

RESEARCH ARTICLE

Open Access

# From SNP co-association to RNA co-expression: Novel insights into gene networks for intramuscular fatty acid composition in porcine

Yulixias Ramayo-Caldas<sup>1,2\*</sup>, Maria Ballester<sup>1,2</sup>, Marina RS Fortes<sup>3</sup>, Anna Esteve-Codina<sup>2</sup>, Anna Castelló<sup>1,2</sup>, Jose L Noguera<sup>4</sup>, Ana I Fernández<sup>5</sup>, Miguel Pérez-Enciso<sup>1,6</sup>, Antonio Reverter<sup>7</sup> and Josep M Folch<sup>1,2\*</sup>

## Abstract

**Background:** Fatty acids (FA) play a critical role in energy homeostasis and metabolic diseases; in the context of livestock species, their profile also impacts on meat quality for healthy human consumption. Molecular pathways controlling lipid metabolism are highly interconnected and are not fully understood. Elucidating these molecular processes will aid technological development towards improvement of pork meat quality and increased knowledge of FA metabolism, underpinning metabolic diseases in humans.

**Results:** The results from genome-wide association studies (GWAS) across 15 phenotypes were subjected to an Association Weight Matrix (AWM) approach to predict a network of 1,096 genes related to intramuscular FA composition in pigs. To identify the key regulators of FA metabolism, we focused on the minimal set of transcription factors (TF) that explored the majority of the network topology. Pathway and network analyses pointed towards a trio of TF as key regulators of FA metabolism: *NCOA2*, *FHL2* and *EP300*. Promoter sequence analyses confirmed that these TF have binding sites for some well-known regulators of lipid and carbohydrate metabolism. For the first time in a non-model species, some of the co-associations observed at the genetic level were validated through co-expression at the transcriptomic level based on real-time PCR of 40 genes in adipose tissue, and a further 55 genes in liver. In particular, liver expression of *NCOA2* and *EP300* differed between pig breeds (Iberian and Landrace) extreme in terms of fat deposition. Highly clustered co-expression networks in both liver and adipose tissues were observed. *EP300* and *NCOA2* showed centrality parameters above average in the both networks. Over all genes, co-expression analyses confirmed 28.9% of the AWM predicted gene-gene interactions in liver and 33.0% in adipose tissue. The magnitude of this validation varied across genes, with up to 60.8% of the connections of *NCOA2* in adipose tissue being validated via co-expression.

**Conclusions:** Our results recapitulate the known transcriptional regulation of FA metabolism, predict gene interactions that can be experimentally validated, and suggest that genetic variants mapped to *EP300*, *FHL2*, and *NCOA2* modulate lipid metabolism and control energy homeostasis in pigs.

**Keywords:** Pig, Fatty acid, Gene network, Transcription factor, Co-association, Co-expression

\* Correspondence: yulixias@gmail.com; JosepMaria.Folch@uab.cat

<sup>1</sup>Centre de Recerca en Agrigenòmica (CRAG), Consorci CSIC-IRTA-UAB-UB, Campus UAB, Bellaterra 08193, Spain

<sup>2</sup>Departament de Ciència Animal i dels Aliments, Facultat de Veterinària, Universitat Autònoma de Barcelona, Bellaterra 08193, Spain

Full list of author information is available at the end of the article

## Background

Fatty acids (FA) are a major energy source and important constituents of cell membranes, playing a relevant role as cellular signaling molecules in various metabolic pathways, including metabolic diseases [1]. Environmental and genetic effects determining FA composition in pigs have been the subject of many studies. Supporting a genetic influence on FA composition moderate to high heritability estimates have been reported [2,3]. However, the molecular process controlling FA composition and metabolism is far from being fully understood. Technological, nutritional and organoleptic properties of pork meat quality are highly dependent on lipid content and FA composition [4-6]. Thus, elucidating this molecular process could aid improve meat quality for healthy human consumption and increase knowledge of FA metabolism, underpinning metabolic diseases. Pigs are important models for metabolic diseases such as obesity, type II diabetes (T2D) and atherosclerosis [7-10].

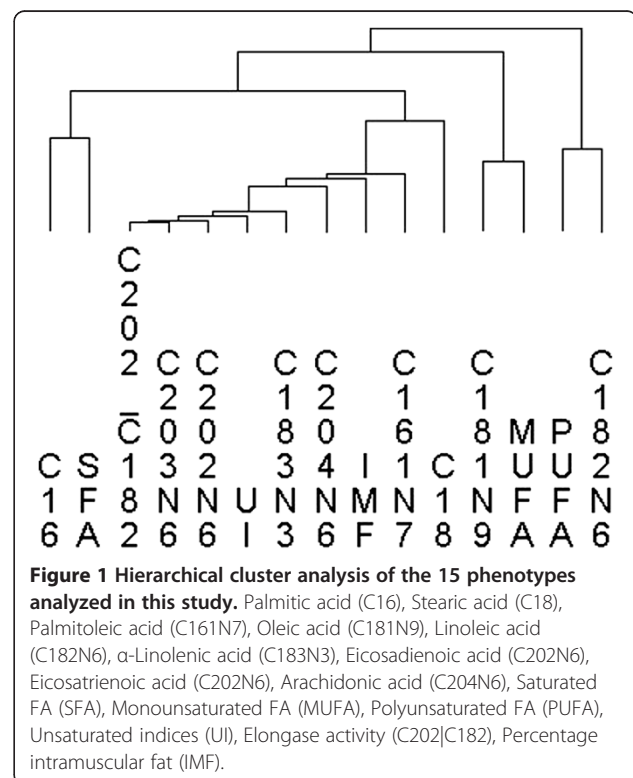
Molecular pathways controlling lipid metabolism are highly interconnected. Also, they interact with other related pathways, such as carbohydrate metabolism and energy homeostasis pathways. Together, these pathways and its interactions constitute an essential metabolic network for homeostatic control and normal organism development [11]. In this context, a system biology approach focused on the connections and functional interactions between genes that underpin these metabolic pathways is an attractive alternative to the classical “single-gene-single-trait” approach found in most genome-wide association studies (GWAS) using single nucleotide polymorphisms (SNP).

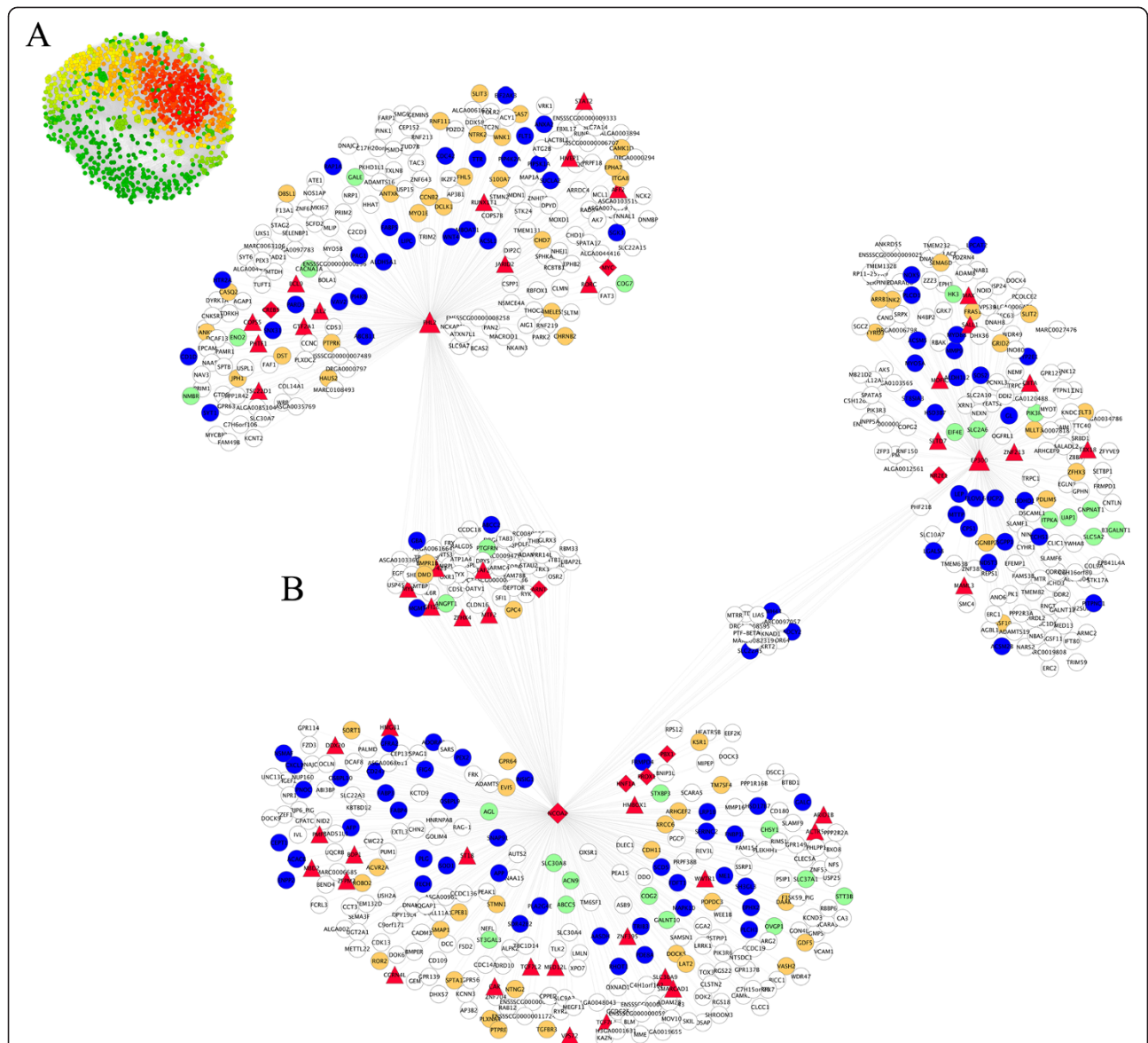
The main goal of this study was to employ a previously described system biology approach termed Association Weight Matrix (AWM) [12] and, based on a SNP-to-SNP co-association evidence, infer a gene network for intramuscular (IMF) FA composition in pigs. This multi-trait approach was applied to data from 15 phenotypes related to FA composition and metabolism from an Iberian x Landrace intercross. Iberian pigs are a local Mediterranean breed extreme for obesity and appetite [13], whereas Landrace is a lean international breed. The analysis of the predicted gene network revealed key transcription factors that are network hubs and would be critical to determining meat quality, FA composition and controlling energy homeostasis. Finally, we experimentally validated some of the AWM network predictions using real-time PCR gene co-expression analyses in adipose and liver tissues.

