Identifying Molecular Pathways and Candidate Genes Associated With Knob Traits by Transcriptome Analysis in the Goose (Anser Cygnoides)

Wangyang Ji
Key Laboratory of Animal Genetics, Breeding and Molecular Design of Jiangsu Province, Yangzhou University, Yangzhou,225009, China

Qi Xu
Key Laboratory of Animal Genetics, Breeding and Molecular Design of Jiangsu Province, Yangzhou University, Yangzhou,225009, China

Wenming Zhao (wmzhao@yzu.edu.cn)
Key Laboratory of Animal Genetics, Breeding and Molecular Design of Jiangsu Province, Yangzhou University, Yangzhou,225009, China

li E Hou
Key Laboratory of Animal Genetics, Breeding and Molecular Design of Jiangsu Province, Yangzhou University, Yangzhou,225009, China

Xiaoya Yuan
Key Laboratory of Animal Genetics, Breeding and Molecular Design of Jiangsu Province, Yangzhou University, Yangzhou,225009, China

Tiantian Gu
Key Laboratory of Animal Genetics, Breeding and Molecular Design of Jiangsu Province, Yangzhou University, Yangzhou,225009, China

ZhuoYu Chen
Key Laboratory of Animal Genetics, Breeding and Molecular Design of Jiangsu Province, Yangzhou University, Yangzhou,225009, China

Yu Zhang
Key Laboratory of Animal Genetics, Breeding and Molecular Design of Jiangsu Province, Yangzhou University, Yangzhou,225009, China

Yang Zhang
Key Laboratory of Animal Genetics, Breeding and Molecular Design of Jiangsu Province, Yangzhou University, Yangzhou,225009, China

Guohong Chen
Key Laboratory of Animal Genetics, Breeding and Molecular Design of Jiangsu Province, Yangzhou University, Yangzhou,225009, China

Research Article

Keywords: Anser cygnoides, Yangzhou geese, KEGG pathway

Posted Date: March 6th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-273735/v1

Abstract

Anser cygnoides has a spherical crest on the beak roof, which is described as knob. However, the mechanisms affecting knob morphology are unclear. Here, we investigated the phenotypic characteristics and molecular basis of knob-size differences in Yangzhou geese. The bony crest of konb was found in the frontal area of the skull, rather than hump of upper beak. Although the knob length, width, and height varied greatly in geese with different knob phenotypes, growth of the bony crest was mainly reflected in the length and height, but not the width. Histological analysis showed that knob skin in large-size knobs geese have a greater length in the stratum corneum, stratum spinosum, and stratum reticular. Transcriptome profiling revealed 415 differentially expressed genes involved in knob growth and development. In addition, GO enrichment and KEGG pathway analysis revealed 455 significant GO terms and 210 enriched KEGG pathways. We focused on the TGF-β-signaling and thyroid hormone synthesis-signaling KEGG pathways. Geese with larger knobs had increased *ADCY3*, *TSHR*, *DCN*, and *BMP5* mRNA-expression levels, suggesting that both pathways (and the associated genes) mediate knob growth and development. Our data provide comprehensive molecular determinants of knob size, which can potentially be used to promote the genetic improvement of goose knobs to meet consumer preferences.

Introduction

Many birds have a distinctive bony crest on their heads. Most birds have a bony crest on their cranium. Some birds show helmet-like casques on the dorsal surface of the neurocranium (e.g. *Casuarius spp*), and some birds display inflated bulges in the frontal area, immediately caudal to the nasofrontal hinge (e.g., *Balearica*). Only a few birds present protuberances of the upper beak (e.g., *Pauxi unicornis*)[1]. Both Chinese and African geese (*Anser cygnoides*) also possess a spherical crest across the beak roof, which is described as a knob. However, whether the knob is a hump of the upper beak or a frontal crest requires further investigation.

Although all domesticated Chinese and African geese have a characteristic knob-like crest across the base of the bill near the forehead, a small crest is found in wild species, such as in swan geese. The knob has been greatly exaggerated by artificial selection in Chinese and African goose breeds [2]. The knob size is relatively larger in males than in females and in adults than in juveniles. The knob trait is inherited in an incompletely dominant manner. Hence, knob sizes vary among different breeds and within different breeds. Shitou geese, a large goose bred in China, have an average knob size of 40mm, with a maximum of 55 mm and a minimum of approximately 30 mm. In contrast, the average knob size of Magang geese is approximately 28 mm. However, the genetic basis underlying differences in knob sizes is not fully understood.

The knob, as an ornamental trait, is well developed by the time of sexual maturity and provides an identifier of sexual maturity. Moreover, the knobs influence first impressions of customers in China when making purchase decisions, with a large knob phenotypic size generally being preferred. However, the morphological structures of knobs and the mechanisms underlying phenotypic variation remain unclear.

In this study, adult Yangzhou geese with large or small knobs were selected, and knob morphologies and histologies were observed. Furthermore, the genetic basis of knob-size differences was investigated by RNA sequencing. The results provide an alternative strategy for the genetic improvement of goose knobs to meet consumer preferences.

Materials And Methods

- **Ethics Statement**

Our study was carried out in compliance with the ARRIVE (Animal Research: Reporting of In Vivo Experiments) guidelines. All animal experiments were were approved by with the Institutional Animal Care and Use Committee of Yangzhou University (approval number: 151-2014). Procedures were strictly performed in accordance with the Regulations for the Administration of Affairs Concerning Experimental Animals (Yangzhou University, China, 2012) and the Standards for the Administration of Experimental Practices (Jiangsu, China, 2008). We also confirm that we have done all efforts to minimize the suffering of animals.

- **Animal Samples Collection**

Approximately 500 380-day-old, healthy male Yangzhou geese were raised at Yangzhou Tiange Goose Industry Co., Ltd., their knob sizes were measured, and the three geese with largest (be set as group L) or smallest knob sizes (be set as group s) were selected for experiments. The experimental animals were humanely euthanized, and tissues were obtained around the center line of the goose knob. A mixture of skin and bony crest was immediately placed in liquid nitrogen and stored in a −80°C freezer for subsequent RNA isolation. A portion of the skin tissues were taken for paraffin sections, and the remaining parts were used for skull taxidermy.

- **Histological Observation**

The skin samples were fixed in 4% paraformaldehyde. After 24 h, the samples were placed in an embedding cassette and rinsed with running water (to remove the fixative from the tissue) for 30 min, and the samples were dehydrated in a graded ethanol series. A JB-P5 tissue-embedding machine (Wuhan Junjie Electronics Co., Ltd., Wuhan, China) was used for paraffin embedding at 70°C. The paraffin blocks were cut (RM2016, Germany) along the horizontal axis into 3 μm-thick sections and stained with hematoxylin and eosin (HE) according to standard protocols. The knob skin was examined under an upright optical microscope (Nikon, Tokyo, Japan), and image acquisition and analysis were performed with a DS-U3 Imaging system (Nikon, Tokyo, Japan).

- **Skull Specimen Making**

All skulls were put in a moist and warm plastic bag to accelerate decomposition and skeletonization. Each individual skull was classified with a label, the skulls were collected after 4 months, and then they

were cleaned with running water and ethyl alcohol. A toothbrush was used to remove all the rotten muscle residues. Finally, the skulls were dried naturally.

RNA Sequencing (RNA-seq) and Bioinformatics Analysis

- **RNA Extraction**

Total RNA was extracted from a mixture of skin and bone crest of geese knob in the large (L) and small (S) groups. Using the TRIzol® Reagent (Animal RNA Purification Reagent for animal tissues; Invitrogen) following the manufacturer's recommendations., and genomic DNA was removed using DNase I (TaKaRa). Total RNA purity,concentration and integrity of each samples were estimated by using a Nanodrop 2000 instrument (Thermo Scientific, Wilmington, DE, USA) and an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA), and only high-quality RNA samples were used to construct the sequencing library[20].

