

RESEARCH

Open Access



# Transcriptomic analysis reveals the dynamic changes of transcription factors during early development of chicken embryo

Liqin Liao<sup>1,2,3,4</sup>, Ziqi Yao<sup>1,2</sup>, Jie Kong<sup>1,2,3</sup>, Xinheng Zhang<sup>1,2,3</sup>, Hongxin Li<sup>1,2,4</sup>, Weiguo Chen<sup>1,3</sup> and Qingmei Xie<sup>1,2,3,4\*</sup>

## Abstract

**Background:** The transition from fertilized egg to embryo in chicken requires activation of hundreds of genes that were mostly inactivated before fertilization, which is accompanied with various biological processes. Undoubtedly, transcription factors (TFs) play important roles in regulating the changes in gene expression pattern observed at early development. However, the contribution of TFs during early embryo development of chicken still remains largely unknown that need to be investigated. Therefore, an understanding of the development of vertebrates would be greatly facilitated by study of the dynamic changes in transcription factors during early chicken embryo.

**Results:** In the current study, we selected five early developmental stages in White Leghorn chicken, *gallus gallus*, for transcriptome analysis, cover 17,478 genes with about 807 million clean reads of RNA-sequencing. We have compared global gene expression patterns of consecutive stages and noted the differences. Comparative analysis of differentially expressed TFs (FDR < 0.05) profiles between neighboring developmental timepoints revealed significantly enriched biological categories associated with differentiation, development and morphogenesis. We also found that Zf-C2H2, Homeobox and bHLH were three dominant transcription factor families that appeared in early embryogenesis. More importantly, a TFs co-expression network was constructed and 16 critical TFs were identified.

**Conclusion:** Our findings provide a comprehensive regulatory framework of TFs in chicken early embryo, revealing new insights into alterations of chicken embryonic TF expression and broadening better understanding of TF function in chicken embryogenesis.

**Keywords:** Embryogenesis, Chicken, Transcription factors, Gene expression, Network analysis

## Introduction

Transcription factors (TFs) interpret the genome directly, and are responsible for decoding DNA sequences [1]. It is reported that transcriptional factors are key components of cells that control gene expression, determining how the cells function [2]. Transcription factors acting

as conductor orchestrate complex regulatory networks of gene expression. A deeper understanding of the common transcription factors and their shared interaction by analyzing a set of coregulated or differentially expressed genes can provide insight into the pathways underlying such expression patterns [3]. Embryonic development involves a mass of cells achieving specific cell identities depending on morphogen gradients and the activation of transcription factors (TFs) [4]. Embryos in the early stages of their development show transcriptional activities that are different from those occurring later. Normally, changes in the gene expression are regulated by

\*Correspondence: qmx@scau.edu.cn

<sup>1</sup> Heyuan Branch, Guangdong Provincial Laboratory of Lingnan Modern Agricultural Science and Technology, College of Animal Science, South China Agricultural University, Guangzhou 510642, China  
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

transcription factors, which play crucial roles in biological processes such as cell proliferation, cell differentiation.

Successful embryo development is dependent on the early stages of embryogenesis and the proper activation of the genome. For example, T-box factors are an ancient family of transcription factors that govern gene expression patterns that are critical for embryonic development [5], such as Tbx5 and Tbx4 binding with LMP-4 with important roles in vertebrate limb and heart development [6]. The transcription factors fork-head box (Fox) is commonly conserved in organisms varying from yeast to humans [7]. In the chicken reproduction development, Fox family is a prominent regulator for development of testis or ovarian [8, 9]. Moreover, it is considered critical to identify regulatory elements within the promoter region in order to understand the mechanism underlying transcriptional regulation in specific cell types [10], such as Sox11 activating Prox1 expression through multiple regulatory elements to promote chicken embryonic neurogenesis [11], transcription factor Sox2 binding with Cpbd1 to regulate the formation of chicken spermatogonial stem cells [12].

Chicken is one of the most important commercial species as well as a model organism for biological and medical research (chicken genomics). An increasingly efforts to character transcripts in chicken by RNA-sequencing have provided key insights into function of the chicken genome, such as the transcriptome analysis of early embryo to distinct gene clusters with specific morphological changes [13], revealing the chicken specific signaling pathways and gradually analogous gene expression via zygotic genome activation (ZGA) by RNA-sequencing [14], as well as analysis of transcriptome-wide m6A methylation modification pattern in the gonads of chicken embryos [15]. The study of embryogenesis is critical for a comprehensive understanding of the gene expression patterns and underlying biological changes during early embryonic developmental stages of an organism. The transcriptome profiling of chicken embryos creates an opportunity to advance our understanding of the molecular regulation of embryo development. Nevertheless, researches about transcription factors in chicken genome mainly focus on studying functions of specific factors, such as: 1) the fact that chicken NANOG, SOX2, and POUV expression varies dramatically at different stages shows that chickens have a distinctive pluripotent circuitry and may be crucial in the early development of pluripotency; 2) Chicken C/EBP has the ability to directly bind to and activate the PPAR gene promoter, which is one of the primary controllers of adipogenesis [16, 17]. However, the whole transcription factors landscape of early chicken embryo remains unclear. Here, we focused on early chicken embryo development underlying its

diverse transcription factors and investigated the distribution and expression pattern of TFs.

In this study, we used RNA-sequencing to systematically investigate the expression profiles of all annotated transcription factors of chicken during early development stages. Five early developmental stages, including 1, 2, 3, 4 and 5 days after fertilization, were selected for transcriptome sequencing and analysis. We have identified differentially expressed genes (DEGs) between neighboring developmental stages. Identifying key genes and pathways involved in the regulation of embryonic development was achieved by analyzing differentially expressed transcription factors (DE-TFs) across five stages of development. The DE-TFs were used to conduct Gene Ontology (GO) enrichment analysis to reveal the biological functions. Importantly, this is the first comprehensive regulatory framework for transcription factors in early embryogenesis in chickens, highlighting the dynamics of TFs expression at the early stages of embryo.

## Materials and methods

### Ethics statement

All of the experimental protocols involved in animal care and sample collection were approved by the Animal Ethics Committee at the South China Agricultural University, China (approval ID: SYXK-2022-0136).

### Embryos collection and RNA extraction

Fertilized eggs from *White Leghorns* were purchased from Guangdong Wen's DaHuaNong Biotechnology Co., Ltd. The eggs were incubated at 37.5 °C and 65% relative humidity in an automated egg incubator, rotating every 6 h. Embryos were collected at the following times point: 24 h, 48 h, 72 h, 96 h, and 120 h, with three biological replicates for each embryonic stage, labeled Em1d-Em5d. Total RNA was extracted using Trizol reagent kit (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. The RNA concentration and purity were measured using the Nano-Drop 2000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA). RNA quality was assessed on an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA).

### Library construction and sequencing

Constructing cDNA library was performed as previous studies following the instructions of the manufacturer provided by the GENE-DENOVO Biotechnology [18–20]. Briefly, after total RNA was extracted, eukaryotic mRNA was enriched by Oligo(dT) beads, while prokaryotic mRNA was depleted by removing rRNA by RiboZero™ Magnetic Kit (Epicentre, Madison, WI, USA). Then the enriched mRNA was fragmented into short fragments using fragmentation buffer and was reverse

transcribed into cDNA with random primers. Second-strand cDNA were synthesized by DNA polymerase I, RNase H, dNTP and buffer. Then the cDNA fragments were purified with QiaQuick PCR extraction kit (Qiagen, Venlo, The Netherlands), end repaired, A base added, and ligated to Illumina sequencing adapters. The ligation products were size selected by agarose gel electrophoresis, PCR amplified, and sequenced using Illumina Novaseq6000. Library construction and sequencing reactions were conducted at GENE-DENOVO Biotechnology Co., Ltd (Guangzhou, China). The raw RNA-seq data is available at NCBI (PRJNA850787).

### Transcriptome assembly

Reads were further filtered according to the following rules to obtain high-quality clean reads by fastp (version 0.18.0). Firstly, deleting adapter-containing reads; secondly, readings with more than 10% unknown nucleotides (N) are also removed; thirdly, we removed all reads with terminal poly A; lastly, eliminating low quality reads (containing more than 50% number of bases with mass value  $Q \leq 20$ ). The short reads alignment tool Bowtie2 was used to compare the clean reads to the ribosome database of the species [21]. After comparative analysis based on the chicken genome (GRCg6a) using the HISAT2 software [22], we re-constructed the transcriptome by StringTie and then counted the expression of each gene via RSEM [23, 24].

### Gene expression analysis

Gene expression was presented with fragments per kilobase of transcript per million fragments mapped (FPKM). Principal component analysis was used to assess sample repeatability. The DESeq2 tool was used to perform differential expression analysis between the five stages. Genes with FDR (false discovery rate)  $\leq 0.05$  and Fold Change  $\geq 2$  were considered as DEGs between two stages. Simultaneously, the ggplot2 software was used to carry out a hierarchical cluster analysis of differentially expressed genes. (<http://www.r-project.org/>). The final lists of unique genes were used for further analysis after duplicate and missing values were removed.

### Detection of TFs in the list of DEGs

To identify the TFs that have differentially expression levels as they go from one stage to the next, we performed Hidden Markov Model scan (hmmScan) to compared the lists of DEGs with the Animal Transcription Factor DataBase [25]. Raw data for DEGs and DE-TFs can be found in the supplementary files.

### Network construction and analysis

All DE-TFs and their target genes were applied to construct the co-expressed network. Protein-protein networks were constructed by extracting the information regarding TFs interactions from STRING database [26]. Cytoscape [27] software were used to visualize and analyze the networks. Moreover, hub TFs were analyzed by KEGG and shown by Sankey plot.

### Functional annotation of TFs

The Gene Ontology (GO, <http://www.geneontology.org/>) terms for biological process, cellular component, and molecular function categories [28], as well as Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (<https://www.kegg.jp/kegg/>) [29–31], were enriched based on the OmicShare online tool with default parameters (<https://www.omicshare.com/>).  $P$ -value  $< 0.05$  were considered to be significantly enriched.

### Data validation by quantitative real-time PCR

Embryonic gene expression analysis for 16 selected hub TFs, based on RNA-seq results, was validated by Quantitative real-time polymerase chain reaction (qRT-PCR). qRT-PCR was performed with an CFX96™ Real-Time system (BIO-RAD, USA) using the SYBR Green qPCR Master Mix (Bimake, China) according to the manufacturer's instructions. The primers were designed by Primer Premier5 software. GAPDH was used as the internal reference, and the sequences of the gene-specific primers are listed in Table 1. The comparative Ct method ( $2^{-\Delta\Delta C_t}$  method) was used to calculate the relative gene expressions of the samples, which were normalized using the GAPDH mRNA level.

### Statistical analysis

Relative expression differences between consecutive stages were calculated, and a t-test was performed in GraphPad Prism 7 (GraphPad Software, San Diego, CA, USA). The differences were considered to be statistically significant at a  $P$ -value  $< 0.05$ .