## Results

Genotyping data from 48,119 SNPs in 144 backcross pigs (25% Iberian x 75% Landrace) was employed for GWAS of fatty acid related traits in the *Longissimus dorsi* muscle. For all 15 phenotypes, estimated SNP

additive effects were standardized (z-scores) by subtracting the mean and dividing by the phenotype-specific standard deviation. After applying a series of selection criteria (see Methods), a total of 1,096 SNPs were retained to build the AWM matrix. Correlations between phenotypes were calculated using AWM columns (standardized SNP effects across traits) and were visualized as a hierarchical tree cluster, in which strong positive and negative correlations are displayed as proximity and distance, respectively (Figure 1). The observed cluster distribution is in concordance with the physiological similarities and relationships among FA. Hence, palmitic acid with saturated FA (SFA), oleic with monounsaturated FA (MUFA), and linoleic with polyunsaturated FA (PUFA) cluster together (Figure 1). Linoleic acid and PUFA are clearly differentiated from other FAs. This result can be explained by the inability of mammals to synthesize linoleic and  $\alpha$ -linoleic FAs, which must be provided by the diet. Gene interactions were predicted using pair-wise correlation analysis of the SNP effects across pair-wise rows of the AWM. Hence, the AWM predicted gene interactions based on significant co-association between SNPs. In the network, every node represents a gene (or SNP), whereas every edge connecting two nodes represents a significant interaction. In total, 111,198 significant edges (or 18.5% of all the possible edges) between the 1,096 nodes were identified as significant by the PCIT algorithm [14] (Figure 2A). For every node we computed the total number of connections





**Figure 2** Co-association network based on the AWM approach. (A) Entire network with 1,096 nodes (i.e., genes or SNPs) and 111,198 interactions. The color spectrum ranges from green to red for low and high density, respectively. (B) Subset of the network showing the best trio of transcription factors: *NCOA2*, *EP300* and *FHL2*. Node color corresponding with the functional classification of the *in-silico* predicted target gene as follows: TF (red), lipid metabolism process (blue), carbohydrate metabolisms (green), development process (orange) and finally, white nodes represent genes with others functional classification. Node shape indicates classification as: diamond (TF involved in lipid metabolism), triangle (TF), ellipse (other genes).

based on significant interactions. Table 1 lists the ten most connected nodes and Additional file 1: Table S1 their positional concordance with fat-related QTL deposited in the Pig QTL Database.

Gene ontology (GO) and pathway enrichment analyses were performed to gain insight into the predicted gene network. Overrepresented GO terms in the network included: “Cellular component organization” ( $P = 4.02 \times 10^{-6}$ ,  $FDR = 3.95 \times 10^{-2}$ ), “Cellular component organization or biogenesis” ( $P = 7.34 \times 10^{-6}$ ,  $FDR = 3.6 \times 10^{-2}$ ),

“Cell projection morphogenesis” ( $P = 9.59 \times 10^{-5}$ ,  $FDR = 9.42 \times 10^{-2}$ ), “Fatty acid metabolic process” ( $P = 5.89 \times 10^{-4}$ ,  $FDR = 1.03 \times 10^{-2}$ ), “Glycerolipid metabolic process” ( $P = 1.2371 \times 10^{-3}$ ,  $FDR = 1.66 \times 10^{-2}$ ), “Sphingolipid metabolic process” ( $P = 7.45 \times 10^{-4}$ ,  $FDR = 1.16 \times 10^{-2}$ ) and “Unsaturated fatty acid biosynthetic process” ( $P = 2.13 \times 10^{-3}$ ,  $FDR = 2.27 \times 10^{-2}$ ). Additional file 2: Table S2 provides the full list of overrepresented GO terms. Pathway analyses revealed an enrichment for “Regulation of actin cytoskeleton (hsa04810),”

**Table 1 Description of the ten most connected nodes in the co-association network**

SNP/Gene	Illumina Chip SNP	Associated Traits	Connections	Consequence
ALGA0061664	ALGA0061664	3	376	Intergenic variant
SLC30A9	H3GA0024739	1	373	Intronic variant
SEMA3F	DIAS0001129	3	370	Intronic variant
ARHGEF2	ASGA0021047	3	368	Downstream variant
NTRK3	MARC0045253	3	367	Intronic variant
ZFH4	ALGA0025325	1	366	Intronic variant
SLC22A3	ALGA0117149	6	365	Intronic variant
ARMC4	ALGA0059185	4	356	Intronic variant
C9orf171	ASGA0008154	6	355	Intronic variant
PPP2R2A	DIAS0004697	5	353	Splice region variant

“Focal adhesion (hsa04510)”, “Pathways in cancer (hsa05200)”, “Chemokine signalling pathway (hsa04062)”, “Phosphatidylinositol signalling system (hsa04070)” and “Inositol phosphate metabolism (hsa00562)” (Additional file 3: Table S3).

To identify potential regulators of the above-mentioned pathways and GO categories, we focused on TF found in the gene network. We applied an information lossless approach that explored the 64,824 possible trios among the available 74 TF (see Methods and Additional file 4: Table S4 for complete list of TF) and identified the TF trio that spanned most of the network topology with minimum redundancy. These three TF were: *Nuclear receptor coactivator 2 (NCOA2, alias TIF2)*, *E1A binding protein p300 (EP300, alias p300)* and *four and a half LIM domains 2 (FHL2, alias SLIM-3)*. Interestingly, the promoter region of these TF contain binding sites for some well-known TF that are considered as important regulators of lipid and carbohydrate metabolism such as: *SREBP-1*, *PPARG*, *PPAR-α*, *HNF1A*, *HNF4-α*, *ER-α* and *GR-α*. In the predicted network, a total of 730 genes show co-association with the three key TF (Figure 2B). A detailed examination of the most representative pathways related to these 730 predicted target genes showed a significant over-representation for “HIF-1 signaling pathway (hsa04066)”, “Acute myeloid leukemia (hsa05221)”, “Colorectal cancer (hsa05210)”, “Renal cell carcinoma (hsa05211)” and “Type II diabetes mellitus (hsa04930)” (Additional file 5: Figure S1). Admittedly, some of the above-mentioned GO terms and pathways could have been expected from a network predicted from GWAS of FA-related phenotypes and this gives confidence in the reliability of the results. Others, however, were unexpected and might lead to new insights on FA physiology.

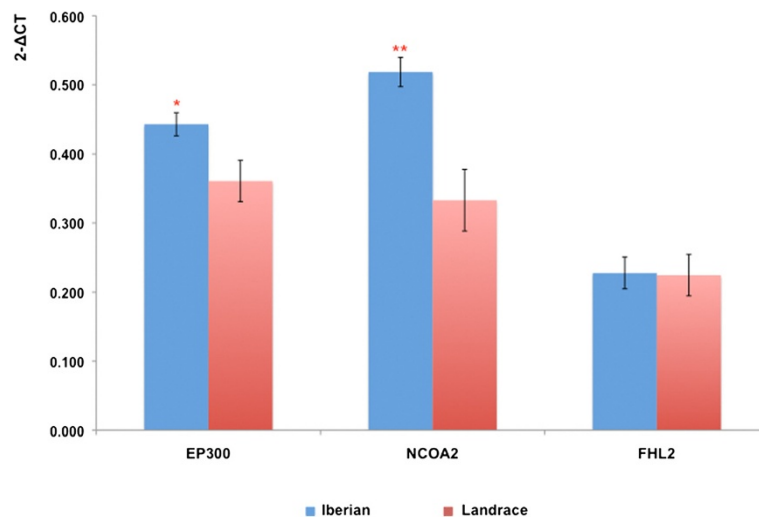
#### Experimental validation: From co-association to co-expression analysis in liver and adipose tissues

The expression of the three TF across *Longissimus dorsi* muscle (LD), adipose and liver tissues was explored. In

concordance with previous results suggesting that highly connected TF are in general broadly expressed across tissues [15], the three TF were expressed across all the studied tissues. Further, a comparison between Iberian and Landrace pig breeds revealed significant increase fold changes (FC) in the liver of Iberian pigs for the expression of *NCOA2* (FC = 1.56,  $P < 0.01$ ) and *EP300* (FC = 1.23,  $P < 0.05$ ) (Figure 3).

The expression patterns of 43 genes in liver and 40 genes in adipose tissue were successfully measured across 55 backcross animals. In liver, the expression data of twelve additional genes were also included in the co-expression analysis (see Methods). Co-expression analysis revealed highly connected networks in both liver and adipose tissue, suggesting strong functional interconnections among the studied genes. Topology of liver co-expression network showed 55 nodes connected by 425 edges (Additional file 6: Figure S2A) and in adipose tissue 40 nodes and 261 edges were observed (Additional file 6: Figure S2B). Network parameters such as average degree (Deg) and average distance (AvD<sub>G</sub>) were slightly higher in liver co-expression network compared to adipose tissue network (Deg<sub>Liver</sub> = 15.45 AvD<sub>G</sub> = 1.81 vs Deg<sub>Adipose</sub> = 13.05 and AvD<sub>G</sub> = 1.75). Based on network centrality, the relevance of individual genes differs within each network. For example, topological properties of the liver co-expression network suggest an important role for *ARNT* in the regulation of hepatic lipogenic and gluconeogenesis activity, and these findings agree with published results [16,17]. It should be noted that *BCL9* showed the highest centrality value in the liver co-expression network (Additional file 6: Figure S2A). In addition, degree analysis showed that *BCL9*, *EP300*, *PBX1*, *SIRT1*, *PIP5K1A* and *ARNT* were the most central genes in the liver co-expression network. However, in the adipose co-expression network, degree analysis suggested that *ANK2*, *NCOA2*, *SIRT1*, *EIF4E*, *HMBX1* are the most central genes (Additional file 6: Figure S2B). When analysing a sub-network of the liver co-expression





**Figure 3** Results of the liver differential expression analysis comparing the best TF trio in the Iberian and Landrace breeds.

network, formed only by the same 40 genes included in the adipose co-expression network, five genes (*BCL9*, *EP300*, *PBX1*, *PIP5K1A*, and *SIRT1*) were still the most central genes and this finding underscores their relevant role in the function and structure of the liver co-expression network.