- **Library preparation for sequencing**

An RNA sequencing library was prepared using the TruSeq™ RNA Sample Preparation Kit from Illumina (San Diego, CA) following the manufacturer's recommendations. In brief, mRNA was purified via Poly (A) selection with oligo(dT) cellulose. and then fragmented in fragmentation buffer. Continually, double-stranded cDNA was synthesized using a Super Script Double-Stranded cDNA Synthesis Kit (Invitrogen, CA) with random hexamer primers (Illumina). The synthesized cDNA was subjected to end-repaired, phosphorylation and the A-tailed according to library-construction protocol of Illumina. Libraries were size-selected for cDNA target fragments of 300 base pairs (bp) on a 2% low-range ultra agarose gel, followed by PCR amplification. Finally, the amplified fragments were sequenced with an Illumina HiSeq Xten/NovaSeq 6000 sequencer according to the manufacturer's instructions[21].

- **Read Mapping**

The raw paired-end reads were trimmed and quality-controlled using the default parameters of the SeqPrep (https://github.com/jstjohn/SeqPrep) and sickle (https://github.com/najoshi/sickle) software programs. Clean reads were then separately aligned to the reference genome in orientation mode, using HISAT2 software (http://ccb.jhu.edu/software/hisat2/index.shtml) [22]. The mapped reads of each sample were assembled using a reference-based approach with StringTie software (https://ccb.jhu.edu/software/stringtie/index.shtml?t=example) [23].

- **Differential Expression Analysis and Functional Enrichment**

To identify differentially expressed genes (DEGs) between two different samples, the expression levels of each transcript were calculated using the transcripts-per-million reads (TPM) method. RSEM software (http://deweylab.biostat.wisc.edu/rsem/) [24] was used to quantify gene abundances. Essentially, differential-expression analysis was performed using the DESeq2 program, [25] where an adjusted P-value

(P-adjust) of ≤ 0.05, and DEGs with absolute fold-changes (FCs) of >2, a P-adjust value of ≤ 0.05 (DESeq2), and a P-adjust value of ≤ 0.001 (DESeq2) were considered DEGs with statistical significance. In addition, functional-enrichment analysis was performed with the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases to determine which DEGs were significantly enriched for GO terms and metabolic pathways at a Bonferroni-corrected P-value of ≤0.05, when compared with the whole-transcriptome background. GO functional enrichment and KEGG-pathway analysis were performed using the Goatools (https://github.com/tanghaibao/Goatools) and KOBAS (http://kobas.cbi.pku.edu.cn/home.do) software programs [26].

- **RNA-seq Data Validation by Real-time Quantitative PCR (RT-qPCR) analysis** Thirteen DEGs were randomly selected to verify the RNA-seq data by RT-qPCR, including adenylate cyclase 3 (*ADCY3*), angiotensinogen, serpin peptidase inhibitor, clade A, member 8 (*AGT*); bone morphogenetic protein 5 (*BMP5*); decorin (*DCN*); deiodinase, iodothyronine, type III (*DIO3*), fibrillin 1 (*FBN1*), insulin-like growth factor 1 (*IGF1*), lumican (*LUM*), natriuretic peptide C (*NPPC*); osteoglycin (*OGN*); secreted protein, acidic, cysteine-rich (*SPARC*); steroid-5-alpha-reductase, alpha polypeptide 2 (*SRD5*α2); and thyroid stimulating hormone (*TSHR*). Total RNA extracted from each bare, a mixture of skin and bone crest of geese knob was subjected to RT-qPCR analysis. Single-strand cDNA was synthesized using approximately 5 mg of total RNA and a Revertaid^{TM} First Strand cDNA Synthesis Kit (Fermentas, Fermentas China Co., Ltd., China), and the resulting cDNA was diluted five-fold. RT-qPCR analysis was performed using SYBR Green Real-time PCR Master Mix (TOYOBO, Osaka, Japan) and an ABI 7500 Real-Time PCR System (Applied Biosystems, Foster City, CA). Each 10 μL reaction contained 5 μL SYBR Green Real time PCR Master Mix, 0.4 μL forward primer (10 μM), 0.4 μL reverse primer (10 μM), 2 μL cDNA, and 2.2 μL distilled water. The RT-qPCR program was 50°C for 2 min; 95°C for 2 min; followed by 40 cycles of 95°C for 15 s and 40 cycles of 60°C for 15 s; 95°C for 15 s; 60°C for 1 min; 95°C for 15 s. Relative mRNA-expression values were calculated using the $2^{-\Delta\Delta}$Ct method with β-actin serving as an internal reference gene. The sequences of the primers used to amplify each of the 13 mRNAs are shown in Table S1.

Results

Phenotype, Histological, and Bone Structure Observations.

Morphological, histological, and anatomical structures were observed to determine the histological and bone structures of goose knobs. Morphological analysis showed that the knob accentuated the facial contours of the Yangzhou geese (Fig. 1A, 1B). The average knob length, width, and height were approximately 31.5 mm, 36.5 mm, 32.4 mm at 380 days of age, respectively, and these parameters varied significantly ($P < 0.05$). The maximum length, width, and height were approximately 36.4 mm, 40.3 mm, and 37.2 mm, and the corresponding minimum values were close to 26.6 mm, 32.8 mm, and 27.6 mm, respectively (Fig. 1D). Specific measurements standards are shown in Fig. 1C, and these standards were applied when measuring the bony crest.

Anatomically, the knob was found to include the skin derivative and bony crest. Although the knob was located at the base of the beak, it was joined to a typically frontal crest, rather than on the upper beak (Fig. 2A–2C). Unlike the whole knob, the bony crest varied greatly in terms of the length and height ($P < 0.05$), but not the width (Fig. 2D). Furthermore, increased crest sizes did not correlate with the hardness of the peripheral bones. Instead, the outer layer of the bone became thinner as the hump became larger, in a manner reminiscent of blowing a balloon. In addition, bony pores were often observed in large bony crests.

Based on the observations and measurements above, we analyzed the correlation between the knob size and bony crest size. The knob length, width, and height each significantly correlated with the same parameters in the bony crest (Table 1).

Histological measurements were performed to determine the thickness of five layers in the skin, including the stratum corneum (Sc), stratum spinosum (Ss), stratum epidermis (El), stratum reticular (Sr), and stratum corium (Sc). The stratum corneum is composed of flat, dead cells (corneocytes) embedded in a matrix of lipids[3], which were observed as highly eosinophilic cells with few nuclei following HE staining. Denser structures and significant enrichment for eosinophilic substances were observed in skin tissues from large knobs versus those from small knobs. Greater thicknesses of stratum corneum, stratum spinosum, and stratum reticular were observed in geese with large knobs than in those with small knobs ($P < 0.05$) (Fig. 3A–3C).

RNA Library Construction and Sequencing

In this study, an average of 50,136,474 raw reads was obtained from the L and S samples, and the average number of clean reads was 49,500,375. All downstream analyses were based on high-quality, clean sequence data. The error rates were all less than 0.025%. Approximately 92.01%–96.37% of the clean reads in the libraries were mapped to the Anser cygnoides reference genome and the percentage of phred quality scores of >20 (Q20) was >98% in all samples. The sequence read statistics are summarized in Table S3.

DEG Screening in Geese With Different Knob Sizes

DEGs were screened in large and small goose knobs. Scatter and volcano plots showed variations in mRNA-expression levels between geese with small (S group) and large (L group) knobs (Fig. 4A, 4B). We identified 415 differentially expressed mRNAs (Fig. 4C and Table S4; FC ≥ 2), including 357 upregulated genes and 58 downregulated genes. Most upregulated genes were related to hormone secretion, skeletal system development, and tissue composition, including adrenoceptor alpha 1A (*ADRA1A*); *AGT*; bromodomain, testis-specific (*BRDT*); prostaglandin F2 receptor inhibitor (*PTGFRN*); relaxin/insulin-like family peptide receptor 3 (*RXFP3*); *SRD5a2*; *TSHR*; *IGF1*; insulin-like growth factor 2 (*IGF2*), insulin-like growth factor binding protein 4 (*IGFBP4*); insulin-like growth factor binding protein 5 (*IGFBP5*); bone-morphogenetic protein/retinoic acid inducible neural-specific 3 (*BRINP3*); calcium channel, voltage-dependent subunit alpha1 H (*CACNA1H*); calneuron 1 (*CALN1*); collagen-related genes (*COL12A1*,

COL15A1, *COL1A2*, *COL20A1*, *COL21A1*, *COL26A1*, *COL3A1*, *COL5A2*, *COL6A1*, and *COL6A2*); *DIO3*; fibulin 1 *(FBLN1*); fibulin 2 (FBLN2); fibulin 7 (*FBLN7*); *FBN1*; fibrillin 2 (*FBN2*); fibroblast growth factor 7 (*FGF7*); fibroblast growth factor 9 (*FGF9*); *NPPC*; natriuretic peptide receptor 1 (*NPR1*); and *OGN*. The downregulated DEGs accounted for a small proportion of the total. These genes are known to be mainly related to enzymes, such as aldehyde dehydrogenase 3 family (*ALDH3B1*); mitogen-activated protein kinase 15 (MAP3K15); monooxygenase, DBH-like 1 (*MOXD1*); and protein kinase, X-linked (*PRKX*).