## Results

### Global view of transcriptome during chicken early development

To better understand regulation of chicken early development, we performed a comparative transcriptomic analysis. Transcriptome sequencing resulted in a total 813 million raw data for all samples. After removing reads of adapter, reads of poly A and low-quality with a quality score  $< 20$ , more than 807 million high-quality reads were remained for further analysis. Reads

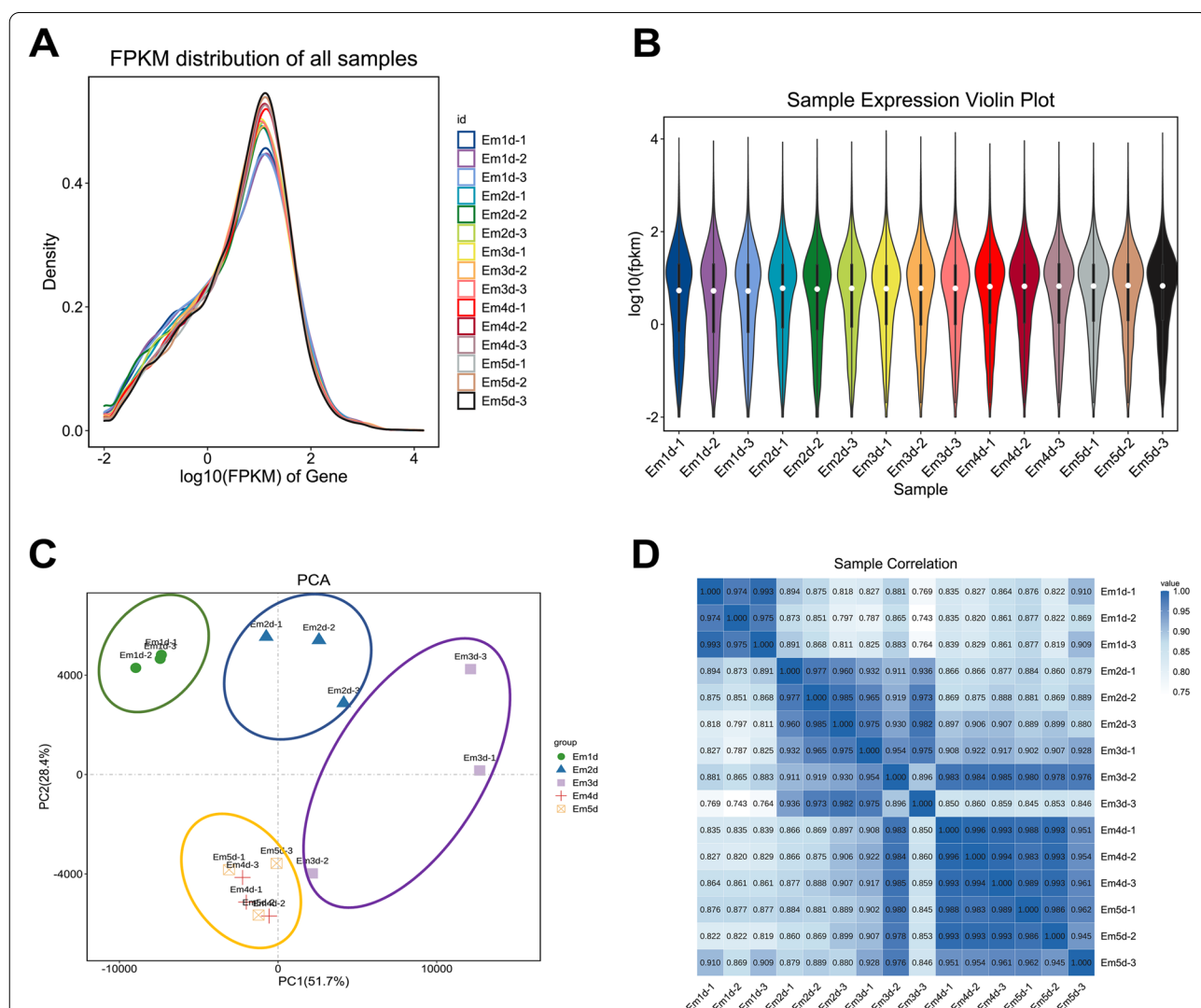
**Table 1** List of primer sequences used in qRT-PCR

| Gene   | Primer   | Accession No.  | Product length (bp) | Annealing temperature (°C) |
|--------|--|----------------|---------------------|----------------------------|
| POU5F3 | F: GGAGCAGTTTGCCAAAGGA<br>R: GAAAGCGCAGATGGTGT               | NM_001309372.2 | 124                 | 61                         |
| NANOG  | F: CAGGTGAAGACGTGGTTT<br>R: AGCCCTGGTGAATGTAG                | NM_001146142.2 | 148                 | 61                         |
| SOX10  | F: TCTGAAGACCACCACCTCCCTC<br>R: CTTGACCTTGCCCATCTCCATTC      | NM_204792.2    | 99                  | 60                         |
| CDX2   | F: AAAACAGGACGAAAGGACAAATACCG<br>R: GGTGATATAGCGGCTGTAGTGAAC | NM_204311.2    | 92                  | 60                         |
| ISL1   | F: TCCGAGGGTTCATCAGGGTTTGG<br>R: TGGGTTGCTGCTGCTGAAGTTG      | NM_205414.2    | 92                  | 60                         |
| PAX6   | F: CTATCCCGATGTTTGGAGAGAG<br>R: CTGGGAGTGTGCTGGCTTGTGTC      | NM_205066.2    | 150                 | 61                         |
| SOX2   | F: AAAACCAAGACCCTGATGAAGA<br>R: ATCCCATAGCCTCCGTTG           | NM_205188.3    | 175                 | 61                         |
| OLIG2  | F: GCTTCAAGTCTCGTCGTGTC<br>R: CGGCTCCGTCATCTGCTTCTTG         | NM_001031526.1 | 80                  | 61                         |
| SMAD3  | F: AGACGGCACATCGGAAGGAGAG<br>R: AATGGCACTGTACTAAGGCATCTC     | NM_204475.2    | 81                  | 61                         |
| GATA6  | F: AGAGAGCACAGTCCCGAAAGC<br>R: ACACCAAGTATCTCGCTGACG         | NM_205420.2    | 121                 | 61                         |
| SOX9   | F: GCTGTGGAGGCTGCTGAATGAG<br>R: CCTGCGTGGTTGGTACTTGTAGTC     | NM_204281.2    | 112                 | 60                         |
| HNF-4A | F: GCCAACCTCAACACCTCCAACAG<br>R: TCCTCAGGCTCCTCTGAAGAAG      | NM_001030855.2 | 133                 | 61                         |
| MKX    | F: GAACACAGTCAGGCAACCAGACC<br>R: CCATCTTCAGAGCACGAGTCAITCAC  | XM_015282064.4 | 120                 | 60                         |
| THRB   | F: AATCAGTGCCAGGAATGTCGCTTC<br>R: GCCCTTGTGTATCCAACACC       | XM_015282064.4 | 82                  | 61                         |
| NFIA   | F: GGAGCCGTTCTATACAAGCCAAAGG<br>R: GAAAGCGGAGGACTGCTGAAAC    | NM_205273.2    | 141                 | 60                         |
| ZBTB16 | F: TCCACCGCAACAGTCAGCATTAC<br>R: CGTAGACGAGGTATCCAAGTCTTC    | NM_001321488.2 | 126                 | 61                         |
| GAPDH  | F: GCCATCACGCCACACAGA<br>R: TTTCCCCACAGCCTTAGCA              | NM_204305.2    | 120                 | 60 / 61                    |



from each sample were aligned to the chicken reference genome (Supplementary Figure S1). The average number of raw data, filter data, GC content, number of mapped reads and mapping rate for samples are shown in Supplementary Table S1. From each stage, a total number of 93.90–95.21% reads were successfully mapped. Approximately 80% of transcripts exhibited great gene coverage (Supplementary file Figure S2). The number of genes displayed saturation tendencies, and all samples were distributed in a homogeneous and random manner. (Supplementary Figure S3). Original

gene read counts were normalized using the FPKM (Fragments per kilo-base of exon per million fragments mapped) method. Figure 1A represents the FPKM distribution of mRNAs, while Fig. 1B depicts the expression of different samples as a violin chart (Fig. 1B). Principal components analysis is useful for exploring the distance relationship between samples. The 15 samples were divided into four parts, which showed satisfactory repeatability and strong clustering associated with development stage, excluding sample Em3d-2 (Fig. 1C). To be clear, although the principal component



**Fig. 1** Overview of gene expression during early development in chicken. **A** The density distribution of mRNAs was according to log10 (FPKM); **B** The 15 Samples expression (Em1d-1, Em1d-2, Em1d-3, Em2d-1, Em2d-2, Em2d-3, Em3d-1, Em3d-2, Em3d-3, Em4d-1, Em4d-2, Em4d-3, Em5d-1, Em5d-2, Em5d-3) violin plot, which was replaced by log10 (FPKM). **C** Principal components analysis reveal strong clustering associated with different stages of embryonic development. **D** Sample relationship heatmap plot revealed exact stage of development except for Em3d-2. Dark blue represents strong correlation and light blue represents weak correlation, each column and row correspond to one sample's relationships with the other 15 samples including itself

analysis shows that sample Em3-2 is more similar to day4 and day5, the correlation analysis presents a greater convincing result that Em3d-1, Em3d-2 and Em3d-3 are good replicates with  $>0.85$  Pearson correlation coefficient. Additionally, with the low degree of outlier that would not affect the following analysis, we did not eliminate the sample Em3d-2. Then, we established a relationship cluster heatmap plot to reflect the relationship between samples intuitively (Fig. 1D). Data showed a reliable clustering effect, which ensured the veracity of the subsequent analysis except for the sample Em3d-2.

#### Identification of DEGs during early development of chicken

To investigate embryonic development alterations in the gene expression pattern during the early stages, differential gene expression analysis was conducted among the five developmental stages in chicken using the software package DESeq2. Generally, the expression of 18,325 distinct genes was identified, including 847 novel genes. The highest number of expressed genes (15,398) occurred on day 5 of embryo, while Em1d sample contained the lowest number of expressed genes (14,536) (Fig. 2A). Subsequently, DEGs ( $FDR < 0.05$  and Fold Changes  $> 2$ ) were identified by comparing two consecutive developmental stages. The number of DEGs varied from 267 (251 upregulated and 16 downregulated) between 5 and 4-day of embryo, to 2920 (2081 upregulated and 839 downregulated) between 2 and 1 day of embryo (Fig. 2B). Interestingly, up-regulation dominated the genes expression patterns in all comparisons, except for the transition from Em3d to Em4d stages, while 51% of genes showed down-regulation. Additionally, Hierarchical clustering of DEGs, based on log<sub>2</sub>-transformed expression values, was able to cluster these stages into distinct groups (Fig. 2C). Unexpectedly, stages Em1d and Em2d were clustered together in one group, while Em3d, Em4d and Em5d were grouped in a separate cluster, indicating that a major shift occurred in that situation.

#### Transcription factor expression patterns during early development of chicken embryos

To visualize the landscape of transcription factors at the genome-wide level, we have constructed a CIRCOS diagram (Fig. 3A). From the results, a total of 1134 TFs (Supplementary Table S2) were distributed in 32 normal chromosomes and 2 sex chromosomes (Z and W), where 41 TFs were located in Z chromosome but only 6

in W chromosome. The fact that TFs were abundant in the left hemisphere suggests that their location in the genome was not random. Then, to explore the different contributions of TFs in different stages of early embryonic development, we identified multiple TFs in variation of expression (Fig. 3B). The most different expression TFs (DE-TFs) change was observed in the transition from Em2d to Em3d, while fewer and fewer counts of DE-TFs are getting involved in later stages, where expression of only 27 TFs changes during the transition from Em4d to Em5d. Furthermore, to investigate TFs that express commonly between successive stages in embryo development throughout the early embryonic period, we performed Venn on DE-TFs at different stages. Figure 3C shows that 32 DE-TFs are expressed from Em1d to Em4d, while 5 TFs from Em2d to Em5d. More importantly, transcription factor OSR2 and EOMES were observed that significantly different expressing among all stages, from Em1d to Em5d. Additionally, 164 TFs showed constant and highly expressed through all stages (Supplementary Table S3).