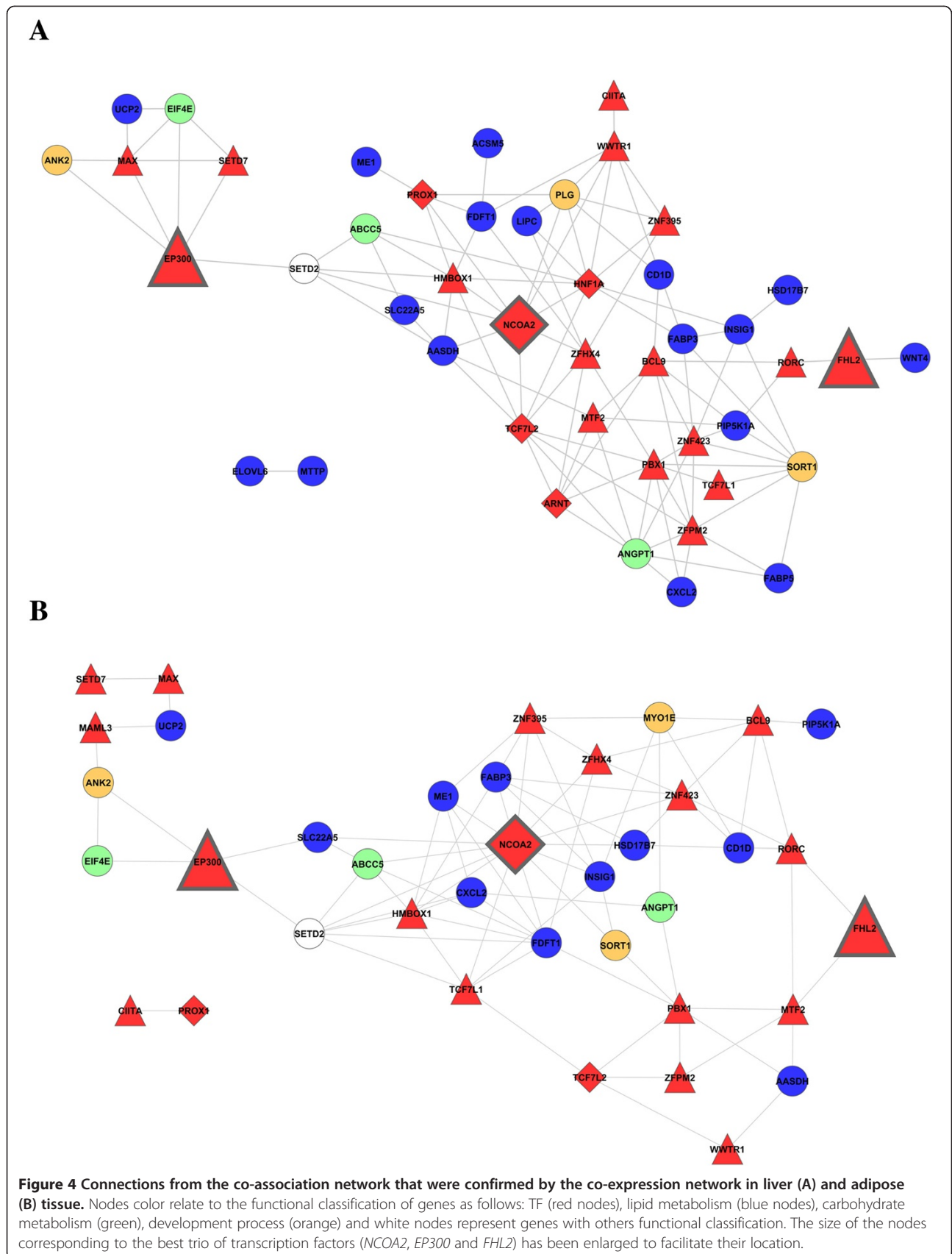
Beyond the study of the topological properties of the liver and adipose tissue co-expression networks, we were concerned with those, if any gene interactions predicted via SNP co-association were corroborated through co-expression analyses. In line with recent results in yeast [18], we observed that interacting loci could jointly regulate the co-expression patterns of pairs of genes. For the first time in a not model species, co-expression analyses confirmed gene-gene interactions predicted based on SNP co-association. However, the magnitude of this validation varied in a tissue-specific manner. For instance, with respect to the liver module formed by 48 AWM nodes and 359 edges (based on co-association analysis) we observed that 28.9% (104/359) of the predicted gene-gene interactions were validated by the co-expression results. Whereas in the adipose tissue, the observed percentage of the AWM validated interactions was slightly higher representing 33.0% of the possible combinations (Figure 4B). When we limited this comparison to the intersecting 39 genes included in both co-expression networks, the proportion of the AWM gene-gene interactions validated in liver (29.5%) was still lower than in adipose tissue (33.0%). Comparing both networks, we observed that approximately 35.7% (or 30 out of 84) of the interactions validated in the adipose tissue were also validated in the liver co-expression analysis (Additional file 7: Table S5). Interestingly, these always co-associated and co-expressed genes belong to biological processes related to lipid metabolism including: Negative Regulation

of Fat Cell Differentiation (*INSIG1*, *TCF7L2*, *ZFPM2*), Androgen Receptor Signalling Pathway (*EP300*, *FHL2*, *NCOA2*), Response to Hormone Stimulus (*ABCC5*, *ANGPT1*, *FABP3*, *EP300*, *SORT1*, *FHL2*) and Lipid Metabolic Process (*PBX1*, *INSIG1*, *FABP3*, *FDFT1*, *PIP5K1A*, *MAX*, *AASDH*).

When we focused on the best TF trio, we observed that 60.8% (or 14 out of 23) of the interactions of *NCOA2* predicted by the AWM co-association network were corroborated in the co-expression network of the adipose tissue. This percentage dropped to 34.6% (or 9 out of 26) in the co-expression network of the liver tissue. For *EP300*, 44.4% (or 4 out of 9) of the AWM predicted interactions were observed in the adipose co-expression network and 41.6% (5 out of 12) in the liver co-expression network. Finally, for *FHL2* we observed the lowest percentage of validated interactions: 20.0% (or 2 out of 10) in adipose tissue and 14.3% (2 out of 14) in liver (Table 2).

## Discussion

Molecular processes controlling FA metabolism are highly interconnected and linked with related pathways, such as lipid, carbohydrate and energy metabolism. In fact, FA are a major energy source and together with several factors, such as total energy intake, dietary fat/carbohydrate ratio, or glucose and/or insulin concentration, regulate *de novo* lipogenesis [19,20]. As a consequence, it is expected that at the selected threshold ( $P < 0.035$ ) our best trio of TF (*NCOA2*, *EP300*, *FHL2*) show co-association with a large number of genes and other TF relevant for lipid, carbohydrate and energy metabolism. For instance, 39 of the predicted target genes via SNP co-association (Additional file 8: Table S6) have been recently reported in two large-scale meta-analysis studies for plasma lipids in humans [21,22]. Interestingly, many of these genes, including our



**Table 2 Concordance validation between the co-association and the co-expression networks for the best TF trio and in adipose and liver tissues**

Tissue	TF	Connections in the AWM co-association network <sup>A</sup>	Connections in the qPCR co-expression network <sup>A</sup>	Validated connections	% Validation
Adipose	NCOA2	23	21	14	60.8
	EP300	9	19	4	44.4
	FHL2	10	9	2	20.0
Liver	NCOA2	26	18	9	34.6
	EP300	12	28	5	41.6
	FHL2	14	13	2	14.3

<sup>A</sup>Connections deemed significant according to the PCIT algorithm.

TF trio and other FA relevant genes, would have been missed by traditional single-trait GWAS due to the lack of an acceptably significant association level (i.e.  $P > 0.05$  after correction for multiple testing). As noted before [12] and confirmed by this study, AWM points to new candidate genes, TF and gene interactions via exploring SNP co-associations across multiple traits beyond the one-dimensional approach for identifying genes affecting single traits. However, results should be interpreted with caution due to the limited sample size used in our study (144 pigs), which reduces the power to identify small effects and may introduce spurious results. Therefore, these TF might regulate other important genes for IMF FA composition not represented in this network and false positive results may be included in the network. However, only the SNPs associated with a large number of phenotypes were included in the AWM analysis and, due the multi-trait nature of the AWM methodology, the probability that the same SNP was associated with several phenotypes by chance is much lower than the probability of being associated with a single phenotype.

In the predicted network, *NCOA2*, a key TF regulating energy homeostasis [20,23] and adipogenesis [24], showed co-association with a total of 326 genes, including relevant TF and genes associated with lipid and carbohydrate metabolisms, such as *PROX1*, *PBX1*, *ARNT*, *MYB*, *MTF2*, *TCF7L1*, *SCD5*, *ABCC2*, *INSIG1*, *ACACB*, *FABP4*, *FABP3*, *ME1*, *AASDH*, *ABCC5* and *SORT1*. A role for *PROX1* in the control of energy homeostasis has been proposed [25]. Moreover, association of SNPs mapped to *PROX1* and *SLC30A8* with fasting glucose levels and increased risk for T2D has been reported in humans [26]. Both *PROX1* and *SLC30A8*, together with other T2D risk loci (*IL6R*, *TCF7L2*, *HNF1A*) and 21 genes reported as associated with plasma lipids in humans [22] were predicted as target genes of *NCOA2* in our study. Co-expression analysis in adipose tissue validated 60.8% of the *NCOA2* co-association target genes, including *INSIG1* ( $r_{\text{co-expression}} = 0.68$ ), *FDFT1* ( $r_{\text{co-expression}} = 0.70$ ), *SETD2* ( $r_{\text{co-expression}} = 0.59$ ) and *ABCC5* ( $r_{\text{co-expression}} = 0.65$ ). In liver, 34.6% of the predicted targets of *NCOA2* were validated, including

the above-mentioned *PROX1* ( $r_{\text{co-expression}} = 0.48$ ), *HNF1A* ( $r_{\text{co-expression}} = 0.56$ ) and *TCF7L2* ( $r_{\text{co-expression}} = 0.50$ ). It should be noted that previous studies in pigs show a correlation between *NCOA2* expression ( $r = 0.605$ ,  $P < 0.01$ ) and IMF content of LD muscle [24]. Also, *NCOA2* was reported as modulating an AWM-network predicted for puberty in cattle [27], which included fat deposition measurements as traits related to puberty. Furthermore, knockout *NCOA2*<sup>-/-</sup> mice are protected against obesity, showing lean phenotype and decreased expression of genes involved in the uptake and storage of FA [20]. A decreased expression of genes required for FA synthesis in liver tissue of *NCOA2*<sup>-/-</sup> mice was observed [28]. In agreement with these previous results and the phenotypic difference in fat deposition between Iberian and Landrace breeds, a significant higher activity of *NCOA2* in the liver of Iberian pigs was detected (FC = 1.56,  $P < 0.01$ ) relative to Landrace pigs (Figure 3).