GO Enrichment and KEGG Pathway Analyses for DEGs

To further elucidate the functional roles of DEGs in determining the knob size, we performed GO and KEGG pathway-enrichment analysis for the DEGs using the online programs, Goatools and KOBAS. The DEGs were categorized into three main GO categories, namely biological process, cellular component, and molecular function. We identified 455 significantly enriched GO terms (Table S4) in 94 molecular functions, 52 cellular components, and 309 biological processes (P-value ≤ 0.05). The cellular component category mainly relates to extracellular components and chemical composition. Among the hormone-related GO terms, 13 were associated with insulin, and four were thyroid hormone-related. *ACTA2*, *ACTA1*, and *ACTG2* are enriched in pathways related to tissue migration and mesenchymal migration. Some DEGs were associated with calcium ion binding, such as *SULF1*, *SPARC*, *FBN1*, and *FBN2*.

In addition, the 10 genes with the greatest significant differences (|log2 FC|) for each GO term were plotted (Fig. 5). GO-term analysis showed that significantly upregulated DEGs in tissues from large knobs were highly related to certain categories, including extra-cellular structure organization, extra-cellular matrix organization, tissue migration, collagen fibril organization, mesenchyme migration, and skeletal system development. Furthermore, we mapped the DEGs in the KEGG pathway database and classified all pathways into five categories.

We mapped the 415 DEGs to 210 KEGG pathways, and the 20 most enriched pathways are shown in the bubble chart in Fig. 6 The associated KEGG pathway terms include protein digestion and absorption, PI3K-Akt signaling pathway, and ECM-receptor-interaction. We also identified many signaling hormone-related pathway terms, such as relaxin signaling pathway; thyroid hormone synthesis (including *CREB3L1*, *GPX7*, *ADCY3*, and *TSHR*); endocrine resistance (including *CDKN2C,IGF1*, and *ADCY3*); cortisol synthesis and secretion (including *CREB3L1*, *ADCY3*, and *CACNA1H*); insulin secretion (including *ADCY3* and *CREB3L1*); pancreatic secretion (including *ATP2B2* and *ADCY3*); insulin resistance (including *GFPT2* and *CREB3L1*); steroid hormone biosynthesis (including *SRD5a2*); dopaminergic synapse (including *GRIA3* and *CREB3L1*); parathyroid hormone synthesis, secretion and action (including *CREB3L1* and *ADCY3*); renin secretion (including *NPR1* and *AGT*); and TGF-β signaling pathway (including *DCN*, *FMOD*, *LTBP1*, *TGFB2*, and *BMP5*).

Protein–Protein Interaction (PPI) Analysis

A PPI network was generated using the STRING program in Cytoscape [4] software, based on the DEGs identified in this study (Fig. 7). *FBN1* and *DCN* were associated with the most genes; thus, they were in the center of the PPI network. The *SPARC*, *IGF1*, *LUM*, and *COL1A2* genes were the four genes with high degree.

Validation of DEGs in Geese With Different Knob Sizes

To validate the RNA-seq results, 13 relatively important DEGs (*ADCY3*, *AGT*, *BMP5*, *DCN*, *DIO3*, *FBN1*, *IGF1*, *LUM*, *NPPC*, *OGN*, *SPARC*, *SRD5a2*, and *TSHR1*) were selected for RT-qPCR analysis. All of these DEGs showed concordant expression patterns between the RNA-seq and RT-qPCR results (Fig. 8).

Discussion

In this study, goose knob morphologies and histologies were observed, and molecular pathways and candidate genes associated with knob phenotypes were identified. The knob included the skin derivative and bony crest. However, the knob cannot be strictly classified as a skin derivative, as the outer layer of the knob is composed of cornified skin. Phenotypically, the skin of a knob is much thicker than normal skin tissue. In addition, the knob is characterized by a cranial bony crest, whose size is highly correlated with that of the knob, suggesting that the bony crest hump determines the knob phenotype. However, not all Chinese geese have distinct bony crests, such as the Magang goose, whose knobs are soft and contain a mound of fleshy fat with tiny humps. In this study, we observed that the knobs had great variations in terms of the length, width, and height, although the width of the bony crest showed little variation. The length and height of the bony crest were well beyond the range of the frontal bone, but much of the width of the bony crest coincides with the width of the frontal bone. These results indicate that the different thickness of the skin in a knob affects the overall knob size and that the width of the bone crest is not related to the overall knob size.

Anatomically, the bony crest of the goose is located on the upper beak and is separated from the beak by the nasofrontal junction. Observations of the goose skull showed that the bony crest of the goose is formed by a combination of paired humps above the nasofrontal junction, which suggests that the knob may form from the inflated humps in the frontal area, rather than from the beak. Many birds in the Anseriformes order show the same pattern, such as *Melanitta* [1] and *Cygnus olor*.[5] These results suggest that the bony crest is also an important factor affecting the knob size.

Histologically, our results also confirmed that significant differences occurred in skin tissue between the large and small knobs. The knobs included the stratum corneum, stratum spinosum, and stratum reticular, and a greater length was observed in geese with large knobs than in those with small knobs. The thicker stratum corneum helps to protect the skin. The stratum reticular is rich in thick, elastic fiber bundles and collagen fiber bundles. The chicken comb, another skin derivative, is similar to the goose knob, which also contains abundant collagen. Data from previous studies have shown that the chemical components of combs have many industrial uses and great medicinal value [6,7]. The thickness of the

stratum reticular may directly affect the content of the chemical components in skin tissue, considering that large knobs have higher elastin and collagen contents. However, further research is needed to better define the main components of knobs.

Based on the above observations, we selected a mixture of skin and bone as the sample for further study. We conducted transcriptome profiling based on six Yangzhou geese with extreme goose knob sizes. Most DEGs upregulated in group L are involved in skeletal growth development, hormone synthesis, and secretion, such as *NPPC* and *SRD5a2*. *NPPC* is a C-type natriuretic peptide 3 that has been shown to be important in cartilage homeostasis and endochondral bone formation [8-10]. *SRD5a2* helps regulate testosterone and cortisol metabolism [11]. *SRD5a2* promotes the development of secondary sexual activity, and a lack of *SRD5a2* activity decreases dihydrotestosterone levels in male secondary sexual organs and the serum of vertebrates [12]. Hence, high *NPPC* and *SRD5a2* expression in the L group might promote bony crest growth and sexual maturity.

Six bone-related GO terms (skeletal system development, osteoblast differentiation, chondrocyte morphogenesis involved in endochondral bone morphogenesis, embryonic skeletal system development, bone mineralization involved in bone maturation, and bone trabecula formation) were significantly enriched, and the associated enriched genes included SULF1, VCAN, IGF1, FGF9, and *BMP5*, among others. The FGF and BMP proteins are known to play important roles in skeletal growth. Exogenous recombination of *BMP5* can induce the formation of ectopic bone tissue in normal soft tissue [13,14]. KEGG pathway analysis showed that BMP5 was also in the TGF-β signaling pathway. In this study, *BMP5* expression was downregulated in the goose knob tissues of group S, indicating that *BMP5* may delay the growth of the bony crest by blocking bone formation. KEGG pathway analysis also showed that these DEGs were associated with thyroid hormones. *DIO3* was found in the thyroid hormone signaling pathway, and *DIO3* is critical for the maturation and function of the thyroid axis [15,16]. Another thyroid-related pathway, i.e., thyroid hormone synthesis (including *TSHR*, *CREB3L1*, *GPX7*, and *ADCY3*), also appeared to be crucial in this study, as these genes showed significantly higher expression levels in the L group than in the S group.

