

# Genome-Wide Phylogenetic Analysis, Expression Pattern, and Transcriptional Regulatory Network of the Pig C/EBP Gene Family

Evolutionary Bioinformatics  
Volume 17: 1–14  
© The Author(s) 2021  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/11769343211041382



Chaoxin Zhang<sup>1,2\*</sup>, Tao Wang<sup>1,2\*</sup>, Tongyan Cui<sup>1,2</sup>, Shengwei Liu<sup>1,2</sup>,  
Bing Zhang<sup>1,2</sup>, Xue Li<sup>1,2</sup>, Jian Tang<sup>1,2</sup>, Peng Wang<sup>3</sup>,  
Yuanyuan Guo<sup>1,2</sup> and Zhipeng Wang<sup>1,2,4</sup>

<sup>1</sup>College of Animal Science and Technology, Northeast Agricultural University, Harbin, China.

<sup>2</sup>Bioinformatics Center, Northeast Agricultural University, Harbin, China. <sup>3</sup>HeiLongJiang provincial Husbandry Department, Harbin, China. <sup>4</sup>DaBeiNong Group, Beijing, China.

**ABSTRACT:** The CCAAT/enhancer binding protein (C/EBP) transcription factors (TFs) regulate many important biological processes, such as energy metabolism, inflammation, cell proliferation etc. A genome-wide gene identification revealed the presence of a total of 99 C/EBP genes in pig and 19 eukaryote genomes. Phylogenetic analysis showed that all C/EBP TFs were classified into 6 subgroups named C/EBP $\alpha$ , C/EBP $\beta$ , C/EBP $\delta$ , C/EBP $\epsilon$ , C/EBP $\gamma$ , and C/EBP $\zeta$ . Gene expression analysis showed that the C/EBP $\alpha$ , C/EBP $\beta$ , C/EBP $\delta$ , C/EBP $\gamma$ , and C/EBP $\zeta$  genes were expressed ubiquitously with inconsistent expression patterns in various pig tissues. Moreover, a pig C/EBP regulatory network was constructed, including C/EBP genes, TFs and miRNAs. A total of 27 feed-forward loop (FFL) motifs were detected in the pig C/EBP regulatory network. Based on the RNA-seq data, gene expression patterns related to FFL sub-network were analyzed in 27 adult pig 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. 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.

**KEYWORDS:** C/EBP family, pig, phylogenetic analysis, transcriptional regulatory network, expression pattern, feed-forward loop

**RECEIVED:** February 23, 2021. **ACCEPTED:** August 2, 2021.

**TYPE:** Original Research

**FUNDING:** The authors would like to thank 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).

**DECLARATION OF CONFLICTING INTERESTS:** The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**CORRESPONDING AUTHOR:** Zhipeng Wang, NO.600, Changjiang Street, Xiangfang District, Harbin, 150030, China. Email: wangzhipeng@neau.edu.cn

## Introduction

The 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 superfamily.<sup>1</sup> The C/EBP family consists of 6 genes: C/EBP *alpha* (C/EBP $\alpha$ ), C/EBP *beta* (C/EBP $\beta$ ), C/EBP *delta* (C/EBP $\delta$ ), C/EBP *epsilon* (C/EBP $\epsilon$ ), C/EBP *gamma* (C/EBP $\gamma$ ), and C/EBP *zeta* (C/EBP $\zeta$ ).<sup>2</sup>

The first C/EBP protein was identified in rat liver nuclei.<sup>3</sup> Five additional C/EBP TFs had been identified in humans and rats.<sup>2</sup> 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<sup>4</sup> identified the C/EBP $\alpha$ , C/EBP $\beta$ , and C/EBP $\delta$  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/EBP $\alpha$ , C/EBP $\beta$ , and C/EBP $\delta$  genes. Thus, understanding of the evolution and function of the C/EBP genes requires genome-wide comparative studies.

\* These authors contributed equally to this work.

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

Pig and human are highly similar in anatomy, genetics, immunology and physiology, which makes the pig become a valuable animal model to study human complex diseases.<sup>9</sup> It has been reported that C/EBP $\alpha$ , C/EBP $\beta$ , and C/EBP $\delta$  genes regulate cell proliferation and differentiation in pigs' liver, fat and muscle tissue, respectively.<sup>10–17</sup> In addition, C/EBP $\beta$  and C/EBP $\delta$  genes regulate the proliferation and differentiation of pig granulosa cells too.<sup>18</sup> The C/EBP $\zeta$  gene regulates endoplasmic reticulum stress-mediated apoptosis which has been found in many tissues of pigs.<sup>19–24</sup> But the expression pattern of C/EBP gene family in pigs' tissues also has not been systematically analyzed.



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

In order to further study the function of the pig C/EBP family members, we identified C/EBP genes in pig and others 19 eukaryotic organisms, such as fruit fly, mouse, human and so on, and analyzed physicochemical properties of the proteins, conservative motifs, gene structure, and phylogenetic relationships. We also analyzed the expression patterns of the porcine C/EBP genes, and constructed a transcriptional regulatory network containing the C/EBP genes to investigate the functions of the porcine C/EBP genes. Our study found the C/EBP gene family was divided into 6 main branches, and C/EBPs has high sequence conservation and functional polymorphism. Except *C/EBPε* gene, the C/EBP genes were expressed ubiquitously with inconsistent expression patterns in various pig tissues. This 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 20 eukaryotic organisms were downloaded from the ENSEMBL database (<http://asia.ensembl.org/index.html>).<sup>24</sup> The eukaryotic organisms included nematode (*Caenorhabditis elegans*), fruit fly (*Drosophila melanogaster*), sea squirt (*Ciona intestinalis*), 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 HMMER (v3.0) software<sup>25</sup> 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>)<sup>26</sup> and the SMART tool (<http://smart.embl-heidelberg.de/>)<sup>27</sup> 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/>)<sup>28</sup> 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/>).<sup>29</sup> 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 v2.0) (<http://gsds.gao-lab.org/>).<sup>30</sup>

### Phylogenetic analysis

In this study, we used MEGA (v10.0)<sup>31</sup> to construct Maximum likelihood (ML), Neighbor-Joining (NJ) and Minimum-Evolution (ME) phylogenetic tree of C/EBP TFs based on Jones-Taylor-Thornton model and Games distribution, with 1000 replicate bootstrap tests, and set a cut-off bootstrap value of 65 to define the clades. Representations of the phylogenetic tree were constructed using the iTOL tool (<https://itol.embl.de/>).<sup>32</sup>

### 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.cabr.utoronto.ca/>).<sup>33</sup> The genomic sequences 2 kb upstream of the translation start site (TSS) and 3'-UTR were extracted by the R package biomaRt<sup>34</sup> for all coding genes. The TFBSTools package of R language<sup>35</sup> 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/>).<sup>36</sup> The miRanda software<sup>37</sup> was used to predict the miRNA targets with a threshold TotScore value of 120. The Cytoscape (v3.7.2) software<sup>38</sup> 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 (Mei and Yu, 2018)<sup>39</sup> 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.<sup>40</sup> The tissues included fat, thyroid, lymph, ovary, uterus, breast, spleen, lung, liver, placenta, adrenal gland, brain, gall bladder, gut, heart, kidney, longissimus dorsi, nasopharynx, oesophagus, pancreas, prostate, retina, salivary gland, spinal cord, stomach, testis, and uterus. According to description of Zhao et al, purification of the total RNA from the mixture of equally unrelated pig pool tissues was performed.