Another TF predicted as critical for FA regulation was *EP300*, which encodes the adenovirus *E1A-associated cellular p300 transcriptional co-activator protein*. It functions as histone acetyltransferase that regulates transcription by chromatin remodelling. Via histone acetyltransferase activity, *EP300* regulates the transcription of liver X receptor (*LXR*) [29]. *EP300* is also required for adipocyte differentiation through the regulation of peroxisome proliferator-activated receptor gamma (*PPARG*) [30]. Remarkably, *EP300* has been reported as transcriptional co-activator of estrogen receptor (*ER*), hepatocyte nuclear factor 4  $\alpha$  (*HNF4- $\alpha$* ), aryl hydrocarbon receptor nuclear translocator (*ARNT*) and hepatocyte Nuclear Factor-1  $\alpha$  (*HNF1A*) [31-33]. All these above-mentioned TF co-regulated by *EP300* (*PPARG*, *LXR*, *HNF4*, *HNF1A*, *ER*, *ARNT*) influence lipid and carbohydrate metabolisms and have been extensively studied in this context [17,34-41]. Among the 180 AWM-predicted target genes for *EP300*, there are 30 genes known to be involved in lipid metabolism including *ARNT* a member of the HIF-1 pathway. *ARNT* is a relevant TF regulating hepatic gluconeogenesis and lipogenic gene expression [16]. Interestingly, we observed a significant co-expression between *ARNT* and *EP300* ( $r = 0.61$ ) in

the liver network. Additionally, other genes related to carbohydrate and lipid metabolism were predicted as *EP300* AWM-target genes. These included: *ADCY2*, *MMP9*, *ECHS1*, *ARRB1*, *EIF4E*, *ANK2*, *NR2E1*, *SLC2A6*, *SLC5A2*, *LEP*, *ELOVL6*, *MTTP*, *ACSM5*, *UCP2* and *CYP2E1* (for a full list see Additional file 9: Table S7). Similarly to *NCOA2*, a significant higher expression of *EP300* in the liver of Iberian pigs was detected ( $FC = 1.23$ ,  $P < 0.05$ ) in comparison with Landrace pigs (Figure 3). Our results, predicting targets for *EP300* and studying their co-expression contributes to the knowledge on lipid and carbohydrate metabolism. It is well known that TF require co-regulators to modify and epigenetically remodel chromatin structure to facilitate the basal transcriptional machinery. *EP300* is a chromatin remodeling gene opening new possibilities to study the roll of epigenetic modifications in the regulation of pork meat quality and the molecular control of energy homeostasis.

The third key TF was *FHL2*, an evolutionarily conserved gene that can interact with an important range of proteins from different functional classes, including receptors, signal transducers, TF and cofactors [42]. *FHL2* plays an important role as molecular transmitter linking various signalling pathways to transcriptional regulation. For instance, *FHL2* is involved in the co-activation of human androgen receptor (*AR*), *ER* and peroxisome proliferator-activated receptor alpha (*PPAR $\alpha$* ) [42-44]. In addition, *FHL2* mediates interaction with  $\beta$ -catenin and promotes myoblast C2C12 differentiation in mice [45]. The gene *B-cell CLL/lymphoma 9 (BCL9)*, an activator of the Wnt/ $\beta$ -catenin [46] and *Wingless-type MMTV integration site family, member 4 (WNT4)* was among the 251 targets predicted for *FHL2* in our network. The growth factor *WNT4* is a member of the Wnt signaling pathway involved in developmental processes and relevant for gonad development and sex-determination [47]. Liver expression analyses provided supporting evidence for the predicted interaction between *FHL2* and *WNT4*, as a significant co-expression ( $r = 0.44$ ,  $P < 0.001$ ) was observed. Other genes and TF associated with development process, lipid and carbohydrate metabolism, such as *FHL5*, *MYO1E*, *MYB*, *RORC*, *JARID2*, *ZFH4*, *WNK1*, *LIPC*, *CREB5*, *CDC42*, *ACSL1*, *FABP5*, *ABCB11*, *FLT1* and *HTR2A* were also predicted as targets of *FHL2* according to the co-association network. *FHL2* was not differentially expressed in the comparison between Iberian and Landrace pigs. Also, *FHL2* showed a proportion of validated interactions in the co-expression analysis (20% adipose tissues and 14.3% in liver) lower than for the other two TF, *NCOA2* and *EP300*. These somewhat less promising results could be a consequence of the tissue-specific activity of *FHL2*, as it has been reported for the co-activation of *AR* [43].

Although, some gene to gene interactions predicted by the AWM approach were not corroborated by the co-expression analysis, the possibility of these interactions occurring in other spatial temporal and/or tissues cannot be ruled out, or indeed manifesting their joint effect through other means than co-expression. TF and their target genes interact in a temporal and tissue dependent manner, so the examination of networks spanning multiple tissues is critical to highlight interactions that could otherwise be unknown from individual tissue analysis [48]. In spite of this tissue/time limitation, two of the three TF from the best trio (*EP300* and *NCOA2*) showed higher than average centrality values in both liver and adipose tissue co-expression networks. Moreover, we observed a significant co-expression between *NCOA2* and *EP300* in the liver network with some other TF considered master regulators of the lipid metabolism. For instance, *NCOA2* was significantly co-expressed with *PPAR $\alpha$*  ( $r = 0.39$ ,  $P < 0.01$ ), *HNF1A* ( $r = 0.56$ ,  $P < 0.001$ ) and *HNF4 $\alpha$*  ( $r = 0.36$ ,  $P < 0.01$ ), and *EP300* was co-expressed with *PPARD* ( $r = 0.38$ ,  $P < 0.01$ ) and *HNF1A* ( $r = 0.64$ ,  $P < 0.001$ ) (Additional file 6: Figure S2 A, B). The liver plays a central role in maintaining overall energy balance by controlling lipid and carbohydrate metabolism. In pigs, the liver is the primary site of *de novo* cholesterol synthesis and fatty acid oxidation and, together with adipose tissue, has a crucial role in regulating lipid metabolism [49,50]. All these observations, together with the higher expression of *NCOA2* and *EP300* observed in the liver of the Iberian pigs compared with Landrace pigs, suggest a relevant role of these genes in the hepatic transcriptional regulation of lipid metabolism in pigs.

Overall, our GWAS and network predictions, supported by literature and co-expression analysis in liver and adipose tissue, suggest a co-operative role for the three TF (*NCOA2*, *EP300*, *FHL2*) in the transcriptional regulation of IMF, FA composition and the control of energy homeostasis in pigs. We hypothesize that these TF mediate a highly inter-connected regulatory cascade including pathways such as HIF-1, *AR*, *ER* and Wnt/ $\beta$ -catenin that seem pivotal for lipid metabolism. The role of these pathways in the transcriptional regulation of lipid metabolism is a subject of intense studies [17,38,39,51-54]. A functional cooperation between the three TF in the modulation of these pathways is evident from our results and supported by literature evidence. For example, according to String database [55,56] (<http://string-db.org/>), experimental data confirmed that protein-protein interaction exists among, *EP300*, *NCOA2*, *FHL2*, *AR* and *ESR1* (Additional file 10: Figure S3). In addition, *EP300* and *NCOA2* take part on the *AR* and *ER* pathways and both, *NCOA2* and *FHL2* are *AR* co-regulators [43,57,58]. Studying the combined effect of *NCOA2*, *EP300*, and *FHL2* in the regulation of



specific genes will lead to new knowledge related to FA pathways.

The most overrepresented pathway corresponding to the 730 AWM-predicted target genes of the three TF was HIF-1 (Additional file 5: Figure S1). The HIF-1 pathway is central to adaptive regulation of cellular energy metabolism; by regulating the expression of glycolytic enzymes and hepatic lipid metabolism [17,54,59,60]. Our liver co-expression analysis supports previously reported evidence [16] for the relevance of *ARNT* gene (member of HIF pathway) in the hepatic lipogenic gene expression. Additionally, HIF-1 $\alpha$ , which is another member of HIF pathway and  $\beta$ -catenin co-ordinately enhance AR transactivation. The interaction between  $\beta$ -catenin and both HIF-1 and AR pathways has been documented [61-63]. Moreover,  $\beta$ -catenin is a ligand-dependent co-activator of AR and a functional cooperation in the synergistic activation of AR-mediated transcription among *EP300*, *FHL2* and  $\beta$ -catenin have been reported [64]. Ours results showing the interactions between the three key TF, recapitulate these pathways interactions that are known mammalian biology, extending its significance to pigs.

## Conclusions

In summary, our results suggest that common genetic variants mapped to (or in *linkage disequilibrium* with) *EP300*, *FHL2* and *NCOA2* together with other candidate genes including *ARNT*, *BCL9*, *SIRT1*, *PBX1*, *PROX1*, *HNF1A*, *SLC30A8*, *TCF7L2* and *ANK2* modulate lipid metabolism and control energy homeostasis in pigs. Furthermore, epistatic predicted interactions between TF and their target genes are likely to contribute to the complex inheritance of FA composition and related polygenic traits (lipid metabolism and energy homeostasis). It is generally accepted that metabolic diseases such as obesity and T2D are linked to disturbance of energy homeostasis or homeostatic imbalance. It should be noted that among the 730 predicted target genes, an overrepresentation of genes from the T2D pathway was observed (Additional file 5: Figure S1). Also, 39 of the 730 genes are known to control plasma lipid content in humans [21,22].

Further studies will be required to elucidate the specific cellular and molecular processes of interaction among the three TF and its target genes that determine FA composition and control energy homeostasis in pigs. The implications of research in this area are broad, ranging from applications from pork meat quality to modeling mammal biology.

## Methods

### Phenotypic traits, animals and genotypes

Data from 144 pigs (25% Iberian  $\times$  75% Landrace), representing 26 full-sib families, from backcrossing five F1 males with 26 Landrace sows was utilized. Details about

the management conditions and the phenotype information have been previously reported [65-67]. For this study and based on an previous principal components analysis [66] we selected 15 of the total 48 traits representing the most informative phenotypes within the dataset. Nine of the 15 traits were related to IMF fatty acid (FA) composition in LD muscle, seven correspond to indices of FA metabolism and the last one is the IMF percentage (Additional file 11: Table S8). The Porcine SNP60K BeadChip (Illumina) [68] was used to genotype a total 197 pigs, including the 144 phenotyped animals and the founder population. Quality control excluded SNPs with minor allele frequency < 5% and with call rate < 95%. A subset of 48,119 SNPs were retained for subsequent analysis, in addition, previously detected polymorphisms in the *MTTP*, *FABP4*, *FABP5*, and *ELOVL6* genes were also tested [67,69,70]. The genomic coordinates of the SNP correspond to the *Sus scrofa* genome sequence assembly (Sscrofa10.2, August 2011) [71] and were annotated using as reference the pig assembly 10.2 [ftp://ftp.ncbi.nlm.nih.gov/genomes/Sus\_scrofa/GFF/].