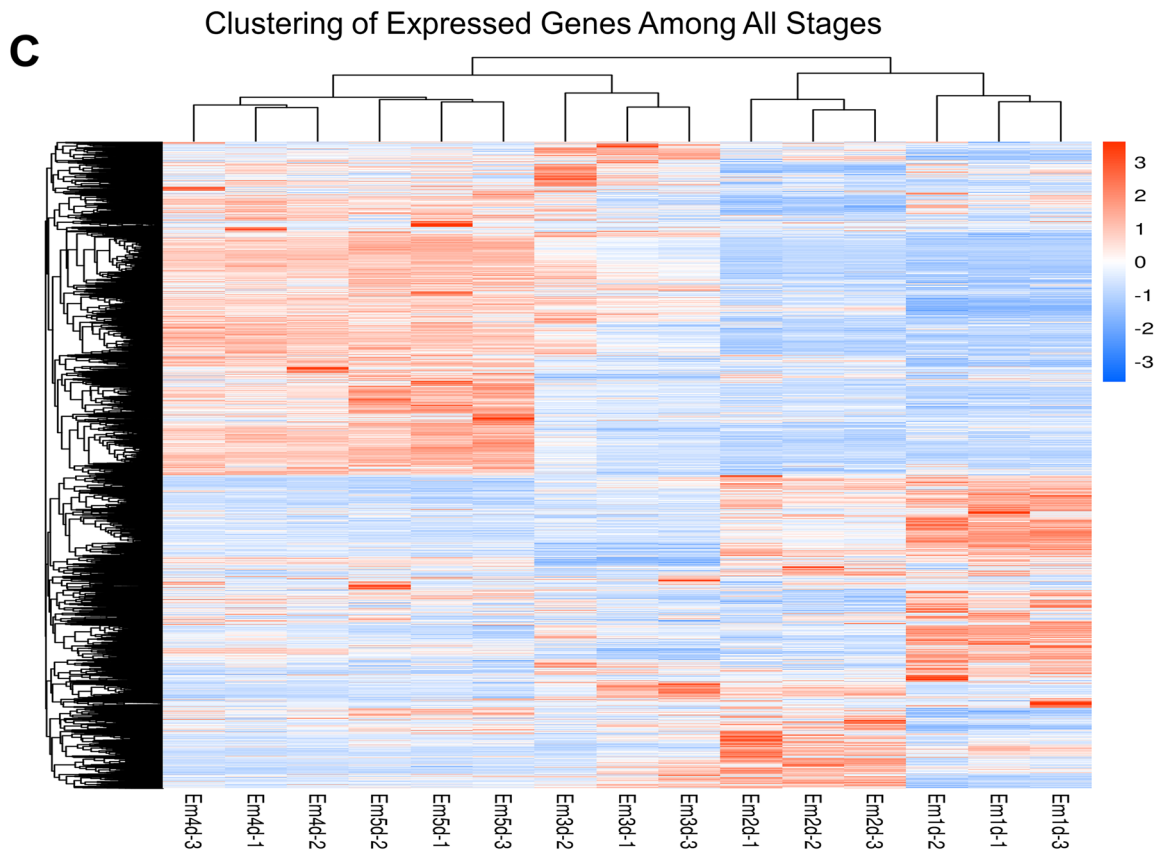
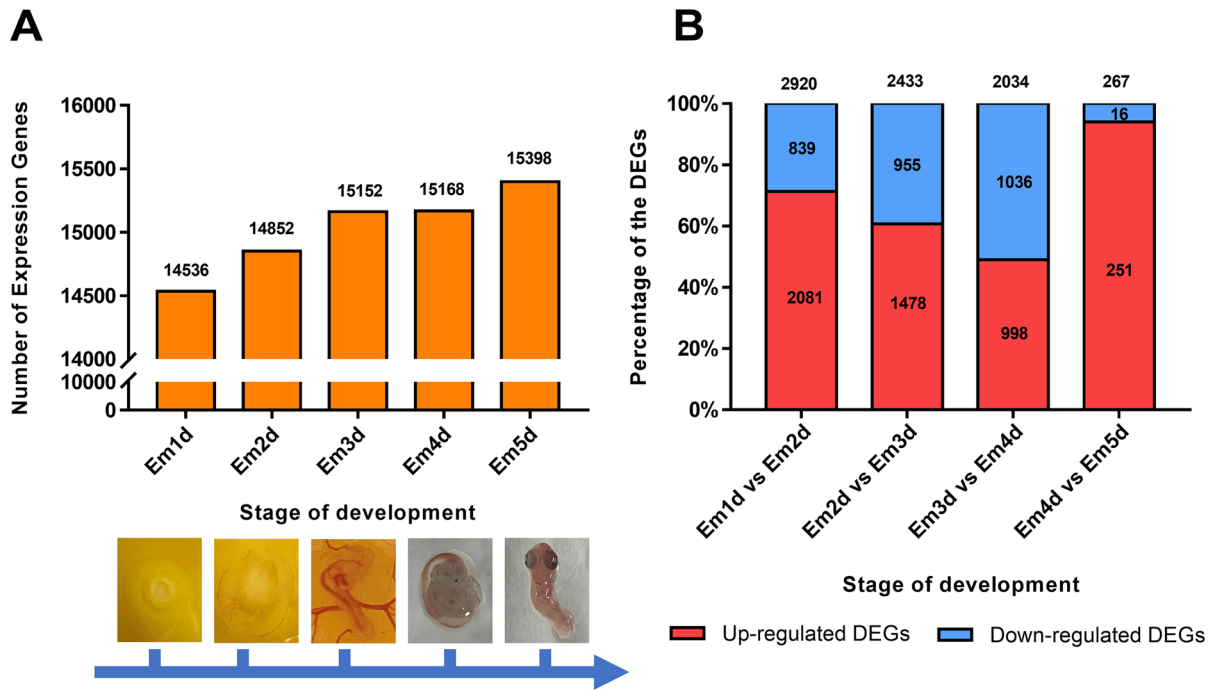
Then, we profiled the time series analysis to illustrate the dynamic changes of TFs. All TFs were clustered into 15 trends, of which three trends appeared significant ( $P < 0.05$ ) (Fig. 4A). The time-series line of differential gene expression is shown in Fig. 4B. The overall TF expression trend was classified as either rising or falling. Generally, a total of 232 DE-TFs were significantly enriched in always up-regulation trend (profile 14) and 55 DE-TFs were consistently down regulation (profile 0) (Fig. 4C). These findings demonstrate the gene expression status of embryo development in the early stages.

#### Function enrichment analysis of differential expressed transcription factors

Cluster analysis and GO enrichment analysis were used to explore the differential expressed TFs. As shown in Fig. 5, two distinct clusters were observed when the later stage compared with the previous stage, indicating significant differences in the regulation of transcription factor expression. Therefore, we performed GO enrichment analysis between the biological processes for up-regulated and down-regulated TFs groups separately. As we known, the transcription processes, biosynthetic processes, and binding processes are the main processes controlled by these TFs, thus we have excluded these annotations from the list of biological processes. The top 10 enrichment significant terms ( $P < 0.05$ ) in

(See figure on next page.)

**Fig. 2** The differential expression analysis of genes. **A** Mean number of expressed genes of three replicates identified at each development stage. **B** The number of differentially expressed genes (DEGs) for comparison of each stage with the next stage. **C** Cluster analysis of gene expression. Embryos at different stages were clustered into two distinct clusters. One cluster contains replicates of Em1d and Em2d, two-stages. While Em3d, Em4d, and Em5d into the second cluster. This indicates a major shift in the gene expression from 3 stage onward



**Fig. 2** (See legend on previous page.)

the Biological Process section were displayed in Fig. 5. Obviously, early stages of embryonic development have a large number of biological processes, while a limited terms are identified at later stages. For instance, at the transition from Em1d to Em2d stages (Fig. 5A), about 30% of up regulated TFs are related to skeletal system development, epithelium development, nervous system development and embryonic morphogenesis (Supplementary Table S4), especially containing cell fate commitment related TFs (*NKX2-5*, *PRRX1*, *LEUTX*, *SOX9*, *SOX8*, *NR2F2*, *SATB2*, *TBX5*, *HOXD10*, *PROX1*, *PAX6*, *ZNF521*, *NR221*, *GCM1*, *PITX1*, *AR*, *FOXA1*, *GLI3*), while 20% of down regulated TFs are related to embryonic morphogenesis and epithelium development including *LHX1*, *GBX2*, *MSX1*, *EOMES*, *OTX1*, *ZIC3*, *SOX17*, *SALL4*, *SP9*, *SCX* (Supplementary Table S4). In the comparison between Em2d and Em3d stages (Fig. 5B), we found that the counts of up regulated TFs in limb development, tube development, brain development, head development and nervous system development are enormous growth (Supplementary Table S4). In contrast, down regulated TFs were involved in reproductive system development and placenta development (*GATA2*, *HNFA1*, *OVOL2*, *PRDM1*, *GCM1*, *GATA4*, *ARID5B*, *FOXA1*, *VDR*, *TBX3*). Gonad development and sex differentiation such as *LHX9*, *SOX9*, *SOX8*, *OSR1*, *HOXA10*, *FOXL2*, *AR*, *NHLH2*, *HOXA11*, *ZFPM2* were up regulated during Em3d to Em4d stages, while TFs (*HNFA1*, *FOXA2*, *HAND1*, *PITX2*) controlling mesenchyme development were observed that down regulated. Moreover, fewer significant different expressed TFs were detected at the stages from Em4d to Em5d, *MYOD1*, *THRB*, *NR4A2*, *RORB*, *EOMES*, *TBR1*, *SOX14*, *OSR2*, *RUNX2*, *NFATC1*, *HELT* involved in cell differentiation were up-regulated, while down regulated *LIN28A* and *SALL4* were enriched in stem cell population maintenance (Supplementary Table S4).

Additionally, KEGG enrichment analysis and GO enrichment analysis were performed to investigate the TFs with the same expression pattern in a time line (Fig. 4D-F). From the KEGG and GO results, we found that the trend profile 14 was enriched in regulating pluripotency of stem cell and cell differentiation including myoblast differentiation, stem cell differentiation and neuron differentiation (Fig. 4D). Especially, MAPK signaling pathway and Hippo signaling pathway both involved in differentiation and stemness. However, in the trend profile 0, TFs were mainly enriched in stem

cell population maintenance and blastocyst development (Fig. 4F).

#### Dominant transcription factor families in early embryo development

Based on structure of DNA-binding domains that are important evolutionary units mediating the specificity of the TF-DNA interaction, transcription factors can be grouped into different families [32]. According to our data, we analyzed the distribution of TF families of DE-TFs at five stages in embryos and found that there were different distributions in the top three largest TF families. The bubble plot (Fig. 6A) showed that *zf-C2H2*, Homeobox and bHLH are three dominant TF families (Supplementary Table S5). Interestingly, *zf-C2H2*, as best known and largest TF family in human [33], is also represent the major class of chicken transcription factors. On the other hand, however, we found that Homeobox family occupied the largest portion and was expressed during the whole stages, while bHLH family contain fewer TFs expressing mostly occurred in Em4d and Em5d (Fig. 6B-D).

