Gene Co-expression Analysis to Characterize Genes Related to Marbling Trait in Hanwoo (Korean) Cattle

Dajeong Lim[1,2], Seung-Hwan Lee[1], Nam-Kuk Kim[3], Yong-Min Cho[1], Han-Ha Chai[1], Hwan-Hoo Seong[1] and Heebal Kim[2,*]

[1] National Institute of Animal Science, RDA, Suwon, Korea
[2] Department of Food and Animal Biotechnology, Seoul National University, Seoul, Korea
[3] National Agricultural products Quality management Service(NAQS), Seoul, Korea

ABSTRACT: Marbling (intramuscular fat) is an important trait that affects meat quality and is a casual factor determining the price of beef in the Korean beef market. It is a complex trait and has many biological pathways related to muscle and fat. There is a need to identify functional modules or genes related to marbling traits and investigate their relationships through a weighted gene co-expression network analysis based on the system level. Therefore, we investigated the co-expression relationships of genes related to the 'marbling score' trait and systemically analyzed the network topology in Hanwoo (Korean cattle). As a result, we determined 3 modules (gene groups) that showed statistically significant results for marbling score. In particular, one module (denoted as red) has a statistically significant result for marbling score (p = 0.008) and intramuscular fat (p = 0.02) and water capacity (p = 0.006). From functional enrichment and relationship analysis of the red module, the pathway hub genes (IL6, CHRNE, RB1, INHBA and NPPA) have a direct interaction relationship and share the biological functions related to fat or muscle, such as adipogenesis or muscle growth. This is the first gene network study with *m.logissimus* in Hanwoo to observe co-expression patterns in divergent marbling phenotypes. It may provide insights into the functional mechanisms of the marbling trait. (**Key Words:** Gene Co-expression Network, Marbling, Hanwoo)

INTRODUCTION

The meat quality of cattle is determined by intramuscular fat deposition (marbling) (Lee et al., 2007) and could be improved by functional genomic studies of genetic factors. Beef is graded according to the amount of marbling since marbling makes beef more tender, flavorful, and juicy. It is one of the main factors used to determine beef quality grade in the United States (USDA, 1989), Japan (JMGA, 1988), and Korea. Several countries identify meat quality challenges, such as marbling, meat tenderness, carcass weight, muscling, and fat cover. All of these areas must be considered to provide consumers with high-quality products. In particular, marbling refers to the appearance of white flecks or streaks of adipose tissue between the bundles of muscle fibres in bovine skeletal muscle (Harper et al., 2001). It is driven through the development of adiposes in combination with declinging muscle growth (Hocquette et al., 2010). From one point of view, marbling

might be an interaction among fat development, connective tissue or blood vessels. Kokta et al. (2004) reported the inreaction between myogenic cells and adipocytes to determine the rate and extent of myogenesis and adipogenesis during animal growth. Three genes were identified as being significantly correlated with bovine skeletal muscle based on microarray data from a gene network (Reverter et al., 2006). Jiang et al. (2009) reported that the genetic network was associated with 19 economically important beef traits. Recently, candidate genes for the marbling trait and their relationships were identified from the protein-protein interaction networks (Lim et al., 2011). Kim et al. (2011) identified the relationship between the expression of heat shock protein β1 (HSPB1) and its regulator genes from the gene network analysis in intramuscular fat of Hanwoo (Kim et al., 2011). These results reflect the fact that many biological pathways or interactions occur between muscle and fat within the skeletal muscle. Therefore, the study of marbling differences needs to analyze the complex interactions between biological pathways or genes from the network level.

Gene expression data have been used to successfully identify relationships between genes involved in biological mechanisms and to predict targetable genetic components associated with complex traits or disease states. Several studies have also shown that mRNA levels of candidate genes are heritable, affecting genetic analysis (Brem et al., 2002; Wayne and McIntyre, 2002; Schadt et al., 2003). Many complex traits in animals, such as disease susceptibility, development, and agricultural product quality, are controlled by interactions among several genes combined with environmental influences. Furthermore, patterns of covariation in the expression of multiple loci can be used to build networks that show relationships between genes and functional traits. These networks provide information on the genetic control of complex traits and can help identify causal genes that affect gene function, rather than gene expression (Haley et al., 2006). System-oriented approaches using gene expression data have been applied by animal geneticists to investigate livestock traits (Nobis et al., 2003; Donaldson et al., 2005; Smith et al., 2007), resulting in the identification and characterization of economically important causal *trans*-acting genes within QTL regions. These *trans*-acting regions share a common biological function (e.g., similar gene ontology function, metabolic pathway, transcriptional co-regulation) (Schadt et al., 2003; Gibson and Weir, 2005; Subramanian et al., 2005).

A weighted gene co-expression network is a gene correlation network created from expression profiling, with each gene having several neighbors (Peter and Steve, 2008). Gene co-expression network (GCN) is useful for identifying genes that control quantitative phenotypes and has been used as a "primary screen", to identify novel genes related to traits from thousands of possible genes. Gene expression networks serve as an effective approach for finding hub genes that have key regulatory roles. Fuller et al. (2007) demonstrated that two types of gene co-expression network analysis can find a body-weight-related gene from weighted gene co-expression network analysis (WGCNA). WGCNA analysis is applied in several research fields such as diseases (Ghazalpour et al., 2006; Miller et al., 2008), complex traits (Ghazalpour et al., 2006) and specific tissues (Oldham et al., 2006; Dewey et al., 2011).

In this study, we reported the gene co-expression network analysis of marbling trait-related genes in *m. longissimus* with divergent marbling phenotypes, and suggest evidences for the biological significance of highly connected genes in Hanwoo (Korean cattle).

MATERIALS AND METHODS

Microarray data processing

We used microarray experiments from intramuscular muscle samples of Korean Cattle (Hanwoo) in our previous study, related to the beef marbling study (Lee et al., 2010).

Briefly, ten steers each from a low-marbled group (7.4±2.4%) and a high-marbled group (23.7±5.6%) were used in this study (Table 1). All arrays were processed to determine the robust multiarray average (RMA) (Irizarry et al., 2003) using the "affy" software package (Gautier et al., 2004). Expression values were computed in detail from raw CEL files by applying the RMA model of probe-specific correction for perfect-match probes. These corrected probe values were then subjected to quantile normalization, and a median polish was applied to compute one expression measure from all probe values. Resulting RMA expression values were log2-transformed.