Subsequently, a PPI network was constructed with the DEGs to identify relationships between genes. Through this analysis, we obtained a list of central genes, including *FBN1*, *DCN*, *SPARC*, *IGF1*, *LUM,* and some collagen-related genes. *FBN1* is an extracellular glycoprotein that is widely distributed in ocular connective tissues [17]. *DCN* and *LUM* are involved in regulating the diameter and growth of collagen fibrils. *DCN* is a small leucine-rich proteoglycan and a component of the extracellular matrix that is thought to be responsible for the structure, tissue organization, and surface properties of fibrils. In previous studies, *DCN* was isolated from chicken sternal cartilage. As biglycan has not been detected in birds, it is speculated that the *DCN* subtype might partially replace the functionality of biglycan. Although the core protein of biglycan is partially homologous to that of *DCN*, the core proteins of *DCN* have a stronger affinity for collagen and, thus, interact more with collagen [18]. Because *DCN* interacts strongly with collagen, we believe that the higher *DCN* expression in large knobs may explain the thicker stratum

reticular mentioned above. *SPARC* is expressed mainly in bones and adipocytes, and is a secreted extracellular matrix protein [19]. Interestingly, in the L group, all of these DEGs were upregulated and might function to promote thicker cornified skin in knob.

This preliminary study provides transcriptome analysis of skin tissues from geese with differing extreme knob sizes. Our results suggest that the candidate genes were associated with phenotypic differences between these two groups. GO and KEGG analyses revealed that the DEGs were involved in some hormone-related signaling pathways, and the PPI network showed that the DEGs were associated with the chemical components of knobs. Taken together, our findings suggest that the TGF-β signaling pathway and thyroid hormone synthesis signaling pathway (including *ADCY3*, *TSHR*, *DCN*, and *BMP5*) might participate in knob growth and development. Our data provide comprehensive molecular information regarding determinants of knob size, which could promote the genetic improvement of goose knobs to meet consumer preferences.

Declarations

Acknowledgments:

This research was supported by the earmarked fund for modern agro-industry technology research system (CARS-42-3), and the plant and animal breeding project of Jiangsu province (PZCZ201735),We appreciate the Yangzhou Tiange Goose Industry Co., Ltd. for providing the experimental animals.

Author Contributions:

The experiments were designed by W.Y.J., Q.X. and W.M.Z; The experimentswere performed by W.Y.J., L.E.H., T.T.G., and Z.Y.C.; The data were analyzed by X.Y.Y., Y.Z.(Yu Zhang), Y.Z.(Yang Zhang) and G.H.C.; The paper was wrote by W.Y.J., Q.X and W.M.Z.

Competing Interest:

The authors declare no competing interest.

References

1. Mayr, G. A survey of casques, frontal humps, and other extravagant bony cranial protuberances in birds. *Zoomorphology* **137**, 457–472 (2018).
2. Mead, D. *4 Domesticated geese and ducks—and allied species*, (Sulawesi Language Alliance, 2013).
3. Champagne, A.M., Pigg, V.A., Allen, H.C. & Williams, J.B. Presence and persistence of a highly ordered lipid phase state in the avian stratum corneum. *Front. Mol. Biosci.* **221**, jeb.176438 (2018).
4. Li, S. et al. Identification of an eight-gene prognostic signature for lung adenocarcinoma. *Cancer. Manag. Res.* **10**, 3383-3392 (2018).

5. Horrocks, N., Perrins, C. & Charmantier, J. Seasonal changes in male and female bill knob size in the mute swan *Cygnus olor*. *Avian. Biol.* **40**, 511-519 (2010).
6. Mukhtar, N. & Khan, S.H. Comb: An important reliable visual ornamental trait for selection in chickens. *Worlds Poultry Science Journal.* **68**, 425-434 (2012).
7. Nakano, T., Imai, S., Koga, T. & Sim, J.S. Light microscopic histochemical and immunohistochemical localisation of sulphated glycosaminoglycans in the rooster comb and wattle tissues. *J Anat.* **189**, 643-650 (1996).
8. Chusho, H. et al. Dwarfism and early death in mice lacking C-type natriuretic peptide. *Natl. Acad. Sci.* **98**, 4016-4021 (2001).
9. Wood, A.R.e.a. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.* **46**, 1173-1186 (2014).
10. Gudbjartsson, D.F.e.a. Many sequence variants affecting diversity of adult human height. *Nat. Genet.* **40**, 609-615 (2008).
11. Ramos, L., Vilchis, F., Chávez, B. & Mares, L.J. Mutational analysis of SRD5A2: From gene to functional kinetics in individuals with steroid 5α-reductase 2 deficiency. *Steroid. Biochem. Mol. Biol.* **200**, 105691 (2020).
12. Robitaille, J. & Langlois, V.S. Consequences of steroid-5α-reductase deficiency and inhibition in vertebrates. *Gen. Comp. Endocrinol.* **290**, 113400 (2020).
13. Pathi, S., Rutenberg, J.B., Johnson, R.L. & Vortkamp. Interaction of Ihh and BMP/Noggin signaling during cartilage differentiation. *A. Dev. Biol* **209**, 239-253 (1999).
14. Storm, E. & Kingsley, D.M. GDF5 coordinates bone and joint formation during digit development. *Dev. Biol.* **209**, 11-27 (1999).
15. Dunn, I.C. et al. Diurnal and photoperiodic changes in thyrotrophin-stimulating hormone β expression and associated regulation of deiodinase enzymes (DIO2, DIO3) in the female juvenile chicken hypothalamus. *J. Neuroendocrinol.* **29**(2017).
16. Renthlei, Z., Hmar, L. & A., K.T. High temperature attenuates testicular responses in tree sparrow (Passer montanus). *Gen. Comp. Endocrinol.* **301**, 113654 (2021).
17. C.M., d.l.C., A., P.-M., C., M.-A., E., M.-S. & J., P.-M. Immunohistochemical localization of fibrillin-1 protein in the cells of chick corneal and conjunctival epithelia during pre- and postnatal development. *Ophthalmic. Res.* **41**, 106-111 (2009).
18. Blaschke, U., Hedbom, E. & Bruckner, P. Distinct isoforms of chicken decorin contain either one or two dermatan sulfate chains. *J. Biol. Chem.* **271**, 30347-30353 (1996).
19. Hu, L. et al. SPARC promotes insulin secretion through down-regulation of RGS4 protein in pancreatic β cells. *Sci. Rep* **10**, 17581 (2020).
20. Liu, C. et al. De Novo Transcriptome Sequencing Analysis of Goose (Anser anser) Embryonic Skin and the Identification of Genes Related to Feather Follicle Morphogenesis at Three Stages of Development. *Int J Mol Sci* **19**(2018).

21. Li, B. et al. Identification of candidate genes associated with porcine meat color traits by genome-wide transcriptome analysis. *Sci. Rep* **6**, 35224 (2016).
22. Kim, D., Langmead, B. & Salzberg, S.L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357-360 (2015).
23. Pertea, M. et al. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290-295 (2015).
24. Dewey, C.N. & Li., B. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC. Bioinformatics.* **12**, 323-323 (2011).
25. Love, M., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome. Biol.* **15**, 550 (2014).
26. Wu, J., Mao, X., Cai, T., Luo, J. & Wei, L. KOBAS server: a web-based platform for automated annotation and pathway identification. *Nucleic. Acids. Res.* **34**, W720-724 (2006).

Table

Table 1. Correlation coefficients between knob and bony crest dimensions

Trait	Knob length	Knob width	Knob height
Bony crest length	0.884**	0.489	0.300
Bony crest width	0.593	0.862**	0.145
Bony crest height	0.818**	0.081	0.833**

Note: Significant differences are indicated with two asterisks ($P < 0.01$).