#### Network construction and analysis of hub transcription factors at each stage of embryonic development

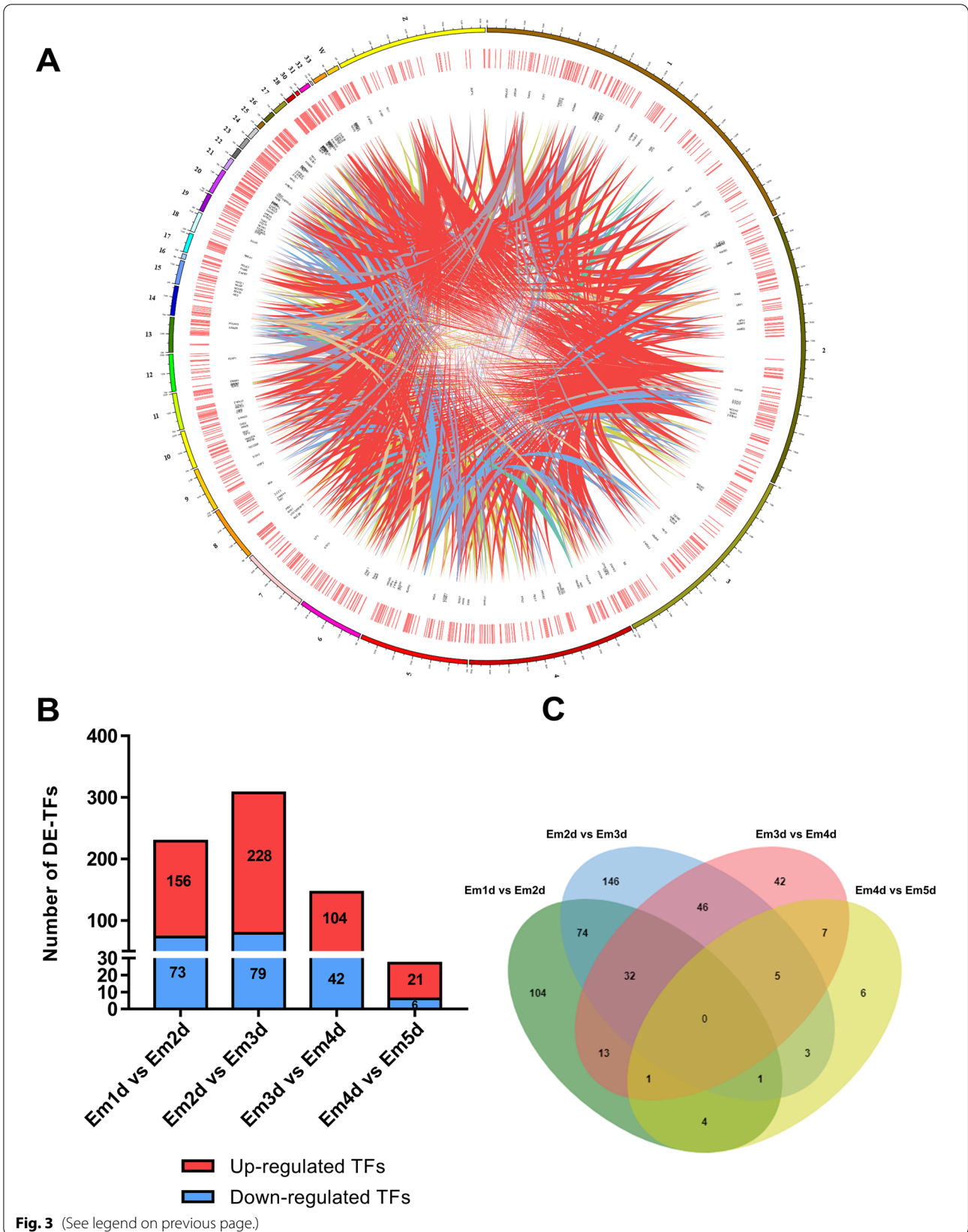
To further identify the function of the co-expressed TFs in different stages and investigate the hub TFs, we have constructed co-expression network. Additionally, a core regulatory networks (Fig. 7) were extracted from the whole network analysis through MCODE algorithm. During the transition from Em1d to Em2d stage, we have detected *EOMES*, *POU5F3*, *PAX6*, *SOX9*, *GATA4*, *NKX2-5*, *OTX2* and *SOX10* as key factors for regulation of TFs (Fig. 7A). The network analysis showed that *GATA4* has the highest number of interactions with other TFs and highly expressed in Em1d stage. Importantly, *POU5F3*, *NANOG* and *CDX2* were also detected as hub genes in the core network (Fig. 8A).

As shown in Fig. 7B, *SOX2*, *OTX2*, *SOX9*, *ISL1*, *FOXG1*, *PAX2* and *PAX6* play a key role during transition from Em2d to Em3d stage, which all were up-regulated. However, the core regulatory network analysis at these stages indicated that *ISL1*, *PAX6*, *SOX2* and *OLIG2* are the hub proteins (Fig. 8B).

When it comes to transition of Em3d to Em4d, the embryos are mostly regulated by *SMAD3*, *MOYD1*, *SOX9*, *GATA2*, *GATA6* and *EOMES* with the highest number of connections (Fig. 7C). On the other hand, the core regulatory network detected not only *SMAD3*, *SOX9*

(See figure on next page.)

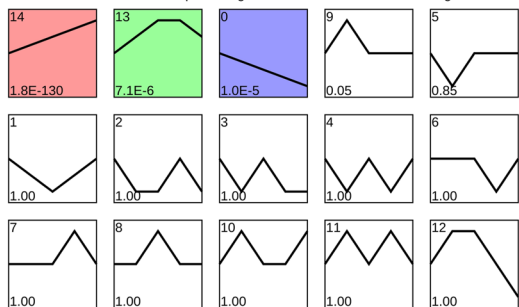
**Fig. 3** Analysis and detection of DE-TFs during early chicken embryogenesis. **A** CIRCOS visualization of TFs at the genome-wide level; from outside to inside: Karyotype of the chicken genome, expression of TFs, specific TFs symbol and linkage of TF family. **B** DE-TFs were identified during the transition of the embryo from Em1d to Em5d. The red color indicates the up-regulated TFs, while the blue color indicates the down-regulated TFs. **C** The status of common TFs involved in each transition. DE-TF: differentially expressed TF; TF: transcription factor



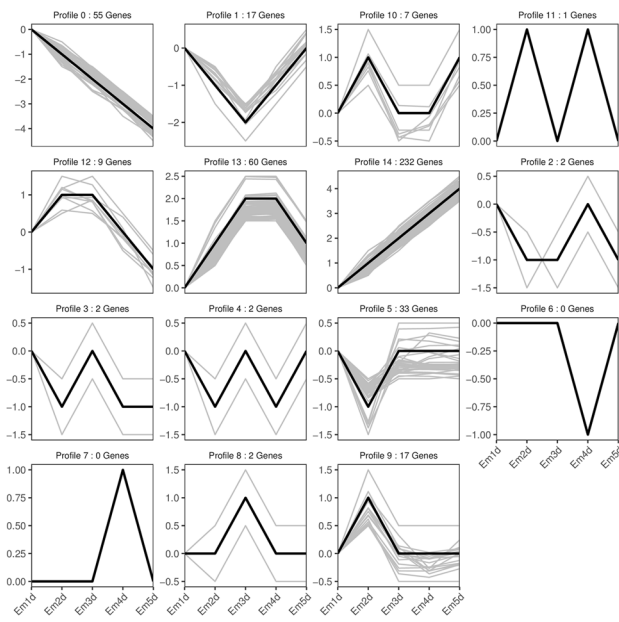


**A**

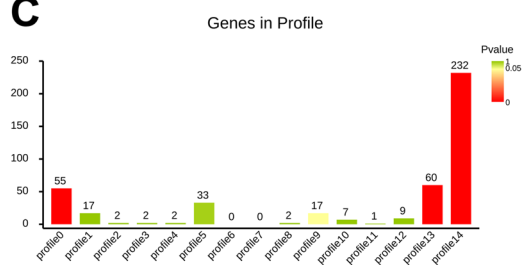
Profiles ordered based on the pvalue significance of number of Genes assigned versus expected



