Genome-wide phylogenetic analysis, expression pattern, and transcriptional regulatory network of the pig C/EBP gene family

Chaoxin Zhang  
Northeast Agricultural University

Tao Wang  
Northeast Agricultural University

Shengwei Liu  
Northeast Agricultural University

Bing Zhang  
Northeast Agricultural University

Xue Li  
Northeast Agricultural University

Jian Tang  
Northeast Agricultural University

Peng Wang  
Heilongjiang provincial Husbandry Dapartment

Yuanyuan Guo  
Northeast Agricultural University

Zhipeng Wang ( wangzhipeng@neau.edu.cn )  
School of animal science! Northeast Agricultural University  https://orcid.org/0000-0002-6041-8906

Research

Keywords: EBP family, Pig, Phylogenetic analysis, Transcriptional regulatory network, Expression pattern, Feed-forward loop

DOI: https://doi.org/10.21203/rs.3.rs-96827/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

Background: The vertebrate C/EBP transcription factors regulate many important biological processes, such as cell proliferation, differentiation, signal transduction, inflammation, and energy metabolism. The first C/EBP protein was identified in rat liver nuclei. Development of sequencing technology resulted in identification of the C/EBP genes in various species. In this study, a bioinformatics approach was used to determine the distribution of the members of the C/EBP family in vertebrates. A phylogenetic tree was constructed to analyze the C/EBP genes in vertebrates. Based on RNA-seq data, the expression patterns of pig C/EBP members in various tissues were analyzed. In addition, a gene transcription regulatory network was constructed with pig C/EBP members as the core.

Results: We identified a total of 92 C/EBP genes in 17 vertebrate genomes. Phylogenetic analysis showed that all C/EBP TFs were classified into two groups; group I contained C/EBPβ TFs, and group II contained the remaining C/EBP TFs. The C/EBPa, C/EBPβ, C/EBPδ, C/EBPγ, and C/EBPζ genes were expressed ubiquitously with inconsistent expression patterns in various tissues. Moreover, a pig C/EBP regulatory network was constructed, including C/EBP genes, TFs, and miRNAs. A total of 39 FFL motifs were detected in the pig C/EBP regulatory network. Based on the RNA-seq data, gene expression patterns related to this FFL sub-network were analyzed in 27 adult Duroc tissues. Certain FFL motifs may be tissue specific. Functional enrichment analysis indicated that C/EBP and its target genes are involved in many important biological pathways.

Conclusions: These results provide valuable information that clarifies the evolutionary relationships of the C/EBP family and contributes to the understanding of the biological function of C/EBP genes.

Introduction

The CCAAT/enhancer binding protein (C/EBP) family includes important regulators that activate or inhibit gene transcription in many tissues and is a subfamily of the basic leucine zipper (bZIP) transcription factor (TF) superfamily [1]. The C/EBP family consists of 6 genes: C/EBP alpha (C/EBPa), C/EBP beta (C/EBPβ), C/EBP delta (C/EBPδ), C/EBP epsilon (C/EBPe), C/EBP gamma (C/EBPγ), and C/EBP zeta (C/EBPζ) [2]. The first C/EBP protein was identified in rat liver nuclei [3]. Five additional C/EBP TFs had been identified in humans and rats [2]. Advances in the sequencing techniques resulted in the identification of the C/EBP genes in various species and investigation of the evolutionary relationship among the C/EBP genes. Qiu et al (2018) identified the C/EBPa, C/EBPβ, and C/EBPδ genes genome-wide in 20 vertebrate genomes, including Atlantic salmon (Salmo salar), Japanese frog (Nanorana pleskei), southern painted turtle (Chrysemys picta bellii), duck (Anas platyrhynchos), and human (Homo sapiens), and analyzed the evolutionary relationships among the C/EBPa, C/EBPβ, and C/EBPδ genes [4]. Thus, understanding of the evolution and function of the C/EBP genes requires genome-wide comparative studies.

Transcription factors play an important regulatory role at the transcriptional level by activating or repressing target genes to regulate gene expression [5]. The C/EBP TFs in the vertebrates regulate many important biological processes, such as cell proliferation, differentiation, signal transduction, inflammation, and energy metabolism [1, 4]. C/EBPa, C/EBPβ, and C/EBPδ are involved in the regulation of adipocyte differentiation [6] and mammary development [7]. In the pulmonary epithelium, C/EBPa regulates proliferation and differentiation dependent gene expression, whereas C/EBPβ and C/EBPδ regulate the expression of differentiation markers and are involved in the responses to injury and hormones [8].

The complex gene transcription regulatory network includes transcriptional regulation and post-transcriptional regulation [9]. MicroRNAs (miRNAs) are short noncoding RNAs that regulate gene expression at the post-transcriptional level by selectively targeting complementary mRNAs for degradation. MiR31 directly binds to the 3’-untranslated region (3’-UTR) of C/EBPa and inhibits its expression to mediate the adipogenic differentiation of human adipose tissue-derived mesenchymal stem cells [10]. In mouse dendritic cells, miR369 suppresses C/EBPβ to regulate the inflammatory cascade in the chronic inflammatory response [11]. Transcription factors and miRNAs often synergize in the regulation the same target gene, defined as feed-forward loop (FFL). For example, miR223-LMO2-C/EBPβ FFL plays an important role in human hematopoiesis [12].
In this study, we identified C/EBP genes in 17 vertebrates and analyzed physicochemical properties of the proteins, conservative motifs, gene structure, and phylogenetic relationships. Subsequently, the expression patterns of the C/EBP genes were analyzed. A transcriptional regulatory network containing the C/EBP genes was constructed to investigate the functions of the C/EBP genes. Our study provides a comprehensive analysis of the pig C/EBP gene family and important information for subsequent investigation of the functions in pig.

