for six genes using (grey) and not using (black) genotypes of the respective genes as fixed effects. x axis—map position in cM; y axis—F-ratio from interval mapping. Physical location of the gene is indicated by an arrowhead.

successful mapping of eQTL in rats, mice or humans with different sample sizes (Morley *et al.*, 2004; Doss *et al.*, 2005; Hubner *et al.*, 2005; MacLaren and Sikela, 2005; Stranger *et al.*, 2005). Mapping of eQTL to the specific gene indicates that *cis* changes are responsible for the different expression levels, whereas mapping positions of eQTL that are different from the positions of the corresponding genes indicate *trans* regulation. These data allow the derivation of regulatory networks for genes. However, to link an eQTL to the genetic background of a classical phenotypic trait of interest, it is necessary to establish a relationship between the variation of that classical phenotypic trait and its corresponding 'pQTL' (phenQTL) position on one hand and either the expression level of that particular transcript or the mapping position of the eQTL on the other hand. Genes that are differentially expressed at variable levels of a trait of interest, where eQTL mapping indicates *cis* regulation for the genes, are more likely to represent the 'cause,' that is, the genetic background of the trait of interest, whereas the differential expression of genes showing *trans* regulation may represent the 'effect,' that is, pathways that are affected by causal

variation. There are numerous publications highlighting the possible merits of 'genetical genomics' applications in livestock breeding (Pomp *et al.*, 2004; Walsh and Henderson, 2004; Kadarmideen *et al.*, 2006; Kadarmideen and Reverter, 2007) and/or addressing issues of data evaluation and experimental design *in silico* (Perez-Enciso *et al.*, 2003, 2004, 2007; De Koning and Haley, 2005; De Koning *et al.*, 2007; Cardoso *et al.*, 2008). Experimental data from genetic investigations of gene expression in farm animals are still limited. For the first time, results from an eQTL study performed with 74 animals of a F₂ pig population, with a focus on drip loss in meat, were presented by our group (Ponsuksili *et al.*, 2008a). Here, we present results from genetic analyses of the transcriptional level in *M. longissimus dorsi* muscle in a much larger sample of individuals using RT-PCR data along with a comparison to microarray data. Although a global microarray analysis allows the exploration of functional and regulatory gene networks reflecting interactions at the level of DNA, RNA or proteins, an analysis of a subset of genes is relevant to narrow down a QTL region through the targeting positional candidate genes.

The expression trait can be inferred from the proximity of the gene physical location to the eQTL. Moreover, the potential to distinguish *cis* and *trans* eQTL, and thus the number of either kind of eQTL, depends on the resolution of the map. Choices of *cis*-window sizes vary across eQTL mapping studies, ranging from 10 kb in yeast (Brem *et al.*, 2002) to 20 Mb in mouse (Bystrykh *et al.*, 2005). In this study, the criterion for a predicted *cis*-acting eQTL was that the interval of microsatellite markers flanking the genes on the genetic maps overlaps with the interval flanked by the cross points of the F-ratio plot for a particular eQTL, with the corresponding chromosome-wide 5% significance level. This relaxed measure was chosen to avoid the exclusion of any true *cis*-acting eQTL that may have been mapped non-precisely because of the low resolution of the genetic map.

The analysis of eQTL for transcripts with expression levels that correlate with the low-heritable trait of drip loss revealed a proportion of *trans* eQTL that is substantially higher than that in other studies of humans and model organisms with no focus on low-heritable traits (Ponsuksili *et al.*, 2008a). This is confirmed here, where only one out of six genes, which were differentially expressed at high vs low drip loss (VTN, SSC.25503.1.S1.at, CYP3A, LYZ) and pH at 24 h post mortem (LAMA4 and KERA), showed a *cis*-acting eQTL (SSC.25503.1.S1.at). This suggests that the low heritability (that is, the high dependency on non-genetic factors) of some traits coincides with the fact that they are under the control of several *trans*-regulated genes, that is, genes that are controlled by other genes. These genes represent members of a regulatory network that might contribute to the flexibility of the organism in regards to expressing low-heritable traits in response to different stimuli. Moreover, it was shown that the expression level of the *trans*-regulated genes was less heritable than that of *cis*-regulated genes (Petretto *et al.*, 2006). Overall, this implies that the expression of genes associated with traits that have low heritability, like drip loss or pH, is largely dependent on *trans* regulation and is low in heritability. As we observed in the previous study, a number of differentially expressed genes with extremely high or low drip loss or pH phenotypes belong to the functional group 'response to stimuli' (Ponsuksili *et al.*, 2008b). The trait-associated expression levels of these genes may be an effect of the external stimuli rather than a cause of phenotypic variation in the target trait. Variations in proteins or RNAs that interact with the *cis*-regulatory sequences of these genes (*trans*-acting effects) may lead to alterations in their expression that are not because of genetic polymorphisms within the regulatory sequences, such as promoters (*cis*-acting). Expression QTL mapping, with its potential to categorize *cis* and *trans* effects, provides the means to discriminate between 'effect' and 'cause' with respect to trait-associated differential expression.