After quality control 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<sup>41</sup> 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 (v1.0.4, Linux ×86\_64).<sup>42</sup> The TBtools software<sup>43</sup> was used to visualize the heatmaps of the gene expression profiles in 27 tissues.

The tissue specificity index ( $\tau$ ) of the C/EBP genes and target genes for C/EBP 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.<sup>44</sup>

#### *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 pig and human sequences were downloaded from the Ensembl database.<sup>24</sup> The dN/dS and dN+dS values of each gene were calculated. In addition, the dN/dS and dN+dS mean values of the target genes of each C/EBP gene were compared using Kolmogorov-Smirnov (KS) test.

#### *Functional enrichment analysis*

The Database for Annotation, Visualization, and Integrated Discovery (DAVID) (<https://david.ncifcrf.gov/>)<sup>45</sup> was used to annotate the functions of C/EBPs and their target genes with all protein-encoding genes on pig genome as the background genes set, which includes gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis.

## Results

### *Identification of C/EBP genes in the eukaryotes*

A total of 99 C/EBP genes were identified in pig and 19 eukaryote 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/EBPα*, *C/EBPβ*, *C/EBPδ*, *C/EBPε*, *C/EBPγ*, and *C/EBPζ*. In particular, the *C/EBPα* and *C/EBPγ* genes of all species are located on the same chromosome within a distance of less than 100 kb, except for nematode, fruit fly, *Xenopus tropicalis* (see Table S1).

Various physicochemical properties of each C/EBP TF were calculated and are shown in Table 2 and Table S1. The C/EBPζ protein contains the highest number of amino acids, and the C/EBPγ protein consists of the least number of amino acids. The aliphatic index, a measure of thermostability, ranged from 60.36 to 77.10. There are significant differences ( $P < 2e^{-16}$ ) in the aliphatic index among C/EBP TFs. The GRAVY values of C/EBP proteins are negative indicating hydrophilic properties. The GRAVY values vary from -1.7940 to -0.2800. The pI values of C/EBPβ and C/EBPε proteins in each species are higher than 7, and the pI values of C/EBPζ are less than 7. The pI values of C/EBPγ proteins in each species are higher than 7, except for nematode. The C/EBPα protein of nematode, sea squirt, zebrafish is acidic, and that of other

species are alkaline. The C/EBPδ protein is acidic in 1/3 species and alkaline in the remaining 2/3 species.

### *Phylogenetic relationship analysis*

To analyze the phylogenetic relationships between 99 C/EBP genes in 20 species, an unrooted maximum likelihood phylogenetic tree was constructed (Figure 1). All C/EBP TFs were classified into 6 groups and was named *C/EBPα*, *C/EBPβ*, *C/EBPδ*, *C/EBPε*, *C/EBPγ*, and *C/EBPζ*. 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 artiodactyl includes pig, cattle, and goat; the order carnivore includes cat and dog; and the order rodents (rat and mouse) genes are clustered together. We also constructed the un-rooted neighbor-joining tree and minimum-evolution tree shown in Figures S1 and S2. The topological structure of NJ and ME tree were similar with ML tree.

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 Figure 2. The number of exons in *C/EBPζ* genes, which contained 15 or 16 exons, was higher than that in the other C/EBP genes. Most of the *C/EBPα*, *C/EBPβ*, *C/EBPδ*, *C/EBPε*, and *C/EBPγ* 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 5. Motif 10 is unique motifs in *C/EBPζ* proteins and may be associated with the clade-specific functions of the C/EBP proteins (see Table S2).

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

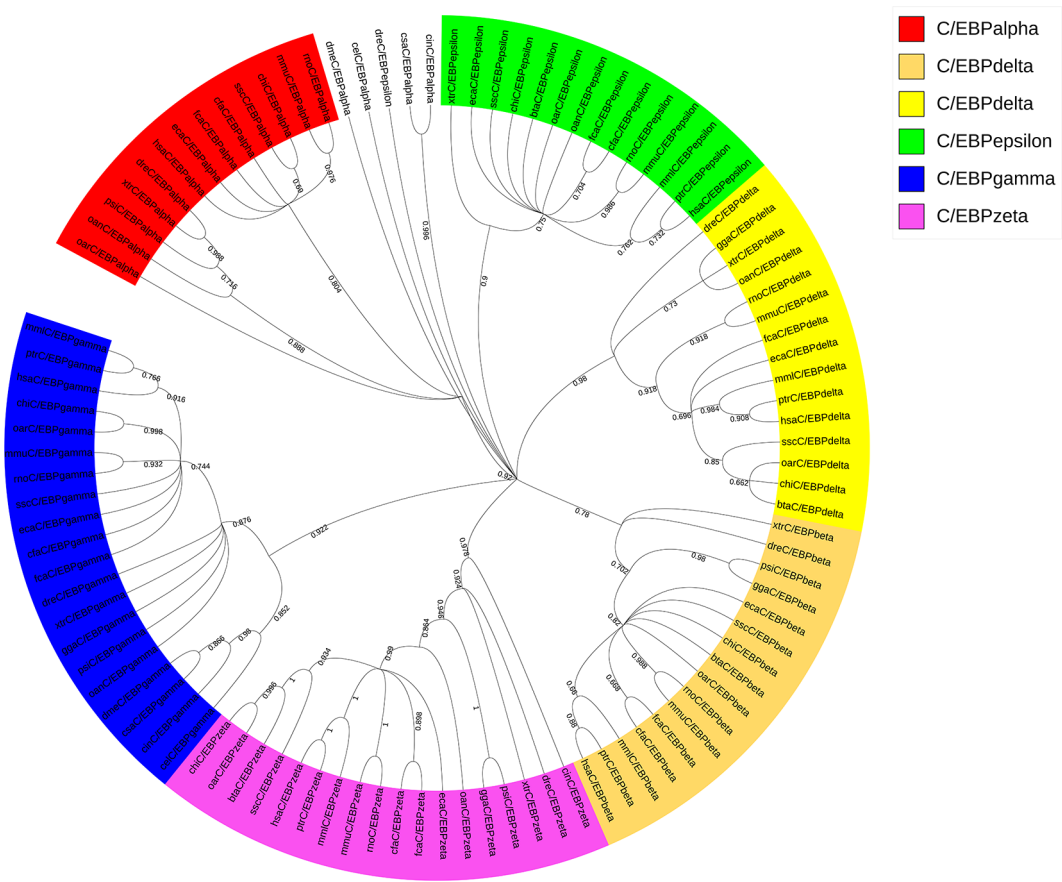
The expression of C/EBP genes was compared in 27 adult Duroc pig tissues. The tissue specificity index ( $\tau$ ) of the *C/EBPα*, *C/EBPβ*, *C/EBPδ*, *C/EBPε*, *C/EBPγ*, and *C/EBPζ* are 0.90, 0.73, 0.68, 0.92, 0.52 and 0.45, respectively. The  $\tau$  value varies from 0 for housekeeping genes to 1 for tissue-specific genes. So, the *C/EBPα* and *C/EBPε* genes may be named as tissue-specific genes, and others C/EBP genes may be the housekeeping genes. The 6 C/EBP genes from pig were classified into 2 groups based on the cluster analysis of the C/EBP gene expression levels in various tissues, as shown in Figure 3. Group I contains only the *C/EBPε* gene, and other genes are included in group II. The expression levels of the *C/EBPε* gene were low in all tissues (FPKM < 5), and the gene is specific expressed in the intestine, salivary gland, thyroid, uterus, and lymph. Although *C/EBPα*, *C/EBPβ*, *C/EBPδ*, *C/EBPγ*, and *C/EBPζ* are expressed ubiquitously, the expression patterns in various tissues are inconsistent. According to the cluster analysis, group II was classified to 3 sub-groups, including sub-group I (*C/EBPβ* and *C/EBPδ* genes),