Materials And Methods

Identification of C/EBP TFs

The General Feature Format Version 3 (GFF3) profiles and the genome and protein sequence files of 17 vertebrates were downloaded from the ENSEMBL database (http://asia.ensembl.org/index.html) [13]. The vertebrates included zebrafish (Danio rerio), tropical xenopus (Xenopus tropicalis), Chinese soft-shell turtle (Pelodiscus sinensis), chicken (Gallus gallus), platypus (Ornithorhynchus anatinus), cat (Felis catus), dog (Canis familiaris), horse (Equus caballus), pig (Sus scrofa), cow (Bos taurus), goat (Capra hircus), sheep (Ovis aries), mouse (Mus musculus), rat (Rattus norvegicus), macaque (Macaca mulatta), chimpanzee (Pan troglodytes), and human (Homo sapiens). The hmmBuild tool of HMMER 3.0 [14] was used to build the Hidden Markov Model (HMM) of the C/EBP family. The hmmsearch tool search for all putative TFs of the C/EBP family was performed (cutoff value <1E-20). The Conserved Domain Database of NCBI (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) [15] and the SMART tool (http://smart.embl-heidelberg.de/) [16] were used to determine the conserved domains of each C/EBP protein. The longest protein of each C/EBP TF was retained.

Physicochemical parameters, exon-intron structure, and motif analysis

ProtParam (http://web.expasy.org/protparam/) [17] was used to analyze the physicochemical parameters of the C/EBP proteins, including molecular weight (MW), aliphatic index, isoelectric point (pI), and grand average of hydropathicity (GRAVY). The conserved motifs of the C/EBP proteins were identified by the online platform MEME (http://meme-suite.org/) [18]. The exon and intron structures of the C/EBP genes were obtained from the ENSEMBL gene annotation information. The diagrams of the gene structures were generated using the Gene Structure Display Server (GSDS) 2.0 (http://gsds.gao-lab.org/) [19].

Phylogenetic analysis

Maximum likelihood (ML) phylogenetic tree of C/EBP TFs was constructed using the Jones-Taylor-Thornton with Gamma Distributed (JTT+G) substitution model and 1,000 replicate bootstrap tests, and a cutoff bootstrap value of 65 was set to define the clades using MEGA (v10.0) [20]. Representations of the phylogenetic tree were constructed using the iTOL tool (https://itol.embl.de/) [21].

Construction of the transcriptional regulatory networks of the pig C/EBP genes

The Position Weight Matrix (PWM) of transcription factors for Sus scrofa was downloaded from the cisbp database (http://cisbp.ccb.r.utoronto.ca/) [22]. The genomic sequences 2 kb upstream of the translation start site (TSS) and 3'-UTR were extracted by the R package biomaRt [23] for all coding genes. The TFBSTools package of R language [24] was used to predict the target genes of C/EBPs and the transcription factors regulating C/EBP genes with a threshold relScore value of 0.85. The sequences of miRNAs were downloaded from the miRBase database (http://www.mirbase.org/) [25]. The miRanda software [26] was used to predict the miRNA targets with a threshold TotScore value of 120. Cytoscape 3.7.2 [27] was used to visualize the gene regulatory network, and the NetworkAnalyzer tool was used to calculate the network topology properties, including the clustering coefficient, network centralization, and network heterogeneity. The R software package basicTrendline [28] was used to analyze the network node degree distribution.

Expression analysis of the pig C/EBP genes and target genes
The raw RNA sequencing (RNA-seq) data obtained in 27 adult Duroc pig tissue samples were downloaded from the NCBI Sequence Read Archive with the BioProject number PRJNA392949 \[29\]. The tissues included fat, thyroid, lymph, ovary, uterus, breast, spleen, lung, liver, placenta, etc. Purification of the total RNA from the mixture of equally unrelated pig pool tissues was performed.

After QC step with FastQC v0.11.8 (http://www.bioinformatics.babraham.ac.uk/ projects/fastqc/), clean data were mapped and genome indexed with Hisat 0.1.6-beta 64-bit \[30\] to the pig genome (Sus scrofa 11.0). To obtain the expression levels of the genes across 27 tissues, the fragments per kilobase of exon model per million mapped reads (FPKM) values were calculated using Stringtie 1.0.4 (Linux x86 64) \[31\]. The TBtools software \[32\] was used to visualize the heatmaps of the gene expression profiles in 27 tissues.

The tissue specificity index (τ) of the C/EBP genes was also investigated. The tissue specificity index is defined as

\[
\tau = \frac{\sum_{i=1}^{N} (1-x_i)}{N-1},
\]

where \(x_i\) is the expression profile component normalized by the maximal component value and \(N\) is the number of tissues \[33\].

**Functional enrichment analysis**

The Database for Annotation, Visualization, and Integrated Discovery (DAVID) (https://david.ncifcrf.gov/) \[34\] was used to annotate the functions of C/EBPs and their target genes, which includes gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis.

**Results**

**Identification of C/EBP genes in the vertebrates**

A total of 92 C/EBP genes were identified in 17 vertebrate genomes. The number of C/EBP genes in each species was slightly different as shown in Table 1. In total, the C/EBP family included \(C/EBP\alpha\), \(C/EBP\beta\), \(C/EBP\delta\), \(C/EBP\epsilon\), \(C/EBP\gamma\), and \(C/EBP\zeta\). In particular, the \(C/EBP\alpha\) and \(C/EBP\gamma\) genes of all species are located on the same chromosome within a distance of less than 100 kb, except for Xenopus tropicalis.

Various physicochemical properties of each C/EBP TF were calculated and are shown in Table 2 and Table S1. The \(C/EBP\zeta\) protein contains the highest number of amino acids and has the highest molecular weight, and the \(C/EBP\gamma\) protein consists of the least number of amino acids with the lowest molecular weight. The aliphatic index, a measure of thermostability, ranged from 60.36 to 76.17. There are significant differences (p<2e-16) in the aliphatic index between C/EBP TFs. The GRAVY values of C/EBP proteins are negative indicating hydrophilic properties. The GRAVY values vary from -1.4060 to -0.4360. The \(pI\) values of \(C/EBP\beta\), \(C/EBP\epsilon\), and \(C/EBP\gamma\) proteins in each species are higher than 7, and the \(pI\) values of \(C/EBP\zeta\) are less than 7. The \(C/EBP\alpha\) protein of zebrafish is acidic, and \(C/EBP\alpha\) proteins of other species are alkaline. The \(C/EBP\delta\) protein is acidic in 1/3 species and alkaline in the remaining 2/3 species.

**Phylogenetic relationship analysis**

To analyze the phylogenetic relationships between 92 C/EBP genes in 17 species, an unrooted maximum likelihood phylogenetic tree was constructed as shown in Fig 1. All C/EBP TFs were classified into two groups. Group I contains 16 vertebrate animal \(C/EBP\beta\) TFs and was named \(C/EBP\beta\). Group II contains the remaining C/EBP TFs, which can be divided into five clades. Following the nomenclature, we named the clades of group II as \(C/EBP\alpha\), \(C/EBP\delta\), \(C/EBP\epsilon\), \(C/EBP\gamma\), and \(C/EBP\zeta\).