**B**

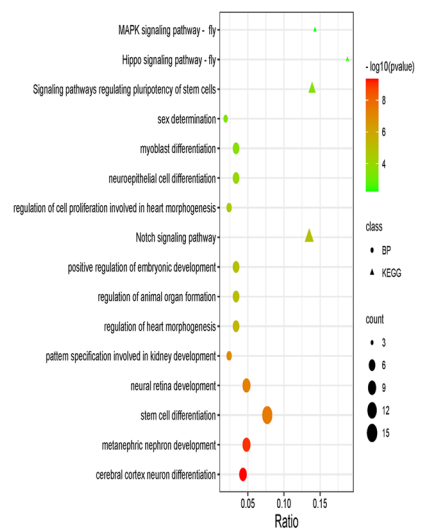


**C**



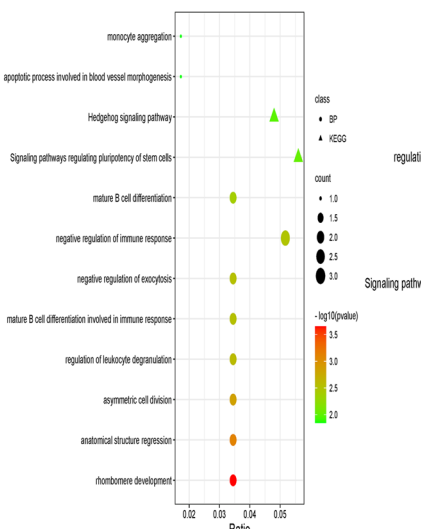
**D**

TF-Trend-profile 14



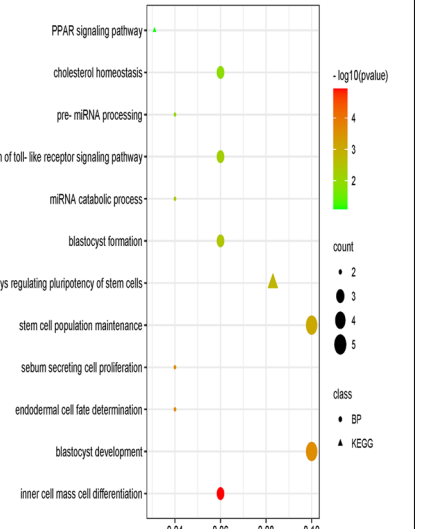
**E**

TF-Trend-profile 13



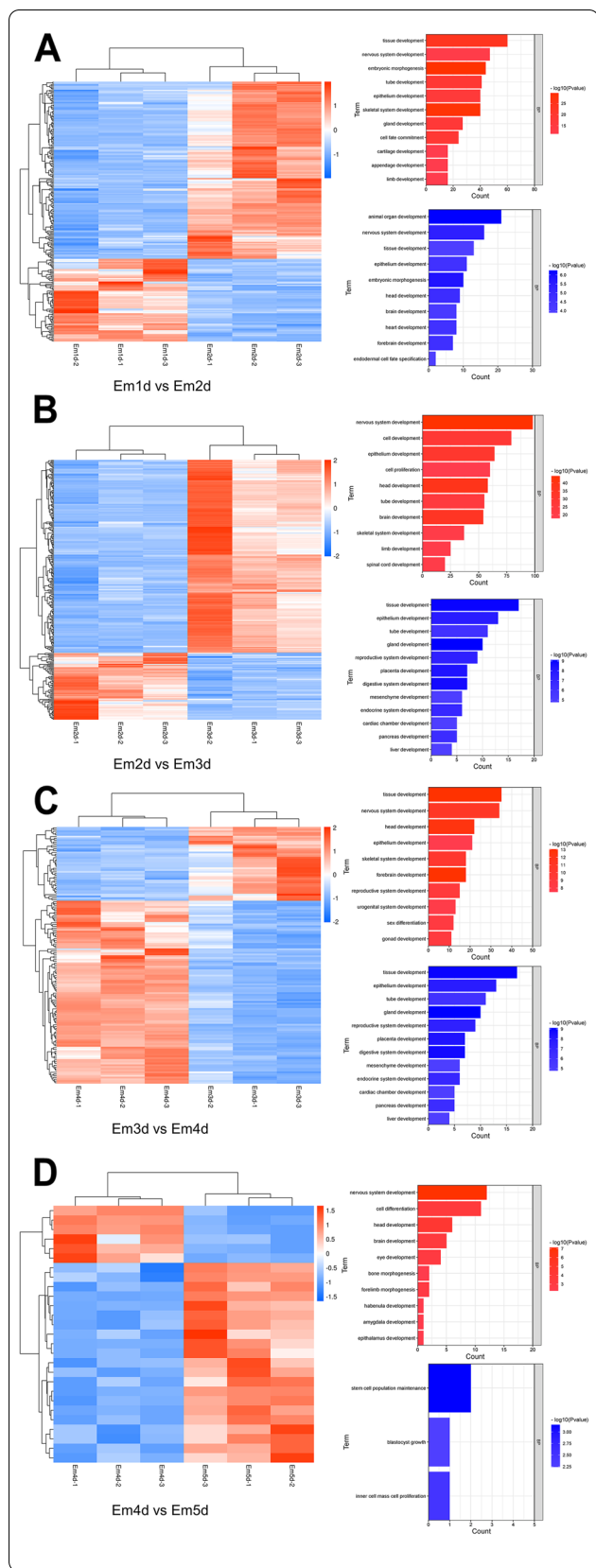
**F**

TF-Trend-profile 0



**Fig. 4** The sample time series analysis of DE-TFs. **A** Distribution trend of differential transcription factors, color means significant difference ( $P < 0.05$ ), gray means not significant ( $P > 0.05$ ); **B** The time series line of differential transcription factors. **C** Representative terms of GO and KEGG using profile-14 TFs. **D** Representative terms of GO and KEGG using profile-13 TFs. **E** Representative terms of GO and KEGG using profile-0 TFs. KEGG pathway database was used to analyze the data





**Fig. 5** Clustering based on DE-TFs and the annotation of the DE-TFs. Distinct clusters in all stages are evident based on the expression of DE-TFs. GO of TFs was also provided alongside each cluster. **A** Transition from Em1d to Em2d, **B** transition from Em2d to Em3d, **C** transition from Em3d to Em4d, **D** transition from Em4d to Em5d. DE-TF: differentially expressed TF; GO: gene ontology; TF: transcription factor

and *GATA6* as hub genes, but also *HNF4A* and *CDX2* (Fig. 8C). Moreover, the pattern of expression during this transition is that *SMAD3* and *SOX9* are up-regulated while *GATA6*, *HNF4A* and *CDX2* are down-regulated.

In the last period, the least differential expressed TFs resulted in that *NIFA*, *THRB*, *MKX*, *OSR2* and *ZBTB16* are detected as hub genes for both PPI network and core regulatory network (Figs. 7D and 8D). Besides, only *ZBTB16* was down-regulated.

Nevertheless, the top significantly enriched pathways particular to the hub TFs include the signaling pathways regulating pluripotency of stem cells, cell cycle, FOXO signaling pathway, AMPK signaling pathway, Hippo signaling pathway and cAMP signaling pathway ect. Also, the network of key pathways was constructed and was displayed in Fig. 9B. We identified two clusters with the predominant clusters belonging to regulation of pluripotency of stem cells and cell cycle signaling pathways as depicted in Fig. 9A. From Fig. 9A, it is shown that TFs such as *NANOG*, *POU5F3*, *SOX2*, *ISL1*, and *PAX6* were the one which are involved in regulation of pluripotency of stem cells, whereas TFs such as *SOX9*, *SMAD3*, *CDX2*, *ZBTB16*, and *HNF4A* were the one associated with cell cycle signaling pathways.

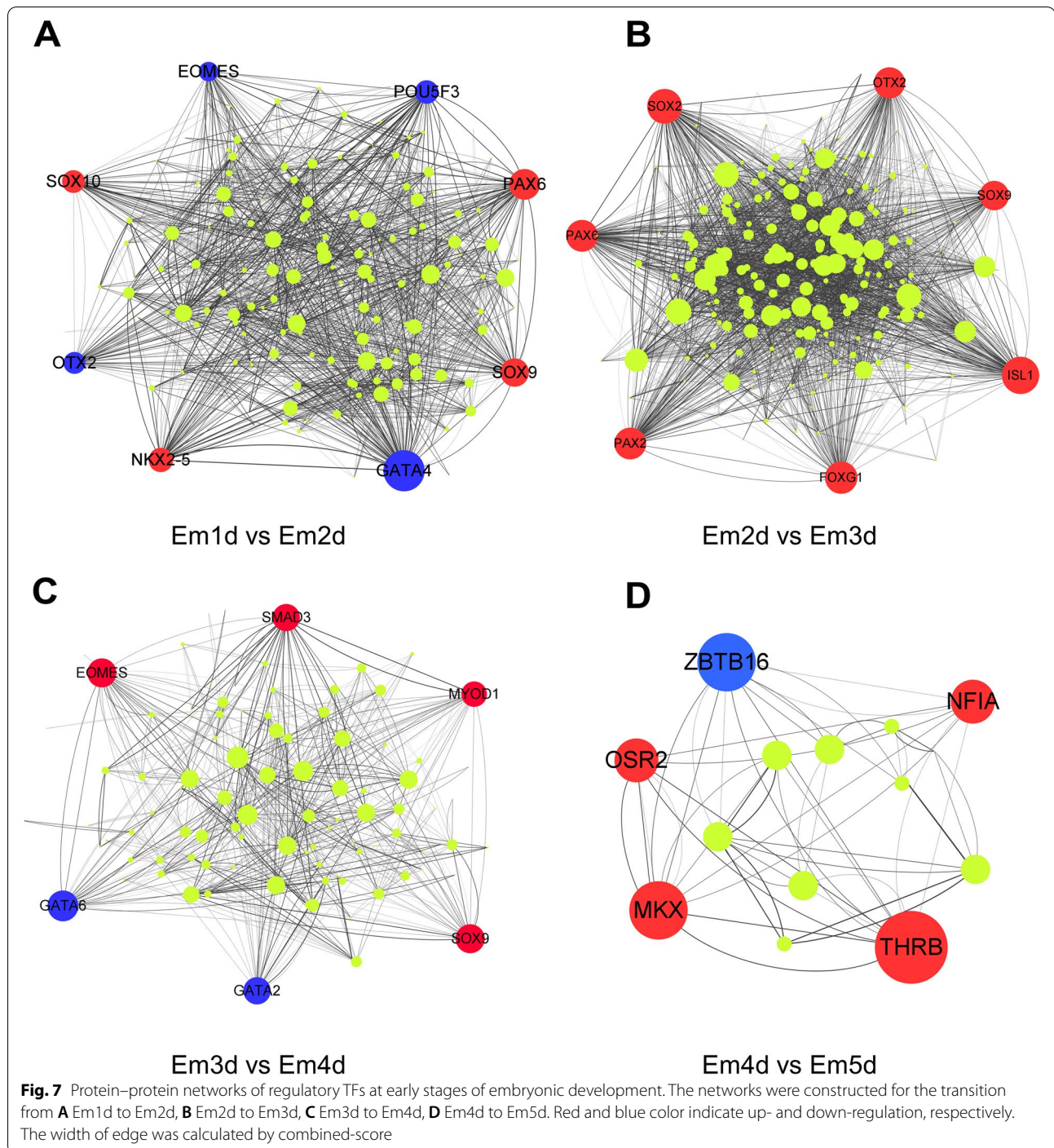
**Validation of the hub TFs in embryonic development by RT-qPCR**

To validate the 16 selected hub TFs at different stages during early embryonic development, RT-qPCR was conducted to illustrated the gene expression shown in Fig. 10. Differences in embryonic TF expression at each stage profiled by RNA-seq results were confirmed for all of 16 genes by qPCR (*P* value < 0.05). Evidently, comparable patterns and similar trends in gene expression could be observed for the key TFs. These findings could validate the specific role of these TFs.

**Discussion**

Chicken have long been regarded as an ideal model for virology, physiological and behavioral traits, immunology, biotechnology and developmental biology [34–38]. In light of the importance of the chicken to human societies around the world, genetic diversity and gene regulatory of the chicken (*Gallus gallus*) is of great interest [39]. Since that vast majority of biological processes,

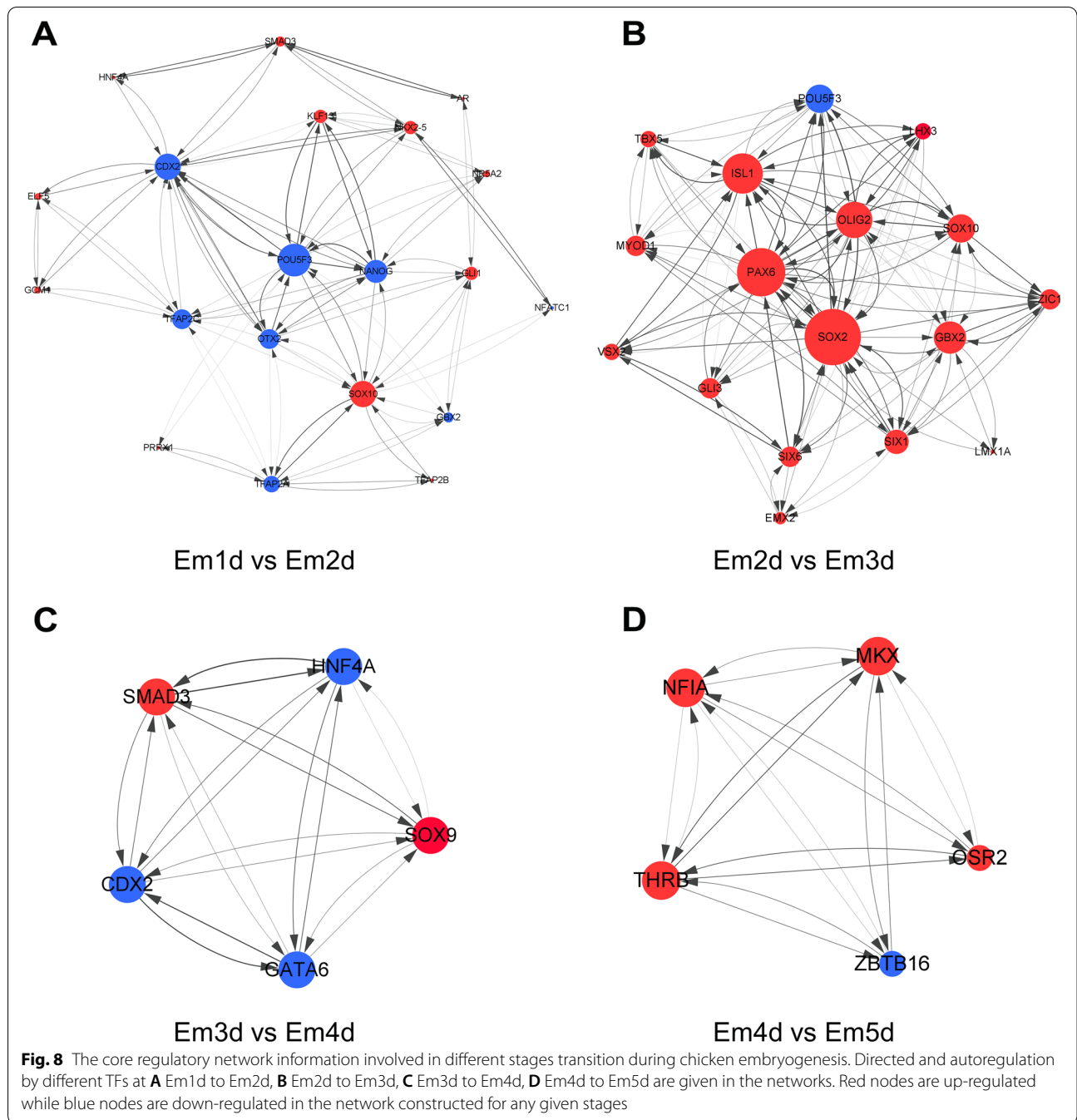




to Homeobox family (Table 2). Accumulating evidences show that the bHLH factors correlate with multipotent and proliferative state and regulate fate determination of somatic cells into neurons [53–55]. More importantly, the cranio-caudal polarity, as well as that of specific cell groups within the somites, is determined by transcription