Despite numerous differences between microarray data and quantitative RT-PCR data, the designation of transcripts as *cis* regulated was relatively consistent. This finding suggests that *cis* regulation is a stable characteristic of individual transcripts, which is detectable, independent of the number of animals or environmental influences. In contrast, for *trans* eQTL the analysis of a high number of individuals by real-time RT-PCR and

a more limited number of individual samples by microarrays did not show any common results. The finding may be explained by the presence of type 1 or type 2 errors, which may occur with any of the approaches because of the small number of individuals, limited mapping resolution and/or small effects from the underlying eQTL. Petretto *et al.* (2006) reported that *trans* eQTL generally have effects that are smaller in size and have higher false discovery rates than *cis* eQTL. In our study, estimates of the effects of eQTL detected by microarrays were much larger than those of eQTL derived from real-time RT-PCR. A bias exists, when estimating QTL effects, towards an overestimation of the phenotypic variances associated with the QTL at lower numbers of individuals (Beavis effect) (Beavis, 1998; Xu, 2003). An important stipulation in expression genetics is that the size of the mapping panel is an important determinant of success. The number of animals in a study has a significant role in the identification of *trans* effects, but it has a relatively small effect on the discovery of *cis* eQTL. The observation that *cis* eQTL was more consistently detected than *trans* eQTL was also made by comparing the results from different studies of eQTL with different numbers of animals and different tissue types (Brem *et al.*, 2002; Yvert *et al.*, 2003; Morley *et al.*, 2004; Göring *et al.*, 2007). Göring *et al.* (2007) found that none of their *trans* eQTL could be confirmed between their study and a previous study of Morley *et al.* (2004).

We also attempted to identify SNPs in genes that could give rise to strong *cis* eQTL in a porcine population. Typically, *cis*-acting eQTL are considered to include SNPs that are within 100 kb upstream and downstream of the gene that is affected by that eQTL. On the basis of earlier work involving human genes with promoter polymorphisms, it has been estimated that a sizable fraction of *cis*-variants (about one-third of the SNPs in promoters) may alter gene expression (Hoogendoorn *et al.*, 2003; Buckland, 2004). One study recently reported on the investigation of *cis*-regulatory variations in 3'-UTRs for a set of genes showing *cis*-acting regulation in a panel of mouse recombinant congenic strains (Lee *et al.*, 2006). In farm animals, annotation of the promoter regions for most genes is incomplete. Accordingly, our SNPs are located in the coding region or at the 5'- or 3'-UTR. It is not likely that these SNPs are responsible for differential expressions, that is, causal SNPs. However, in the F₂ population, regions of linkage disequilibrium are often large and thus SNPs may serve as linkage disequilibrium markers for any causal SNP. SNPs, in genes that give rise to strong *cis* eQTL in a F₂ population, have a role in disrupting eQTL when they are used as fixed effects (except for the SNP in ZDHHC5). In our study, one haplotype of the AHNAK gene and two other single SNPs are associated with the variation in their expression in the F₂ population. The identification of causative polymorphisms and the determination of their functional roles are even more challenging because there are several different molecular mechanisms through which mRNA levels in cells can be regulated. Such putative candidate polymorphisms that affect the transcription of a given gene may be prioritized for experimental validation, and hypotheses can be generated for the possible biological roles of candidate regulatory SNPs based on an examination of the data.

In summary, this is to our knowledge the first report to address the relative sensitivity of the detection of *cis* and *trans* eQTL effects, based on the sample size and the methods used to detect expression levels. Sample size has a significant role with regard to statistical power, not only in the detection of eQTL but also in the determination of the kind of eQTL (*cis*, *trans* effects), as shown in this study. *Cis* regulation was more consistent than *trans* eQTL. The expression level of genes associated with low-heritable traits is predominantly controlled by *trans* eQTL.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

We thank Joana Bittner and Annette Jugert for their excellent technical help. This research was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG; FOR 753).