Weighted gene co-expression network analysis

We selected the 4,000 most varying probes for the generation of a weighted gene co-expression network. We calculated correlations between the gene expression profiles of each pair of genes using Pearson's correlation coefficients (denoted as *r*). Then, the correlation measures were transformed into a connection strength using power adjacency function. The power adjacency function $c = |cor(x_i, x_j)|^\beta$ was used to construct a weighted network as the connection strength between two genes. The weighted network represented "soft" thresholding that weighed each connection as a continuous number [0, 1]. We selected a soft threshold beta (β) = 18 according to scale free topology criterion. A major advantage of weighted networks is that highly robust results are obtained with regard to the choice of the parameter beta (β). A major aim of co-expression network analysis is to determine subsets of nodes (modules) that are tightly connected to each other. To organize genes into modules, we used a module identification method based on a topological overlap dissimilarity measure (Ravasz et al., 2002) in conjunction with a clustering method, which detected biologically meaningful modules. The topological overlap of two nodes refers to their relative interconnectedness. The topological overlap matrix (TOM) $\Omega = [\omega_{ij}]$ provides a similarity measure, which has proven

Table 1. Summary statistics of tissue sample for weighted gene co-expression network analysis

Group	Animal ID	Marbling score	IMF content (%)
Low	509	2	7.11
	537	2	6.02
	554	3	4.88
	670	3	7.36
	691	3	12.04
High	527	7	24.35
	547	7	32.49
	586	7	16.56
	589	7	26.24
	632	7	18.81

useful in biological networks (Ye and Godzik, 2004), where $l_{ij} = \sum_u a_{iu}a$ and $k_i = \sum_u a$ is the node connectivity as follows:

$$\omega_{ij} = \frac{l_{ij} + a_{ij}}{\min(k_i, k_j) + 1 - a_{ij}}$$

In the case of our network, equals the number of nodes to which both i and j are connected. To identify modules, we used TOM-based dissimilarity d_{ij}^w ($d_{ij}^w = 1 - \omega_{ij}$) in a hierarchical cluster analysis. Each module represents a group of genes with similar expression profiles across the samples and the expression profile pattern is distinct from those of other modules.

Connectivity and module membership

A weighted gene co-expression network identified gene modules for biological significance. Because gene modules may correspond to biological pathways, focusing the analysis on modules (and their highly connected intramodular hub genes) amounts to a biologically meaningful data reduction scheme. Highly correlated module genes are represented and summarized by their first principal component (which is referred to as the module eigengene (ME)). The ME isused to define measures of module membership (MM) which quantify how close a gene is to a given module. MM measures allow one to annotate all genes on the array and to screen for disease related intramodular hub genes. We used the intramodular connectivity $K^q(i)$ that is biologically more meaningful than the whole network connectivity (Saris et al., 2009). It is calculated from the sum of connection strengths between a particular gene and all other genes in the module $K^q(i) = \sum_{j \in q, j \neq 1} \left| Cor(x^i, x^j) \right|^\beta$, where q denotes a specific module. We also used the $MM^q(i,)$ which is the correlation of the ME and the gene expression profile. As explained in detail in (Horvath and Dong, 2008), the MM of gene i in module q can be defined $MM^q(i) = Cor(x_i, ME^q)$, where larger absolute values mean greater similarity between a gene x_i and the q-th module eigengene. The statistical significance of MM (denoted as p MM red) is carried out from the correlation test p-value of the WGCNA package. Finally, we can identify genes that have a high significance for marbling score as well as high MM in interesting modules using the gene significance (GS) and MM measures (Peter and Steve, 2008). We first defined a measure of GS that is obtained from the correlation between the gene and the trait. The higher the i-th gene's |GS(i)|, the greater its biological significance. For the i-th genes, we identified GS for marbling score (denoted as GS marbling score) as the absolute value of the Student t-test statistic for

testing differential expression between high- and low-marbled groups. We defined a measure of module significance (denoted as p.MM.red) as the eigengene significance that is the correlation between the ME and the expression profiles.

Functional enrichment analysis

We performed functional enrichment analysis in given modules that were associated with marbling score enrichment in the Gene Ontology or KEGG pathway terms, using the Database for Annotation Visualization and Integrated Discovery (DAVID) tool (http://david.abcc. ncifcrf.gov/). It computes a fisher's exact test p-value. Functional relationships of our genes of interest were used in the Pathway studio program (Stratagene, La Jolla, CA, USA) (Nikitin et al., 2003). We investigated the common regulators and targets of the significant genes in the modules.

RESULTS AND DISCUSSION

Weighted gene co-expression network analysis

We used WGCNA in a first attempt to identify marbling score associated coexpression modules and their key functions. A weighted gene co-expression network was constructed using expression data from the high- and low marbled groups, utilizing the 4,000 most varying transcripts from the 24,128 transcripts present on the array. To find modules of highly correlated genes, we used average linkage hierarchical clustering, which uses the TOM as dissimilarity. We were able to identify 17 distinct modules (except for the "grey" module, which is not grouped into any module) for groups of genes with high topological overlap. Figure 1 shows the co-expression modules ranging in size from 41 (lightcyan) to 1,024 (turquoise) genes. The mean overall connectivity is 24.6, and ranged from 8.83 (midnightblue) to 34.44 (turquoise). Detailed information about all genes and their network properties are calculated (data not shown).

Detection of co-expression modules related to marbling score

The coexpression modules correspond to branches and are color-coded (black, blue, brown, cyan, green, greenyellow, lightcyan, magenta, midnightblue, pink, purple, red, salmon, tan, turquoise, and yellow module). We identify modules that are significantly associated with the measured phenotypic traits. We found that the module significance measures in the three modules (red, tan and lightcyan) were significantly correlated (Supplementary data 1). The red module (referred as MEred) was the most significant (correlation with marbling score $r = 0.77$, correlation p = 0.008) for marbling score. It also showed

(A)

(B)
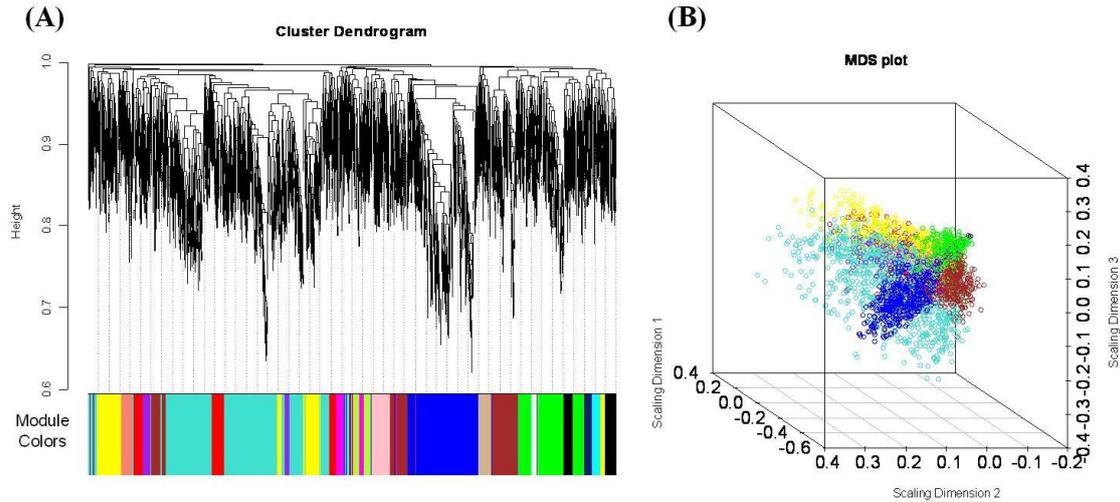
Figure 1. (A) Hierarchical clustering of marbling score related genes and visualization of gene modules. The colored bars (below) directly consistent with the module (color) for the clusters of genes. Distance between genes is shown as height on the y-axis. (B) Multi-dimensional scaling plot of the weighted network. Genes are represented by a dot and colored by module membership. The distance between each gene is indicated by their topological overlap. This representation provides that how the module is related to the rest of the network, and how closely two modules are linked.

significant results for intramuscular fat ($r = 0.72$, p = 0.02) and water capacity ($r = 0.79$, p = 0.006). Figure 2(A) shows red module significance against all traits. The tan module (referred to as MEtan) is significantly associated with three phenotypic traits: marbling score ($r = 0.68$, p = 0.03), intramuscular fat ($r = 0.74$, p = 0.01) and meat color CIE L

Module-trait relationships

Supplementary data 1. The relationship between a module and a trait. Each row corresponds to a module eigengene, column to a trait. Each cell contains the corresponding correlation and p-value. The table in color-coded by correlation according to the color legend.