Figures

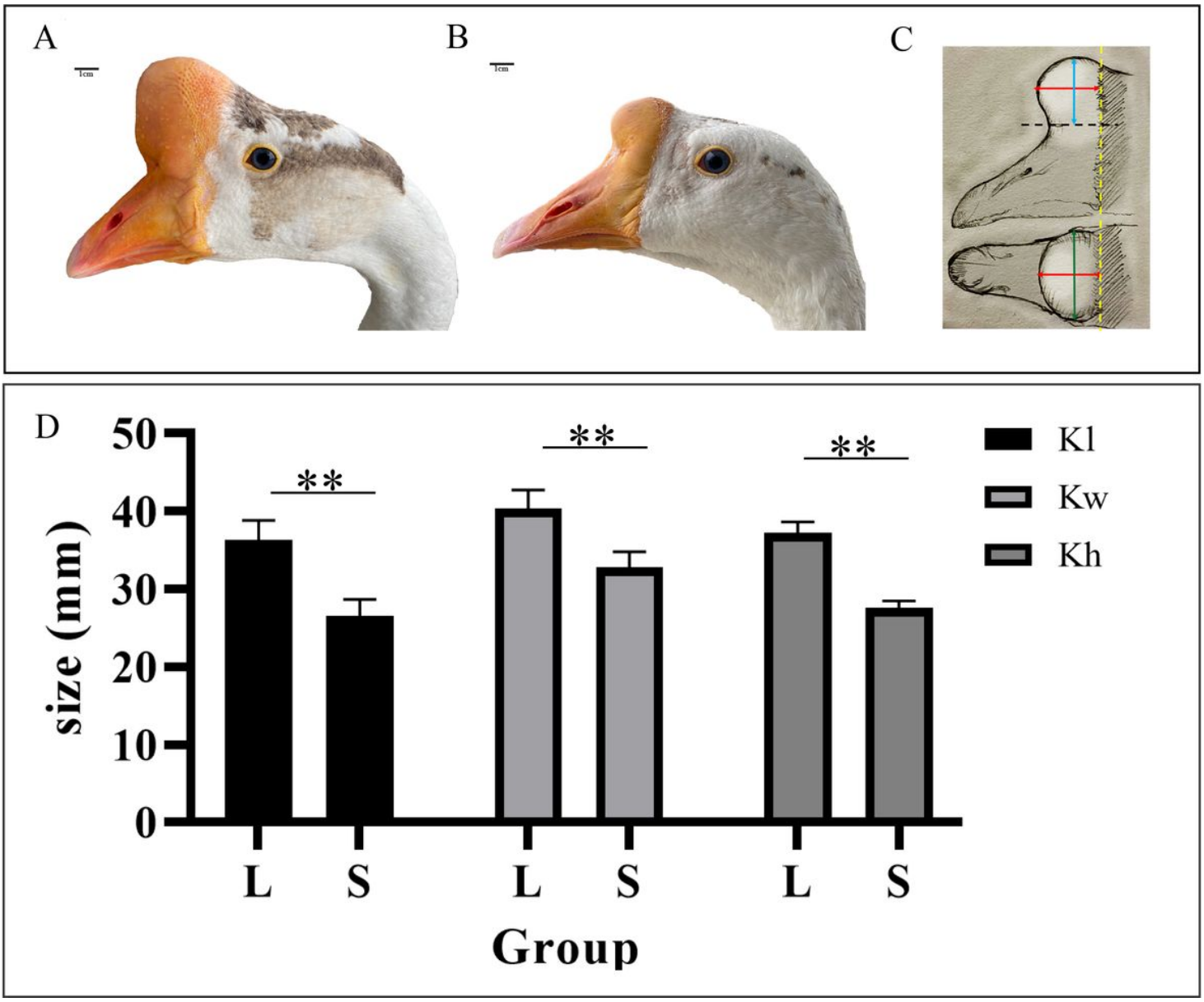

Figure 1

Phenotypes of knobs with different sizes in Yangzhou geese. (A, B) Clear differences in knob sizes were observed at 380 days of age between geese with large (L) knobs (A) and geese with small (S) knobs (B). (C) Schematic depiction of the method used to measure knob sizes. The black dotted line represents fronto-nasal junction. The feather cover area is shown to the right of the yellow dotted line. The red line (parallel to the fron-to-nasal junction) represents the length of the knob. The green line (seen from the top view) represents the width. The blue line (perpendicular to fronto-nasal junction) represents the height. (D) Comparison of knob sizes in the L and S groups. Significant differences are indicated with two asterisks ($P < 0.01$). Kl: knob length, Kw: knob width, Kh: knob height.

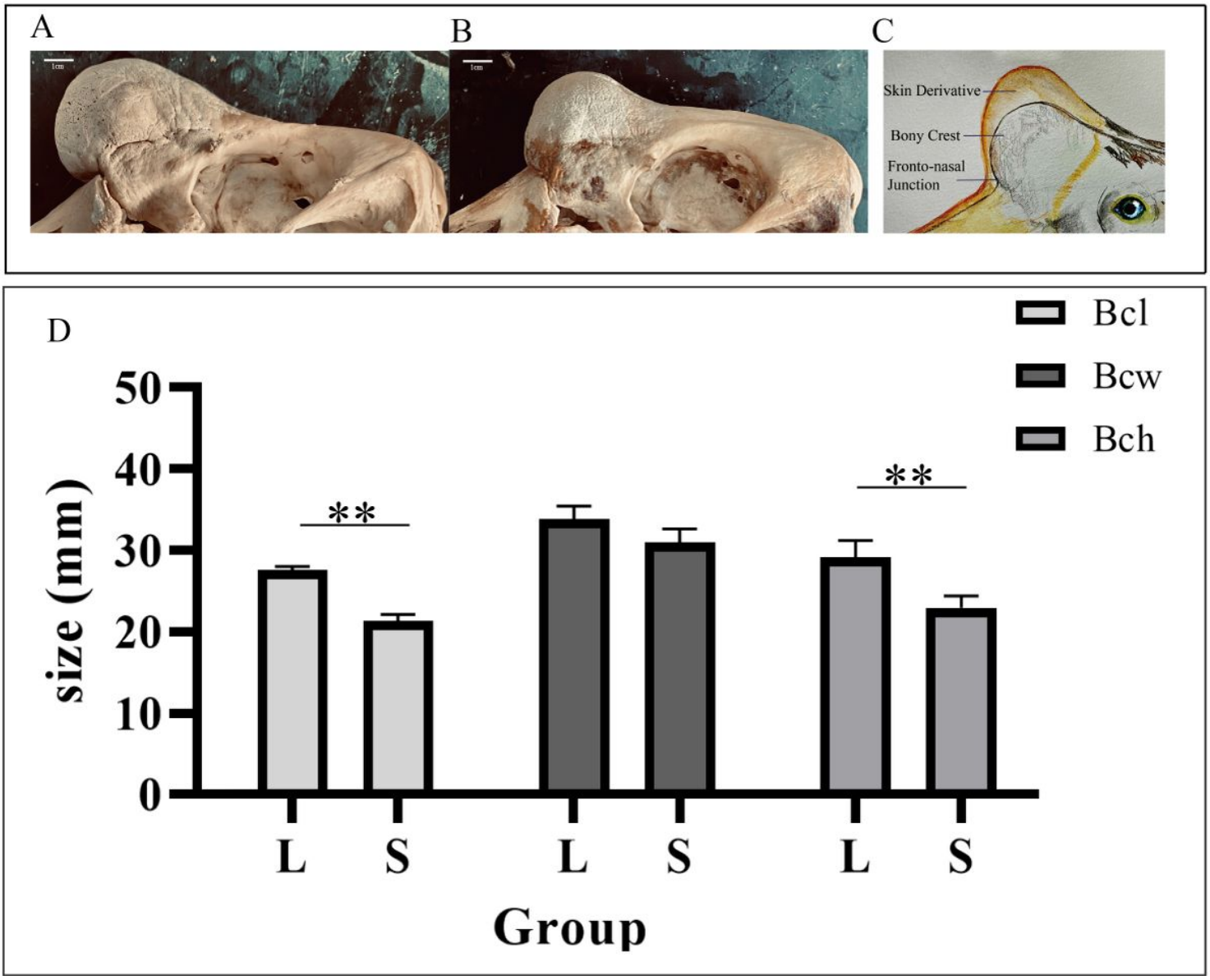

Figure 2

Bone crests of Yangzhou geese with different anatomical knob sizes. (A, B) Clear differences were observed in the bone crest size at 380 days of age between geese in the L (A) and S (B) groups. (C) Schematic representation of the anatomical knob structure. The black dotted line represents the fronto-nasal junction. The feather cover area is shown to the right of the yellow dotted line. The red line (parallel to the fronto-nasal junction) represents the length of the knob. The green line (seen from the top view) represents the width. The blue line (perpendicular to fronto-nasal junc-tion) represents the height. (D) Comparison of bone crests of geese in the L and S groups. Significant differences are indicated with two asterisks ($P < 0.01$). Bcl: bony crest length , Bcw: bony crest width, Bch: bony crest height.