C/EBP genes were detected in most vertebrates indicating that the C/EBP family members originated in the early stage of vertebrate evolution. In each clade, C/EBP genes of the same order tend to cluster together indicating higher similarity to each other than to other orders of C/EBP genes. C/EBP orthologs of the order primates includes human, macaque, and chimpanzee; the order artiodactyls includes pig, cattle, and goat; the order carnivore includes cat and dog; and the order rodents (rat and mouse) genes are clustered together.
To investigate the structural features of the C/EBP members, the gene structure and conserved motifs were evaluated by the phylogenetic analysis, as shown in Fig 2. The number of exons in C/EBP genes, which contained 15 or 16 exons, was higher than that in the other C/EBP genes. Most of the C/EBPa, C/EBPβ, C/EBPδ, C/EBPe, and C/EBPy genes contained one or two exons. A total of 10 conserved motifs were identified in C/EBP proteins. All C/EBP TFs contain motif 2, motif 3, and motif 6. Motifs 8 and 10 are unique motifs in C/EBP proteins and may be associated with the clade-specific functions of the C/EBP proteins.

**Expression analysis of the C/EBP genes**

The expression of C/EBP genes was compared in 27 adult Duroc pig tissues. The 6 C/EBP genes from pig were classified into two groups based on the cluster analysis of the C/EBP gene expression levels in various tissues, as shown in Fig 3. Group I contains only the C/EBPs gene, and other genes are included in group II. The expression levels of the C/EBP gene were low in all tissues (FPKM<5), and the gene is expressed only in the intestine, salivary gland, thyroid, uterus, and lymph. Although C/EBPa, C/EBPβ, C/EBPδ, C/EBPe, and C/EBPy are expressed ubiquitously, the expression patterns in various tissues are inconsistent; the tissue specific index values (r) are 0.862, 0.700, 0.654, 0.499 and 0.433, respectively. C/EBPa was expressed at the high levels in the thyroid, liver, lung, and adipose tissues. C/EBPβ was expressed at the high levels in the thyroid, adrenal gland, lung, adipose, liver, and ovary tissues. C/EBPδ was expressed at the high levels in the thyroid, gall bladder, ovary, and uterus. C/EBPe and C/EBPy are widely expressed in other tissues at similar levels.

Additionally, the results indicate that the expression patterns in certain tissues, such as brain and spinal cord in the nervous system and ovary and uterus tissues in the female reproductive system, are similar.

**Construction of a transcriptional regulatory network of the C/EBP gene family**

The C/EBP family is an important family of transcription factors that regulate the expression of the target genes by binding to the promoter regions to maintain the normal physiological processes in vivo. According to the PWMs of the C/EBPa, C/EBPβ, C/EBPδ, C/EBPe, and C/EBPy genes from the cisbp database, we predicted 4,662, 3,164, 8,383, 7,278, and 1,604 target genes regulated by these genes, respectively (see Table 3). A total of 10,270 target genes are regulated by the C/EBP genes for a total of 25,091 regulatory relationships. Binding sites for other transcription factors and miRNAs are present in the regulatory regions of the C/EBP genes.

In this study, the C/EBP genes are predicted to be regulated by 423 TFs forming 1,582 regulatory relationships; mir503 and mir7140 are predicted to regulate C/EBPβ and C/EBPy, respectively. Additionally, C/EBPβ and C/EBPy are regulated by C/EBPa, C/EBPδ, and C/EBPe and C/EBPβ and C/EBPe, respectively. Interestingly, C/EBPβ self-regulation is also predicted. Thus, we constructed a pig C/EBP regulatory network (summarized in Fig 4) that includes C/EBP genes, TFs, miRNAs, and target genes. These genes were defined as a node, and the distribution of node degree approximately follows the power-law distribution indicating that the gene regulatory network is a scale-free network. Certain network concepts, including the clustering coefficient, network centralization, and network heterogeneity, were calculated to be 0.1890, 0.8160, and 25.1730, respectively.

**FFLs related to the C/EBP genes**

A total of 27 miRNA-FFL motifs were identified in the pig C/EBP regulatory network. According to the sequences, we predicted that mir503 and mir7140 regulate C/EBPβ and C/EBPy, respectively. The mir503 and C/EBPβ genes coregulate 14 target genes forming 14 mir503-C/EBPβ-target gene FFL motifs, and mir7140 and C/EBPy coregulate 11 target genes, including 11 mir7140-C/EBPy-target gene FFL motifs. Additionally, mir503-ELF3-C/EBPβ and mir7140-ARID5B-C/EBPy motifs were identified. The mir503-ELF3-C/EBPβ motif is involved in the mir503 and ELF3 coregulation of the C/EBPβ gene, and the mir503 gene targets the ELF3 and C/EBPβ genes. The mir7140-ARID5B-C/EBPy motif is included in 3 regulatory relationships: mir7140→ARID5B, ARID5B→C/EBPβ, and mir7140→C/EBPy. The C/EBPβ-binding sites in 5'-untranslated region (5'-UTR) of C/EBPβ, and C/EBPβ and C/EBPy co-regulate 10 target genes forming 10 TF-FFL motifs. Thus, the combinations of all FFL motifs were used to construct the FFL sub-network (see Fig 5).
Based on the data of RNA-seq, gene expression patterns related to this FFL sub-network were analyzed in 27 adult Duroc tissues. The results indicate that target genes regulated by C/EBPβ and C/EBPγ have variable expression patterns in various tissues. The ATP synthase F1 subunit alpha (ATP5F1A) gene is expressed ubiquitously, and the glutamate decarboxylase-like protein 1 (GADL1) and Slit-Robo GTPase activating protein 3 (SRGAP3) genes are expressed at the high levels in the muscle and brain, respectively. We suggest that the C/EBPβ-C/EBPγ-GADL1 FFL motif may play an important role in the brain. Some FFL motifs may be tissue-specific. Based on the target genes expression pattern, we estimated that the number of FFL motifs in each tissue may be significantly different (see Table 4); however, miRNA expression patterns were not evaluated in the present study.