### Ethics statement

Animal care and procedures were performed following national and institutional guidelines for the Good Experimental Practices and approved by the Ethical Committee of the Institution (IRTA- Institut de Recerca i Tecnologia Agroalimentàries).

### Statistical analysis

The GWAS was performed using Qxpack 5.0 software [72]. The additive effect of a SNP on each trait was estimate by mixed model [73,74] following the model:

$$y_{ij} = X\beta + Zu + s_{j,k} + e_{ij},$$

where:  $y_{ij}$  represents the vector of observations from the  $i^{\text{th}}$  pig at the  $j^{\text{th}}$  trait ;  $X$  is the incidence matrix relating fixed effects in  $\beta$  with observation in  $y_{ij}$ ;  $Z$  is the incidence matrix relating random additive polygenic effects in  $u$  with observation in  $y_{ij}$ ;  $s_{j,k}$  represents the additive association of the  $k^{\text{th}}$  SNP on the  $j^{\text{th}}$  trait and  $e_{ij}$  is the vector of random residual effects. Fixed effects included in  $\beta$  were, sex (two levels), batch (five levels) and carcass weight as covariate. Polygenic effects were treated as random and distributed as  $N(0, A\sigma_u)$  where  $A$  is a numerator of kinship matrix. Then, the allele substitution effect of the  $i^{\text{th}}$  SNP on the  $j^{\text{th}}$  trait was z-score standardized and employed to constructing the AWM [12].

An R script, available from the authors, was written to automate the process of building an AWM. Palmitoleic acid (C16:1 (n-7)) was used as the key phenotype and the procedure described by Fortes and colleagues [12]

was followed, but we introduced a few modifications, specifically regarding the  $P$ -value threshold for selecting SNP from GWAS. The  $P$ -value threshold was chosen by exploring the sensitivity of the data instead of simply accepting the nominal  $P < 0.05$ . We took advantage of the biological knowledge concerning TF related to the analyzed traits and used it as *a priori* information.

In essence, instead of applying a hard-coded nominal  $P$ -value of, for instance, 0.05 or 0.01 or 0.001, we employed a knowledge-based approach to identify the  $P$ -value threshold at which the information content, in terms of fatty acid regulation, is maximized. To this effect, we mined the literature and relevant databases to compile a list of 340 TF of which 34 were known to be related to FA metabolism. The distribution of these 34 TF relative to the entire set of 340 was explored at various  $P$ -value thresholds. The  $P$ -value threshold that maximized the number of FA-related TF was used as the optimal  $P$ -value to apply when developing the Association Weight Matrix. Quite importantly, others in the past have employed a knowledge-based approach to identify the critical  $P$ -value threshold. More recently, and in the context of GWAS, Yang et al. [75] used an approach similar to ours to find the  $P$ -value that maximise the correlation between the proportion of significant SNPs and the heritability across 47 traits.

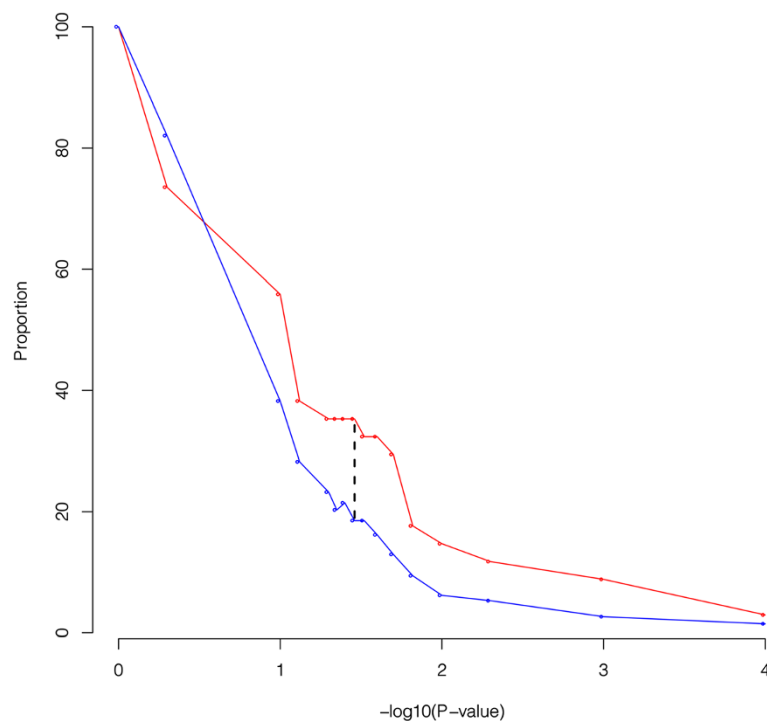
In detail, the process for choosing the threshold was as follows:

**Step1:** A total of 340 TF were located within 2.5 Kb of a SNP and therefore included in the initial dataset. For all these TF included in our dataset, those that are well known key regulators of the lipid metabolism were initially selected.

**Step2:** For each gene, those involved in the lipid metabolism and also reported in the census of human TF by Vaqueriza et al. [76] were included.

**Step3:** The Human Protein Reference Database (HPRD) and the Biomolecular Object Network Databank (BIND) were mined. Then other TF that have been reported to interact with some of the TF retained in the two previous steps were selected. After these first 3 steps, a total of 34 TF were retained (Additional file 12: Table S9).

**Step4:** Subsequently we compare the distribution of the 34 TF at different  $P$ -values from  $P = 1$  to  $P = 10^{-4}$  versus the distribution of the total number of TF included in the AWM (340). As a result, we have chosen  $P < 0.035$  as the threshold. This specific  $P$ -value maximizes the difference between both groups of TF (Figure 5), imposing an informed bias towards lipid metabolism to the network.



**Figure 5** Sensitivity analysis of the 34 lipid-related TF at different  $P$ -values (from  $P < 1$  to  $P < 10^{-4}$ ) against the distribution of the total 340 TF included in the dataset.

After defining the threshold of  $P < 0.035$ , the selection of SNPs for building AWM continued. Those SNPs that were either associated ( $P < 0.035$ ) with palmitoleic acid or with any  $\geq 3$  traits, and were located either  $\leq 2,500$  bp to or  $\geq 850$  kb from the nearest annotated gene (Sscrofa10.2 assembly), were selected to build the AWM matrix. PermutMatrix software [77] was employed to visualize hierarchical clustering of traits (AWM columns) and genes (AWM rows) using Euclidean distance and the Average linkage method. To identify and report gene-gene or gene-SNP interactions we used PCIT algorithm [14]. Cytoscape software [78] was used to visualize the gene network and also to perform overrepresented GO terms analysis, using BiNGO plugin [79]. Node centrality values and network topological parameters were calculated using CentiScaPe plugin [80]. Pathway enrichment analysis were performed using FATIGO tool from BABELOMICS [81,82]. Further, pathways analyses of the 730 predicted target genes (co-associated with the key TF) were performed using ClueGO, Cytoscape plugin [83]. Pathway information was retrieved from the KEGG (<http://www.genome.jp/kegg/>) and BioCarta (<http://www.biocarta.com/>) databases. In all cases, the cut-off for considering a significance overrepresentation was established by Benjamini & Hochberg multiple testing correction of the  $P$ -value ( $FDR < 0.05$ ) [84].

### Expression and co-expression analysis

In order to provide supporting evidence for the *in-silico* AWM-network predictions we obtained and explored gene expression data by reverse transcription quantitative Real-Time PCR (RT-qPCR). The expression pattern of the 3 Key TF (*NCOA2*, *EP300*, *FHL2*) in LD muscle, liver and adipose tissues was tested in two phenotypically divergent breeds for fat deposition traits (Iberian and Landrace which are also the founders of our studied population, five animals per breed). Finally, liver and adipose co-expression analyses of 55 (43 from the present study, and twelve: *ACSM5*, *APOA2*, *ARNT*, *CYP7A1*, *FABP5*, *FADS3*, *HNF4a*, *LIPC*, *MTTP*, *PPARA*, *PPARD* and *ELOVL6* genes from Ballester et al., 2013 submitted) and 40 genes, respectively, were performed using the PCIT algorithm [14] in 55 backcross animals. Since sex differences in liver transcriptome have been reported [85] only females were considered in the co-expression analyses of both tissues.

From the 55 genes explored in the liver co-expression analysis, 48 were present in the AWM network. The remaining seven were incorporated due to their biological relevance, including three well-know TF related to lipid metabolism (*PPAR $\alpha$* , *PPARD*, *HNF4a*) and four genes related to lipid metabolism (*SIRT1*, *FADS3*, *APOA2*, *CYP7A1*). Similarly, from the 40 genes employed in the adipose co-expression analysis, 39 were present in the

AWM network. The one gene out, *SIRT1*, was also included due to its relevant controlling lipolysis [86,87] and promoting fat mobilization in white adipose tissue [88].

Total RNA was obtained from liver, muscle and adipose tissues using the RiboPure kit (Ambion), following the manufacturer's recommendations. RNA was quantified using the NanoDrop ND-1000 spectrophotometer (NanoDrop products) and the RNA integrity was assessed by Agilent Bioanalyzer-2100 (Agilent Technologies). Approximately, one microgram of total RNA was reverse-transcribed into cDNA using the High-Capacity cDNA Reverse Transcription kit (Applied Biosystems) in 20  $\mu$ l of reactions, following the manufacturer's instructions.

To analyze the expression pattern of the 3 key transcription factors, an ABI PRISM 7900 Sequence Detection System (Applied Biosystems) in combination with FastStart Universal Sybr green master (Rox; Roche Applied Science) was used. PCR amplifications were performed in a total reaction volume of 20  $\mu$ l containing 5  $\mu$ l of cDNA diluted 1:25. All primers were used at 300 nM. The thermal cycle was 10 min at 95°C, 40 cycles of 15 s at 95°C and 1 min at 60°C. A dissociation curve was drawn for each primer pair to assess the specificity of the amplification. Three reference genes (*ACTB*, *HPRT1*, *TBP*) frequently used in RT-qPCR experiments were tested as endogenous controls. Using the GeNorm software [89], the *ACTB* and *TBP* genes were selected as the best endogenous controls for all tissues. After ensuring the possibility to use the  $2^{-\Delta\Delta CT}$  method [90], data was analyzed using the RQ manager v1.2.1 and the DataAssist™v3.0 softwares (Applied Biosystems). The  $2^{-\Delta CT}$  values were used to compare our data.