Lim et al. (2013) Asian-Aust. J. Anim. Sci. 26:19-29

23

(A)

(B)
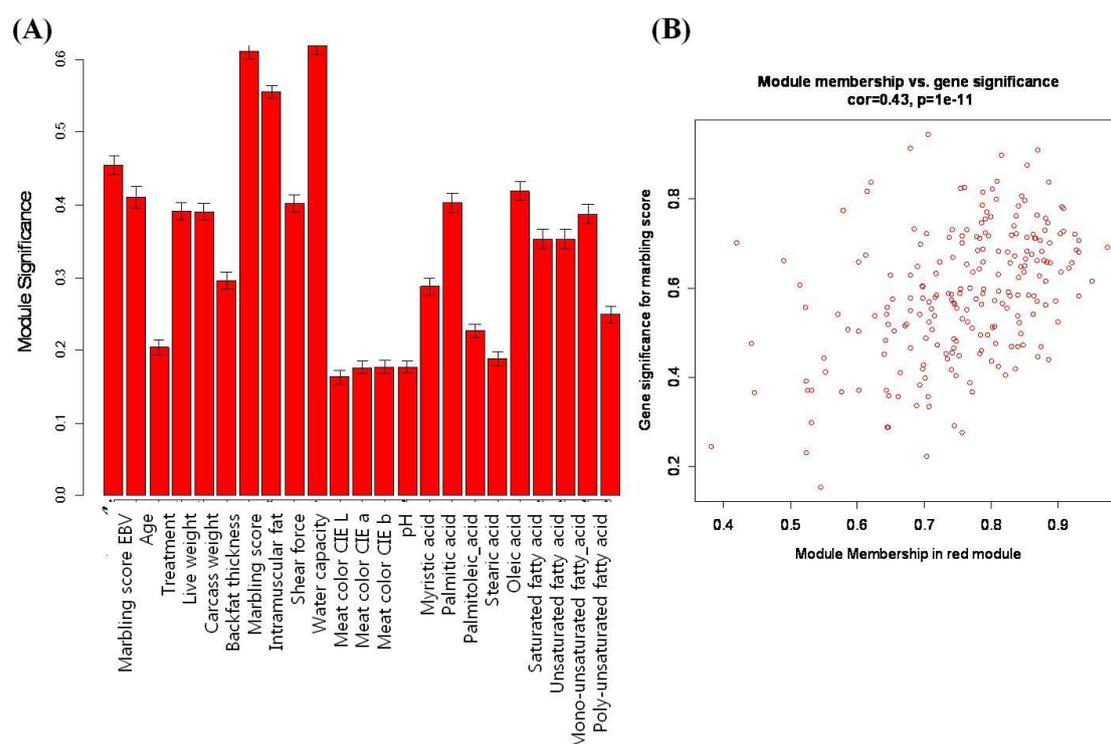
Figure 2. The relationship between the red module and several traits. (A) Module significance is defined as the mean of the absolute value of the correlation coefficient for all genes within a module. (B) The correlation between module membership (MM) and gene significance (GS). There is a highly significant correlation between GS and MM in the red module.

($r = 0.62$, p = 0.05). The red and tan modules were related to marbling score, intramuscular fat. Generally, intramuscular fat is often called an indicator of marbling, because they are highly correlated. The genetic and phenotypic correlations between them were 0.69 to 0.74 and 0.7, respectively (Park et al., 1994; Crews et al., 2003). The lightcyan module is related only to marbling score ($r = 0.66$, p = 0.04). We also investigated the relationship of the MEs to other phenotypic variables. Table 2 shows the modules that have significant p-values against the types of phenotypes.

As detailed in the Methods section, we calculated a measure of MM that can define each module. Large absolute values of MEred(i), MEtan(i) or MElightcyan(i) indicate the gene is closed to the red, tan or lightcyan

module. In contrast, if MMred(i) is closed to 0, then *i*th gene is uncorrelated with the red module eigengene and is unlikely to be part of the red module. We also quantify the association of individual genes with the marbling score trait in each module by determining GS as the absolute value of the correlation between the gene and the trait. Figure 2(B) shows a relationship between the GS and MM in the red module ($r = 0.43$, p = 1e-11). However, there is no significant result ($r = 0.079$, p = 0.44) between GS and MM in the tan module. This implies that hub genes of the red module also tend to be highly correlated with marbling score. We reported 84, 17 and 2 probes that have significant results (p≤0.05) with the GS and the MM against the marbling score in the red, tan and lightcyan module, respectively. Network properties of the top-ranking genes

Table 2. The significant relationship between modules and phenotipyc variables

Module Eigengene	Significant traits (correlation r, p-value)
MEred	Marbling score ($r = 0.77$, p = 0.009), Intramuscular fat ($r = 0.72$, p = 0.02), Water holding capacity ($r = 0.79$, p = 0.006)
MEtan	Marbling score ($r = 0.66$, p = 0.03), Intramuscular fat ($r = 0.74$, p = 0.01), Meat color CIE L ($r = 0.62$, p = 0.05)
MElightcyan	Marbling score ($r = 0.66$, p = 0.04)
MEsamon	Shear force ($r = 0.67$, p = 0.04)
MEyellow	Shear force ($r = 0.69$, p = 0.03)
MEblack	Intramuscular fat ($r = 0.69$, p = 0.03), Meat color CIE b ($r = 0.63$, p = 0.05)
Megreenyellow	Age ($r = 0.75$, p = 0.01)

are shown in Table 3. For example, glomulin, FKBP associated protein (GLMN), showed the most significant result for marbling score (r = 0.95, p.GS.marbling score = 3.69e-5) in the red module. This is involved in differentiation of vascular smooth muscle cells (VSMC) (McIntyre et al., 2004) and indicated as a marker of VSMC. According to Davies et al. (2005), the generation of lipid-filled VSMC resulted from either adipocyte differentiation

Table 3. The significant genes in the red, tan and lightcyan modules for marbling score