**The dN and dS analysis of the C/EBP genes and target genes**

The data on the nonsynonymous (dN) and synonymous (dS) substitution rates between the human and pig sequences were downloaded from the Ensembl database. The dN/dS values of the C/EBP genes ranged from 0.02 to 0.21 indicating that pig C/EBP genes underwent purifying selection. The dN+dS value of the C/EBPδ gene was 1.52, which was higher than that of five other C/EBP genes (0.31 ~ 0.67) indicating that the C/EBPδ gene evolved rapidly and had an increased mutation rate.

The dN+dS mean values of the target genes of C/EBPα, C/EBPβ, C/EBPδ, C/EBPε, and C/EBPγ are 0.58, 0.48, 0.56, 0.53, and 0.51, respectively. The dN/dS mean values of the target genes are 0.166, 0.175, 0.17, 0.177, and 0.166, respectively. The dN/dS and dN+dS mean values of the target genes of each C/EBP gene were compared using Kolmogorov-Smirnov (KS) test. The results indicate that the dN/dS distributions of the target genes of C/EBPα are similar to that of C/EBPβ, C/EBPε and C/EBPγ (p<0.05), respectively. The very low dN/dS values suggest strong negative selection on all C/EBP genes, which may remain due to genetic drift or persistence. The dN+dS value distribution of the C/EBPα target genes is similar to that of the C/EBPδ target genes and is significantly higher than that of other C/EBP target genes (P<0.05). The results indicate that the target genes of C/EBPα appear to be evolving rapidly.

**Functional enrichment analysis of the C/EBP genes and target genes**

We used the DAVID software to analyze the functions of the pig C/EBP genes. The results indicate that the functions are associated with many biological processes, including macrophage differentiation (GO: 0030225), inner ear development (GO: 0048839), positive regulation of osteoblast differentiation (GO: 0045669), transcriptional misregulation in cancer pathways (ssc05202), and tuberculosis pathways (ssc05152) (Table 5).

The functional enrichment analysis of the target genes regulated by the C/EBP genes showed that the target genes of C/EBPα, C/EBPβ, C/EBPε, and C/EBPγ are associated with nucleoplasm (GO: 0005654) and extracellular exosome (GO: 0070062). The target genes of C/EBPδ and C/EBPγ are involved in the transforming growth factor beta (TGFβ) receptor signaling pathway. The target genes of C/EBPδ are involved in the platelet-derived growth factor receptor signaling (GO: 0048008) (Table 6).

**Discussion**

Improved genome sequencing and annotation enabled identification of all C/EBP TFs in the vertebrate genomes. This study identified all C/EBP TFs in 17 vertebrates and found that the C/EBP genes are ubiquitous in the majority of the vertebrates. The molecular phylogenetic tree showed that the C/EBP family members originated in the early stage of vertebrate evolution. Similarly, Qiu et al. analyzed the evolution of the C/EBPα, C/EBPβ, and C/EBPδ genes in 20 vertebrate genomes and demonstrated that the C/EBP TFs originated early in vertebrate evolution [4].

According to the phylogenetic tree, the C/EBPβ orthologs are clustered into a single group, and other C/EBP genes are clustered into another group. However, Qiu et al. analyzed the phylogenetic trees of C/EBPα, C/EBPβ, and C/EBPδ homologs in 2 fish, 2 amphibian, 2 reptile, 11 avian, and 3 mammalian species. The results indicated that the C/EBPα and C/EBPβ genes are clustered into one group. This difference may be due to different species and members of the C/EBP family.
Based on the data of RNA-seq, the $C/EBP\alpha$, $C/EBP\beta$, $C/EBP\delta$, $C/EBP\gamma$, and $C/EBP\zeta$ genes are expressed in all 27 adult Duroc tissues, and $C/EBP\epsilon$ is expressed only in 5 tissues. Summers et al. (2020) created a pig gene expression atlas based on the meta-analysis and demonstrated that $C/EBP\epsilon$ is expressed only in a few tissues, and other C/EBP genes are expressed in the majority of tissues [35]. Uhlen et al. (2015) studied the expression patterns of the C/EBP genes in 32 human tissues and demonstrated that the $C/EBP\beta$, $C/EBP\delta$, $C/EBP\gamma$, and $C/EBP\zeta$ genes are expressed ubiquitously [36].

The differences in the exon-intron, motif structures and gene expression patterns among various clades provide some indications that the C/EBP genes may have a variety of physiological functions. Our data indicate that $C/EBP\alpha$ and $C/EBP\beta$ are expressed at the high levels in the adipose, ovarian, and hepatic tissues. Chen et al (2017) demonstrated that $C/EBP\alpha$ is the lipogenic marker of the porcine adipocytes that promotes adipocyte differentiation [37]. Wang et al. (2015) reported that $C/EBP\beta$ is expressed in porcine adipocytes to regulate adipocyte differentiation [38]. Gillio-Meina et al. (2005) reported that $C/EBP\alpha$ and $C/EBP\beta$ in the ovary regulate the differentiation and proliferation of granulosa cells [39]. Tang et al. (2015) demonstrated that $C/EBP\beta$ regulates insulin-like growth factor 1 (IGF1) expression in the crossbred porcine (Yorkshire * Erhualian pigs) liver during prenatal and postnatal development [40].

The C/EBP genes perform many important biological functions. Functional enrichment analysis demonstrated that pig C/EBP genes are involved in the inner ear development and positive regulation of osteoblast differentiation biological processes and the tuberculosis pathway. C/EBP genes are also involved in various biological functions. For example, $C/EBP\zeta$ is involved in cochlear cell apoptosis in rats [41] and may coregulate the inner ear development with GATA binding protein 2 (GATA2) [42]. Hata et al (2005) reported that $C/EBP\beta$ promotes the differentiation of mouse mesenchymal cells into the osteoblast lineage in cooperation with Runt-related transcription factor 2 (Runx2) transcription factor [43]. Allele-specific induction of Interleukin 1 alpha (IL-1\(\alpha\)) expression by $C/EBP\beta$ contributes to increased tuberculosis susceptibility in humans [44].

Complex regulatory mechanisms of gene expression in eukaryotes control the development, physiology, and pathology [45]. In this study, we constructed a C/EBP gene regulatory network to provide some information to explain the regulatory mechanisms of the pig C/EBP genes. The results indicate that the C/EBP gene regulatory network is a typical scale-free network. A gene regulatory network influencing triple-negative breast tumors [46] and a human TF and miRNA regulatory network [47] were reported to correspond to approximately scale-free topology.