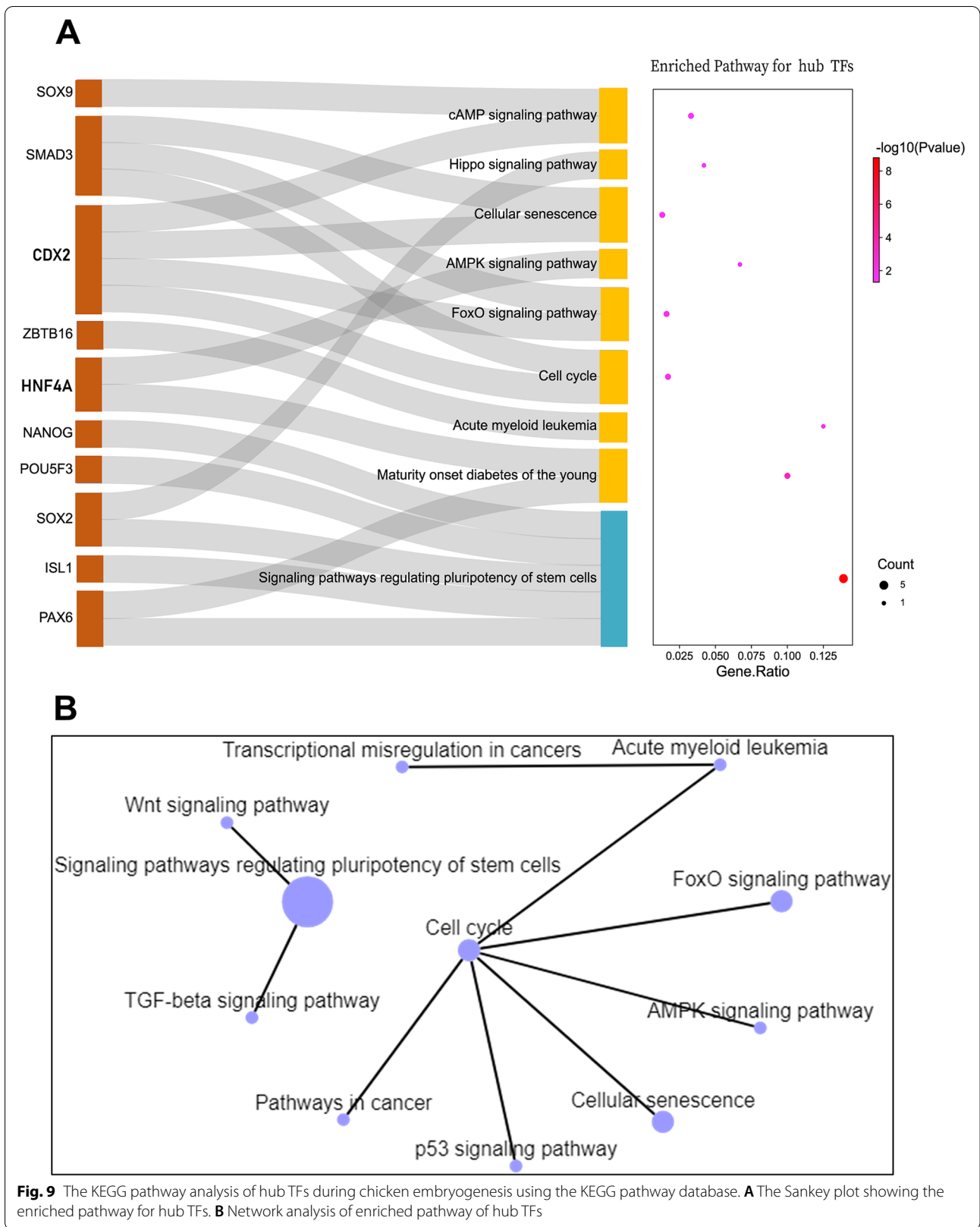
factors of the bHLH and homeodomain type. According to our study, it is found that the bHLH factors were highly expressed in Em4d and Em5d, which have more responsibility for nervous system development. Additionally, 164 constant and highly expressed TFs were observed in all stages, indicating that these TFs are common and necessary in development (Supplementary Table S3).





Embryonic development related TFs have different regulatory effects at different stages of development. Simultaneously, there are significantly change in TF expression at different developmental times. Therefore, time series analysis was utilized to characterize TF expression and disclose the law of embryonic development at various stages. Subsequently, differentially expressed TFs are clustered into three mainly trend profiles. Different TFs in the same trend were analyzed for their involvement in

the same biological process using functional enrichment analysis. Multiple development-related terms were considerably enriched when the GO and KEGG analysis was applied to the increasing trend, such as MAPK signaling pathway, Hippo signaling pathway, PPAR signaling pathway and pathways regulating pluripotency of stem cells. Notably, it was discovered that active p38-MAPK signaling is required for blastocyst development [56]. Interestingly, not only the involvement of the FGF/MAPK



| TF     | Technology | Em1d  | Em2d  | Em3d  | Em4d  | Em5d  | FDR RNA-seq/<br>P value qPCR |              |              |              |
|--------|------------|-------|-------|-------|-------|-------|------------------------------|--------------|--------------|--------------|
|        |            |       |       |       |       |       | Em1d vs Em2d                 | Em2d vs Em3d | Em3d vs Em4d | Em4d vs Em5d |
| NANOG  | RT-qPCR    | Red   | White | Blue  | White | White | 1.11E-03                     | 9.62E-06     | —            | —            |
|        | RNA-seq    | Red   | White | Blue  | White | White | 8.71E-11                     | 4.86E-13     | —            | —            |
| SOX9   | RT-qPCR    | Blue  | Blue  | White | Red   | Red   | 2.81E-03                     | 4.93E-05     | 5.79E-04     | —            |
|        | RNA-seq    | Blue  | Blue  | White | Red   | Red   | 3.22E-04                     | 1.48E-03     | 4.30E-10     | —            |
| CDX2   | RT-qPCR    | Red   | White | Blue  | White | White | 1.14E-05                     | 1.03E-06     | 2.50E-05     | —            |
|        | RNA-seq    | Red   | White | Blue  | White | White | 7.92E-04                     | 2.51E-05     | 1.72E-11     | —            |
| SMAD3  | RT-qPCR    | White | Red   | White | White | White | 7.36E-05                     | —            | 8.72E-04     | —            |
|        | RNA-seq    | Blue  | White | Blue  | Red   | Red   | 7.74E-23                     | —            | 1.94E-18     | —            |
| SOX10  | RT-qPCR    | Blue  | White | White | Red   | White | 3.60E-04                     | 8.54E-04     | —            | —            |
|        | RNA-seq    | Blue  | White | White | Red   | White | 3.94E-11                     | 3.76E-06     | —            | —            |
| PAX6   | RT-qPCR    | Blue  | White | White | Red   | White | 4.54E-05                     | 8.00E-03     | —            | —            |
|        | RNA-seq    | Blue  | White | White | Red   | White | 2.90E-16                     | 6.70E-04     | —            | —            |
| SOX2   | RT-qPCR    | White | Blue  | White | Red   | Blue  | —                            | 3.50E-03     | —            | —            |
|        | RNA-seq    | Blue  | Blue  | White | Red   | White | —                            | 2.07E-11     | —            | —            |
| NFIA   | RT-qPCR    | White | White | White | White | Red   | —                            | —            | 6.31E-05     | 1.93E-04     |
|        | RNA-seq    | White | White | White | White | Red   | —                            | —            | 2.58E-11     | 8.97E-06     |
| ISL1   | RT-qPCR    | White | Blue  | White | Red   | White | —                            | 2.45E-03     | —            | —            |
|        | RNA-seq    | Blue  | Blue  | White | Red   | White | —                            | 2.58E-06     | —            | —            |
| GATA6  | RT-qPCR    | Red   | White | White | Blue  | White | —                            | —            | 3.47E-06     | —            |
|        | RNA-seq    | Red   | White | White | Blue  | White | —                            | —            | 1.26E-05     | —            |
| THRB   | RT-qPCR    | Red   | White | White | White | White | —                            | —            | —            | 4.01E-02     |
|        | RNA-seq    | Blue  | White | White | White | Red   | —                            | —            | —            | 1.20E-02     |
| HNF4A  | RT-qPCR    | Red   | White | White | Blue  | White | —                            | —            | 2.85E-05     | —            |
|        | RNA-seq    | Red   | Red   | White | Blue  | White | —                            | —            | 1.80E-05     | —            |
| ZBTB16 | RT-qPCR    | White | White | White | White | White | —                            | —            | —            | 3.20E-05     |
|        | RNA-seq    | Blue  | White | Red   | White | Blue  | —                            | —            | —            | 1.40E-09     |
| POU5F3 | RT-qPCR    | Red   | White | White | White | White | 3.61E-04                     | 1.18E-04     | —            | —            |
|        | RNA-seq    | Red   | White | White | White | White | 1.33E-17                     | 2.01E-18     | —            | —            |
| OLIG2  | RT-qPCR    | White | Red   | White | White | White | —                            | 1.41E-05     | —            | —            |
|        | RNA-seq    | Blue  | Blue  | Red   | White | White | —                            | 2.32E-10     | —            | —            |
| MKX    | RT-qPCR    | White | Blue  | White | White | Red   | —                            | 2.82E-03     | 8.53E-04     | 8.11E-05     |
|        | RNA-seq    | White | Blue  | White | White | Red   | —                            | 3.43E-04     | 1.21E-02     | 2.98E-07     |

**Fig. 10** The validation of selected hub TFs by RT-qPCR: Heatmap are constructed of qPCR and RNA-seq data for 16 selected genes. The relationship between the relative expression levels of RNA-seq and qPCR data (mean-centered log2 expression values) are shown as a heatmap. The color red denotes higher gene expression levels, whereas the color blue denotes lower gene expression levels when compared to the mean of all samples, respectively (from 2 over 0 to -2). Statistical differences are represented by FDR for the RNA-seq data and by P value for qPCR data ( $P < 0.05$ )

signaling pathway in early neural crest induction during gastrulation has been elucidated, and it also plays many roles in the formation of ectodermal tissues [57]. The HIPPO signaling pathway is highly conserved across animal species ranging from drosophila to mouse [58]. Additionally, Hippo signaling is important in early embryonic development and positively or negatively regulates development of multiple tissues/ organs [59]. Besides, increasing evidences highlight the functional importance of PPAR related gene expression during embryonic development and the maintenance of embryonic stem cells' pluripotent state [43, 60]. Notwithstanding, the mechanisms by which signaling pathways influences development of embryo are not entirely clear, and further studies are needed to supplement the gap.