Probe	Gene symbol	Gene title	p.GS.marbling score	MMred	p.MM.red	Module
Bt.25404.2.S1_at	GLMN	Glomulin, FKBP associated protein	0.00	0.71	0.02	Red
Bt.22038.1.S1_a_at	RARS	Arginyl-trna synthetase	0.00	-0.68	0.03	
Bt.3670.1.A1_at	-	-	0.00	0.87	0.00	
Bt.6338.1.S1_at	FERMT2	Fermitin family homolog 2 (Drosophila)	0.00	-0.81	0.00	
Bt.1020.1.S1_at	CLK1	CDC-like kinase 1	0.00	-0.85	0.00	
Bt.17136.1.A1_at	-	-	0.00	-0.81	0.00	
Bt.7018.1.S1_at	TXNDC13	Thioredoxin domain containing 13	0.00	0.89	0.00	
Bt.18891.1.S1_at	ERGIC3	ERGIC and golgi 3	0.00	0.76	0.01	
Bt.27173.1.S1_at	C10H15orf44	Chromosome 15 open reading frame 44 ortholog	0.00	-0.75	0.01	
Bt.28784.1.A1_at	-	-	0.00	0.80	0.01	
Bt.5194.3.S1_a_at	WBP1	WW domain binding protein 1	0.00	0.79	0.01	
Bt.26240.1.S1_at	FHIT	Fragile histidine triad gene	0.00	0.84	0.00	
Bt.11239.3.S1_at	SPG7	Spastic paraplegia 7 (pure and complicated autosomal recessive)	0.01	-0.81	0.00	
Bt.6611.1.S1_at	-	-	0.01	-0.85	0.00	
Bt.23995.1.A1_at	STK38L	Serine/threonine kinase 38 like	0.01	0.79	0.01	
Bt.20287.2.S1_at	SHF	Src homology 2 domain containing F	0.01	0.84	0.00	
Bt.19321.1.A1_at	-	-	0.01	0.91	0.00	
Bt.9267.1.A1_at	APOBEC3B	Apolipoprotein B mrna editing enzyme, catalytic polypeptide-like 3B	0.01	-0.91	0.00	
Bt.26711.2.S1_at	LRRC20	Leucine rich repeat containing 20	0.01	-0.79	0.01	
Bt.13637.1.A1_at	SULF2	Sulfatase 2	0.01	-0.87	0.00	
Bt.14036.1.S1_at	PCNT	Pericentrin	0.01	-0.84	0.00	
Bt.24716.1.S1_at	-	-	0.01	0.80	0.01	
Bt.2507.1.S1_at	SFRS10	Splicing factor, arginine/serine-rich 10 (transformer 2 homolog, Drosophila)	0.01	-0.88	0.00	
Bt.20134.1.S1_at	CPN1	Carboxypeptidase N, polypeptide 1	0.01	0.79	0.01	
Bt.21563.2.A1_at	SLC8A3	Solute carrier family 8 (sodium/calcium exchanger), member 3	0.01	0.88	0.00	
Bt.27673.1.A1_at	-	-	0.01	0.78	0.01	
Bt.17725.1.A1_at	-	-	0.01	-0.83	0.00	
Bt.5807.1.S1_at	-	-	0.02	0.68	0.03	Red
Bt.28732.1.S1_s_at	LOC407199	T cell receptor delta chain	0.02	-0.76	0.01	
Bt.28732.1.S1_at	TRD	T-cell receptor delta chain	0.02	-0.74	0.01	
Bt.7484.1.S1_at	PLEKHG2	Pleckstrin homology domain containing, family G (with RhoGef domain) member 2	0.02	0.88	0.00	
	LRG1	Leucine-rich alpha-2-glycoprotein 1	0.02	-0.91	0.00	
Bt.13062.1.A1_at	COL9A1	Collagen, type IX, alpha 1	0.02	0.79	0.01	
Bt.20189.1.S1_at	FTSJD2	FtsJ methyltransferase domain containing 2	0.02	-0.90	0.00	
Bt.27184.1.S1_at	HISPPD2A	Histidine acid phosphatase domain containing 2A	0.02	-0.84	0.00	
Bt.5892.1.S1_at	C6orf25	Chromosome 6 open reading frame 25	0.02	0.70	0.02	
Bt.27974.1.S1_at	NRG1	Neuregulin 1	0.02	0.93	0.00	
Bt.26693.1.S1_at	-	-	0.02	0.86	0.00	
Bt.4189.1.S1_a_at	GHRHR	Growth hormone releasing hormone receptor	0.02	0.83	0.00	
Bt.20361.1.S1_at	FBXL20	F-box and leucine-rich repeat protein 20	0.02	-0.80	0.01	
Bt.29696.1.A1_at	FGFR2	fibroblast growth factor receptor 2	0.02	0.76	0.01	
Bt.16351.1.A1_at	WDR20	WD repeat domain 20	0.02	-0.78	0.01	
Bt.3233.1.A1_at	CIAO1	Cytosolic iron-sulfur protein assembly 1 homolog (S. cerevisiae)	0.02	-0.74	0.02	
Bt.28236.1.A1_at	ATP4A	ATPase, H+/K+ exchanging, alpha polypeptide	0.02	0.85	0.00	
Bt.17742.1.A1_at	-	-	0.02	0.87	0.00	
Bt.20225.1.S1_at	DTNBP1	Dystrobrevin binding protein 1	0.02	-0.86	0.00	
BtAffx.1.9.S1_at	NPPA	Natriuretic peptide precursor A	0.02	0.78	0.01	
Bt.19219.1.S1_at	-	-	0.02	-0.87	0.00	
Bt.5386.1.S1_at	COBRA1	Cofactor of BRCA1	0.02	-0.93	0.00	
Bt.25049.1.S1_at	TRAM1L1	Translocation associated membrane protein 1-like 1	0.02	0.76	0.01	
Bt.13929.2.S1_at	DPH3	DPH3, KTI11 homolog (S. cerevisiae)	0.02	-0.89	0.00	
Bt.25510.1.S1_at	LOC513740	Hypothetical LOC513740	0.02	-0.69	0.03	