In the regulatory gene network, transcription factors play an important role at the gene transcription level due to the regulation of certain target genes. Each TF contains different conservation motifs to target various genes. In this study, the target genes of the porcine C/EBP TFs were different. In the Database of Human Transcription Factor Targets [48], the target genes of the C/EBP TFs are also different. Comparison with the target genes of the human C/EBP TFs indicated that porcine $C/EBP\delta$ regulates more target genes, but the number of the target genes of other porcine C/EBP TFs is less than that in humans. Selection of different target genes during evolution to meet the needs of normal homeostatic development and growth or to adapt to various environmental stresses may differ in different species.

The biological network is composed of motifs, and the FFL is an important motif. In this study, 39 FFLs were detected in the C/EBP genes of the transcriptional regulatory sub-network. The FFL-containing C/EBP genes play important roles in multiple tissues. Sun et al. (2010) reported that $\text{miR223}{\rightarrow}C/EBP\beta\rightarrow LMO2$ FFL in human myeloid cells can regulate cell proliferation and maintain normal differentiation and development [12]. Ponomarev et al. (2011) have reported that $\text{MiR124-C/EBP}\alpha$-$PU.1$ FFL is related to the transition between an activated phenotype and a quiescent state of macrophages [49]. Shi et al. (2018) reported that 28 FFLs of $C/EBP\alpha$, $C/EBP\beta$, and $C/EBP\delta$ may be associated with occurrence of human hypertrophic cardiomyopathy [50].

**Conclusion**

In the present study, we identified all C/EBP TFs of 17 vertebrate species and demonstrated that the C/EBP genes are ubiquitous in the majority of the vertebrates. All C/EBP genes were classified into six subgroups with similar number of gene exons and motif compositions within each subgroup. The porcine C/EBP genes are expressed generally, except $C/EBP\epsilon$ gene,
and participate in many important biological functions. The C/EBP gene regulatory network was constructed to provide information on the regulatory mechanisms of the pig C/EBP genes. A total of 39 FFL motifs were identified with variable expression patterns in various tissues. Some FFL motifs may be tissues-specific. Thus, our study provides potential functional insight into the roles of the C/EBP genes in pig.

**Abbreviations**

τ: tissue specificity index; 3′-UTR: 3′-untranslated region; 5′UTR: 5′-untranslated region; ARID5B: AT-rich interactive domain-containing protein 5B; ATPSF1A: ATP synthase F1 subunit alpha; BP: biological processes; bZIP: basic-leucine zipper; C/EBP: CCAAT/enhancer binding proteins; CC: cell component; DAVID: Database for Annotation, Visualization, and Integrated Discovery; dN: nonsynonymous; dS: synonymous; ELF3: E-74 like factor 3; FFL: feed-forward loop; FPKM: fragments per kilobase of exon model per million mapped reads; GADL1: glutamate decarboxylase-like protein 1; GATA2: GATA binding protein 2; GFF3: General Feature Format Version 3; GO: gene ontology; GRAVY: grand average of hydropathicity; GSDS: Gene Structure Display Server; HMM: Hidden Markov model; IL-1b: Interleukin 1 beta; IGF1: insulin-like growth factor 1; JTT + G: Jones-Taylor-Thorton with Gamma; KEGG: Kyoto Encyclopedia of Genes and Genomes; KS: Kolmogorov-Smirnov; LMO2: LIM domain only 2; MF: molecular function; miRNA: MicroRNA; ML: Maximum likelihood; MW: molecular weight; PI: isoelectric point; PU.1: Spi transcription factor; PWM: Position Weight Matrix; RNA-seq: RNA sequencing; Runx2: Runt-related transcription factor 2; SRGAP3: SLIT-ROBO Rho GTPase activating protein 3; TF: transcription factor; TGFβ: transforming growth factor beta; TSS: translation start site.

**Declarations**

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Availability of data and material**

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

**Competing interests**

The authors declare that they have no competing interests.

**Funding**

This work was financially supported by Natural Science Foundation of China (No. 32070571), the Academic Backbone Project of Northeast Agricultural University (No.15XG14), NEAU Research Founding for Excellent Young Teachers (2010RCB29).

**Authors’ contributions**

ZPW and CXZ conceived the study, and ZPW, CXZ, TW and SWL participated in its design. BZ, XL, JT and PW were involved in the acquisition of data, CXZ and TW performed all data analysis. ZPW and CXZ drafted the manuscript, and SWL, YYG, BZ, XL, JT and PW contributed to the writing and editing. All authors read and approved the final manuscript.

**Acknowledgements**

Not applicable.
Authors' details

1College of Animal Science and Technology, Northeast Agricultural University, Harbin 150000, China. 2Bioinformatics Center, Northeast Agricultural University, Harbin 150000, China. 3HeiLongJiang provincial Husbandry Dapartment, Harbin 150000, China.

References

1. Ramji DP, Foka P. CCAAT/enhancer-binding proteins: structure, function and regulation. Biochem J. 2002;365(Pt 3):561-575. doi:10.1042/BJ20020508.

2. Lekstrom-Himes J, Xanthopoulos KG. Biological role of the CCAAT/enhancer-binding protein family of transcription factors. J Biol Chem. 1998;273(44):28545-28548. doi:10.1074/jbc.273.44.28545.

3. Graves BJ, Johnson PF, McKnight SL. Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene. Cell. 1986;44(4):565-576. doi:10.1016/0092-8674(86)90266-7.

4. Qiu J-M, Wang W-X, Hu S-Q, Wang Y-S, Sun W-Q, Hu JW, et al. Molecular cloning, characterization and expression analysis of C/EBP α, β and δ in adipose-related tissues and adipocyte of duck (Anas platyrhynchos). Comp Biochem Physiol B Biochem Mol Biol. 2018;221-222:29-43. doi:10.1016/j.cbpb.2018.04.004.

5. Lambert M, Jambon S, Depauw S, David-Cordonnier MH. Targeting Transcription Factors for Cancer Treatment. Molecules. 2018;23(6):1479. Published 2018 Jun 19. doi:10.3390/molecules23061479.