**Table 1.** Genomic distribution of C/EBP genes on eukaryote genomes.

TAXONOMY				GENES							TOTAL
PHYLUM)	SUBPHYLUM	CLASS	ORDER	SPECIES	C/EBP $\alpha$	C/EBP $\beta$	C/EBP $\delta$	C/EBP $\epsilon$	C/EBP $\gamma$	C/EBP $\zeta$	
Nematoda	Rhabditis	Nematoda	Coccidiales	nematode	1	0	0	0	1	0	2
Arthropoda	Hexapoda	Insecta	Diptera	fruit fly	1	0	0	0	1	0	2
Chordate	Urochordata	Asciacea	Enterogona	sea squirt	1	0	0	0	1	1	3
	Vertebrates	Fishes	Cypriniformes	Zebrafish	1	1	1	1	1	1	6
		Amphibians	Anuras	Tropical xenopus	1	1	1	1	1	1	6
		Reptiles	Chelonias	Chinese soft-shell turtle	1	1	0	0	1	1	4
		Birds	Galliformes	Chicken	0	1	1	0	1	1	4
		Mammalias	Monotrematas	Platypus	1	0	1	1	1	1	5
			Carnivoras	Cat	1	1	1	1	1	1	6
				Dog	1	1	0	1	1	1	5
			Perissodactylas	Horse	1	1	1	1	1	1	6
			Artiodactylas	Pig	1	1	1	1	1	1	6
				Cow	0	1	1	1	0	1	4
				Goat	1	1	1	1	1	1	6
				Sheep	1	1	1	1	1	1	6
			Rodents	Mouse	1	1	1	1	1	1	6
				Rat	1	1	1	1	1	1	6
			Primates	Macaque	0	1	1	1	1	1	5
				Chimpanzee	0	1	1	1	1	1	5
				Human	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)
<i>C/EBP<math>\alpha</math></i>	135~449	34.26	5.72~10.78	61.53	-0.82
<i>C/EBP<math>\beta</math></i>	126~348	32.16	8.3~10.37	60.74	-0.63
<i>C/EBP<math>\delta</math></i>	99~302	26.99	5.43~10.73	60.36	-0.73
<i>C/EBP<math>\epsilon</math></i>	146~460	30.65	9.02~10.29	67.94	-0.74
<i>C/EBP<math>\gamma</math></i>	100~465	18.36	5.65~11.26	64.61	-1.03
<i>C/EBP<math>\zeta</math></i>	612~1074	117.00	5.31~8.36	77.10	-0.68



**Figure 1.** The unrooted maximum likelihood phylogenetic tree of the C/EBP proteins from 20 eukaryotes, with 1000 bootstrap replicates. The numbers on the lines indicate bootstrap support greater than 65%. All C/EBP TFs were classified into 6 groups and was named *C/EBP $\alpha$* , *C/EBP $\beta$* , *C/EBP $\delta$* , *C/EBP $\epsilon$* , *C/EBP $\gamma$* , and *C/EBP $\zeta$* .

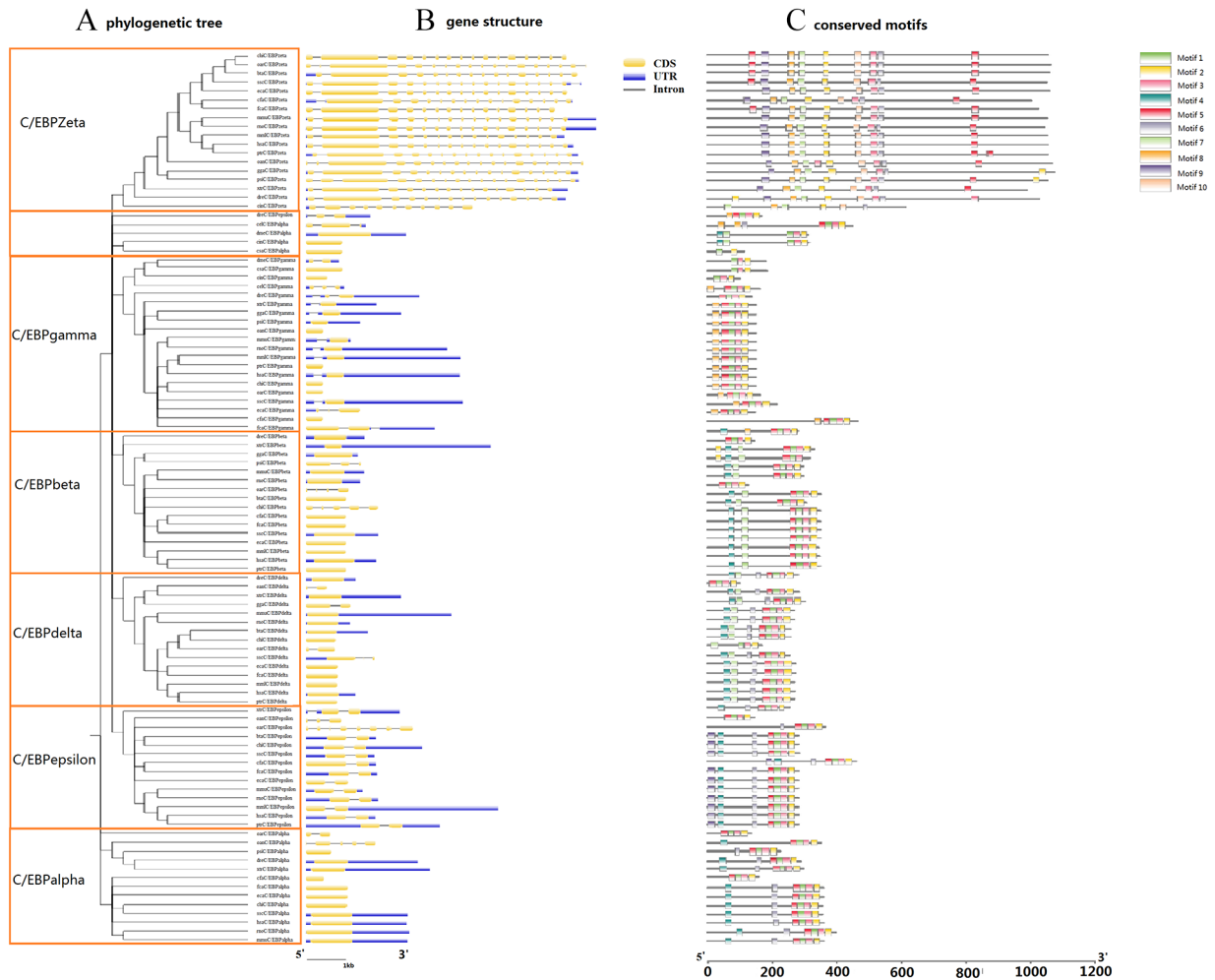
sub-group II (*C/EBP $\alpha$*  gene), sub-group III (*C/EBP $\gamma$*  and *C/EBP $\zeta$*  genes). *C/EBP $\alpha$*  was expressed ubiquitously on 27 pig tissues, and specific expressed in the thyroid, liver, lung, and adipose tissues. *C/EBP $\beta$*  was expressed at the high levels in the thyroid, adrenal gland, lung, adipose, liver, and ovary tissues. *C/EBP $\delta$*  was expressed at the high levels in the thyroid, gall bladder, ovary, and uterus. *C/EBP $\gamma$*  and *C/EBP $\zeta$*  are widely expressed at similar levels.