**Table 3.** The significant genes in the red, tan and lightcyan modules for marbling score (Continued)

Probe	Gene symbol	Gene title	p.GS.marbling score	MM red	p.MM.red	Module
Bt.12404.1.S1_at	CLPTM1L	CLPTM1-like	0.03	-0.97	0.00	
Bt.17824.2.A1_at	VPS36	Vacuolar protein sorting 36 homolog (S. cerevisiae)	0.03	-0.79	0.01	
Bt.568.1.S1_at	IBSP	Integrin-binding sialoprotein	0.03	0.93	0.00	
Bt.28987.1.S1_at	-	-	0.03	0.78	0.01	
Bt.20942.1.S1_at	-	-	0.03	-0.81	0.00	
Bt.20198.1.S1_at	TUBGCP3	Tubulin, gamma complex associated protein 3	0.03	-0.87	0.00	
Bt.23135.1.S1_at	TAGLN2	Transgelin 2	0.03	0.85	0.00	
Bt.26658.2.S1_at	SLC46A1	Solute carrier family 46 (folate transporter), member 1	0.03	0.93	0.00	
Bt.8592.1.S1_a_at	PABPC1L	Poly(A) binding protein, cytoplasmic 1-like	0.03	0.82	0.00	
Bt.8262.1.A1_at	-	-	0.03	-0.86	0.00	
Bt.28716.2.S1_at	LOC532698	Hypothetical protein LOC532698	0.03	-0.84	0.00	
Bt.27339.1.A1_at	MME	Membrane metallo-endopeptidase	0.03	-0.74	0.01	
Bt.18789.2.A1_at	ATF7	Activating transcription factor 7	0.04	0.85	0.00	
Bt.11542.1.A1_at	-	-	0.04	0.87	0.00	
Bt.286.1.S1_at	CACNA1B	Calcium channel, voltage-dependent, N type, alpha 1B subunit	0.04	0.88	0.00	
Bt.18809.1.A1_at	SLC22A23	Solute carrier family 22, member 23	0.04	0.79	0.01	Red
Bt.21688.1.S1_at	LOC100196901	Hypothetical LOC100196901	0.04	-0.72	0.02	
Bt.6348.2.S1_at	DENND1A	DENN/MADD domain containing 1A	0.04	0.88	0.00	
Bt.25454.1.A1_at	-	-	0.04	0.79	0.01	
Bt.26290.2.S1_a_at	IPO4	Importin 4	0.04	0.92	0.00	
Bt.27284.1.S1_at	EIF4H	Eukaryotic translation initiation factor 4H	0.04	-0.89	0.00	
Bt.24979.1.S1_at	CD1E	CD1e molecule	0.04	-0.83	0.00	
Bt.9785.1.S1_at	-	-	0.04	-0.76	0.01	
Bt.20768.1.S1_at	LOC529859	Similar to KIAA1632	0.04	-0.85	0.00	
Bt.26228.1.A1_at	-	-	0.04	0.84	0.00	
Bt.27244.1.A1_at	---	-	0.04	0.69	0.03	
Bt.9562.1.S1_at	SCN5A	Sodium channel, voltage-gated, type V, alpha subunit	0.04	0.92	0.00	
Bt.16757.1.S1_at	DCP1A	DCP1 decapping enzyme homolog A (S. cerevisiae)	0.04	0.78	0.01	
Bt.28733.1.S1_at	ZNF397	Zinc finger protein 397	0.05	-0.90	0.00	
Bt.12288.1.S1_at	NPBWR1	Neuropeptides B/W receptor 1	0.05	0.77	0.01	
Bt.20833.1.S1_at	NHLRC2	NHL repeat containing 2	0.05	0.85	0.00	
Bt.13608.1.A1_at	-	-	0.05	0.72	0.02	
Bt.18127.1.A1_at	WDR87	WD repeat domain 87	0.05	0.76	0.01	
Bt.4220.1.S1_at	PDPR	Pyruvate dehydrogenase phosphatase regulatory subunit	0.05	0.85	0.00	
Bt.28106.1.S1_at	-	-	0.04	-0.96	1.59E-05	
Bt.13948.1.S1_at	-	-	0.04	0.71	0.02	
Bt.4250.2.S1_at	MAP4	Microtubule-associated protein 4	0.04	-0.70	0.03	
Bt.15740.1.A1_at	TPD52L1	Tumor protein D52-like 1	0.036	0.76	0.01	
Bt.2132.1.A1_at	MCPH1	Microcephalin 1	0.03	-0.89	0.00	
Bt.19567.2.S1_at	-	-	0.03	-0.65	0.04	
Bt.11916.1.S1_at	LOC615412	Similar to BAI1-associated protein 2-like 1	0.03	0.75	0.01	
Bt.524.1.S1_at	IL12A	Interleukin 12A (natural killer cell stimulatory factor 1, cytotoxic lymphocyte maturation factor 1, p35)	0.03	0.84	0.00	Tan
Bt.5561.1.S1_at	MC4R	Melanocortin 4 receptor	0.02	0.88	0.00	
Bt.12854.1.S1_at	-	-	0.02	-0.80	0.01	
Bt.16977.1.A1_at	-	-	0.02	0.85	0.00	
Bt.15432.1.A1_at	-	-	0.02	0.76	0.01	
Bt.11730.1.A1_at	-	-	0.02	0.90	0.00	
Bt.26030.1.A1_at	-	-	0.00	0.74	0.01	
Bt.11794.1.S1_at	HIGD1A	HIG1 domain family, member 1A	0.00	0.84	0.00	
Bt.7131.2.S1_at	PLDN	Pallidin homolog (mouse)	0.00	0.64	0.05	Lightcyan
Bt.27873.1.S1_at	-	-	0.00	-0.65	0.04	

or direct promotion of lipogenesis as the result of LXR/SREBP1c activation in humans (Davies et al., 2005). Neuregulin 1 (NRG1), integrin-binding sialoprotein (IBSP) and solute carrier family 46, member 1 (SLC46A1) have the largest module membership (MM = 0.93) in the red module. These genes also show significant p-values for gene significance against the marbling score phenotype.

Pathway and GO analysis for the red module

We performed functional enrichment analysis for the red module according to the GS and MM measurement. GO and biological pathway analysis were used to search for the biological significance or functional relationship of the significant genes associated with marbling score. We explored the functional relationship (expression, regulation and direct interaction) in the red module using the pathway studio program. Out of 15 pathway annotated genes, 8 muscle-related genes (NRG1, RB1, JUN, CHRNE, CXCL10, IL6, SRF and FGFR2) have a direct relationship in the pathway analysis (Figure 3). These genes have significant p-value (p<0.05) for MM or GS in red module for marbling score. The NRG family have been observed to stimulate myotube formation and muscle specific gene expression (Florini et al., 1996; Lebrasseur et al., 2003) and facilitate glucose uptake that is an important factor for improving marbling in adipocyte of beef cattle (Suarez et al., 2001). Activation of NRG/ErbB signaling may also mediate one or more adaptive growth and metabolic responses of skeletal muscle to exercise. Fibroblast growth factor