6. Ross SE, Hemati N, Longo KA, Bennett CN, Lucas PC, Erickson RL, et al.. Inhibition of adipogenesis by Wnt signaling. Science. 2000;289(5481):950-953. doi:10.1126/science.289.5481.950.

7. Zahnow CA. CCAAT/enhancer binding proteins in normal mammary development and breast cancer. Breast Cancer Res. 2002;4(3):113-121. doi:10.1186/bcr428.

8. Cassel TN, Nord M. C/EBP transcription factors in the lung epithelium. Am J Physiol Lung Cell Mol Physiol. 2003;285(4):L773-L781. doi:10.1152/ajplung.00023.2003.

9. Graves P, Zeng Y. Biogenesis of mammalian microRNAs: a global view. Genomics Proteomics Bioinformatics. 2012;10(5):239-245. doi:10.1016/j.gpb.2012.06.004.

10. Liu Y-T, Wang Y-Q, He X-F, Zhang S, Wang K, Wu H-J, et al. LncRNA TINCR/miR-31-5p/C/EBP-α feedback loop modulates the adipogenic differentiation process in human adipose tissue-derived mesenchymal stem cells. Stem Cell Res. 2018;32:35-42. doi:10.1016/j.scr.2018.08.016.

11. Galleggiante V, De Santis S, Liso M, Verna G, Sommella E, Mastronardi M, et al. Quercetin-Induced miR-369-3p Suppresses Chronic Inflammatory Response Targeting C/EBP-β. Mol Nutr Food Res. 2019;63(19):e1801390. doi:10.1002/mnfr.201801390.

12. Sun W, Shen W-W, Yang S, Hu F, Li H-H, Zhu T-H. miR-223 and miR-142 attenuate hematopoietic cell proliferation, and miR-223 positively regulates miR-142 through LMO2 isoforms and CEBP-β. Cell Res. 2010;20(10):1158-1169. doi:10.1038/cr.2010.134.

13. Cunningham F, Achuthan P, Akanni W, Allen J, Amode MR, Armean IM, et al. Ensembl 2019. Nucleic Acids Res. 2019;47(D1):D745-D751. doi:10.1093/nar/gky1113

14. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. HMMER web server: 2018 update. Nucleic Acids Res. 2018;46(W1):W200-W204. doi:10.1093/nar/gky1113

15. Lu S-N, Wang J-Y, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, et al. CDD/SPARCLE: the conserved domain database in 2020. Nucleic Acids Res. 2020;48(D1):D265-D268. doi:10.1093/nar/gkz991.

16. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, et al. The Pfam protein families database in 2019. Nucleic Acids Res. 2019;47(D1):D427-D432. doi:10.1093/nar/gky995.

17. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins M R, Appel R D, et al. Protein Identification and Analysis Tools on the ExPASy Server. In John M, Walker, editors: The Proteomics Protocols Handbook, Humana Press (2005). p. 571-607.
18. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009;37:W202-W208. doi:10.1093/nar/gkp335.

19. Hu B, Jin J-P, Guo A-Y, Zhang H, Luo J-C, Gao G. GDSS 2.0: an upgraded gene feature visualization server. Bioinformatics. 2015;31(8):1296-1297. doi:10.1093/bioinformatics/btu817.

20. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol. 2018;35(6):1547-1549. doi:10.1093/molbev/msy096.

21. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 2019;47(W1):W256-W259. doi:10.1093/nar/gkz239.

22. Weirauch MT, Yang A, Albu M, Cote AG, Montenegro-Montero A, Drewe P, et al. Determination and inference of eukaryotic transcription factor sequence specificity. Cell. 2014;158(6):1431-1443. doi:10.1016/j.cell.2014.08.009.

23. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat Protoc. 2009;4(8):1184-1191. doi:10.1038/nprot.2009.97.

24. Tan G, Lenhard B. TFBSTools: an R/bioconductor package for transcription factor binding site analysis. Bioinformatics. 2016;32(10):1555-1556. doi:10.1093/bioinformatics/btw024.

25. Kozomara A, Birgaoanu M, Griths-Jones S. miRBase: from microRNA sequences to function. Nucleic Acids Res. 2019;47(D1):D155-D162. doi:10.1093/nar/gky1141.

26. Enright AJ, John B, Gaul U, Tuschi T, Sander C, Marks DS. MicroRNA targets in Drosophila. Genome Biol. 2003;5(1):R1. doi:10.1186/gb-2003-5-1-r1.

27. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498-2504. doi:10.1101/gr.1239303.

28. Mei W-P and Yu G-C. basicTrendline: Add Trendline and Confidence Interval of Basic Regression Models to Plot. R package version 2.0.3. 2018. https://CRAN.R-project.org/package=basicTrendline.

29. Zhao P-J, Zheng X-R, Feng W,Wang H-F, Wang H-M, Ning C, et al. Proling long noncoding RNA of multi-tissue transcriptome enhances porcine noncoding genome annotation. Epigenomics. 2018;10(3):301-320. doi:10.2217/epi-2017-0149.

30. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12(4):357-360. doi:10.1038/nmeth.3317.

31. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015;33(3):290-295. doi:10.1038/nbt.3122.

32. Chen C-J, Chen H, Zhang Y, Thomas HR, Frank MH, He Y-H, et al. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. Mol Plant. 2020;13(8):1194-1202. doi:10.1016/j.molp.2020.06.009.

33. Yanai I, Benjamin H, Shmoish M, Chalifa-Caspi V, Shklar M, Ophir R, et al. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. Bioinformatics. 2005;21(5):650-659. doi:10.1093/bioinformatics/bti042.

34. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44-57. doi:10.1038/nprot.2008.211.

35. Summers KM, Bush SJ, Wu C, Su AI, Muriuki C, Clark EL, et al. Functional Annotation of the Transcriptome of the Pig, Sus scrofa, Based Upon Network Analysis of an RNAseq Transcriptional Atlas. Front Genet. 2020;10:1355. Published 2020 Feb 14. doi:10.3389/fgene.2019.01355.

36. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. Science. 2015;347(6220):1260419. doi:10.1126/science.1260419.

37. Chen X-L, Luo Y-L, Jia G, Liu G-G, Zhao H, Huang Z-Q. The effect of arginine on the Wnt/β-catenin signaling pathway during porcine intramuscular preadipocyte differentiation. Food Func. 2017;8(1):381-386. doi:10.1039/c6fo01452f.