References

- Beavis WD (1998). QTL analyses: power, precision, and accuracy. In: Paterson AH (ed). *Molecular Dissection of Complex Traits*. CRC Press: New York, pp 145–162.
- Brem RB, Yvert G, Clinton R, Kruglyak L (2002). Genetic dissection of transcriptional regulation in budding yeast. *Science* **296**: 752–755.
- Buckland PR (2004). Allele-specific gene expression differences in humans. *Hum Mol Genet* **2**: 255–260.
- Bystrykh L, Weersing E, Dontje B, Sutton S, Pletcher MT, Wiltshire T *et al.* (2005). Uncovering regulatory pathways that affect hematopoietic stem cell function using ‘genetical genomics’. *Nat Genet* **37**: 225–232.
- Cardoso FF, Rosa GJ, Steibel JP, Ernst CW, Bates RO, Tempelman RJ (2008). Selective transcriptional profiling and data analysis strategies for expression quantitative trait loci mapping in outbred F2 populations. *Genetics* **180**: 1679–1690.
- Cheung VG, Spielman RS, Ewens KG, Weber TM, Morley M, Burdick JT (2005). Mapping determinants of human gene expression by regional and genome-wide association. *Nature* **437**: 1365–1369.
- Churchill GA, Doerge RW (1994). Empirical threshold values for quantitative trait mapping. *Genetics* **138**: 963–971.
- De Koning D-J, Haley C (2005). Genetical genomics in human and model organisms. *Trends Genet* **21**: 377–381.
- De Koning D-J, Cabrera CP, Haley CS (2007). Genetical Genomics: combining gene expression with marker genotypes in poultry. *Poult Sci* **86**: 1501–1509.
- Doss S, Schadt EE, Drake TA, Lusis AJ (2005). *Cis*-acting expression quantitative trait loci in mice. *Genome Res* **15**: 681–691.
- Farrall M (2004). Quantitative genetic variation: a post-modern view. *Hum Mol Genet* **13**: R1–R7.
- Göring HH, Curran JE, Johnson MP, Dyer TD, Charlesworth J, Cole SA *et al.* (2007). Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. *Nat Genet* **39**: 1208–1216.
- Green P, Falls K, Crooks S (1990). *Documentation for CRIMAP* version 2.4 Washington University School of Medicine: St. Louis, MO.
- Hoogendoorn B, Coleman SL, Guy CA, Smith K, Bowen T, Buckland PR *et al.* (2003). Functional analysis of human promoter polymorphisms. *Hum Mol Genet* **12**: 2249–2254.
- Hubner N, Wallace CA, Zimdahl H, Petretto E, Schulz H, Maciver F *et al.* (2005). Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. *Nat Genet* **37**: 243–253.
- Jansen RC, Nap JP (2001). Genetical genomics: the added value from segregation. *Trends Genet* **17**: 388–391.
- Jin C, Lan H, Attie AD, Churchill GA, Bulutuglo D, Yandell BS (2004). Selective phenotyping for increased efficiency in genetic mapping studies. *Genetics* **168**: 2285–2293.
- Kadarmideen HN, Reverter A (2007). Combined genetic, genomic and transcriptomic methods in the analysis of animal traits. *CAB Rev Perspect Agric Vet Sci Nutr Nat Resour* **2**: 16.
- Kadarmideen HN, von Rohr P, Janss LLG (2006). From genetical genomics to systems genetics: potential applications in quantitative genomics and animal breeding. *Mamm Genome* **17**: 548–564.
- Knott SA, Elsen JM, Haley CS (1996). Methods for multiple-marker mapping of quantitative trait loci in half-sib populations. *Theor Appl Genet* **93**: 71–80.
- Lander E, Kruglyak L (1995). Genetic dissection of complex traits—guidelines for interpreting and reporting linkage results. *Nat Genet* **11**: 241–247.
- Lee PD, Ge B, Greenwood CM, Sinnett D, Fortin Y, Brunet S *et al.* (2006). Mapping *cis*-acting regulatory variation in recombinant congenic strains. *Physiol Genomics* **25**: 294–302.
- Liu G, Jennen DG, Tholen E, Juengst H, Kleinwächter T, Hölker M *et al.* (2007). A genome scan reveals QTL for growth, fatness, leanness and meat quality in a Duroc-Pietrain resource population. *Anim Genet* **38**: 241–252.
- Liu G, Kim JJ, Jonas E, Wimmers K, Ponsuksili S, Murani E *et al.* (2008). Combined line-cross and half-sib QTL analysis in Duroc-Pietrain population. *Mamm Genome* **19**: 429–438.
- MacLaren EJ, Sikela JM (2005). Cerebellar gene expression profiling and eQTL analysis in inbred mouse strains selected for ethanol sensitivity. *Alcohol Clin Exp Res* **29**: 1568–1579.
- Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, Spielman RS *et al.* (2004). Genetic analysis of genome-wide variation in human gene expression. *Nature* **430**: 743–747.
- Nettleton D, Wang D (2006). Selective transcriptional profiling for trait-based eQTL mapping. *Anim Genet* **37**: 13–17.
- Perez-Enciso M (2004). In silico study of transcriptome genetic variation in outbred populations. *Genetics* **166**: 547–554.
- Perez-Enciso M, Quevedo JR, Bahamonde A (2007). Genetical genomics: use all data. *BMC Genomics* **8**: 69.
- Perez-Enciso M, Toro MA, Tenenhaus M, Gianoly D (2003). Combining gene expression and molecular marker information for mapping complex trait genes: a simulation study. *Genetics* **164**: 1597–1606.
- Petretto E, Mangion J, Dickens NJ, Cook SA, Kumaran MK, Lu H *et al.* (2006). Heritability and tissue specificity of expression quantitative trait loci. *PLoS Genet* **2**: e172.
- Ponsuksili S, Jonas E, Murani E, Phatsara C, Srikanthai T, Walz C *et al.* (2008a). Trait correlated expression combined with expression QTL analysis reveals biological pathways and candidate genes affecting water holding capacity of muscle. *BMC Genomics* **9**: 367.
- Ponsuksili S, Murani E, Phatsara C, Jonas E, Walz C, Schwerin M *et al.* (2008b). Expression profiling of muscle reveals transcripts differentially expressed in muscle that affect water-holding capacity of pork. *J Agric Food Chem* **56**: 10311–10317.
- Pomp D, Allan MF, Wesolowsky SR (2004). Quantitative genomics: exploring the genetic architecture of complex trait predisposition. *J Anim Sci* **82**: 300–312.
- Rockman MV, Wray GA (2002). Abundant raw material for *cis*-regulatory evolution in humans. *Mol Biol Evol* **19**: 1991–2004.
- Schadt EE, Lamb J, Yang X, Zhu J, Edwards S, Guhathakurta D *et al.* (2005). An integrative genomics approach to infer causal associations between gene expression and disease. *Nat Genet* **37**: 710–717.