receptor 2 (FGFR2) is a member of four transmembrane tyrosine kinase receptors and affects skeletal muscle myogenesis (Rhoads et al., 2009). The function of satellite cells during muscle regeneration is regulated by many growth factors and cytokines such as fibroblast growth factor (FGF) and transforming growth factor-β (TGF-β) families, insulin-like growth factors-1 and -2 (IGF-1, IGF-2), hepatocyte growth factor (HGF), and interleukin-6 (IL-6) (Grefte et al., 2007). One of the FGF family, polymorphisms in the FGF8 is associated with carcass quality, growth and feed efficiency in beef cattle (Moore and Marques, 2008). Interleukine 6 (IL6) regulates skeletal muscle differentiation and metabolism. In particular, it increased glucose incorporation into glycogen, glucose uptake, lactate production, and fatty acid uptake and oxidation in humans (Al-Khalili et al., 2006). Retinoblastoma 1 (RB) plays an important role in determining whether myoblasts proliferate or differentiate (Rosenthal and Cheng, 1995). RB family proteins promote adipogenesis by direct interaction with C/EBPs (Chen et al., 1996). Chemokine ligand 10 (CXCL10) is differentially expressed in the *longissimus* tissues from Meishan, Meishan×Large White cross and Large White pigs (Li et al., 2010). Serum response factor (SRF) was shown to be differentially expressed between fat and lean and between different muscles using RT-PCR in chickens and was suggested as a potential regulator of several functional candidates affecting glycogen turnover in the muscle for meat quality (Sibut et al., 2011). Jun oncogene (JUN) is
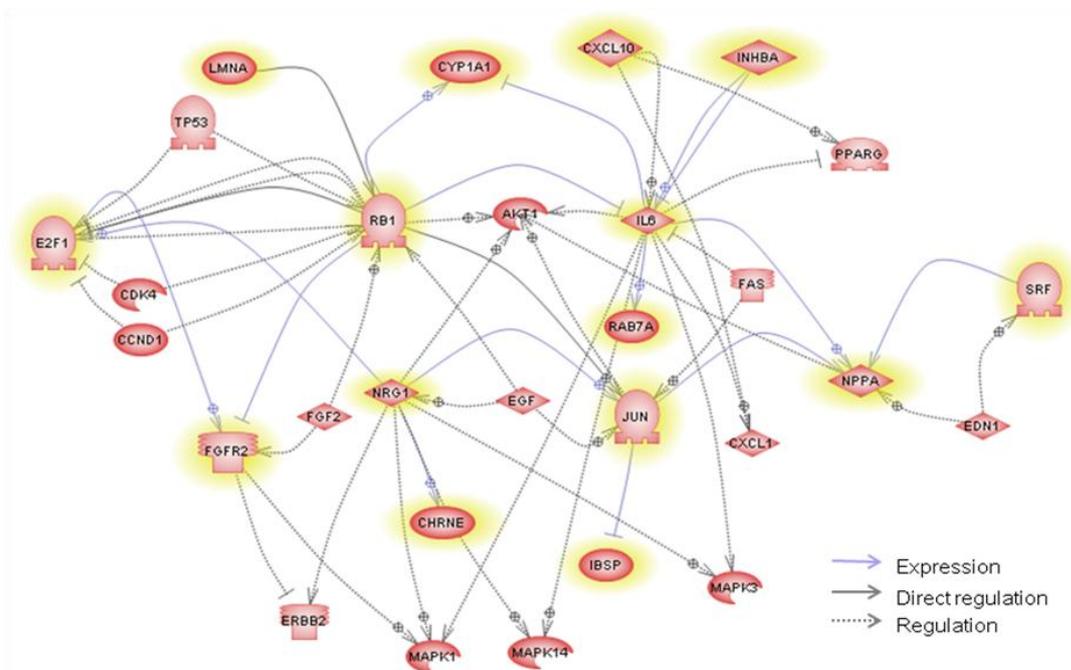
Figure 3. Pathway analysis generated by Pathway studio applied to genes in the red module related to marbling score. The network contains common regulators as well as common targets for the group of direct-interacted genes in order to examine their possible roles. The yellow-highlighted denotes genes in the red module.

Table 4. Gene Ontology terms overrepresented in the red module related to marbling score by DAVID tool

Category	GO terms	p-value	Gene symbols
Biological process	Transmembrane transport (GO:0055085)	0.012	CACNA1B,CNGB1,SCN5A,SLC8A3,KCNH1,SLC46A1, ATP4A,RANBP2,TAP2,NUP85
	Regulation of biological quality (GO:0065008)	0.041	DTNBP1,SLC9A3R1,CACNA1B,NPPA,IL6,TXNDC13,PCSK2,CHRNE,COL9A1,INHBA,APTX,RB1
	Salivary gland morphogenesis (GO:0007435)	0.047	IL6,FGFR2
Molecular function	Protein binding (GO:0005515)	0.029	NPPA,FTSJD2,NUP85,LRRC20,CIAO1,SCN5A,PITPNA, CD3G,IBSP,RANBP2,VAPB,TAGLN2,ATF7,FBXL20,APTX,RB1,FABP5,BAIAP2,LMNA,KCNH1,C22H3orf60,WBP1,JUN,TAF1B,JARID1C,CUL7,CXCL10,INHBA,MYLIP,DTNBP1,VPS45,MGC148992,VPS36,HIT,VPS26A,LRG1,SHF,DPH3,RAB7A,PCSK2,IL6,NPBWR1,FGFR2,CNPY3,TCHP
	RNA binding (GO:0003723)	0.046	ILF3,TRUB2,RBM18,EIF4H,RPS20,CSTF2,SFRS10,APTX
Cellular component	Phosphoglycolate phosphatase activity (GO:0008967)	0.046	PGP,APTX
	Nuclear envelope (GO:0005635)	0.013	TAGLN2,LMNA,IPO4,RANBP2,NUP85
	Nuclear membrane (GO:0031965)	0.027	TAGLN2,LMNA,NUP85
	Endomembrane system (GO:0012505)	0.035	TAGLN2,LMNA,IPO4,RANBP2,ST3GAL2,VPS26A,NUP85,VAPB
	Nuclear pore (GO:0005643)	0.040	IPO4,RANBP2,NUP85

called an activator protein 1 (AP-1) and is known to inhibit myogenic differentiation (Su et al., 1991). It controls the transcription factor involved in myogenesis and those involved in cell proliferation (Li et al., 1992). AP-1 is also one of the transcription factors binding in the promoter of FABP4 with CEBPα (Shin et al., 2009). Recently, the mutation of cholinergic receptor, nicotinic, epsilon (CHREN) is significantly associated with muscle growth in beef cattle from primer-extension assay (Sevane et al., 2011). These results indicate that the genes in the red module may function in regulating muscle growth or fat-related mechanisms and co-expressed genes with similar functions in the module. In addition, we explored regulatory relationships (i.e., common regulators and targets) between 15 direct-interacted genes using pathway studio. The common targets or regulators are shown in Figure 3 with the direct interaction relationship. We found the common regulators based on an assumption that the genes within a similar biological pathway are controlled by common regulators. E2F1 is one gene of the E2F family and is a candidate to be a transcription factor controlling corticotropin releasing hormone (CRH) for marbling and subcutaneous fat depth in beef cattle (Wibowo et al., 2007). In *longissimus* muscle tissue expression during growth in the porcine, E2F1 also showed a significant relationship with differential expressed genes as a transcription factor with myogenin and PAX3 (D'Andrea et al., 2011). In our network analysis, E2F1 is a member of the red module and regulates FGFR2. The FGF family plays a role in cell growth, such as cell proliferation and angiogenesis. The

FGFR2 protein is induced in the mid-to-late G1 phase of the cell cycle by E2F1 (Tashiro et al., 2003).

Finally, we investigated the functional bias of the significant genes according to GO classification and understood the biological significance of the module genes, and determined the putative pathways using DAVID. Table 4 lists the significant gene ontology terms and the representative genes. Due to the incomplete annotation of the bovine genome, 168 of 222 probe sets were annotated (Table 4). In significant GO terms of Biological processes, the regulation of biological quality (GO:0065008) indicates that the process modulates a measuable attribute of an organism or part of an organism, such as size, mass, shape, color, etc. This result is also reflected in the pathway analysis. The term is included in 5 pathway hub genes (IL6, CHRNE, RB1, INHBA and NPPA) of 12 annotated genes by gene ontology. Collagen, type IX, alpha 1 (COL9A1) is detected in the other genes. Significant associations of the COL9A1 gene with body length, depth and width have previously been reported in pigs. Recently, it has also been related to *logissimus* muscle area from assocation analysis in the pig (Fan et al., 2009). These findings suggest that genes in the red module tend to be highly enriched with meat quality and have a potential role to change or control a specific phenotype for animal production.