Additionally, the results indicate that the expression patterns in certain tissues, such as brain and spinal cord in the

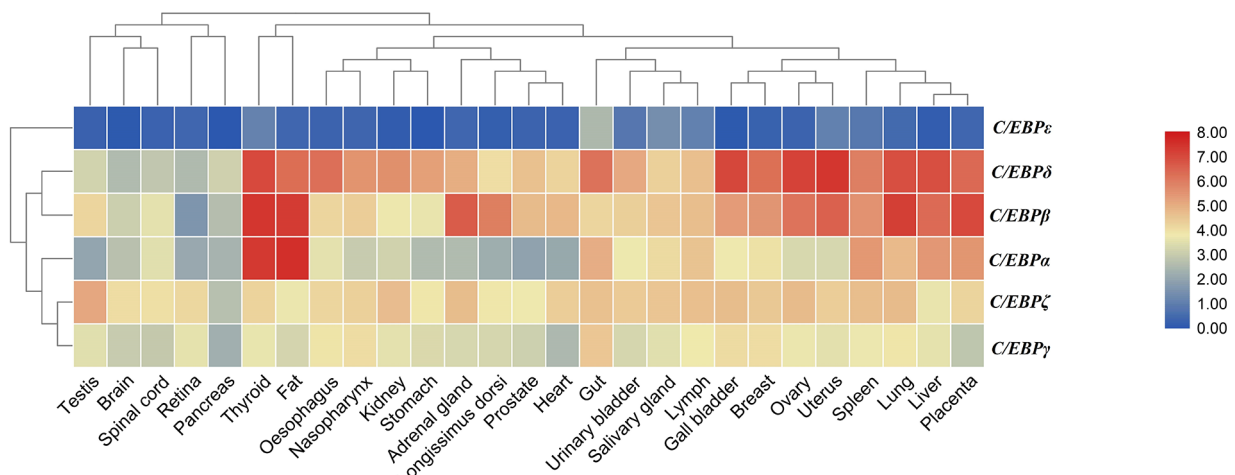
nervous system and ovary and uterus tissues in the female reproductive system, are similar.

### 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



**Figure 2.** The maximum likelihood phylogenetic tree, exon-intron structure, and conserved motifs of C/EBP genes. (A) The phylogenetic tree was constructed based on the C/EBP proteins from 20 eukaryotes. (B) Exon-intron structures of C/EBP genes. Yellow boxes indicate translated regions, blue boxes indicate 5'-UTR and 3'-UTR regions, and black lines indicate introns. (C) The conserved motifs of C/EBP proteins. The motifs, numbered 1–10, are displayed in different colored boxes.

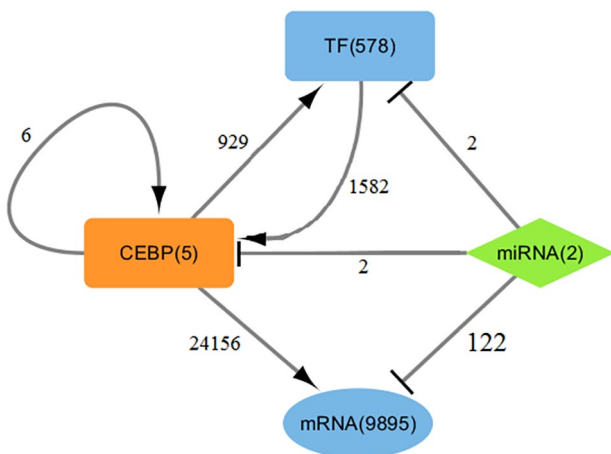


**Figure 3.** Heatmap of C/EBP genes in different tissues of duroc pig. Red represents higher expression levels; blue represents lower expression levels. The color vary from blue for lower expression level to red for higher expression level.

**Table 3.** Regulatory relationships involved in C/EBP family members.

C/EBP	TARGET GENES			FACTORS REGULATING C/EBP	
	TOTAL	NON-TF GENES	TF (GENE FAMILY)	C/EBP	TF (GENE FAMILY) MIMRA
<i>C/EBP<math>\alpha</math></i>	4662	4466	196 (31)	<i>C/EBP<math>\beta</math></i>	311 (29) –
<i>C/EBP<math>\beta</math></i>	3164	3059	105 (34)	<i>C/EBP<math>\beta</math>; C/EBP<math>\gamma</math></i>	293 (29) miR503
<i>C/EBP<math>\delta</math></i>	8383	8063	320 (36)	<i>C/EBP<math>\beta</math></i>	336 (32) –
<i>C/EBP<math>\epsilon</math></i>	7278	7027	221 (35)	<i>C/EBP<math>\beta</math>; C/EBP<math>\gamma</math></i>	310 (30) –
<i>C/EBP<math>\gamma</math></i>	1604	1541	63 (19)	–	325 (32) miR7140

The numbers in brackets is the number of gene families of transcription factors.



**Figure 4.** The transcriptional regulatory network with pig C/EBP family genes as the core. The arrows and represent TFs to regulate their target genes, and the T-lines represent miRNAs to regulate their target genes.

the *C/EBP $\alpha$* , *C/EBP $\beta$* , *C/EBP $\delta$* , *C/EBP $\epsilon$* , and *C/EBP $\gamma$*  genes from the cisbp database, we predicted 4662, 3164, 8383, 7278, and 1604 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.

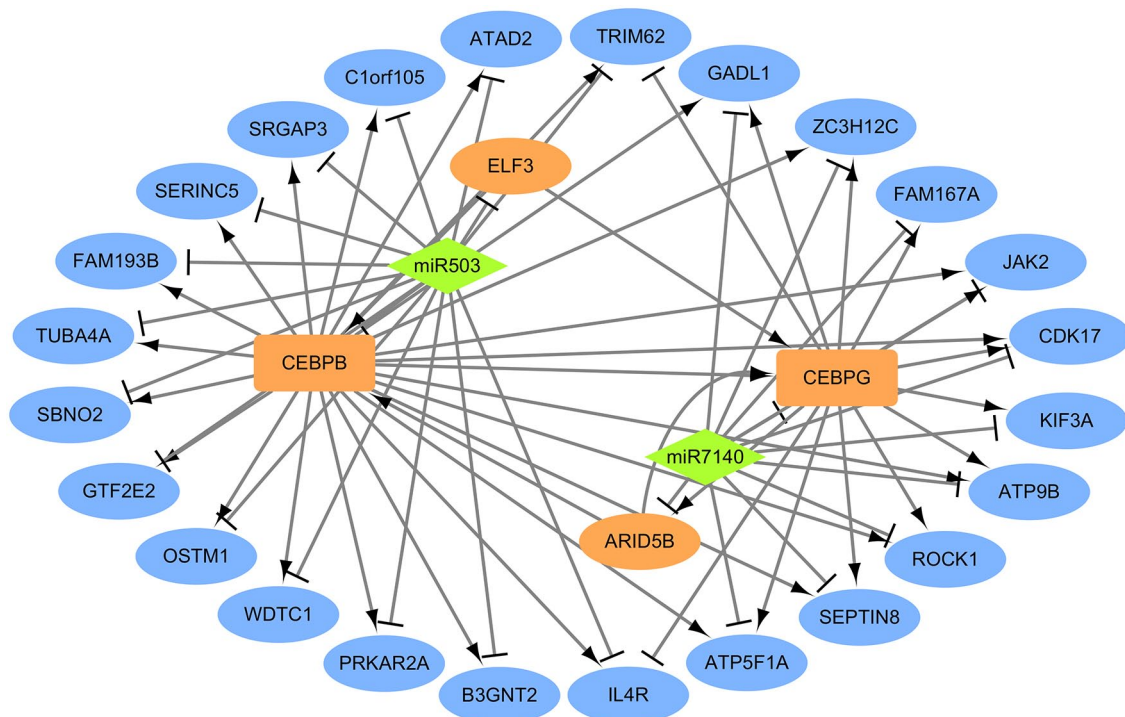
In this study, the C/EBP genes are predicted to be regulated by 423 TFs forming 1582 regulatory relationships; *mir503* and *mir7140* are predicted to regulate *C/EBP $\beta$*  and *C/EBP $\gamma$* , respectively. Additionally, *C/EBP $\beta$*  and *C/EBP $\gamma$*  are regulated by *C/EBP $\alpha$* , *C/EBP $\delta$* , and *C/EBP $\epsilon$*  and *C/EBP $\beta$*  and *C/EBP $\epsilon$* , respectively. Interestingly, *C/EBP $\beta$*  self-regulation is also predicted. Thus, we constructed a pig C/EBP regulatory network (summarized in Figure 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. Certain network concepts, including the clustering coefficient, network centralization, and network heterogeneity, were calculated to be 0.1890, 0.8160, and 25.1730, respectively. Among them, the clustering coefficient