The 48.48 microfluidic dynamic array IFC chip (Fluidigm) was used to analyze the expression of 48 genes (44 target genes and 4 reference genes) in liver and adipose tissue of 55 backcross animals belonging to the same population in which the GWAS was performed. Two  $\mu$ l of 1:5 diluted cDNA was pre-amplified using 2X Taqman PreAmp Master Mix (Applied Biosystems) and 50 nM of each primer pair in 5  $\mu$ l reaction volume, according to the manufacturer's directions. The cycling program was 10 min at 95°C followed by 16 cycles of 15 s at 95°C and 4 min at 60°C. At the end of this pre-amplification step, the reactions were diluted 1:5 (diluted pre-amplification samples). RT-qPCR on the dynamic array chips was conducted on the BioMark™ system (Fluidigm). Five  $\mu$ l sample pre-mix containing 2.5  $\mu$ l of SsoFast EvaGreen Supermix with Low ROX (Bio-Rad), 0.25  $\mu$ l of DNA Binding Dye Sample Loading Reagent (Fluidigm) and 2.25  $\mu$ l of diluted pre-amplification samples (1:16 or 1:64 from the diluted pre-amplification samples from liver and backfat, respectively), as well as 5  $\mu$ l assay mix containing 2.5  $\mu$ l of Assay Loading Reagent (Fluidigm), 2.25  $\mu$ l of DNA Suspension Buffer (Teknova) and 0.25  $\mu$ l of 100  $\mu$ M primer pairs (500 nM in the final reaction) were mixed

inside the chip using the IFC controller MX (Fluidigm). The thermal cycle was 60s at 95°C followed by 30 cycles of 5 s at 96°C and 20s at 60°C. A dissociation curve was also drawn for each primer pair.

Data was collected using the Fluidigm Real-Time PCR analysis software 3.0.2 (Fluidigm) and analyzed using the DAG expression software 1.0.4.11 [91] applying the relative standard curve method (see Applied Biosystems user bulletin #2). Standard curves with a four-fold dilutions series (1/4, 1/16, 1/64, 1/256, 1/1024) of a pool of 10 cDNA samples were constructed for each gene to extrapolate the quantity values of the studied samples. The PCR efficiencies were almost 100% in both tissues for all the assays (Additional file 13: Table S10) with low coefficients of inter-assay variation of threshold cycle (<2.4% in liver and <3.5% in adipose tissue). Of the four endogenous genes tested (*ACTB*, *B2M*, *HPRT1*, *TBP*), *ACTB* and *TBP* were the genes with the most stable expression [89] in both tissues. The normalized quantity values of each sample and assay were used to compare our data.

All the primers used in this study were designed using PrimerExpress 2.0 software (Applied Biosystems) and are shown in Additional file 13: Table S10. Prior to perform the Fluidigm Real-Time PCR, all the assays were tested for PCR specificity in an ABI PRISM 7900 Sequence Detection System (Applied Biosystems) using two-fold dilutions (1/20, 1/200) of a pool of ten cDNA samples and a minus RT control to check the presence of DNA contamination. Melting curve analysis was performed for all the assays.

#### Data availability

The relevant information and full data sets are included as additional files.

#### Additional files

**Additional file 1: Table S1.** Positional concordance among the ten top connected nodes and QTL deposited in the pig QTL database for fatness related traits.

**Additional file 2: Table S2.** Overrepresented GO terms identified in the network using BinGO Cytoscape plugin.

**Additional file 3: Table S3.** Overrepresented pathways identified with Fatigo.

**Additional file 4: Table S4.** Complete list of the 74 available TF in the AWM.

**Additional file 5: Figure S1.** Overrepresented pathways related to the 730 AWM-target genes according ClueGO results. ClueGO visualizes the terms in a functionally grouped annotation network, reflecting the relationships between the terms (based on the similarity of their associated genes). The size of the nodes reflects the statistical significance of the terms. The group leading term is the most significant term of the group.

**Additional file 6: Figure S2.** Gene co-expression network in liver (A) and adipose (B) tissue. Nodes color relate to the functional classification of genes as follows: TF (red nodes), lipid metabolism (blue nodes), carbohydrate metabolism (green), development process (orange) and white nodes represent genes with others functional classification.

**Additional file 7: Table S5.** Predicted AWM gene-gene interactions confirmed by the co-expression analysis in both liver and adipose tissues.

**Additional file 8: Table S6.** List of the 39 AWM-predicted target genes that have been recently reported in two large-scale meta-analysis studies for plasma lipids in humans.

**Additional file 9: Table S7.** List of the 30 genes involved in lipid metabolism predicted as target genes of *EP300*.

**Additional file 10: Figure S3.** Protein-protein interaction among *EP300*, *FHL2* and *NCOA2* with *ESR1* and *AR* inferred from String database.

**Additional file 11: Table S8.** Brief description, mean, standard deviation (SD) and estimated heritability ( $h^2$ ) of the 15 analyzed traits.

**Additional file 12: Table S9.** List of 34 TF retained after Step 3 for choosing the threshold.

**Additional file 13: Table S10.** Primers used in the experimental validation by real-time PCR.

#### Competing interests

The authors declare that they have no competing of interests.

#### Authors' contributions

YR-C, AR and JMF conceived and designed the experiment. JMF was the principal investigator of the project. YR-C, MB, MF, MP-E and AR performed the data analysis. AE annotated the SNPs. YR-C, MB, MF, AR and JM drafted the manuscript. AC, JN, AF, MP-E and JMF collected the samples. AC and MB performed DNA and RNA isolation. MB and AC performed the qPCR and RT-PCR assays. All authors read and approved the final manuscript.

#### Acknowledgments

This work was funded by MICINN project AGL2011-29821-C02 (Ministerio de Economía y Competitividad), and by the Innovation Consolider-Ingenio 2010 Program (CSD2007-00036, Centre for Research in Agrigenomics). We thank Dr. Nick Hudson for insightful suggestions and manuscript proofread. Y. Ramayo-Caldas was funded by a FPU PhD grant from the Spanish Ministerio de Educación (AP2008-01450).

#### Author details

<sup>1</sup>Centre de Recerca en Agrigenòmica (CRAG), Consorci CSIC-IRTA-UAB-UB, Campus UAB, Bellaterra 08193, Spain. <sup>2</sup>Departament de Ciència Animal i dels Aliments, Facultat de Veterinària, Universitat Autònoma de Barcelona, Bellaterra 08193, Spain. <sup>3</sup>The University of Queensland, Queensland Alliance for Agriculture and Food Innovation, Center for Animal Science, Gatton, Queensland 4343, Australia. <sup>4</sup>IRTA, Genètica i Millora Animal, Lleida 25198, Spain. <sup>5</sup>INIA, Mejora Genètica Animal, Madrid 28040, Spain. <sup>6</sup>Institutió Catalana de Recerca i Estudis Avançats, ICREA, Barcelona, Spain. <sup>7</sup>Commonwealth Scientific and Industrial Research Organisation, Division of Animal, Food and Health Sciences, Brisbane, Queensland 4067, Australias.

Received: 19 July 2013 Accepted: 21 March 2014

Published: 26 March 2014

#### References

1. Wakil SJ, Abu-Elheiga LA: **Fatty acid metabolism: target for metabolic syndrome.** *J Lipid Res* 2009, **50**(Supplement):S138-S143.
2. Casellas J, Noguera JL, Reixach J, Diaz I, Amills M, Quintanilla R: **Bayes factor analyses of heritability for serum and muscle lipid traits in Duroc pigs.** *J Anim Sci* 2010, **88**(7):2246-2254.
3. Ntawubizi M, Colman E, Janssens S, Raes K, Buys N, De Smet S: **Genetic parameters for intramuscular fatty acid composition and metabolism in pigs.** *J Anim Sci* 2010, **88**(4):1286-1294.
4. Wood JD, Enser M: **Factors influencing fatty acids in meat and the role of antioxidants in improving meat quality.** *Br J Nutr* 1997, **78**(Suppl 1):S49-S60.
5. Wood JD, Enser M, Fisher AV, Nute GR, Richardson RI, Sheard PR: **Manipulating meat quality and composition.** *Proc Nutr Soc* 1999, **58**(2):363-370.
6. Wood JD, Enser M, Fisher AV, Nute GR, Sheard PR, Richardson RI, Hughes SJ, Whittington FM: **Fat deposition, fatty acid composition and meat quality: A review.** *Meat Sci* 2008, **78**(4):343-358.