CONCLUSION

A major objective of this study was to construct the gene co-expression network and then to find hub modules

or genes associated with the marbling score. Therefore, we attempted to find coexpression patterns associated with marbling in Hanwoo (Korean cattle) by the WGCNA method. As a result, three large co-expression modules were significantly associated with marbling score and intramuscular fat. Among these three modules, we focused on the red module for functional enrichment analysis. This is because the tan and lightcyan modules have not shown a significant correlation between gene significance and module membership in each module. Through the pathway and gene ontology analysis, we consistently observed that hub genes within the red module were predominantly a co-expression group having biological pathways related to skeletal muscle. We noticed overlapping genes from the analysis, and five genes (IL6, CHRNE, RB1, INHBA and NPPA) belonged to a red module. These genes are shared in skeletal muscle related biological pathways that might represent a phenomenon occuring in muscle with highly divergent marbling phenotype as key drives. Our results do not point to a single biological pathway or candidate gene like a standard differential expression analysis. Instead, we find several highly significant biological pathways and patterns of co-expressed genes as key drivers in the marbling score related modules. These results will provide valuable information for the additional biological study of meat quality in Hanwoo (Korean cattle).

REFERENCES

Al-Khalili, L., K. Bouzakri, S. Glund, F. Lonnqvist, H. A. Koistinen and A. Krook. 2006. Signaling specificity of interleukin-6 action on glucose and lipid metabolism in skeletal muscle. Mol. Endocrinol. 20:3364-3375.

Brem, R. B., G. Yvert, R. Clinton and L. Kruglyak. 2002. Genetic dissection of transcriptional regulation in budding yeast. Science 296:752-755.

Chen, P. L., D. J. Riley, Y. Chen and W. H. Lee. 1996. Retinoblastoma protein positively regulates terminal adipocyte differentiation through direct interaction with C/EBPs. Genes Dev. 10:2794-2804.

Crews Jr, D., E. Pollak, R. Weaber, R. Quaas and R. Lipsey. 2003. Genetic parameters for carcass traits and their live animal indicators in Simmental cattle. J. Anim. Sci. 81:1427-1433.

D'Andrea, M., S. Dal Monego, A. Pallavicini, M. Modonut, R. Dreos, B. Stefanon and F. Pilla. 2011. Muscle transcriptome profiling in divergent phenotype swine breeds during growth using microarray and RT PCR tools. Anim. Genet. 42:501-509

Davies, J. D., K. L. Carpenter, I. R. Challis, N. L. Figg, R. McNair, D. Proudfoot, P. L. Weissberg and C. M. Shanahan. 2005. Adipocytic differentiation and liver x receptor pathways regulate the accumulation of triacylglycerols in human vascular smooth muscle cells. J. Biol. Chem. 280:3911-3919.

Dewey, F. E., M. V. Perez, M. T. Wheeler, C. Watt, J. Spin, P. Langfelder, S. Horvath, S. Hannenhalli, T. P. Cappola and E. A. Ashley. 2011. Gene coexpression network topology of cardiac development, hypertrophy, and failure clinical perspective. Circ. Cardiovasc. Genet. 4:26-35.

Donaldson, L., T. Vuocolo, C. Gray, Y. Strandberg, A. Reverter, S. McWilliam, Y. Wang, K. Byrne and R. Tellam. 2005. Construction and validation of a bovine innate immune microarray. BMC Genomics 6:135.

Fan, B., S. Onteru, M. Nikkila, K. Stalder and M. Rothschild. 2009. The COL9A1 gene is associated with longissimus dorsi muscle area in the pig. Anim. Genet. 40:788.

Florini, J. R., D. S. Samuel, D. Z. Ewton, C. Kirk and R. M. Sklar. 1996. Stimulation of myogenic differentiation by a neuregulin, glial growth factor 2. Are neuregulins the long-sought muscle trophic factors secreted by nerves? J. Biol. Chem. 271:12699-12702.

Fuller, T. F., A. Ghazalpour, J. E. Aten, T. A. Drake, A. J. Lusis and S. Horvath. 2007. Weighted gene coexpression network analysis strategies applied to mouse weight. Mamm. Genome 18:463-472.

Gautier, L., L. Cope, B. Bolstad and R. Irizarry. 2004. Affy-analysis of Affymetrix GeneChip data at the probe level. Bioinformatics 20:307-315.

Ghazalpour, A., S. Doss, B. Zhang, S. Wang, C. Plaisier, R. Castellanos, A. Brozell, E. E. Schadt, T. A. Drake and A. J. Lusis. 2006. Integrating genetic and network analysis to characterize genes related to mouse weight. PLoS Genet. 2: e130.

Gibson, G. and B. Weir. 2005. The quantitative genetics of transcription. Trends Genet. 21:616-623.

Grefte, S., A. M. Kuijpers-Jagtman, R. Torensma and J. W. Von den Hoff. 2007. Skeletal muscle development and regeneration. Stem Cells Dev. 16:857-868.

Haley, C. and D. J. de Koning. 2006. Genetical genomics in livestock: potentials and pitfalls. Anim. Genet. 37(Suppl 1):10-12.

Harper, G., D. Pethick, V. Oddy, R. Tume, W. Barendse and L. Hygate. 2001. Biological determinants of intramuscular fat deposition in beef cattle: current mechanistic knowledge and sources of variation. Meat Livest. Australia, Sydney.

Hocquette, J., F. Gondret, E. Baeza, F. Medale, C. Jurie and D. Pethick. 2010. Intramuscular fat content in meat-producing animals: development, genetic and nutritional control, and identification of putative markers. Animal 4:303-319.

Horvath, S. and J. Dong. 2008. Geometric interpretation of gene coexpression network analysis. PLoS Comput. Biol. 4: e1000117.

Irizarry, R. A., B. M. Bolstad, F. Collin, L. M. Cope, B. Hobbs and T. P. Speed. 2003. Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res. 31:e15.

Jiang, Z., J. J. Michal, J. Chen, T. F. Daniels, T. Kunej, M. D. Garcia, C. T. Gaskins, J. R. Busboom, L. J. Alexander and R. W. Wright. 2009. Discovery of novel genetic networks associated with 19 economically important traits in beef cattle. Int. J. Biol. Sci. 5:528-542.

Kim, N. K., D. Lim, S. H. Lee, Y. M. Cho, E. W. Park, C. S. Lee, B. S. Shin, T. H. Kim and D. Yoon. 2011. Heat shock protein B1 and its regulator genes are negatively correlated with intramuscular fat content in the Longissimus thoracis muscle of Hanwoo (Korean Cattle) steers. J. Agric. Food Chem. 25:5657-5664.

Kokta, T., M. Dodson, A. Gertler and R. Hill. 2004. Intercellular signaling between adipose tissue and muscle tissue. Domest. Anim. Endocrinol. 27:303-331.