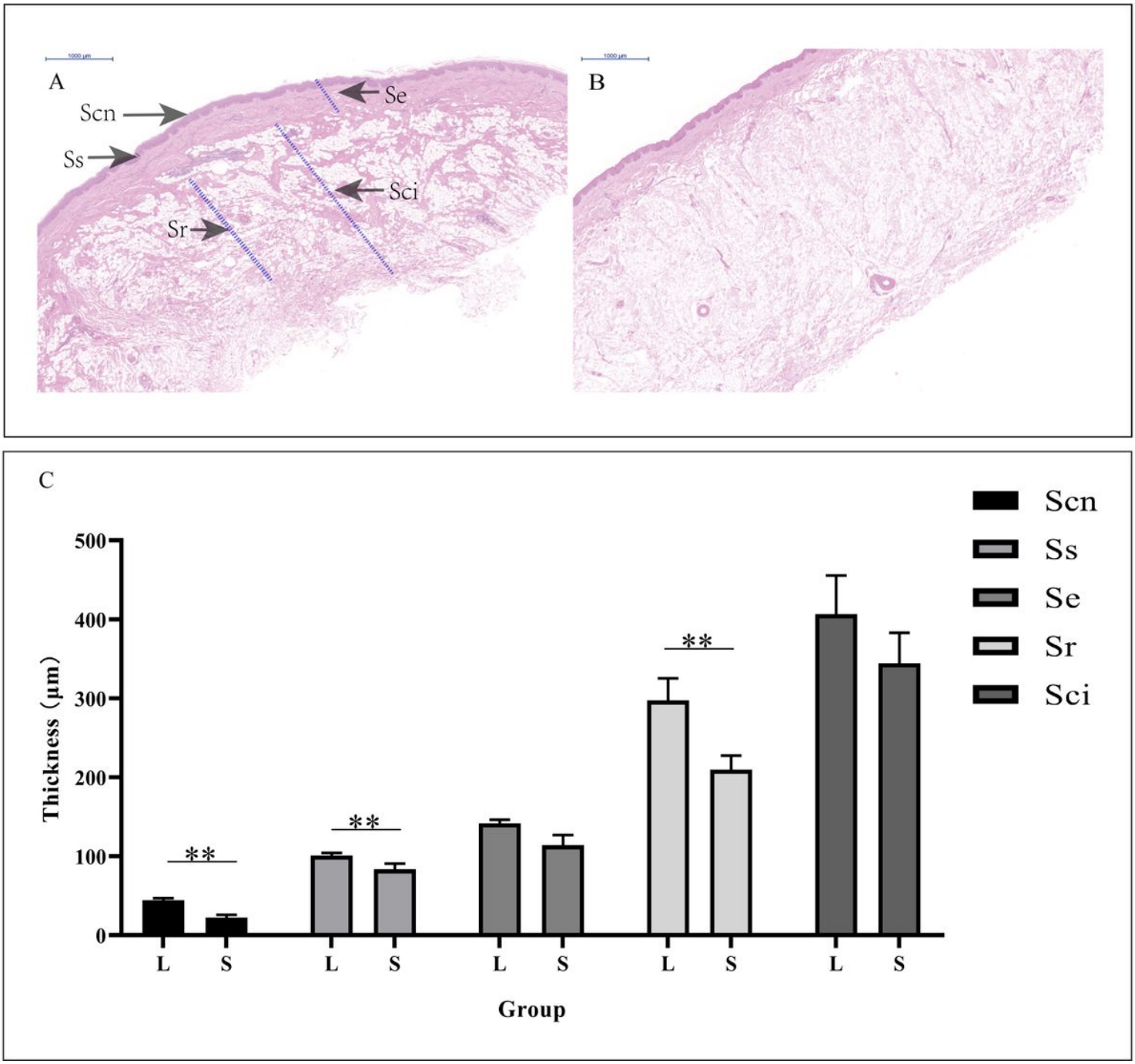

Figure 3

Histological characteristics of knobs with different sizes in Yangzhou geese. (A, B) Histological analysis of skin from geese with large knobs (A) and small knobs (B), based on HE staining. Sc: stratum corneum, Ss: stratum spinosum, Se: stratum epidermis, Sr: stratum reticular, Sc: stratum corium. (C) Comparing the thicknesses of five layers of skin from geese in the L and S groups. Significant differences are indicated with two aster-isks ($P < 0.01$).

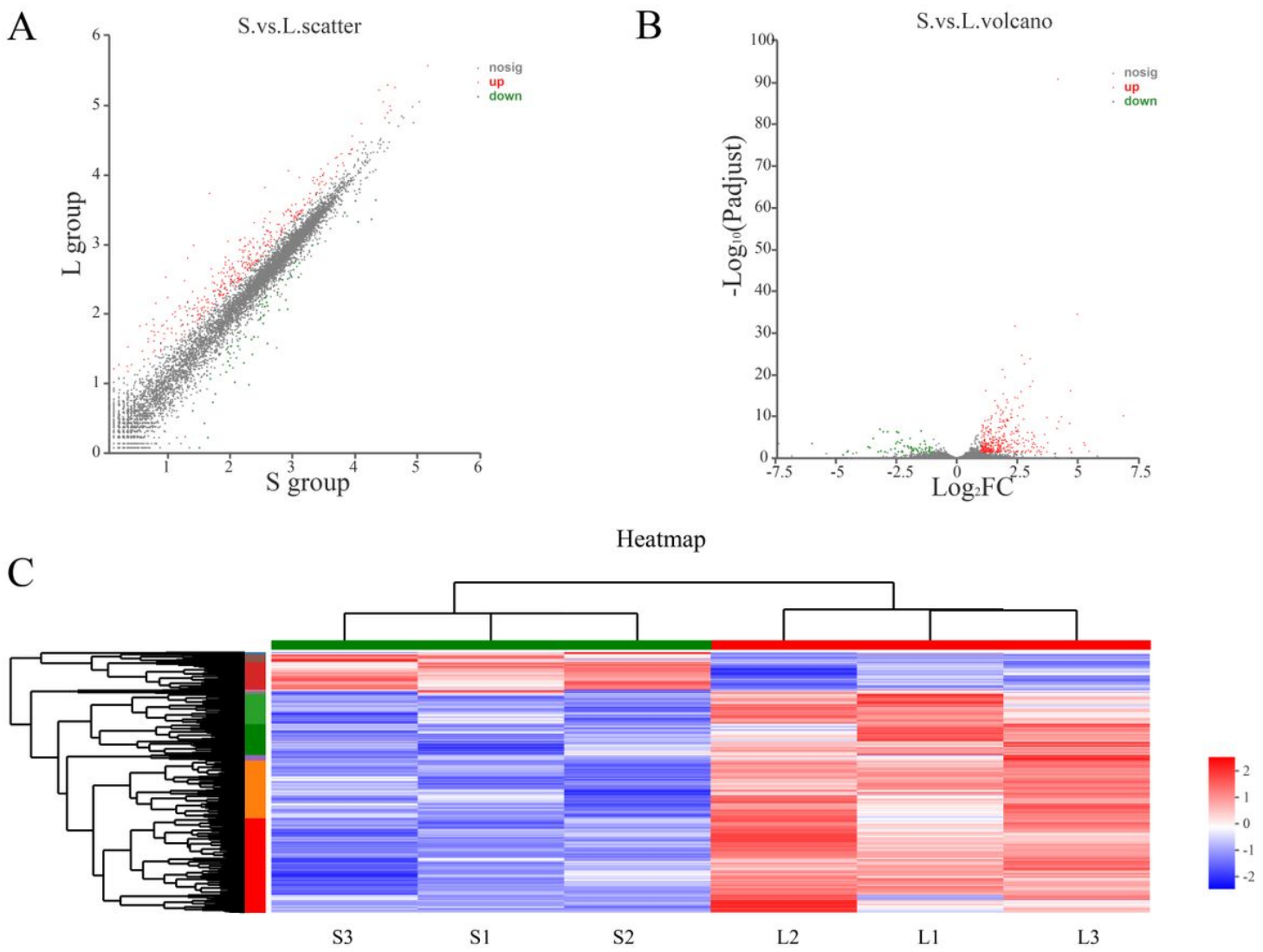

Figure 4

Scatter plot, volcano plot, and heat-map of differentially expressed genes (DEGs). (A) The scatter plot was used to assess variations in gene-expression levels between small and large knobs in geese. (B) Volcano plot showing DEGs. The X-axis represents the log2 of the expression ratio, and the Y-axis represents adjusted P (P-adjust) value. Each dot represents a specific gene. Red dots represent genes expressed at significantly higher levels in geese with large knobs, and green dots represent genes with significantly lower relative expression. Genes showing no significant differences in expression levels are represented with gray dots. (C) Hierarchical-clustering analysis for the transcriptome profiles of knobs in the S versus L groups. The heat-map presents the mean relative abundances of genes with a color scale. FC: fold-change.