38. Wang Y-L, Liu X-Y, Hou L-M, Wu W-J, Zhao S-H, Xiong Y-Z. Fibroblast Growth Factor 21 Suppresses Adipogenesis in Pig Intramuscular Fat Cells. Int J Mol Sci. 2015;17(1):11. Published 2015 Dec 23. doi:10.3390/ijms17010011.
39. Gillio-Meina C, Hui YY, LaVoie HA. Expression of CCAAT/enhancer binding proteins alpha and beta in the porcine ovary and regulation in primary cultures of granulosa cells. Biol Reprod. 2005;72(5):1194-1204. doi:10.1095/biolreprod.104.035246.

40. Tang Y-T, Xiong K, Shen M, Mu Y-L, Li K, Liu H-L. CCAAT-enhancer binding protein (C/EBP) β regulates insulin-like growth factor (IGF) 1 expression in porcine liver during prenatal and postnatal development. Mol Cell Biochem. 2015;401(1-2):209-218. doi:10.1007/s11010-014-2308-8.

41. Zong S-M, Liu T-Y, Wan F-M, Chen P, Luo P, Xiao H-J. Endoplasmic Reticulum Stress Is Involved in Cochlear Cell Apoptosis in a Cisplatin-Induced Ototoxicity Rat Model. Audiol Neurootol. 2017;22(3):160-168. doi:10.1159/000480346.

42. Holley MC, Kneebone A, Milo M. Information for gene networks in inner ear development: a study centered on the transcription factor gata2. Hear Res. 2007;227(1-2):32-40. doi:10.1016/j.heares.2006.04.006.

43. Hata K, Nishimura R, Ueda M, Ikeda F, Matsubara T, Ichida F, et al. A CCAAT/enhancer binding protein beta isoform, liver-enriched inhibitory protein, regulates commitment of osteoblasts and adipocytes. Mol Cell Biol. 2005;25(5):1971-1979. doi:10.1128/MCB.25.5.1971-1979.2005.

44. Zhang G-L, Zhou B-P, Li S-Y, Yue J, Yang H, Wen Y-X, et al. Allele-specific induction of IL-1β expression by C/EBPβ and PU.1 contributes to increased tuberculosis susceptibility. PLoS Pathog. 2014;10(10):e1004426. Published 2014 Oct 16. doi:10.1371/journal.ppat.1004426.

45. Lambert M, Jambon S, Depauw S, David-Cordonnier MH. Targeting Transcription Factors for Cancer Treatment. Molecules. 2018;23(6):1479. Published 2018 Jun 19. doi:10.3390/molecules23061479.

46. Jung HC, Kim SH, Lee JH, Kim JH, Han SW. Gene Regulatory Network Analysis for Triple-Negative Breast Neoplasms by Using Gene Expression Data. J Breast Cancer. 2017;20(3):240-245. doi:10.4048/jbc.2017.20.3.240.

47. Jiang W, Mitra R, Lin C-C, Wang Q, Cheng F-X, Zhao Z-M. Systematic dissection of dysregulated transcription factor-miRNA feed-forward loops across tumor types. Brief Bioinform. 2016;17(6):996-1008. doi:10.1093/bib/bbv107.

48. Zhang Q, Liu W, Zhang H-M, Xie G-Y, Miao Y-R, Xia M-X, et al. hTFtarget: a comprehensive database for regulations of human transcription factors and their targets. Genomics Proteomics Bioinformatics. 2020. doi:10.1016/j.gpb.2019.09.006.

49. Ponomarev ED, Veremeyko T, Barteneva N, Krichevsky AM, Weiner HL. MicroRNA-124 promotes microglia quiescence and suppresses EAE by deactivating macrophages via the C/EBP-α-PU.1 pathway. Nat Med. 2011;17(1):64-70. doi:10.1038/nm.2266.

50. Shi H-B, Li J-Y, Song Q, Cheng L, Sun H-R, Fan W-J, et al. Systematic identification and analysis of dysregulated miRNA and transcription factor feed-forward loops in hypertrophic cardiomyopathy. J Cell Mol Med. 2019;23(1):306-316. doi:10.1111/jcmm.13928.

### Tables

Table 1 Genomic distribution of C/EBP genes on vertebrates genome.
| Taxonomy       | Genes        | Total |
|---------------|--------------|-------|
| Class         | Order        | Species                  |
| Pisces        | Cypriniformes| Zebrafish                | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| Amphibians    | Anuras       | Tropical xenopus         | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| Reptiles      | Chelonias    | Chinese soft-shell turtle| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 4 |
| Fishes        | Galliformes  | Chicken                  | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 4 |
| Mammals       | Monotrematas | Platypus                 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 5 |
|               | Carnivoras   | Cat                      | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
|               |              | Dog                      | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 5 |
|               | Perissodactylas | Horse                | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
|               | Artiodactylas| Pig                      | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
|               |              | Cattle                   | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 4 |
|               |              | Goat                     | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
|               |              | Sheep                    | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
|               | Rodents      | Mouse                    | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
|               |              | Rat                      | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
|               | Primates     | Macaque                  | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 5 |
|               |              | Chimpanzee               | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 5 |
|               |              | Human                    | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |

Table 2 The physicochemical properties of C/EBP proteins.

| Proteins | Number of amino acids | MW (mean) | pI | Aliphatic index (mean) | GRAVY (mean) |
|----------|-----------------------|-----------|----|------------------------|--------------|
| C/EBPα   | 135~395               | 32.81     | 6.81~10.78 | 61.69 | -0.77 |
| C/EBPβ   | 126~348               | 32.16     | 8.3~10.37 | 60.74 | -0.63 |
| C/EBPδ   | 99~302                | 26.99     | 5.43~10.73 | 60.36 | -0.73 |
| C/EBPε   | 146~460               | 30.65     | 9.02~10.29 | 67.94 | -0.74 |
| C/EBPγ   | 137~465               | 19.02     | 9.43~11.26 | 70.16 | -0.95 |
| C/EBPζ   | 998~1074              | 119.81    | 5.31~6.03 | 76.17 | -0.71 |

Table 3 Regulatory relationships involved in C/EBP family members.
Table 4 Distribution of FFLs in various tissues. In the each subnetworks, red represents higher expression levels; blue represents lower expression levels.