- Schadt EE, Monks SA, Drake TA, Lusk AJ, Che N, Colinayo V *et al.* (2003). Genetics of gene expression surveyed in maize, mouse and man. *Nature* **422**: 297–302.
- Seaton G, Haley CS, Knott SA, Kearsley M, Visscher PM (2002). QTL Express: mapping quantitative trait loci in simple and complex pedigrees. *Bioinformatics* **18**: 339–340.
- Srikanchai T, Murani E, Wimmers K, Ponsuksili S (2009). Four loci differentially expressed in muscle tissue depending on water-holding capacity are associated with meat quality in commercial pig herds. *Mol Biol Rep* **37**: 595–601.
- Stranger BE, Forrest MS, Clark AG, Minichiello MJ, Deutsch S, Lyle R. *et al.* (2005). Genome-wide associations of gene expression variation in humans. *PLoS Genet* **6**: e78.
- Walsh B, Henderson D (2004). Microarrays and beyond: what potential do current and future genomics tools have for breeders? *J Anim Sci* **82**: 292–299.
- Wray GA (2007). The evolutionary significance of *cis*-regulatory mutations. *Nat Rev Genet* **8**: 206–216.
- Xu S (2003). Theoretical basis of the Beavis effect. *Genetics* **165**: 2259–2268.
- Yan H, Zhou W (2004). Allelic variations in gene expression. *Curr Opin Oncol* **16**: 39–43.
- Yvert G, Brem RB, Whittle J, Akey JM, Foss E, Smith EN *et al.* (2003). Trans-acting regulatory variation in *Saccharomyces cerevisiae* and the role of transcription factors. *Nat Genet* **35**: 57–64.

Supplementary Information accompanies the paper on Heredity website (<http://www.nature.com/hdy>)