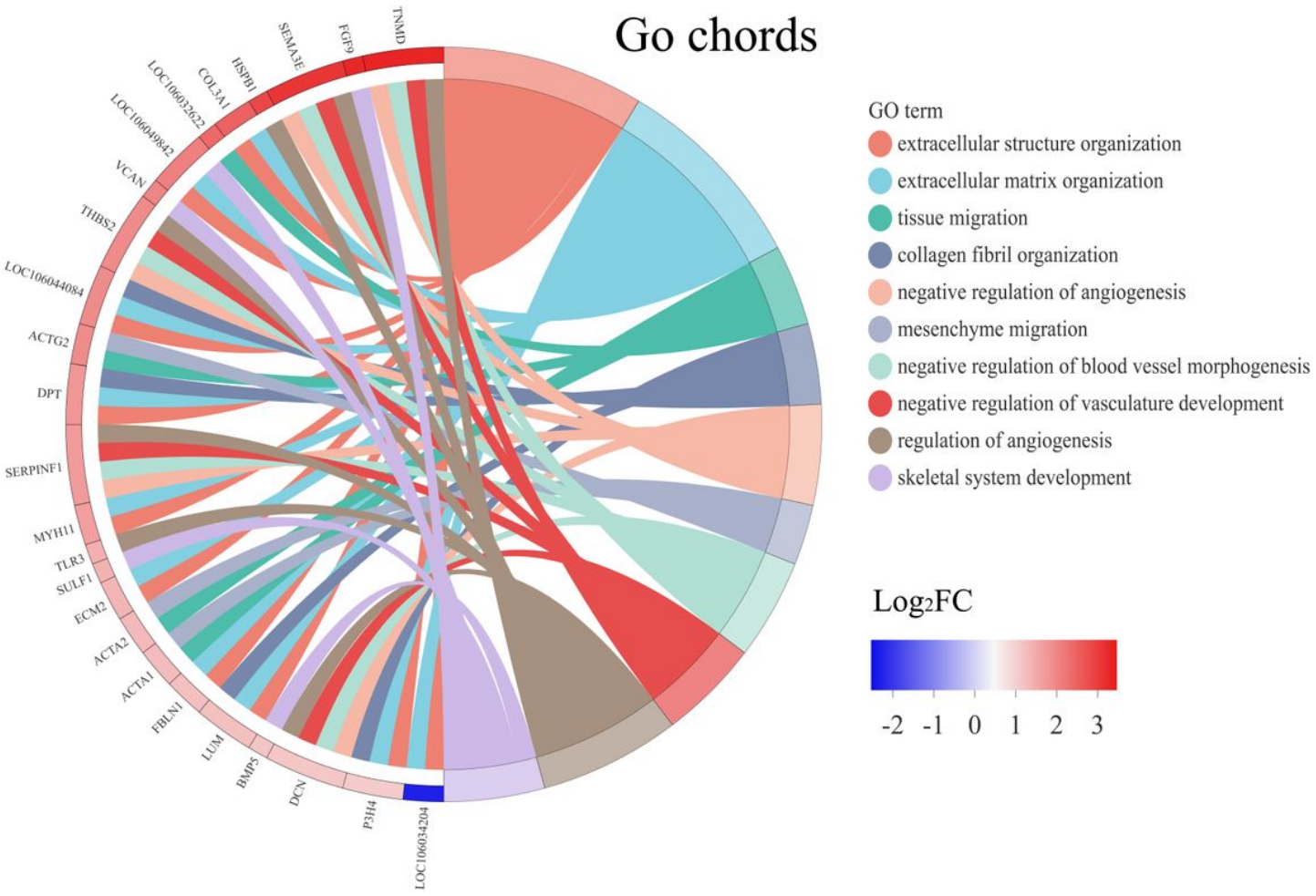

Figure 5

Gene Ontology (GO)-term analysis for DEGs in the skin tissue of different-sized knob. Chords represent a detailed relationship between the ex-pression levels of DEGs (left semicircle perimeter) and their enriched GO terms (right semicircle perimeter). The genes are linked to their assigned terms via colored ribbons. For each gene, the expression values (transcripts-per-million reads) of DEGs in the knob tissue are shown with colored rectangles.

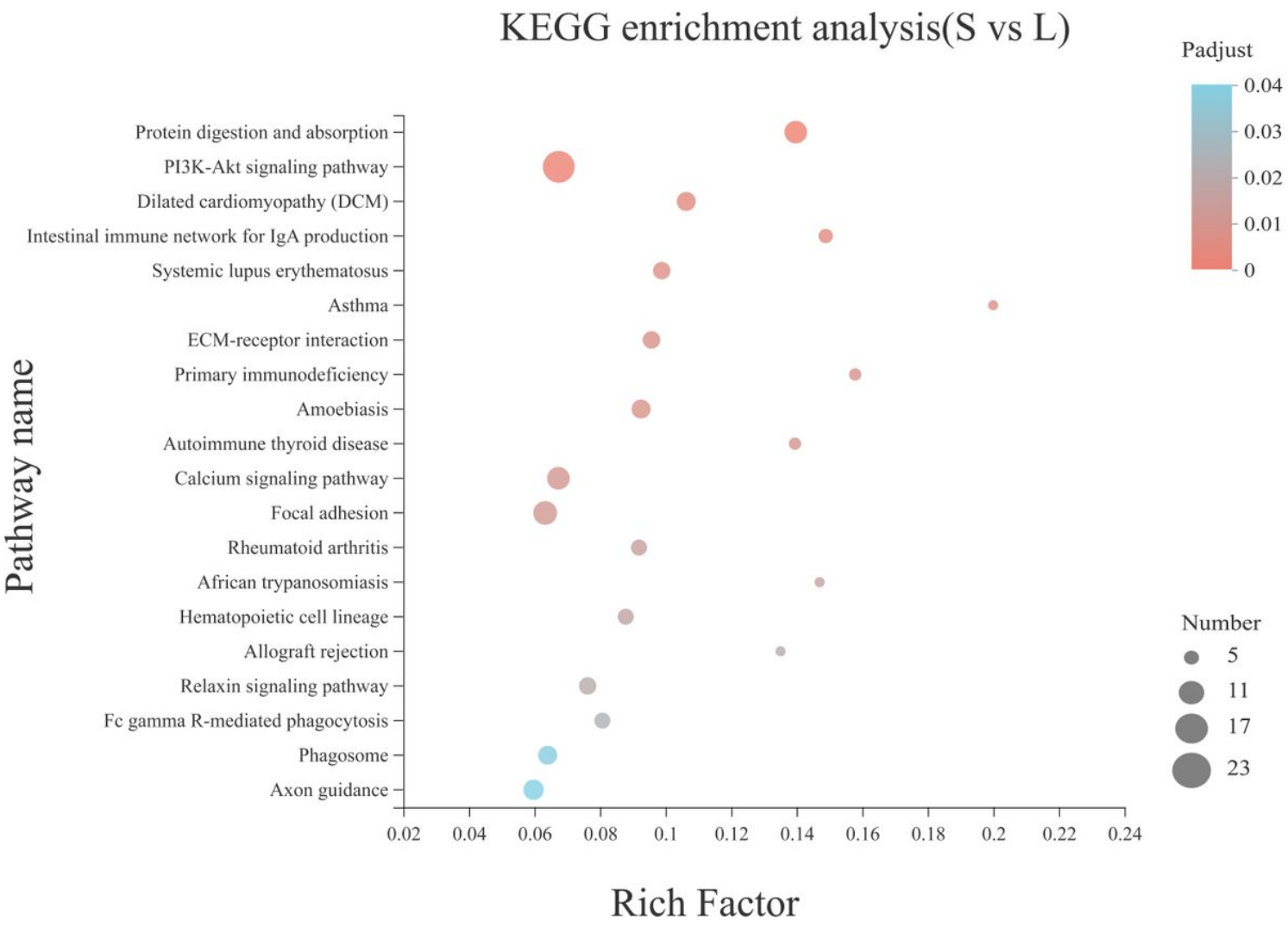

Figure 6

Associations between pathway names and Rich factor ratios. The larger the enrichment coefficient is, the larger the degree of enrichment is. The dot sizes represent the number of genes in each indicated pathway, and the dot colors correspond to different P-adjust ranges. KEGG: Kyoto Encyclo-pedia of Genes and Genomes.