describes the cohesive characteristics of the network. Network centralization and heterogeneity describe the distribution of connectivity (degree) in the network. Generally, if the network is highly heterogeneous, its centralization will be low. Moreover, the higher the heterogeneity of the network, the more uneven the distribution of degree in the network, that is only a few nodes in the network have high connectivity, while most other nodes have low connectivity. So, these results indicated the pig C/EBP regulatory network was a scale-free network.

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

According to the sequence, we predicted that *mir503* and *mir7140* regulate *C/EBP $\beta$*  and *C/EBP $\gamma$* , respectively. The *miR503* and *C/EBP $\beta$*  genes coregulate 14 target genes forming 14 *miR503-C/EBP $\beta$* -target gene FFL motifs, and *miR7140* and *C/EBP $\gamma$*  coregulate 11 target genes, including 11 *miR7140-C/EBP $\gamma$* -target gene FFL motifs. Additionally, *miR503-ELF3-C/EBP $\beta$*  and *miR7140-ARID5B-C/EBP $\gamma$*  motifs were identified. The *miR503-ELF3-C/EBP $\beta$*  motif is involved in the *miR503* and *ELF3* coregulation of the *C/EBP $\beta$*  gene, and the *miR503* gene targets the *ELF3* and *C/EBP $\beta$*  genes. The *miR7140-ARID5B-C/EBP $\gamma$*  motif is included in 3 regulatory relationships: *miR7140*→*ARID5B*, *ARID5B*→*C/EBP $\gamma$* , and *miR7140*→*C/EBP $\gamma$* . So, a total of 27 FFL motifs were identified in the pig C/EBP regulatory network. The combinations of all FFL motifs were used to construct the FFL sub-network (see Figure 5). There are *C/EBP $\beta$* -binding sites in 5'-untranslated region (5'-UTR) of *C/EBP $\gamma$* . Thus, *C/EBP $\beta$*  and *C/EBP $\gamma$*  co-regulated 10 target genes in FFL sub-network.

Based on the data of RNA-seq, gene expression patterns related to this FFL sub-network were analyzed in 27 adult tissues of Duroc pig. We explored the tissue-specific target genes for C/EBP in FFL sub-network. The results indicate that target genes regulated by *C/EBP $\beta$*  and *C/EBP $\gamma$*  have variable expression patterns in various tissues (see Table 4, Figures S3–S29). For example, the ATP synthase F1 subunit alpha (*ATP5F1A*) gene is a target gene of *C/EBP $\beta$*  and *C/EBP $\gamma$* , and



**Figure 5.** Subnetwork composed of 27 FFLs. The arrows and represent TFs to regulate their target genes, and the T-lines represent miRNAs to regulate their target genes. There are 2 miRNA-TF-C/EBP FFL motifs and 25 miRNA-C/EBP-target gene FFL motifs, including *miR7140-C/EBP $\beta$ -ARID5B*. In addition, there are 12 TF-TF-target motifs in this network.

**Table 4.** Statistical table of distribution of FFLs in various tissues.

TISSUES	MIRNA-FFLS		TF-FFLS	
	MIRNA-C/EBP-TARGET	MIRNA-TF-C/EBP	TF-C/EBP $\beta$ -C/EBP $\gamma$	C/EBP $\beta$ -C/EBP $\gamma$ -TARGET
Testis	24	1	1	10
Brain	24	1	1	10
Spinal cord	24	1	1	10
Retina	21	1	1	9
Pancreas	19	2	2	8
Thyroid	22	1	1	9
Fat	21	1	1	9
Oesophagus	21	2	2	9
Nasopharynx	24	2	2	10
Kidney	21	2	2	9
Stomach	21	2	2	9
Adrenal gland	23	1	1	10
Longissimus dorsi	18	1	1	9
Prostate	22	2	2	9
Heart	21	2	2	9
Gut	20	1	1	8

(Continued)



Table 4. (Continued)

TISSUES	MIRNA-FFLS		TF-FFLS	
	MIRNA-C/EBP-TARGET	MIRNA-TF-C/EBP	TF-C/ EBP $\beta$ -C/ EBP $\gamma$	C/EBP $\beta$ -C/ EBP $\gamma$ - TARGET
Urinary bladder	22	2	2	9
Salivary gland	23	2	2	9
Lymph	23	1	1	9
Gall bladder	22	2	2	9
Breast	21	2	2	8
Ovary	22	1	1	9
Uterus	23	2	2	9
Spleen	22	1	1	9
Lung	23	2	2	9
Liver	17	2	2	7
Placenta	21	2	2	9

form a FFL motif which is expressed ubiquitously. The glutamate decarboxylase-like protein 1 (*GADL1*) gene is muscle-specific express gene ( $\tau=0.99$ ). We suggest that the *C/EBP $\beta$ -C/EBP $\gamma$ -GADL1* motif may play an important role in the muscle tissue. 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; 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/EBP $\delta$*  gene was 1.52, which was higher than that of 5 other C/EBP genes (0.31–0.67) indicating that the *C/EBP $\delta$*  gene evolved rapidly and had an increased mutation rate.

The dN+dS mean values of the target genes of *C/EBP $\alpha$* , *C/EBP $\beta$* , *C/EBP $\delta$* , *C/EBP $\epsilon$* , and *C/EBP $\gamma$*  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/EBP $\alpha$*  are similar to that of *C/EBP $\beta$* , *C/EBP $\epsilon$*  and *C/EBP $\gamma$*  ( $P < .05$ ), 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/EBP $\alpha$*  target genes is similar to that of the *C/EBP $\delta$*  target genes and is significantly higher than that of other C/EBP target genes ( $P < .05$ ). The results indicate that the target genes of *C/EBP $\alpha$*  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/EBP $\alpha$* , *C/EBP $\beta$* , *C/EBP $\epsilon$* , and *C/EBP $\gamma$*  are associated with nucleoplasm (GO: 0005654) and extracellular exosome (GO: 0070062). The target genes of *C/EBP $\delta$*  and *C/EBP $\gamma$*  are involved in the transforming growth factor beta (*TGF $\beta$* ) receptor signaling pathway. The target genes of *C/EBP $\delta$*  are involved in the platelet-derived growth factor receptor signaling (GO: 0048008) (Table 6).