Gene regulation networks (GRNs) control a variety of developmental and cellular functions including cell differentiation and cell fates by regulating gene

expression [61]. Transcription factors control the expression of regulatory genes and all other genes by means of regulatory interactions [62, 63]. Therefore, it is important to explore hub TFs in the early embryo development by constructing gene regulation networks. In our networks, we found the main regulators of transition in the early stages from Em1d to Em5d. In our results, *NANOG*, *POU5F3*, and *CDX2* were found that play a pivotal role in the core regulatory network of transition from Em1d to Em2d. Especially, *NANOG* and *POU5F3* are involved not only in fundamental events such as zygotic genome activation (ZGA), but also in the acquisition of pluripotency that occurs at stage EGK.VI to EGK.VIII [16]. However, in our presented data, *NANOG* and *POU5F3* were significantly down regulated thereafter. *CDX2* plays a well-defined role in determining the first lineage decisions and in assigning positional identity during orchestrated



**Table 2** List of detected hub TFs in chicken embryo development

| Symbol | NCBI Gene ID | Chromosome | TF family |
|--------|--------------|------------|-----------|
| POU5F3 | 427,781      | 17         | Pou       |
| NANOG  | 100,272,166  | 1          | Homeobox  |
| SOX10  | 395,573      | 1          | HMG       |
| CDX2   | 374,205      | 1          | Homeobox  |
| ISL1   | 396,383      | Z          | Homeobox  |
| PAX6   | 395,943      | 5          | PAX       |
| SOX2   | 396,105      | 9          | HMG       |
| OLIG2  | 428,612      | 1          | bHLH      |
| SMAD3  | 395,132      | 10         | MH1       |
| GATA6  | 396,390      | 2          | zf-GATA   |
| SOX9   | 374,148      | 18         | HMG       |
| HNF4A  | 419,198      | 20         | RXR-like  |
| MKX    | 771,284      | 2          | Homeobox  |
| THRB   | 396,431      | 2          | THR-like  |
| NFIA   | 396,210      | 8          | NFI       |
| ZBTB16 | 419,759      | 24         | Zf-C2H2   |

process of embryogenesis [64, 65], and is also involved in gut epithelial differentiation and intestinal differentiation [66, 67]. While in the core regulatory network at the stages from Em2d to Em3d, *SOX2*, *OLIG2*, *PAX6*, *ISL1* and *SOX10* are shown interaction with each other, and both participate in central neuronal system development including development of neural crest cell that is important in embryogenesis [68, 69]. In addition, *PAX6* and *ISL1* are required for other neuronal development such as dendrite morphogenesis and pancreatic development [70, 71]. *ISL1* is also known as a marker for cardiac differentiation [72, 73]. Meanwhile, previous studies have shown that *PAX6* is involved in the regulation and development of the eye [74–76]. Network analysis have introduced *SMAD3*, *GATA6* and *SOX9* as hub TFs during the stage from Em3d to Em4d, which are critical players in reproductive development and function [77]. *SMAD3* and *SOX9* was shown highly expression in E4 when PGCs migrate into primitive gonad (develop on ventromedial surface of the embryonic kidney), which promote differentiation of gonad [78–80]. Immunity system and brain development are the main issues in the transition of Em4d to Em5d. For instance, *ZBTB16* regulates innate and innate-like lymphoid lineage development [81, 82]. *THRB* and *NFIA* are both involved in retina and brain development [83, 84]. However, a few studies have investigated the roles of *ZBTB16*, *THRB*, *NFIA* or *MKX* in chicken. Their functions need to be uncovered through further researches.

## Conclusion

This study first analyzed TFs expression pattern from embryonic development stage Em1d to Em5d through RNA-seq, clustering, enrichment and network analysis. Our comprehensive, unbiased analysis of dynamic TFs change during early embryo development in chicken reveals critical regulatory factors and provide new insights into embryogenesis. Collectively, these results offer a basis resource for further studies.

## Abbreviations

DEGs: Differentially expressed genes; TF: transcription factor; EGA: embryonic genome activation; ZGA: zygotic genome activation; Inner cell mass: ICM; FPKM: Fragments per kilo-base of exon per million fragments mapped; PAC: Principal component analysis; FDR: False discovery rate; GO: Gene ontology; KEGG: Kyoto encyclopedia of genes and genomes; qRT-PCR: Quantitative real-time polymerase chain reaction; GRNs: Gene regulation networks.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-022-09054-x>.

**Additional file 1: Figure S1.** Overview of RNA-seq mapping in chicken genome. **Figure S2.** Gene coverage of different samples. **Figure S3.** Sample randomness distribution.

**Additional file 2: Table S1.** Data quality assessment in sequencing.

**Additional file 3: Table S2.** Distribution of TFs in genome.

**Additional file 4: Table S3.** Common TFs highly expressed in all stages.

**Additional file 5: Table S4.** GO enrichment of DE-TF in different stages.

**Additional file 6: Table S5.** Differentially expressed transcription factors (DE-TFs).

## Acknowledgements

We thank the researchers in our laboratory for their assistance in samples collection. We are grateful to Guangzhou Gene-de-novo Biotechnology Co., Ltd for assisting in sequencing and bioinformatics analysis.

## Authors' contributions

Conceptualization: LLQ, ZQY and QMX; methodology: LLQ, ZQY and JK; formal analysis: LLQ and ZQY; funding acquisition: XHZ, HXL, WGC and QMX; investigation: LLQ, ZQY, and JK; supervision: XHZ and QMX; writing—original draft preparation: LLQ; writing—review and editing: XHZ, and QMX. All authors have read and agreed to the published version of the manuscript.

## Funding

This research was supported by grants from Guangdong Provincial Key R&D Program (2020B020222001, 2018B020203001), Guangdong basic and applied basic research fund project named as the China agriculture research system of MOF and MARA (CARS-42-13) (2019B1515210034), the construction project of modern agricultural science and technology innovation alliance in Guangdong province (2021KJ128, 2020KJ128) and the Special Project of National Modern Agricultural Industrial Technology System (CARS-41).

## Availability of data and materials

RNA-Seq raw data in this study are available at NCBI's Sequence Read Archive (SRA) under the BioProject accession PRJNA850787. Additionally, all data generated or analyzed during this study are included in this published article and its supplementary information files.

## Declarations

### Ethics approval and consent to participate

All of the experimental protocols involved in animal care and sample collection were approved by the Animal Ethics Committee at the South China Agricultural University, China (approval ID: SYXK-2022-0136). We have confirmed that all works were followed in strict accordance with the South China Agricultural University Laboratory Animal Welfare and Ethics guidelines and ICLAS Ethical guidelines.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Author details

<sup>1</sup>Heyuan Branch, Guangdong Provincial Laboratory of Lingnan Modern Agricultural Science and Technology, College of Animal Science, South China Agricultural University, Guangzhou 510642, China. <sup>2</sup>Guangdong Provincial Key Lab of Agro Animal Genomics and Molecular Breeding, Guangzhou 510642, China. <sup>3</sup>South China Collaborative Innovation Center for Poultry Disease Control and Product Safety, Guangzhou 510642, P. R. China. <sup>4</sup>Key Laboratory of Animal Health Aquaculture and Environmental Control, Guangzhou, Guangzhou 510642, P. R. China.

Received: 12 July 2022 Accepted: 28 November 2022

Published online: 13 December 2022

## References

- Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, et al. The human transcription factors. *Cell*. 2018;172(4):650–65.
- Vaquerez JM, Kummerfeld SK, Teichmann SA, Luscombe NM. A census of human transcription factors: function, expression and evolution. *Nat Rev Genet*. 2009;10(4):252–63.
- Guy JL, Mor GG. Transcription factor-binding Site Identification and Enrichment Analysis. *Methods Mol Biol*. 2021;2255:241–61.
- Huilgol D, Venkataramani P, Nandi S, Bhattacharjee S. Transcription factors that govern development and disease: an Achilles Heel in cancer. *Genes (Basel)*. 2019;10(10):794.
- Willmer T, Cooper A, Peres J, Omar R, Prince S. The T-Box transcription factor 3 in development and cancer. *Biosci Trends*. 2017;11(3):254–66.
- Krause A, Zacharias W, Camarata T, Linkhart B, Law E, Lischke A, et al. Tbx5 and Tbx4 transcription factors interact with a new chicken PDZ-LIM protein in limb and heart development. *Dev Biol*. 2004;273(1):106–20.
- Golson ML, Kaestner KH. Fox transcription factors: from development to disease. *Development*. 2016;143(24):4558–70.
- Yu X, Yuan Y, Qiao L, Gong Y, Feng Y. The sertoli cell marker FOXD1 regulates testis development and function in the chicken. *Reprod Fertil Dev*. 2019;31(5):867–74.
- Cui C, Han SS, Yin HD, Luo B, Shen XX, Yang FL, et al. FOXO3 is expressed in ovarian tissues and acts as an apoptosis initiator in granulosa cells of chickens. *Biomed Res Int*. 2019;2019:6902906.
- Choi HJ, Jin SD, Rengaraj D, Kim JH, Pain B, Han JY. Differential transcriptional regulation of the NANOG gene in chicken primordial germ cells and embryonic stem cells. *J Anim Sci Biotechnol*. 2021;12(1):40.
- Jacob A, Wust HM, Thalhammer JM, Frob F, Kuspert M, Reiprich S, et al. The transcription factor prospero homeobox protein 1 is a direct target of Sox2. *Biosci Rep*. 2018;38(5):BSR20180707.
- Zhang C, Wang F, Zuo Q, Sun C, Jin J, Li T, et al. Cpod1 promotes chicken SSCs formation with the aid of histone acetylation and transcription factor Sox2. *Biosci Rep*. 2018;38(5):BSR20180707.
- Ren J, Sun C, Clinton M, Yang N. Dynamic Transcriptional Landscape of the early chick embryo. *Front Cell Dev Biol*. 2019;7:196.
- Hwang YS, Seo M, Lee BR, Lee HJ, Park YH, Kim SK, et al. The transcriptome of early chicken embryos reveals signaling pathways governing rapid asymmetric cellularization and lineage segregation. *Development*. 2018;145(6):dev157453.
- Li J, Zhang X, Wang X, Sun C, Zheng J, Li J, et al. The m6A methylation regulates gonadal sex differentiation in chicken embryo. *J Anim Sci Biotechnol*. 2022;13(1):52.
- Han JY, Lee HG, Park YH, Hwang YS, Kim SK, Rengaraj D, et al. Acquisition of pluripotency in the chick embryo occurs during intrauterine embryonic development via a unique transcriptional network. *J Anim Sci Biotechnol*. 2018;9:31.
- Ding N, Gao Y, Wang N, Li H. Functional analysis of the chicken PPAR-gamma gene 5'-flanking region and C/EBPalpha-mediated gene regulation. *Comp Biochem Physiol B Biochem Mol Biol*. 2011;158(4):297–303.
- Wang L, Li S, Xu L, Li Y, Chen H, Chen D. De novo transcriptome sequencing and analysis of the cuttlefish (*Sepiella japonica*) with different embryonic developmental stages. *Anim Biotechnol*. 2021;32(5):602–9.
- Zhou L, Liu Z, Dong Y, Sun X, Wu B, Yu T, et al. Transcriptomics analysis revealing candidate genes and networks for sex differentiation of yesso scallop (*Patinopecten yessoensis*). *BMC Genomics*. 2019;20(1):671.
- Liu C, Sello CT, Sui Y, Hu J, Chen S, Msuthwana P, et al. Characterization of embryonic skin transcriptome in Anser cygnoides at three feather follicles developmental stages. *G3 (Bethesda)*. 2020;10(2):443–54.
- Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357–U54.
- Kim D, Landmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12(4):357–U121.
- Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *Bmc Bioinformatics*. 2011;12:323.
- Perteua M, Perteua 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):290–4.
- Hu H, Miao YR, Jia LH, Yu QY, Zhang Q, Guo AY. AnimalTFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors. *Nucleic Acids Res*. 2019;47(D1):D33–D8.
- Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING database in 2021: customizable protein-protein networks and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res*. 2021;49(D1):D605–D12.
- 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):2498–504.
- Huang da W, Sherman BT, Stephens R, Baseler MW, Lane HC, Lempicki RA. DAVID gene ID conversion tool. *Bioinformatics*. 2008;2(10):428–30.
- Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27–30.
- Kanehisa M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci*. 2019;28(11):1947–51.
- Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res*. 2021;49(D1):D545–D51.
- Charoensawan V, Wilson D, Teichmann SA. Lineage-specific expansion of DNA-binding transcription factor families. *Trends Genet*. 2010;26(9):388–93.
- Najafabadi HS, Mnaimneh S, Schmitges FW, Garton M, Lam KN, Yang A, et al. C2H2 zinc finger proteins greatly expand the human regulatory lexicon. *Nat Biotechnol*. 2015;33(5):555–62.
- Burt D, Pourquie O. Genetics. Chicken genome—science nuggets to come soon. *Science*. 2003;300(5626):1669.
- Stern CD. The chick embryo—past, present and future as a model system in developmental biology. *Mech Dev*. 2004;121(9):1011–3.
- Stern CD. The chick; a great model system becomes even greater. *Dev Cell*. 2005;8(1):9–17.
- Vilches-Moure JG. Embryonic chicken (*Gallus gallus domesticus*) as a model of Cardiac Biology and Development. *Comp Med*. 2019;69(3):184–203.
- Fonseca BB, da Silva MV, de Moraes Ribeiro LN. The chicken embryo as an in vivo experimental model for drug testing: advantages and limitations. *Lab Anim (NY)*. 2021;50(6):138–9.
- Rodrigues T, Brodier L, Matter JM. Investigating neurogenesis in birds. *Methods Mol Biol*. 2020;2092:1–18.