7. Ekser B, Ezzelarab M, Hara H, van der Windt DJ, Wijkstrom M, Bottino R, Trucco M, Cooper DK: **Clinical xenotransplantation: the next medical revolution?** *Lancet* 2012, **379**(9816):672–683.
8. Seki Y, Williams L, Yuguin PM, Charron MJ: **Minireview: Epigenetic programming of diabetes and obesity: animal models.** *Endocrinology* 2012, **153**(3):1031–1038.
9. Bendixen E, Danielsen M, Larsen K, Bendixen C: **Advances in porcine genomics and proteomics—a toolbox for developing the pig as a model organism for molecular biomedical research.** *Brief Funct Genomics* 2010, **9**(3):208–219.
10. Houpt K, Houpt T, Pond W: **The pig as a model for the study of obesity and of control of food intake: a review.** *Yale J Biol Med* 1979, **52**(3):307–329.
11. Hardie DG: **Organismal carbohydrate and lipid homeostasis.** *Cold Spring Harb Perspect Biol* 2012, **4**(5). doi:10.1101/cshperspect.a006031.
12. Fortes MR, Reverter A, Zhang Y, Collis E, Nagaraj SH, Jonsson NN, Prayaga KC, Barris W, Hawken RJ: **Association weight matrix for the genetic dissection of puberty in beef cattle.** *Proc Natl Acad Sci U S A* 2010, **107**(31):13642–13647.
13. Serra X, Gil F, Perez-Enciso M, Oliver MA, Vazquez JM, Gispert M, Diaz I, Moreno F, Latorre R, Noguera JL: **A comparison of carcass, meat quality and histochemical characteristics of Iberian (Guadyerbas line) and Landrace pigs.** *Livestock Prod Sci* 1998, **56**(3):215–223.
14. Reverter A, Chan EK: **Combining partial correlation and an information theory approach to the reversed engineering of gene co-expression networks.** *Bioinformatics* 2008, **24**(21):2491–2497.
15. Ravasi T, Suzuki H, Cannistraci CV, Katayama S, Bajic VB, Tan K, Akalin A, Schmeier S, Kanamori-Katayama M, Bertin N, Carninci P, Daub CO, Forrester AR, Gough J, Grimmond S, Han JH, Hashimoto T, Hide W, Hofmann O, Kamburov A, Kaur M, Kawaji H, Kubosaki A, Lassmann T, van Nimwegen E, MacPherson CR, Ogawa C, Radovanovic A, Schwartz A, Teasdale RD, et al: **An Atlas of Combinatorial Transcriptional Regulation in Mouse and Man.** *Cell* 2010, **140**(5):744–752.
16. Wang XL, Suzuki R, Lee K, Tran T, Gunton JE, Saha AK, Patti M-E, Goldfine A, Ruderman NB, Gonzalez FJ, Kahn CR: **Ablation of ARNT/HIF1 $\beta$  in liver alters gluconeogenesis, lipogenic gene expression, and serum ketones.** *Cell Metab* 2009, **9**(5):428–439.
17. Rankin EB, Rha J, Selak MA, Unger TL, Keith B, Liu Q, Haase VH: **Hypoxia-inducible factor 2 regulates hepatic lipid metabolism.** *Mol Cell Biol* 2009, **29**(16):4527–4538.
18. Wang L, Zheng W, Zhao H, Deng M: **Statistical analysis reveals co-expression patterns of many pairs of genes in yeast are jointly regulated by interacting Loci.** *PLoS Genet* 2013, **9**(3):e1003414.
19. Parks EJ, Krauss RM, Christiansen MP, Neese RA, Hellerstein MK: **Effects of a low-fat, high-carbohydrate diet on VLDL-triglyceride assembly, production, and clearance.** *J Clin Invest* 1999, **104**(8):1087–1096.
20. Picard F, Gehin M, Annicotte J-S, Rocchi S, Champy M-F, O'Malley BW, Chambon P, Auwerx J: **SRC-1 and TIF2 control energy balance between white and brown adipose tissues.** *Cell* 2002, **111**(7):931–941.
21. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, Pirruccello JP, Ripatti S, Chasman DI, Willer CJ, Johansen CT, Fouchier SW, Isaacs A, Peloso GM, Barbalic M, Ricketts SL, Bis JC, Aulchenko YS, Thorleifsson G, Feitosa MF, Chambers J, Orho-Melander M, Melander O, Johnson T, Li X, Guo X, Li M, Shin Cho Y, Jin Go M, Jin Kim Y, et al: **Biological, clinical and population relevance of 95 loci for blood lipids.** *Nature* 2010, **466**(7307):707–713.
22. Asselbergs FW, Guo Y, van Iperen EP, Sivapalaratnam S, Tragante V, Lanktree MB, Lange LA, Almqoguera B, Appelman YE, Barnard J, Baumert J, Beitelshes AL, Bhargale TR, Chen YD, Gaunt TR, Gong Y, Hopewell JC, Johnson T, Kleber ME, Langae TY, Li M, Li YR, Liu K, McDonough CW, Meijs MF, Middelberg RP, Musunuru K, Nelson CP, O'Connell JR, Padmanabhan S, et al: **Large-scale gene-centric meta-analysis across 32 studies identifies multiple lipid loci.** *Am J Hum Genet* 2012, **91**(5):823–838.
23. Duteil D, Chambon C, Ali F, Malivindi R, Zoll J, Kato S, Geny B, Chambon P, Metzger D: **The transcriptional coregulators TIF2 and SRC-1 regulate energy homeostasis by modulating mitochondrial respiration in skeletal muscles.** *Cell Metab* 2010, **12**(5):496–508.
24. Wang X, Chen J, Liu H, Xu Y, Wang X, Xue C, Yu D, Jiang Z: **The pig p160 co-activator family: Full length cDNA cloning, expression and effects on intramuscular fat content in Longissimus Dorsi muscle.** *Domest Anim Endocrinol* 2008, **35**(2):208–216.
25. Dufour CR, Levasseur M-P, Pham NHH, Eichner LJ, Wilson BJ, Charest-Marcotte A, Duguay D, Poirier-Héon JF, Cermakian N, Giguère V: **Genomic Convergence among ERR $\alpha$ , PROX1, and BMAL1 in the Control of Metabolic Clock Outputs.** *PLoS Genet* 2011, **7**(6):e1002143.
26. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, Wheeler E, Glazer NL, Bouatia-Naji N, Gloyn AL, Lindgren CM, Mägi R, Morris AP, Randall J, Johnson T, Elliott P, Rybin D, Thorleifsson G, Steinthorsdottir V, Henneman P, Grallert H, Dehghan A, Hottenga JJ, Franklin CS, Navarro P, Song K, Goel A, Perry JR, Egan JM, Lajunen T, et al: **New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk.** *Nat Genet* 2010, **42**(2):105–116.
27. Fortes MRS, Reverter A, Nagaraj SH, Zhang Y, Jonsson NN, Barris W, Lehnert S, Boe-Hansen GB, Hawken RJ: **A single nucleotide polymorphism-derived regulatory gene network underlying puberty in 2 tropical breeds of beef cattle.** *J Anim Sci* 2011, **89**(6):1669–1683.
28. Jeong J-W, Kwak I, Lee KY, White LD, Wang X-P, Brunicaudi FC, O'ÁoMalley BW, DeMayo FJ: **The genomic analysis of the impact of steroid receptor coactivators ablation on hepatic metabolism.** *Molecular Endocrinology* 2006, **20**(5):1138–1152.
29. Huuskonen J, Fielding PE, Fielding CJ: **Role of p160 coactivator complex in the activation of liver X receptor.** *Arterioscler Thromb Vasc Biol* 2004, **24**(4):703–708.
30. Takahashi N, Kawada T, Yamamoto T, Goto T, Taimatsu A, Aoki N, Kawasaki H, Taira K, Yokoyama KK, Kamei Y, Fushiki T: **Overexpression and ribozyme-mediated targeting of transcriptional coactivators CREB-binding protein and p300 revealed their indispensable roles in adipocyte differentiation through the regulation of peroxisome proliferator-activated receptor  $\gamma$ .** *J Biol Chem* 2002, **277**(19):16906–16912.
31. Torres-Padilla ME, Weiss MC: **Effects of interactions of hepatocyte nuclear factor 4  $\alpha$  isoforms with coactivators and corepressors are promoter-specific.** *FEBS Lett* 2003, **539**(1-3):19–23.
32. Chen D, Huang S-M, Stallcup MR: **Synergistic, p160 coactivator-dependent enhancement of estrogen receptor function by CARM1 and p300.** *J Biol Chem* 2000, **275**(52):40810–40816.
33. Ban N, Yamada Y, Someya Y, Miyawaki K, Ihara Y, Hosokawa M, Toyokuni S, Tsuda K, Seino Y: **Hepatocyte nuclear factor-1  $\alpha$  recruits the transcriptional Co-activator p300 on the GLUT2 gene promoter.** *Diabetes* 2002, **51**(5):1409–1418.
34. Li AC, Glass CK: **PPAR- and LXR-dependent pathways controlling lipid metabolism and the development of atherosclerosis.** *J Lipid Res* 2004, **45**(12):2161–2173.
35. Ulven SM, Dalen KT, Gustafsson J-Ö, Nebb HI: **LXR is crucial in lipid metabolism.** *Prostaglandins Leukot Essent Fatty Acids* 2005, **73**(1):59–63.
36. Palanker L, Tennessen JM, Lam G, Thummel CS: **Drosophila HNF4 regulates lipid mobilization and  $\beta$ -oxidation.** *Cell Metab* 2009, **9**(3):228–239.
37. Marcil V, Seidman E, Sinnett D, Boudreau F, Gendron F-P, Beaulieu J-F, Menard D, Precourt L-P, Amre D, Levy E: **Modification in oxidative stress, inflammation, and lipoprotein assembly in response to hepatocyte nuclear factor 4CE  $\pm$  knockdown in intestinal epithelial cells.** *J Biol Chem* 2010, **285**(52):40448–40460.
38. Alaynick WA: **Nuclear receptors, mitochondria and lipid metabolism.** *Mitochondrion* 2008, **8**(4):329–337.
39. Huss JM, Torra IP, Staels B, Giguère V, Kelly DP: **Estrogen-related receptor  $\alpha$  directs peroxisome proliferator-activated receptor  $\alpha$  signaling in the transcriptional control of energy metabolism in cardiac and skeletal muscle.** *Mol Cell Biol* 2004, **24**(20):9079–9091.
40. Xie X, Liao H, Dang H, Pang W, Guan Y, Wang X, Shyy JY-J, Zhu Y, Sladek FM: **Down-regulation of hepatic HNF4  $\alpha$  gene expression during hyperinsulinemia via SREBPs.** *Mol Endocrinol* 2009, **23**(4):434–443.
41. Liu Y, Qiu DK, Ma X: **Liver X receptors bridge hepatic lipid metabolism and inflammation.** *J Dig Dis* 2012, **13**(2):69–74.
42. Johannessen M, Moller S, Hansen T, Moens U, Ghelue MV: **The multifunctional roles of the four-and-a-half-LIM only protein FHL2.** *Cell Mol Life Sci* 2006, **63**(3):268–284.
43. Muller JM, Isele U, Metzger E, Rempel A, Moser M, Pscherer A, Breyer T, Holubarsch C, Buettner R, Schule R: **FHL2, a novel tissue-specific coactivator of the androgen receptor.** *EMBO J* 2000, **19**(3):359–369.
44. Ciarlo JD, Flores AM, McHugh NG, Aneskievich BJ: **FHL2 expression in keratinocytes and transcriptional effect on PPARC $\pm$ /RXRC $\pm$ .** *J Dermatol Sci* 2004, **35**(1):61–63.