## Discussion

### *The characteristics and phylogenetic relationship analysis of the C/EBP gene family*

Improved genome sequencing and annotation enabled identification of all C/EBP TFs in the eukaryotic organism genomes.

**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	<i>C/EBP<math>\alpha</math></i> , <i>C/EBP<math>\epsilon</math></i> , <i>C/EBP<math>\delta</math></i> , <i>C/EBP<math>\gamma</math></i> , <i>C/EBP<math>\zeta</math></i>
	GO:0000978	RNA polymerase II core promoter proximal region sequence-specific DNA binding	1.40E-06	<i>C/EBP<math>\alpha</math></i> , <i>C/EBP<math>\epsilon</math></i> , <i>C/EBP<math>\delta</math></i> , <i>C/EBP<math>\gamma</math></i> , <i>C/EBP<math>\zeta</math></i>
GO_BP	GO:0030225	Macrophage differentiation	3.10E-03	<i>C/EBP<math>\alpha</math></i> , <i>C/EBP<math>\epsilon</math></i>
	GO:0048839	Inner ear development	9.60E-03	<i>C/EBP<math>\alpha</math></i> , <i>C/EBP<math>\delta</math></i>
	GO:0045669	Positive regulation of osteoblast differentiation	1.40E-02	<i>C/EBP<math>\alpha</math></i> , <i>C/EBP<math>\delta</math></i>
	GO:0051091	Positive regulation of sequence-specific DNA binding transcription factor activity	2.20E-02	<i>C/EBP<math>\beta</math></i>
KEGG_PATHWAY	ssc05202	Transcriptional misregulation in cancer	4.00E-05	<i>C/EBP<math>\alpha</math></i> , <i>C/EBP<math>\beta</math></i> , <i>C/EBP<math>\epsilon</math></i>
	ssc05152	Tuberculosis	3.60E-03	<i>C/EBP<math>\beta</math></i> , <i>C/EBP<math>\gamma</math></i>

**Table 6.** The results of functional enrichment analysis of porcine C/EBP target genes.

	CATEGORY	ID	TERM	P VALUE
<i>C/EBP<math>\alpha</math></i> target genes	GO_CC	GO:0005654	Nucleoplasm	1.37E-09
		GO:0070062	Extracellular exosome	8.76E-09
		GO:0005634	Nucleus	1.81E-05
		GO:0005737	Cytoplasm	5.53E-05
	GO_MF	GO:0044822	Poly(A) RNA binding	1.49E-05
		GO:0003824	Catalytic activity	5.88E-05
KEGG_PATHWAY	ssc01100	Metabolic pathways	4.17E-06	
<i>C/EBP<math>\beta</math></i> target genes	GO_CC	GO:0070062	Extracellular exosome	3.63E-09
		GO:0005654	Nucleoplasm	7.83E-06
	KEGG_PATHWAY	ssc01100	Metabolic pathways	4.21E-06
		ssc00310	Lysine degradation	1.40E-04
<i>C/EBP<math>\delta</math></i> target genes	GO_BP	GO:0001649	Osteoblast differentiation	5.31E-05
		GO:0000122	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:0007179	Transforming growth factor beta receptor signaling pathway	4.21E-04
		GO:0098779	Mitophagy in response to mitochondrial depolarization	6.24E-04
		GO:0001525	Angiogenesis	8.71E-04
		GO:0007420	Brain development	1.06E-03
		GO:1901998	Toxin transport	1.04E-03
		GO:0007507	Heart development	1.49E-03
		GO:0048008	Platelet-derived growth factor receptor signaling pathway	2.18E-03
GO:0007059	Chromosome segregation	2.01E-03		

(Continued)

Table 6. (Continued)

	CATEGORY	ID	TERM	P VALUE
		GO:0007519	Skeletal muscle tissue development	3.54E-03
		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
<i>C/EBP<math>\epsilon</math></i> 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_MF	GO:0044822	Poly(A) RNA binding	2.81E-07
		GO:0005524	ATP binding	8.89E-06
	KEGG_PATHWAY	ssc01100	Metabolic pathways	9.13E-09
<i>C/EBP<math>\gamma</math></i> target genes	GO_CC	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

This study identified all *C/EBP* TFs in pig and others eukaryotic organisms and found that the *C/EBP* genes are ubiquitous in the majority of the vertebrates. There are differences in biochemical parameters among *C/EBP* TFs. These biochemical parameters affect protein solubility, sub-cellular localization and interaction.<sup>46</sup> The physical and chemical properties of different *C/EBPs* also quite indicated the differences in the functionalities of the *C/EBPs*. Some biochemical parameters have very difference in different species. For example, the *C/EBP $\alpha$*  of nematode, sea squirt zebrafish is acidic while in other species is basic. The *C/EBP $\delta$*  is acidic in 1/3 of the species and alkaline in the remaining 2/3. These results indicate that some *C/EBP* genes have low functional conservation among different eukaryotes.<sup>47,48</sup>

In lower eukaryotes, such as nematode, fruit fly, there are *C/EBP $\alpha$*  and *C/EBP $\gamma$*  genes. But in vertebrate species, more *C/EBP* genes distributed on genome. It indicated the *C/EBP* gene family may have been expanded in early stage of vertebrate species evolution. The molecular phylogenetic tree showed that the *C/EBP* family members originated in the early stage of vertebrate evolution. Similarly, Qiu et al<sup>4</sup>

analyzed the evolution of the *C/EBP $\alpha$* , *C/EBP $\beta$* , and *C/EBP $\delta$*  genes in 20 vertebrate genomes and demonstrated that the *C/EBP* TFs originated early in vertebrate evolution.

According to the phylogenetic tree, all *C/EBP* TFs were classified into 6 subgroups. However, Qiu et al analyzed the phylogenetic trees of *C/EBP $\alpha$* , *C/EBP $\beta$* , and *C/EBP $\delta$*  homologs in 2 fish, 2 amphibian, 2 reptile, 11 avian, and 3 mammalian species. The results indicated that the *C/EBP $\alpha$*  and *C/EBP $\beta$*  genes are clustered into 1 group. This difference may be due to different species and members of the *C/EBP* family.