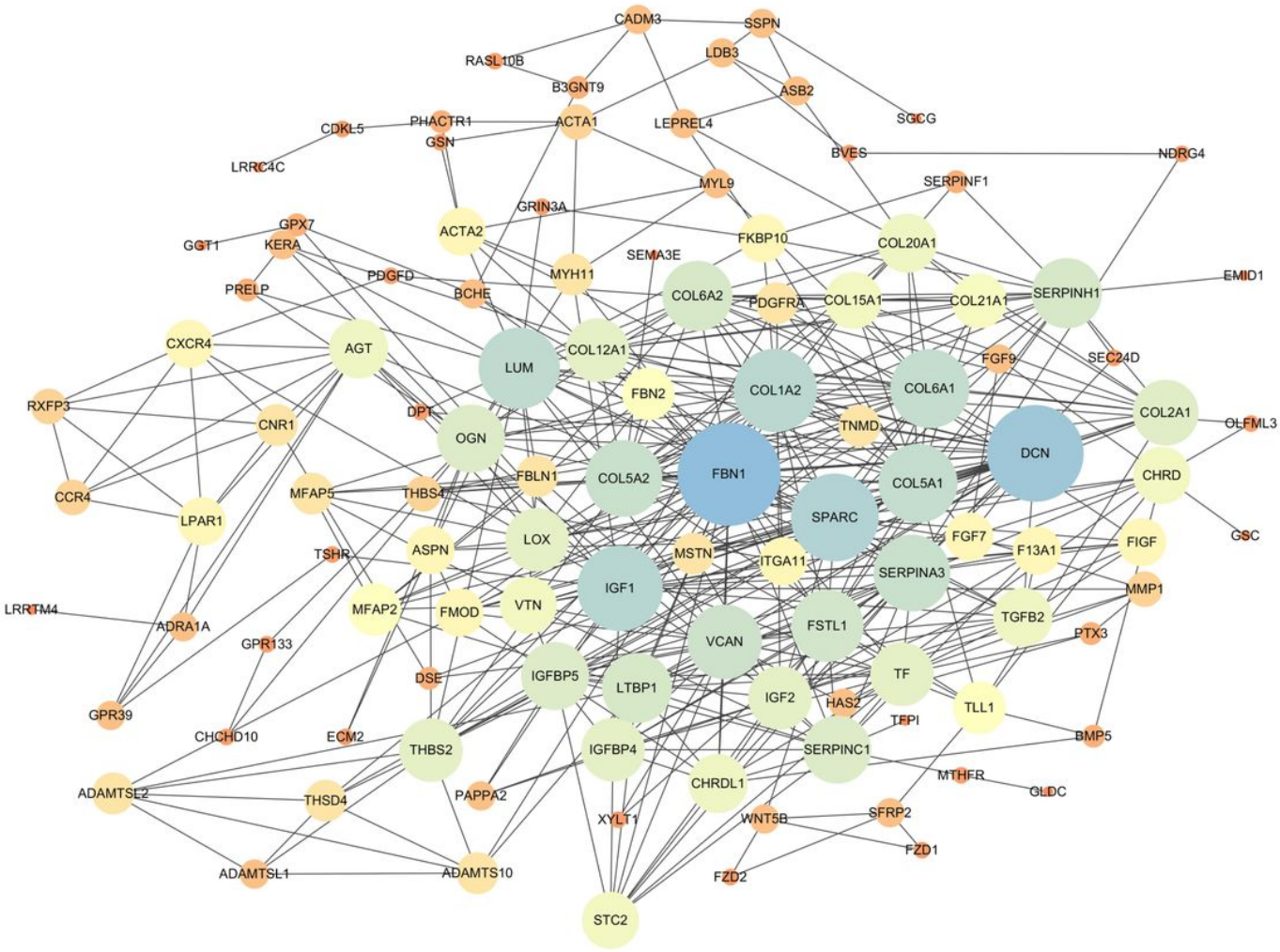

Figure 7

Protein–protein interaction network for the DEGs identified in this study. Nodes represent proteins, and edges represent interaction between pro-teins. The size and color of a given node are proportional to its degree, where the degree is defined as the number of proteins that interact with the node.

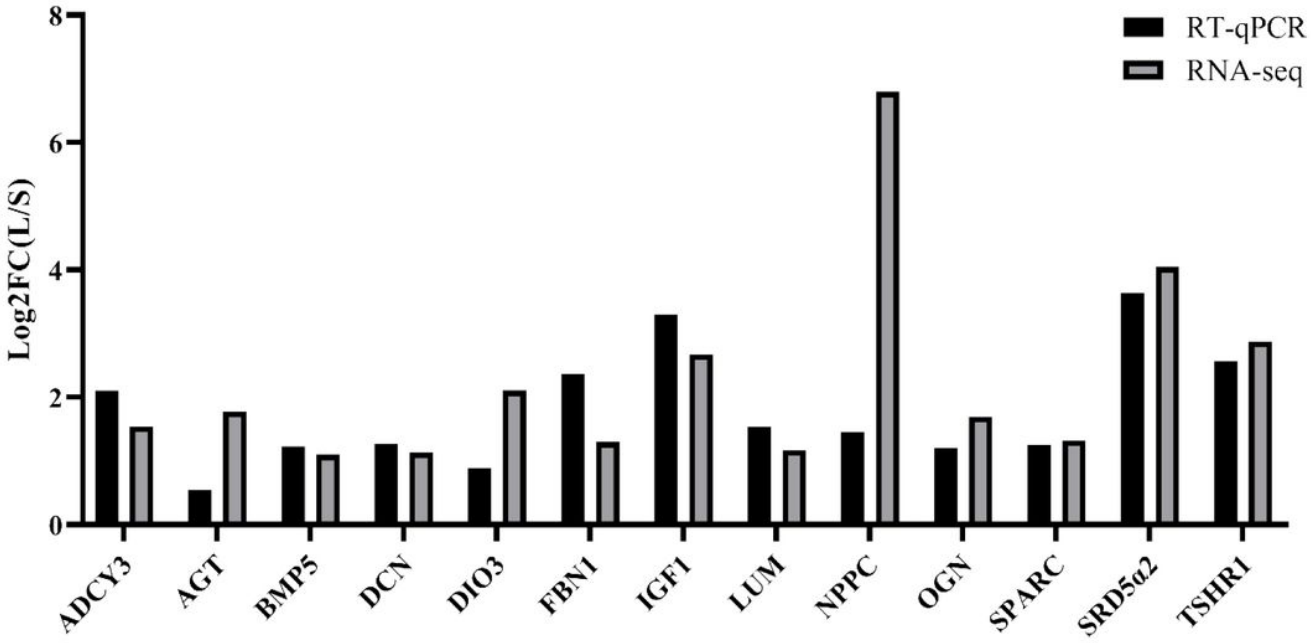

Figure 8

Validation of DEGs: The X-axis represents the selected 13 genes and the Y-axis represents the log2 FC (L/S) values derived from RT-qPCR and RNA-seq data. RT-qPCR: Real-Time quantitative polymerase chain reaction, RNA-seq: RNA sequencing.

Supplementary Files

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

- supplementarytable.docx