40. Simicevic J, Deplancke B. Transcription factor proteomics-Tools, applications, and challenges. *Proteomics*. 2017;17:3–4.
41. Zhao MH, Liang S, Kim NH, Cui XS. MLL2 is essential for porcine embryo development in vitro. *In Vitro Cell Dev Biol Anim*. 2016;52(6):699–704.
42. Hisey E, Ross PJ, Meyers SA. A review of OCT4 functions and applications to equine embryos. *J Equine Vet Sci*. 2021;98:103364.
43. Sidrat T, Khan AA, Idrees M, Joo MD, Xu L, Lee KL, et al. Role of wnt signaling during in-vitro bovine blastocyst development and maturation in synergism with PPARdelta signaling. *Cells*. 2020;9(4):923.
44. Silva PGC, Moura MT, Silva RLO, Nascimento S, Silva JB, Ferreira-Silva JC, et al. Temporal expression of pluripotency-associated transcription factors in sheep and cattle preimplantation embryos. *Zygote*. 2018;26(4):270–8.
45. Pourquie O. The chick embryo: a leading model in somitogenesis studies. *Mech Dev*. 2004;121(9):1069–79.
46. Pourquie O. Somite formation in the chicken embryo. *Int J Dev Biol*. 2018;62(1–2–3):57–62.
47. Oliver B, Misteli T. A non-random walk through the genome. *Genome Biol*. 2005;6(4):214.
48. Chua EHZ, Yasar S, Harmston N. The importance of considering regulatory domains in genome-wide analyses - the nearest gene is often wrong! *Biol Open*. 2022;11(4):bio059091.
49. Al-Naama N, Mackeh R, Kino T. C2H2-Type zinc finger proteins in Brain Development, Neurodevelopmental, and other Neuropsychiatric Disorders: systematic literature-based analysis. *Front Neurol*. 2020;11:32.
50. Razin SV, Borunova VV, Maksimenko OG, Kantidze OL. Cys2His2 zinc finger protein family: classification, functions, and major members. *Biochem (Mosc)*. 2012;77(3):217–26.
51. Mackeh R, Marr AK, Fadda A, Kino T. C2H2-Type zinc finger proteins: evolutionarily old and New Partners of the Nuclear hormone receptors. *Nucl Recept Signal*. 2018;15:1550762918801071.
52. Burglin TR. Homeodomain subtypes and functional diversity. *Subcell Biochem*. 2011;52:95–122.
53. Imayoshi I, Kageyama R. bHLH factors in self-renewal, multipotency, and fate choice of neural progenitor cells. *Neuron*. 2014;82(1):9–23.
54. Kageyama R, Shimojo H, Ohtsuka T. Dynamic control of neural stem cells by bHLH factors. *Neurosci Res*. 2019;138:12–8.
55. Dennis DJ, Han S, Schuurmans C. bHLH transcription factors in neural development, disease, and reprogramming. *Brain Res*. 2019;1705:48–65.
56. Bora P, Gahurova L, Masek T, Hauserova A, Potesil D, Jansova D, et al. p38-MAPK-mediated translation regulation during early blastocyst development is required for primitive endoderm differentiation in mice. *Commun Biol*. 2021;4(1):788.
57. Grocott T, Johnson S, Bailey AP, Streit A. Neural crest cells organize the eye via TGF-beta and canonical wnt signalling. *Nat Commun*. 2011;2:265.
58. Ma S, Meng Z, Chen R, Guan KL. The Hippo Pathway: Biology and Pathophysiology. *Annu Rev Biochem*. 2019;88:577–604.
59. Wu Z, Guan KL. Hippo Signaling in Embryogenesis and Development. *Trends Biochem Sci*. 2021;46(1):51–63.
60. Xie H, Tranguch S, Jia X, Zhang H, Das SK, Dey SK, et al. Inactivation of nuclear wnt-beta-catenin signaling limits blastocyst competency for implantation. *Development*. 2008;135(4):717–27.
61. Peter IS. The function of architecture and logic in developmental gene regulatory networks. *Curr Top Dev Biol*. 2020;139:267–95.
62. Ayers KL, Lambeth LS, Davidson NM, Sinclair AH, Oshlack A, Smith CA. Identification of candidate gonadal sex differentiation genes in the chicken embryo using RNA-seq. *BMC Genomics*. 2015;16:704.
63. Levine M, Davidson EH. Gene regulatory networks for development. *Proc Natl Acad Sci U S A*. 2005;102(14):4936–42.
64. Menchero S, Sainz de Aja J, Manzaneres M. Our first choice: Cellular and genetic underpinnings of Trophoctoderm Identity and differentiation in the mammalian embryo. *Curr Top Dev Biol*. 2018;128:59–80.
65. Chawengsaksophak K. Cdx2 animal models reveal developmental origins of cancers. *Genes (Basel)*. 2019;10(11):928.
66. Kumar N, Tsai YH, Chen L, Zhou A, Banerjee KK, Saxena M, et al. The lineage-specific transcription factor CDX2 navigates dynamic chromatin to control distinct stages of intestine development. *Development*. 2019;146(5):dev172189.
67. Sun X, Yang Q, Rogers CJ, Du M, Zhu MJ. AMPK improves gut epithelial differentiation and barrier function via regulating Cdx2 expression. *Cell Death Differ*. 2017;24(5):819–31.
68. Sock E, Wegner M. Using the lineage determinants Olig2 and Sox10 to explore transcriptional regulation of oligodendrocyte development. *Dev Neurobiol*. 2021;81(7):892–901.
69. Motohashi T, Kawamura N, Watanabe N, Kitagawa D, Goshima N, Kunisada T. Sox10 functions as an inducer of the Direct Conversion of Keratinocytes into neural crest cells. *Stem Cells Dev*. 2020;29(23):1510–9.
70. Sneha P, Thirumal Kumar D, Lijo J, Megha M, Siva R. George Priya Doss C. probing the protein-protein Interaction Network of Proteins causing Maturity Onset Diabetes of the Young. *Adv Protein Chem Struct Biol*. 2018;110:167–202.
71. Pepe GJ, Albrecht ED. Fetal endocrinology/hormones. *Encyclopedia of reproduction*. 2018. p. 406–14.
72. Bai C, Hou L, Zhang M, Wang L, Guan W, Ma Y. Identification and biological characterization of chicken embryonic cardiac progenitor cells. *Cell Prolif*. 2013;46(2):232–42.
73. Eng G, Lee BW, Radisic M, Vunjak-Novakovic G. Cardiac tissue engineering. *Principles of tissue engineering*. 2014. p. 771–92.
74. Trejo-Revels V, McTeir L, Summers K, Rainger J. An analysis of anterior segment development in the chicken eye. *Mech Dev*. 2018;150:42–9.
75. Ravi V, Bhatia S, Shingate P, Tay BH, Venkatesh B, Kleinjan DA. Lampreys, the jawless vertebrates, contain three Pax6 genes with distinct expression in eye, brain and pancreas. *Sci Rep*. 2019;9(1):19559.
76. Grocott T, Lozano-Velasco E, Mok GF, Munsterberg AE. The Pax6 master control gene initiates spontaneous retinal development via a self-organising Turing network. *Development*. 2020;147(24):dev185827.
77. Jakob S, Lovell-Badge R. Sex determination and the control of Sox9 expression in mammals. *FEBS J*. 2011;278(7):1002–9.
78. Saito D, Tamura K, Takahashi Y. Early segregation of the adrenal cortex and gonad in chicken embryos. *Dev Growth Differ*. 2017;59(7):593–602.
79. Sun L, Guo L, Wang J, Li M, Appiah MO, Liu H, et al. Photoperiodic effect on the testicular transcriptome in broiler roosters. *J Anim Physiol Anim Nutr (Berl)*. 2020;104(3):918–27.
80. Estermann MA, Williams S, Hirst CE, Roly ZY, Serralbo O, Adhikari D, et al. Insights into gonadal sex differentiation provided by single-cell transcriptomics in the Chicken embryo. *Cell Rep*. 2020;31(1):107491.
81. Mao AP, Ishizuka IE, Kasal DN, Mandal M, Bendelac A. A shared Runx1-bound Zbtb16 enhancer directs innate and innate-like lymphoid lineage development. *Nat Commun*. 2017;8(1):863.
82. Cheng ZY, He TT, Gao XM, Zhao Y, Wang J. ZBTB transcription factors: key regulators of the development, differentiation and effector function of T cells. *Front Immunol*. 2021;12:713294.
83. El-Hodiri HM, Campbell WA, Kelly LE, Hawthorn EC, Schwartz M, Jalligam-pala A, et al. Nuclear factor I in neurons, glia and during the formation of Muller glia-derived progenitor cells in avian, porcine and primate retinas. *J Comp Neurol*. 2022;530(8):1213–30.
84. Schick E, McCaffery SD, Keblish EE, Thakuridin C, Emerson MM. Lineage tracing analysis of cone photoreceptor associated cis-regulatory elements in the developing chicken retina. *Sci Rep*. 2019;9(1):9358.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