#### *The expression pattern and function of the C/EBP gene family of pig*

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 tissues of adult Duroc pig, and *C/EBP $\epsilon$*  is expressed only in 5 tissues. Uhlén et al<sup>49</sup> 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. Summers et al<sup>50</sup> created a pig gene expression

atlas based on the meta-analysis and demonstrated that *C/EBPε* is expressed only in a few tissues, and other *C/EBP* genes are expressed in the majority of tissues. Shyamsunder et al<sup>51</sup> and Studd et al<sup>52</sup> found the *C/EBPε* gene is a major regulator of myeloid and lymphoid cell differentiation, respectively. Our results found the *C/EBPε* gene shown the exceptionally low expression levels in 27 pig tissues, and it also named as thyroid and lymph tissue-specific genes. So, it indicated that *C/EBPε* gene may only perform functions in specific tissues related to immune system, such as myeloid and lymphoid cell. The differences in the exon-intron, motif structures and gene expression patterns among various clades provided some indications that the *C/EBP* genes may have a variety of physiological functions. The difference number of exons and motifs might be the basis of *C/EBP* genes functional diversity. Functional enrichment analysis demonstrated that pig *C/EBP* genes perform many important biological functions, involved in the inner ear development (GO: 0048839), positive regulation of osteoblast differentiation (GO: 0045669), and the tuberculosis pathway (ssc05152). The *C/EBPζ* gene contained the highest number of exons and motifs, which indicated that it may take some specific roles. Previous studies showed that *C/EBPζ* is involved in cochlear cell apoptosis in rats and may co-regulate the inner ear development with GATA binding protein 2 (*GATA2*).<sup>53,54</sup>

We used cluster analysis to explore similar expression patterns in multiple tissues among *C/EBP* genes. There were similar expression between *C/EBPβ* and *C/EBPδ* genes, belonged to the same sub-group. It has been reported that both *C/EBPβ* and *C/EBPδ* were involved in the chondrocyte differentiation, heart development and early regulation of fat formation.<sup>55-57</sup> The sub-group III includes *C/EBPγ* and *C/EBPζ* genes, which it has been reported that they are important regulators of cellular stress response networks to maintain the basic life activities of cells.<sup>58</sup>

Regulation of gene expression is one of the most important and complex issues in biology. Many studies suggested that tissue specificity of gene expression or gene regulatory network took biologically important roles. In this study, we found the *C/EBPα* may be named as fat tissue-specific gene. Chen et al<sup>59</sup> demonstrated that *C/EBPα* gene is the lipogenic marker of the porcine adipocytes that promotes adipocyte differentiation. In this study, we also calculated the tissue specific index of the *C/EBP* family target genes. Our results indicated that *C/EBPβ* genes may be the housekeeping gene, which are expressed at the high levels in many tissues, as previous studies<sup>11,13,18</sup> had shown. But some of its target genes might name as tissue-specific genes, such as *ELF3*, *GADL1*, *SRGAP3*. Combined with the FFL sub-network, we deem some FFL motifs may be tissue-specific. Sonawane et al<sup>60</sup> reported that transcription factors primarily participate in tissue-specific regulatory processes via alterations in their targeting patterns.

### The transcriptional regulatory network

Complex regulatory mechanisms of gene expression in eukaryotes control the development, physiology, and pathology.<sup>5</sup> 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 key advantage of the scale-free topology is the increased tolerance to random failures, which helps the animals adjust to a stressful environment and maintains *C/EBPs* to perform its biological functions.<sup>61</sup> Furthermore, human TF and miRNA regulatory network were also reported to correspond to approximately scale-free topology.<sup>62,63</sup>

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 each porcine *C/EBP* TFs were different. In the Database of Human Transcription Factor Targets,<sup>64</sup> 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/EBPδ* 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, 27 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 human tissues or cells. Sun et al<sup>65</sup> reported that *miR223->C/EBPβ->LMO2* FFL in human myeloid cells can regulate cell proliferation and maintain normal differentiation and development. Ponomarev et al<sup>66</sup> have reported that *MiR124-C/EBPα-PU.1* FFL is related to the transition between an activated phenotype and a quiescent state of macrophages. Shi et al<sup>67</sup> reported that 28 FFLs of *C/EBPα*, *C/EBPβ*, and *C/EBPδ* may be associated with occurrence of human hypertrophic cardiomyopathy. Although these FFLs has not been reported in pig tissues or cells, we deem that further experiment, such as ChIP-seq, would help identifying this FFL sub-networks, and demonstrating the regulation function of the *C/EBP* gene family in pig.

### Conclusion

In the present study, we identified all *C/EBP* TFs of pig and others eukaryotic organisms and demonstrated the conservation and evolutionary relationship of the *C/EBP* family genes. The porcine *C/EBP* genes are expressed generally, except *C/EBPε* 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. Some FFL motifs may be tissues-specific. Thus, our study provides potential functional insight into the roles of the C/EBP genes in pig.

### Author contribution statement

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

### Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

### ORCID iD

Zhipeng Wang  <https://orcid.org/0000-0002-6041-8906>

### Supplemental material

Supplemental material for this article is available online.