Lebrasseur, N. K., G. M. Cote, T. A. Miller, R. A. Fielding and D. B. Sawyer. 2003. Regulation of neuregulin/ErbB signaling by contractile activity in skeletal muscle. Am. J. Physiol. Cell Physiol. 284:C1149-1155.

Lee, S. H., C. Gondro, J. van der Werf, N. K. Kim, D. Lim, E. W. Park, S. J. Oh, J. Gibson and J. Thompson. 2010. Use of a bovine genome array to identify new biological pathways for beef marbling in Hanwoo (Korean Cattle). BMC Genomics 11: 623.

Lee, S. H., E. W. Park, Y. M. Cho, S. K. Kim, J. H. Lee, J. T. Jeon, C. S. Lee, S. K. Im, S. J. Oh and J. M. Thompson. 2007. Identification of differentially expressed genes related to intramuscular fat development in the early and late fattening stages of hanwoo steers. J. Biochem. Mol. Biol. 40:757-764.

Li, L., J. C. Chambard, M. Karin and E. N. Olson. 1992. Fos and Jun repress transcriptional activation by myogenin and MyoD: the amino terminus of Jun can mediate repression. Genes Dev. 6:676-689.

Li, Y., Z. Xu, H. Li, Y. Xiong and B. Zuo. 2010. Differential transcriptional analysis between red and white skeletal muscle of Chinese Meishan pigs. Int. J. Biol. Sci. 6:350-360.

Lim, D., N. K. Kim, H. S. Park, S. H. Lee, Y. M. Cho, S. J. Oh, T. H. Kim and H. Kim. 2011. Identification of candidate genes related to bovine marbling using protein-protein interaction networks. Int. J. Biol. Sci. 7:992-1002.

McIntyre, B. A. S., P. Brouillard, V. Aerts, I. Gutierrez-Roelens and M. Vikkula. 2004. Glomulin is predominantly expressed in vascular smooth muscle cells in the embryonic and adult mouse. Gene Expr. Patterns 4:351-358.

Miller, J. A., M. C. Oldham and D. H. Geschwind. 2008. A systems level analysis of transcriptional changes in Alzheimer's disease and normal aging. J. Neurosci. 28:1410-1420.

Moore, S. S. and E. F. Marques. 2008. Associations of polymorphisms in the fibroblast growth factor 8 (FGF8) and its haplotypes with carcass quality, growth and feed efficiency in beef cattle, Google Patents.

Nikitin, A., S. Egorov, N. Daraselia and I. Mazo. 2003. Pathway studio-the analysis and navigation of molecular networks. Bioinformatics 19:2155-2157.

Nobis, W., X. Ren, S. P. Suchyta, T. R. Suchyta, A. J. Zanella and P. M. Coussens. 2003. Development of a porcine brain cDNA library, EST database, and microarray resource. Physiol. Genomics 16:153-159.

Oldham, M. C., S. Horvath and D. H. Geschwind. 2006. Conservation and evolution of gene coexpression networks in human and chimpanzee brains. Proc. Natl. Acad. Sci. 103: 17973-17978.

Park, B., A. D. Whittaker, R. K. Miller and D. S. Hale. 1994. Predicting intramuscular fat in beef longissimus muscle from speed of sound. J. Anim. Sci. 72:109-116.

Peter, L. and H. Steve. 2008. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9:599

Ravasz, E., A. Somera, D. Mongru, Z. Oltvai and A. Barabasi. 2002. Hierarchical organization of modularity in metabolic networks. Science 297:1551-1555.

Reverter, A., N. Hudson, Y. Wang, S. Tan, W. Barris, K. Byrne, S. McWilliam, C. Bottema, A. Kister and P. Greenwood. 2006. A gene coexpression network for bovine skeletal muscle inferred from microarray data. Physiol. Genomics 28:76-83.

Rhoads, R. P., M. E. Fernyhough, X. Liu, D. C. McFarland, S. G. Velleman, G. J. Hausman and M. V. Dodson. 2009. Extrinsic regulation of domestic animal-derived myogenic satellite cells II. Domest. Anim. Endocrinol. 36:111-126.

Rosenthal, S. M. and Z. Q. Cheng. 1995. Opposing early and late effects of insulin-like growth factor I on differentiation and the cell cycle regulatory retinoblastoma protein in skeletal myoblasts. Proc. Natl. Acad. Sci. 92:10307-10311.

Saris, C. G., S. Horvath, P. W. van Vught, M. A. van Es, H. M. Blauw, T. F. Fuller, P. Langfelder, J. DeYoung, J. H. Wokke, J. H. Veldink. 2009. Weighted gene co-expression network analysis of the peripheral blood from Amyotrophic Lateral Sclerosis patients. BMC Genomics 10:405.

Schadt, E. E., S. A. Monks, T. A. Drake, A. J. Lusis, N. Che, V. Colinayo, T. G. Ruff, S. B. Milligan, J. R. Lamb and G. Cavet. 2003. Genetics of gene expression surveyed in maize, mouse and man. Nature 422:297-302.

Sevane, N., I. Crespo, J. Canon and S. Dunner. 2011. A Primer-Extension Assay for simultaneous use in cattle Genotype Assisted Selection, parentage and traceability analysis. Livest. Sci. 137:141-150.

Shin, J., B. Li, M. E. Davis, Y. Suh and K. Lee. 2009. Comparative analysis of fatty acid-binding protein 4 promoters: conservation of peroxisome proliferator-activated receptor binding sites. J. Anim. Sci. 87:3923-3934.

Sibut, V., C. Hennequet-Antier, E. Le Bihan-Duval, S. Marthey, M. J. Duclos and C. Berri. 2011. Identification of differentially expressed genes in chickens differing in muscle glycogen content and meat quality. BMC Genomics 12:112.

Smith, G. W. and G. J. Rosa. 2007. Interpretation of microarray data: trudging out of the abyss towards elucidation of biological significance. J. Anim. Sci. 85(13 Suppl):E20-E23.

Su, H. Y., T. J. Bos, F. S. Monteclaro and P. K. Vogt. 1991. Jun inhibits myogenic differentiation. Oncogene 6:1759-1766.

Suarez, E., D. Bach, J. Cadefau, M. Palacin, A. Zorzano and A. Guma. 2001. A novel role of neuregulin in skeletal muscle. Neuregulin stimulates glucose uptake, glucose transporter translocation, and transporter expression in muscle cells. J. Biol. Chem. 276:18257-18264.

Subramanian, A., P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub and E. S. Lander. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. 102:15545-15550.

Tashiro, E., Y. Minato, H. Maruki, M. Asagiri and M. Imoto. 2003. Regulation of FGF receptor-2 expression by transcription factor E2F-1. Oncogene 22:5630-5635.

Wayne, M. L. and L. M. McIntyre. 2002. Combining mapping and arraying: An approach to candidate gene identification. Proc. Natl. Acad. Sci. USA. 99:14903-14906.

Wibowo, T. A., J. J. Michal and Z. Jiang. 2007. Corticotropin releasing hormone is a promising candidate gene for marbling and subcutaneous fat depth in beef cattle. Genome 50:939-945.

Ye, Y. and A. Godzik. 2004. Comparative analysis of protein domain organization. Genome Res. 14:343-353.