45. Martin B, Schneider R, Janetzky S, Waibler Z, Pandur P, Kuhl M, Behrens J, von der Mark K, Starzinski-Powitz A, Wixler V: **The LIM-only protein FHL2 interacts with  $\beta$ -catenin and promotes differentiation of mouse myoblasts.** *J Cell Biol* 2002, **159**(1):113–122.
46. Brack AS, Murphy-Seiler F, Hanifi J, Deka J, Eyckerman S, Keller C, Aguet M, Rando TA: **BCL9 is an essential component of canonical Wnt signaling that mediates the differentiation of myogenic progenitors during muscle regeneration.** *Dev Biol* 2009, **335**(1):93–105.
47. Bernard P, Harley VR: **Wnt4 action in gonadal development and sex determination.** *Int J Biochem Cell Biol* 2007, **39**(1):31–43.
48. Eisen MB, Spellman PT, Brown PO, Botstein D: **Cluster analysis and display of genome-wide expression patterns.** *Proc Natl Acad Sci U S A* 1998, **95**(25):14863–14868.
49. O'Hea EK, Leveille GA: **Significance of adipose tissue and liver as sites of fatty acid synthesis in the Pig and the efficiency of utilization of various substrates for lipogenesis.** *J Nutr* 1969, **99**(3):338–344.
50. Nafikov RA, Beitz DC: **Carbohydrate and lipid metabolism in farm animals.** *J Nutr* 2007, **137**(3):702–705.
51. Willert K, Brown JD, Danenberg E, Duncan AW, Weissman IL, Reya T, Yates JR, Nusse R: **Wnt proteins are lipid-modified and can act as stem cell growth factors.** *Nature* 2003, **423**(6938):448–452.
52. Zhou D, Strakovsky RS, Zhang X, Pan YX: **The skeletal muscle Wnt pathway may modulate insulin resistance and muscle development in a diet-induced obese rat model.** *Obesity (Silver Spring)* 2012, **20**(8):1577–1584.
53. Abiola M, Favier M, Christodoulou-Vafeiadou E, Pichard A-L, Martelly I, Guillet-Deniau I: **Activation of Wnt/ $\beta$ -catenin signaling increases insulin sensitivity through a reciprocal regulation of Wnt10b and SREBP-1c in skeletal muscle cells.** *PLoS ONE* 2009, **4**(12):e8509.
54. Goda N, Kanai M: **Hypoxia-inducible factors and their roles in energy metabolism.** *Int J Hematol* 2012, **95**(5):457–463.
55. Snel B, Lehmann G, Bork P, Huynen MA: **STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene.** *Nucleic Acids Res* 2000, **28**(18):3442–3444.
56. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguéz P, Doerks T, Stark M, Muller J, Bork P, Jensen LJ, von Mering C: **The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored.** *Nucleic Acids Res* 2011, **39**(suppl 1):D561–D568.
57. Heinelein C, Chang C: **Androgen receptor (AR) coregulators: An overview.** *Endocr Rev* 2002, **23**:175–200.
58. Urbanucci A, Waltering K, Suikki H, Helenius M, Visakorpi T: **Androgen regulation of the androgen receptor coregulators.** *BMC Cancer* 2008, **8**(1):219.
59. Hamaguchi T, Iizuka N, Tsunedomi R, Hamamoto Y, Miyamoto T, Iida M, Tokuhisa Y, Sakamoto K, Takashima M, Tamesa T, Oka M: **Glycolysis module activated by hypoxia-inducible factor 1 $\alpha$  is related to the aggressive phenotype of hepatocellular carcinoma.** *Int J Oncol* 2008, **33**(4):725–731.
60. Ke Q, Costa M: **Hypoxia-inducible factor-1 (HIF-1).** *Mol Pharmacol* 2006, **70**(5):1469–1480.
61. Mitani T, Harada N, Nakano Y, Inui H, Yamaji R: **Coordinated action of hypoxia-inducible factor-1  $\alpha$  and  $\beta$ -catenin in androgen receptor signaling.** *J Biol Chem* 2012, **287**(40):33594–33606.
62. Kaidi A, Williams AC, Paraskeva C: **Interaction between  $\beta$ -catenin and HIF-1 promotes cellular adaptation to hypoxia.** *Nat Cell Biol* 2007, **9**(2):210–217.
63. Yang F, Li X, Sharma M, Sasaki CY, Longo DL, Lim B, Sun Z: **Linking  $\beta$ -catenin to androgen-signaling pathway.** *J Biol Chem* 2002, **277**(13):11336–11344.
64. Labalette C, Renard C-A, Neuveut C, Buendia M-A, Wei Y: **Interaction and functional cooperation between the LIM protein FHL2, CBP/p300, and  $\beta$ -catenin.** *Mol Cell Biol* 2004, **24**(24):10689–10702.
65. Ramayo-Caldas Y, Mercade A, Castello A, Yang B, Rodriguez C, Alves E, Diaz I, Ibanez-Escriche N, Noguera JL, Perez-Enciso M, Fernández AI, Folch JM: **Genome-wide association study for intramuscular fatty acid composition in an Iberian x Landrace cross.** *J Anim Sci* 2012, **90**(9):2883–2893.
66. Ramayo-Caldas Y, Mach N, Esteve-Codina A, Corominas J, Castello A, Ballester M, Estelle J, Ibanez-Escriche N, Fernandez AI, Perez-Enciso M, Folch JM: **Liver transcriptome profile in pigs with extreme phenotypes of intramuscular fatty acid composition.** *BMC Genomics* 2012, **13**:547.
67. Corominas J, Ramayo-Caldas Y, Puig-Oliveras A, Perez-Montarelo D, Noguera JL, Folch JM, Ballester M: **Polymorphism in the ELOVL6 gene is associated with a major QTL effect on fatty acid composition in pigs.** *PLoS ONE* 2013, **8**(1):e53687.
68. Ramos AM, Crooijmans RP, Affara NA, Amaral AJ, Archibald AL, Beever JE, Bendixen C, Churcher C, Clark R, Dehais P, Hansen MS, Hedegaard J, Hu ZL, Kerstens HH, Law AS, Megens HJ, Milan D, Nonneman DJ, Rohrer GA, Rothschild MF, Smith TP, Schnabel RD, Van Tassell CP, Taylor JF, Wiedmann RT, Schook LB, Groenen MA: **Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology.** *PLoS ONE* 2009, **4**(8):e6524.
69. Estelle J, Fernandez AI, Perez-Enciso M, Fernandez A, Rodriguez C, Sanchez A, Noguera JL, Folch JM: **A non-synonymous mutation in a conserved site of the MTPP gene is strongly associated with protein activity and fatty acid profile in pigs.** *Anim Genet* 2009, **40**(6):813–820.
70. Mercade A, Perez-Enciso M, Varona L, Alves E, Noguera JL, Sanchez A, Folch JM: **Adipocyte fatty-acid binding protein is closely associated to the porcine FAT1 locus on chromosome 4.** *J Anim Sci* 2006, **84**(11):2907–2913.
71. Groenen MAM, Archibald AL, Uenishi H, Tuggle CK, Takeuchi Y, Rothschild MF, Rogel-Gaillard C, Park C, Milan D, Megens H-J, Li S, Larkin DM, Kim H, Frantz LAF, Caccamo M, Ahn H, Aken BL, Anselmo A, Anthon C, Auvil L, Badaoui B, Beattie CW, Bendixen C, Berman D, Blecha F, Blomberg J, Bolund L, Bosse M, Botti S, Bujie Z, et al: **Analyses of pig genomes provide insight into porcine demography and evolution.** *Nature* 2012, **491**(7424):393–398.
72. Perez-Enciso M, Misztal I: **Oxpk5: old mixed model solutions for new genomics problems.** *BMC Bioinformatics* 2011, **12**:202.
73. Henderson CR: **Best linear unbiased estimation and prediction under a selection model.** *Biometrics* 1975, **31**(2):423–447.
74. Henderson CR: *Applications of Linear Models in Animal Breeding.* Guelph, Ontario, Canada: University of Guelph; 1984.
75. Yang J, Lee T, Kim J, Cho M-C, Han B-G, Lee J-Y, Lee H-J, Cho S, Kim H: **Ubiquitous polygenicity of human complex traits: genome-wide analysis of 49 traits in Koreans.** *PLoS Genet* 2013, **9**(3):e1003355.
76. Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM: **A census of human transcription factors: function, expression and evolution.** *Nat Rev Genet* 2009, **10**(4):252–263.
77. Caraux G, Pinloche S: **PermutMatrix: a graphical environment to arrange gene expression profiles in optimal linear order.** *Bioinformatics* 2005, **21**(7):1280–1281.
78. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: **Cytoscape: a software environment for integrated models of biomolecular interaction networks.** *Genome Res* 2003, **13**(11):2498–2504.
79. Maere S, Heymans K, Kuiper M: **BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks.** *Bioinformatics* 2005, **21**(16):3448–3449.
80. Scardoni G, Pitterlini M, Laudanna C: **Analyzing biological network parameters with CentiScaPe.** *Bioinformatics* 2009, **25**(21):2857–2859.
81. Medina I, Carbonell J, Pulido L, Madeira SC, Goetz S, Conesa A, Tarraga J, Pascual-Montano A, Nogales-Cadenas R, Santoyo J, Garcia F, Marba M, Montaner D, Dopazo J: **Babelomics: an integrative platform for the analysis of transcriptomics, proteomics and genomic data with advanced functional profiling.** *Nucleic Acids Res* 2010, **38**(suppl 2):W210–W213.
82. Al-Shahrour F, Minguéz P, Tarraga J, Medina I, Alloza E, Montaner D, Dopazo J: **FatiGO  $\pm$ : a functional profiling tool for genomic data. Integration of microarray annotation, regulatory motifs and interaction data with microarray experiments.** *Nucleic Acids Res* 2007, **35**(suppl 2):W91–W96.
83. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman W-H, Pagés F, Trajanoski Z, Galon JRM: **ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks.** *Bioinformatics* 2009, **25**(8):1091–1093.
84. Benjamini Y, Hochberg Y: **Controlling the false discovery rate: a practical and powerful approach to multiple testing.** *J Royal Stat Soc Series B (Methodological)* 1995, **57**(1):289–300.
85. Zhang Y, Klein K, Sugathan A, Nassery N, Dombkowski A, Zanger UM, Waxman DJ: **Transcriptional profiling of human liver identifies Sex-biased genes associated with polygenic dyslipidemia and coronary artery disease.** *PLoS ONE* 2011, **6**(8):e23506.
86. Chakrabarti P, English T, Karki S, Qiang L, Tao R, Kim J, Luo Z, Farmer SR, Kandror KV: **SIRT1 controls lipolysis in adipocytes via FOXO1-mediated expression of ATGL.** *J Lipid Res* 2011, **52**(9):1693–1701.
87. Schug TT, Li X: **Sirtuin 1 in lipid metabolism and obesity.** *Ann Med* 2011, **43**(3):198–211.
88. Picard F, Kurtev M, Chung N, Topark-Ngarm A, Senawong T, Machado De Oliveira R, Leid M, McBurney MW, Guarente L: **Sirt1 promotes fat**

mobilization in white adipocytes by repressing PPAR-gamma. *Nature* 2004, **429**(6993):771–776.

89. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F: **Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes.** *Genome Biol* 2002, **3**(7):RESEARCH0034.
90. Livak KJ, Schmittgen TD: **Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method.** *Methods* 2001, **25**(4):402–408.
91. Ballester M, Cerdón R, Folch JM: **DAG Expression: high-throughput gene expression analysis of Real-Time PCR data using standard curves for relative quantification.** *Plos One* 2013, **8**(11):e80385.

doi:10.1186/1471-2164-15-232

**Cite this article as:** Ramayo-Caldas *et al.*: From SNP co-association to RNA co-expression: Novel insights into gene networks for intramuscular fatty acid composition in porcine. *BMC Genomics* 2014 **15**:232.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