### REFERENCES

- Ramji DP, Foka P. CCAAT/enhancer-binding proteins: structure, function and regulation. *Biochem J*. 2002;365:561-575.
- Lekstrom-Himes J, Xanthopoulos KG. Biological role of the CCAAT/enhancer-binding protein family of transcription factors. *J Biol Chem*. 1998;273:28545-28548.
- Graves BJ, Johnson PF, McKnight SL. Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene. *Cell*. 1986;44:565-576.
- Qiu J, Wang W, Hu S, et al. Molecular cloning, characterization and expression analysis of C/EBP  $\alpha$ ,  $\beta$  and  $\delta$  in adipose-related tissues and adipocyte of duck (*Anas platyrhynchos*). *Comp Biochem Physiol B Biochem Mol Biol*. 2018;221:222-29-43.
- Lambert M, Jambon S, Depauw S, David-Cordonnier MH. Targeting transcription factors for cancer treatment. *Molecules*. 2018;23:1479.
- Ross SE, Hemati N, Longo KA, et al. Inhibition of adipogenesis by Wnt signaling. *Science*. 2000;289:950-953.
- Cassel TN, Nord M. C/EBP transcription factors in the lung epithelium. *Am J Physiol Lung Cell Mol Physiol*. 2003;285:L773-L781.
- Graves P, Zeng Y. Biogenesis of mammalian microRNAs: a global view. *Genomics Proteomics Bioinformatics*. 2012;10:239-245.
- Pabst R. The pig as a model for immunology research. *Cell Tissue Res*. 2020;380:287-304.
- Zhang W, Chen Q, Xu L, Cai J, Zhang J. The potential role of PSMA6 in modulating fat deposition in pigs by promoting preadipocyte proliferation and differentiation. *Gene*. 2021;769:145228.
- Wang Y, Liu X, Hou L, Wu W, Zhao S, Xiong Y. Fibroblast growth factor 21 suppresses adipogenesis in pig intramuscular fat cells. *Int J Mol Sci*. 2015;17:11.
- Óvilo C, Benítez R, Fernández A, et al. Longissimus dorsi transcriptome analysis of purebred and crossbred Iberian pigs differing in muscle characteristics. *BMC Genomics*. 2014;15:413.
- Tang Y, Xiong K, Shen M, Mu Y, Li K, Liu H. CCAAT-enhancer binding protein (C/EBP)  $\beta$  regulates insulin-like growth factor (IGF) 1 expression in porcine liver during prenatal and postnatal development. *Mol Cell Biochem*. 2015;401:209-218.
- Kang T, Peng D, Bu G, et al. Transcriptional regulation analysis of FAM3A gene and its effect on adipocyte differentiation. *Gene*. 2016;595:92-98.
- Tada A, Islam MA, Kober AH, et al. Transcriptome modifications in the porcine intramuscular adipocytes during differentiation and exogenous stimulation with TNF- $\alpha$  and serotonin. *Int J Mol Sci*. 2020;21:638.
- Deng B, Wen J, Ding Y, et al. Functional analysis of pig myostatin gene promoter with some adipogenesis- and myogenesis-related factors. *Mol Cell Biochem*. 2012;363:291-299.
- Jing L, Hou Y, Wu H, et al. Transcriptome analysis of mRNA and miRNA in skeletal muscle indicates an important network for differential Residual Feed Intake in pigs. *Sci Rep*. 2015;5:11953.
- 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;72:1194-1204.
- Jiang Q, Chen S, Ren W, et al. Escherichia coli aggravates endoplasmic reticulum stress and triggers CHOP-dependent apoptosis in weaned pigs. *Amino Acids*. 2017;49:2073-2082.
- Li S-J, Ding S-T, Mersmann HJ, Chu C-H, Hsu C-D, Chen C-Y. A nutritional nonalcoholic steatohepatitis minipig model. *J Nutr Biochem*. 2016;28:51-60.
- Tominaga T, Sharma I, Fujita Y, Doi T, Wallner AK, Kanwar YS. Myo-inositol oxygenase accentuates renal tubular injury initiated by endoplasmic reticulum stress. *Am J Physiol Renal Physiol*. 2019;316:F301-F315.
- Qu H, Ajuwon KM. Adipose tissue-specific responses reveal an important role of lipogenesis during heat stress adaptation in pigs. *J Anim Sci*. 2018;96:975-989.
- Li S-J, Liu C-H, Chu H-P, et al. The high-fat diet induces myocardial fibrosis in the metabolically healthy obese minipigs—the role of ER stress and oxidative stress. *Clin Nutr*. 2017;36:760-767.
- Cunningham F, Achuthan P, Akanni W, et al. Ensembl 2019. *Nucleic Acids Res*. 2019;47:D745-D751.
- Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. HMMER web server: 2018 update. *Nucleic Acids Res*. 2018;46:W200-W204.
- Lu S, Wang J, Chitsaz F, et al. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res*. 2020;48:D265-D268.
- El-Gebali S, Mistry J, Bateman A, et al. The Pfam protein families database in 2019. *Nucleic Acids Res*. 2019;47:D427-D432.
- Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. *Protein identification and analysis tools on the ExPASy server*. In: Walker JM, ed. *The Proteomics Protocols Handbook*. Humana Press; 2005:571-607.
- Bailey TL, Boden M, Buske FA, et al. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res*. 2009;37:W202-W208.
- Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G. GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics*. 2015;31:1296-1297.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. 2018;35:1547-1549.
- Letunic I, Bork P. Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res*. 2019;47:W256-W259.
- Weirauch MT, Yang A, Albu M, et al. Determination and inference of eukaryotic transcription factor sequence specificity. *Cell*. 2014;158:1431-1443.
- 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;4:1184-1191.
- Tan G, Lenhard B. TFBSTools: an R/bioconductor package for transcription factor binding site analysis. *Bioinformatics*. 2016;32:1555-1556.
- Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. *Nucleic Acids Res*. 2019;47:D155-D162.
- Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS. MicroRNA targets in *Drosophila*. *Genome Biol*. 2003;5:R1.
- Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13:2498-2504.
- Mei W, Yu G. basicTrendline: add trendline and confidence interval of basic regression models to plot. R package version 2.0.3. 2018. Accessed June 13, 2020. <https://CRAN.R-project.org/package=basicTrendline>
- Zhao P, Zheng X, Feng W, et al. Profiling long noncoding RNA of multi-tissue transcriptome enhances porcine noncoding genome annotation. *Epigenomics*. 2018;10:301-320.
- Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12:357-360.
- 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;33:290-295.
- Chen C, Chen H, Zhang Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant*. 2020;13:1194-1202.
- Yanai I, Benjamin H, Shmoish M, et al. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics*. 2005;21:650-659.
- Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4:44-57.
- Nandi S, Mehra N, Lynn AM, Bhattacharya A. Comparison of theoretical proteomes: identification of COGs with conserved and variable pI within the multimodal pI distribution. *BMC Genomics*. 2005;6:116.

47. Liu M, Dong H, Wang M, Liu Q. Evolutionary divergence of function and expression of laccase genes in plants. *J Genet.* 2020;99:23.
48. Siddiqui MQ, Badmalia MD, Patel TR. Bioinformatic analysis of structure and function of LIM domains of human Zyxin family proteins. *Int J Mol Sci.* 2021;22:2647.
49. Uhlén M, Fagerberg L, Hallström BM, et al. Proteomics. Tissue-based map of the human proteome. *Science.* 2015;347:1260419.
50. Summers KM, Bush SJ, Wu C, 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.
51. Shyamsunder P, Shanmugasundaram M, Mayakonda A, et al. Identification of a novel enhancer of CEBPE essential for granulocytic differentiation. *Blood.* 2019;133:2507-2517.
52. Studd JB, Yang M, Li Z, et al. Genetic predisposition to B-cell acute lymphoblastic leukemia at 14q11.2 is mediated by a CEBPE promoter polymorphism. *Leukemia.* 2019;33:1-14.
53. Zong S, Liu T, Wan F, Chen P, Luo P, Xiao H. Endoplasmic reticulum stress is involved in cochlear cell apoptosis in a cisplatin-induced ototoxicity rat model. *Audiol Neurootol.* 2017;22:160-168.
54. 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;227:32-40.
55. Okuma T, Hirata M, Yano F, et al. Regulation of mouse chondrocyte differentiation by CCAAT/enhancer-binding proteins. *Biomed Res.* 2015;36:21-29.
56. Huang GN, Thatcher JE, McAnally J, et al. C/EBP transcription factors mediate epicardial activation during heart development and injury. *Science.* 2012;338:1599-1603.
57. Merrett JE, Bo T, Psaltis PJ, Proud CG. Identification of DNA response elements regulating expression of CCAAT/enhancer-binding protein (C/EBP)  $\beta$  and  $\delta$  and MAP kinase-interacting kinases during early adipogenesis. *Adipocyte.* 2020;9:427-442.
58. Huggins CJ, Mayekar MK, Martin N, et al. C/EBP $\gamma$  is a critical regulator of cellular stress response networks through heterodimerization with ATF4. *Mol Cell Biol.* 2015;36:693-713.
59. Chen X, Luo Y, Jia G, Liu G, Zhao H, Huang Z. The effect of arginine on the Wnt/ $\beta$ -catenin signaling pathway during porcine intramuscular preadipocyte differentiation. *Food Funct.* 2017;8:381-386.
60. Sonawane AR, Platig J, Fagny M, et al. Understanding tissue-specific gene regulation. *Cell Rep.* 2017;21:1077-1088.
61. Bassett DS, Bullmore ET. Small-world brain networks revisited. *Neuroscientist.* 2017;23:499-516.
62. Le K, Guo H, Zhang Q, et al. Gene and lncRNA co-expression network analysis reveals novel ceRNA network for triple-negative breast cancer. *Sci Rep.* 2019;9:15122.
63. Jiang W, Mitra R, Lin CC, Wang Q, Cheng F, Zhao Z. Systematic dissection of dysregulated transcription factor-miRNA feed-forward loops across tumor types. *Brief Bioinform.* 2016;17:996-1008.
64. Zhang Q, Liu W, Zhang HM, et al. hTFtarget: a comprehensive database for regulations of human transcription factors and their targets. *Genomics Proteomics Bioinformatics.* 2020;18:120-128.
65. Sun W, Shen W, Yang S, Hu F, Li H, Zhu TH. miR-203 and miR-142 attenuate hematopoietic cell proliferation, and miR-223 positively regulates miR-142 through LMO2 isoforms and CEBP- $\beta$ . *Cell Res.* 2010;20:1158-1169.
66. 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- $\alpha$ -PU.1 pathway. *Nat Med.* 2011;17:64-70.
67. Shi H, Li J, Song Q, et al. Systematic identification and analysis of dysregulated miRNA and transcription factor feed-forward loops in hypertrophic cardiomyopathy. *J Cell Mol Med.* 2019;23:306-316.