Due to technical limitations, table 4 is only available as a download in the Supplemental Files section.

Table 5 The results of functional enrichment analysis of porcine C/EBP family members.

| Category | ID    | Term                                                                 | P value   | C/EBP Genes                  |
|----------|-------|----------------------------------------------------------------------|-----------|------------------------------|
| GO_MF    | GO:0001077 | Transcriptional activator activity, RNA polymerase II core promoter proximal region sequence-specific binding | 3.60E-07  | C/EBPα, C/EBPε, C/EBPδ, C/EBPγ, C/EBPζ |
|          | GO:0000978 | RNA polymerase II core promoter proximal region sequence-specific DNA binding | 1.40E-06  | C/EBPα, C/EBPε, C/EBPδ, C/EBPγ, C/EBPζ |
| GO_BP    | GO:0030225 | Macrophage differentiation                                           | 3.10E-03  | C/EBPα, C/EBPε               |
|          | GO:0048839 | Inner ear development                                                | 9.60E-03  | C/EBPα, C/EBPδ               |
|          | GO:0045669 | Positive regulation of osteoblast differentiation                     | 1.40E-02  | C/EBPα, C/EBPδ               |
|          | GO:0051091 | Positive regulation of sequence-specific DNA binding transcription factor activity | 2.20E-02 | C/EBPβ                      |
| KEGG_PATHWAY | ssc05202 | Transcriptional misregulation in cancer                              | 4.00E-05  | C/EBPα, C/EBPβ, C/EBPε       |
|          | ssc05152 | Tuberculosis                                                        | 3.60E-03  | C/EBPβ, C/EBPγ               |

Table 6 The results of functional enrichment analysis of porcine C/EBP target genes.
| Category | ID | Term | \( P \) value |
|----------|----|------|--------------|
| \( C/EBP\alpha \) target genes | GO_CC | GO:0005654 | Nucleoplasm | 1.37E-09 |
| | | GO:0070062 | Extracellular exosome | 8.76E-09 |
| | | GO:005634 | Nucleus | 1.81E-05 |
| | | GO:005737 | Cytoplasm | 5.53E-05 |
| | GO_MF | GO:0044822 | Poly(A) RNA binding | 1.49E-05 |
| | | GO:003824 | Catalytic activity | 5.88E-05 |
| | KEGG_PATHWAY | ssc01100 | Metabolic pathways | 4.17E-06 |
| \( C/EBP\beta \) target genes | GO_CC | GO:0070062 | Extracellular exosome | 3.63E-09 |
| | | GO:005654 | Nucleoplasm | 7.83E-06 |
| | KEGG_PATHWAY | ssc01100 | Metabolic pathways | 4.21E-06 |
| | | ssc00310 | Lysine degradation | 1.40E-04 |
| \( C/EBP\delta \) target genes | GO_BP | GO:001649 | Osteoblast differentiation | 5.31E-05 |
| | | GO:000122 | Negative regulation of transcription from RNA polymerase II promoter | 2.02E-04 |
| | | GO:0045944 | Positive regulation of transcription from RNA polymerase II promoter | 2.91E-04 |
| | | GO:007179 | Transforming growth factor beta receptor signaling pathway | 4.21E-04 |
| | | GO:0098779 | Mitophagy in response to mitochondrial depolarization | 6.24E-04 |
| | | GO:001525 | Angiogenesis | 8.71E-04 |
| | | GO:007420 | Brain development | 1.06E-03 |
| | | GO:1901998 | Toxin transport | 1.04E-03 |
| | | GO:007507 | Heart development | 1.49E-03 |
| | | GO:0048008 | Platelet-derived growth factor receptor signaling pathway | 2.18E-03 |
| | | GO:0007059 | Chromosome segregation | 2.01E-03 |
| | | GO:0007519 | Skeletal muscle tissue development | 3.54E-
| GO:0060325 | Face morphogenesis | 2.18E-03 |
| GO:0070372 | Regulation of ERK1 and ERK2 cascade | 1.89E-03 |
| GO:0045732 | Positive regulation of protein catabolic process | 3.25E-03 |
| GO:0006096 | Glycolytic process | 4.34E-03 |
| GO:0022008 | Neurogenesis | 5.97E-03 |
| GO:0032212 | Positive regulation of telomere maintenance via telomerase | 3.82E-03 |
| GO:0043967 | Histone H4 acetylation | 4.59E-03 |
| GO:0010628 | Positive regulation of gene expression | 8.20E-03 |

| C/EBPε target genes | GO_CC | GO:0070062 | Extracellular exosome | 1.88E-12 |
| GO:0005654 | Nucleoplasm | 1.81E-12 |
| GO:0005737 | Cytoplasm | 2.14E-07 |
| GO:0005829 | Cytosol | 1.75E-07 |
| GO:0016020 | Membrane | 7.11E-06 |

| GO:0044822 | Poly(A) RNA binding | 2.81E-07 |
| GO:0005524 | ATP binding | 8.89E-06 |

| KEGG_PATHWAY | ssc01100 | Metabolic pathways | 9.13E-09 |

| C/EBPγ target genes | GO_CC | GO:0005654 | Nucleoplasm | 3.28E-05 |
| GO:0005654 | Nucleoplasm | 3.28E-05 |
| GO:0070062 | Extracellular exosome | 9.08E-05 |

| GO_BP | GO:0030512 | Negative regulation of transforming growth factor beta receptor signaling pathway | 1.29E-05 |

**Figures**
To analyze the phylogenetic relationships between 92 C/EBP genes in 17 species, an unrooted maximum likelihood phylogenetic tree was constructed as shown in Fig 1.

**Figure 1**

To analyze the phylogenetic relationships between 92 C/EBP genes in 17 species, an unrooted maximum likelihood phylogenetic tree was constructed as shown in Fig 1.
To investigate the structural features of the C/EBP members, the gene structure and conserved motifs were evaluated by the phylogenetic analysis, as shown in Fig 2.
Figure 3

The 6 C/EBP genes from pig were classified into two groups based on the cluster analysis of the C/EBP gene expression levels in various tissues, as shown in Fig 3.

Figure 4

We constructed a pig C/EBP regulatory network (summarized in Fig 4) that includes C/EBP genes, TFs, miRNAs, and target genes.
Figure 5

the combinations of all FFL motifs were used to construct the FFL sub-network.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table.docx
- TableS1.xlsx