

Developmental transcriptome and genes related to skin appendage differentiation in hedgehog (*Atelerix albiventris*)

Hui-Ming Li

Guangdong Institute of Applied Biological Resources

Bi-Ze Yang

Guangdong Institute of Applied Biological Resource

Xiu-Juan Zhang

Guangdong Institute of Applied Biological Resource

Lin-Miao Li

Guangdong Institute of Applied Biological Resource

Hai-Ying Jiang

Guangdong Institute of Applied Biological Resource

Jinping Chen (✉ chenjp@giabr.gd.cn)

Guangdong Institute of Applied Biological Resources <https://orcid.org/0000-0002-0808-6617>

Research

Keywords: adaptive evolution, RNA-Seq, skin disease, differentially expressed genes, molecular biology

Posted Date: December 16th, 2019

DOI: <https://doi.org/10.21203/rs.2.18760/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Hedgehog spines are skin appendages that have evolved as a result of the interaction of their skin with the environment. However, such a differentiation mechanism during skin development leads to a high skin appendage diversity, the origins of which are still not fully understood. Spine-skin and hair-skin offer a natural model for the analysis of the genomic basis for the evolution of epidermal appendage formation. We performed a comparative transcriptomic analysis of hedgehog (*Atelerix albiventris*) at multiple developmental stages, and tried to explore the genetic basis for differentiation and development and the resulting expression of the spine-trait.

Results

A total 15,158 differentially expressed genes (DEGs) were identified. We determined the gene modules and programs corresponding to the various phenotypic traits at different developmental stages by WGCNA analysis. Objective analysis of gene module expression revealed that HIPPO, TGFB, MAPK and Wnt signaling pathways regulate the activation and cell proliferation and differentiation at the skin-appendage development stage. Further, the key genes encoding keratin, FGF, TEAD, and other proteins regulate molecular localization and the cell cycle for hair development and differentiation. Finally, we found a number of highly expressed immune genes in the skin, suggesting that hedgehog spines, unlike pangolin scales, have evolved independently to protect against predators rather than compensate for low autoimmune immunity.

Conclusions

The variability of gene expression profiles of hair-type and spine-type at multiple development stages, and key candidate genes and pathways at the molecular level might provide evidence that can help elucidate the genetic basis of skin appendage development and differentiation of hedgehog.

Background

Perceiving and responding to life-threatening signals and regulating their own morphological characteristics constitute a fundamental challenge for all mammals. Evolution has shaped the ability of organisms to adapt their skin appendages to increase chances of survival. Accordingly, a variety of appendages have evolved on the skin of mammals, either to assist organisms in carrying out specific behaviors or to protect them from predators and pathogenic bacteria [1, 2].

Compared with other animals that possess passive-defense feathers, scales, etc., the hedgehog has evolved into an unusual, nocturnal, spine-covered mammal [3]. Spines potentiate the attacking power and enhance the role of skin appendages as a defense mechanism. During its life cycle, the hedgehog progresses through a series of well-defined stages, from the embryonic, to the non-spine, to the with-spine

stage. These transitions are governed by tightly regulated gene expression at the pre-transcriptional, epigenetic, and translational levels. Spine-skin and hair-skin offer a natural model for the analysis of the genomic basis for the evolution of epidermal appendage formation.

The development of hair follicles on a skin appendage with a complex growth cycle is characterized by the anagen, catagen, and telogen stages, and many key signaling pathways and genes were involved in regulation [4–6]. According to many studies, the WNT/ β -catenin signaling pathway in the dermis may be the first dermal signal [7–9]. Another secreted protein in the follicular placode that plays a major part in epithelial-mesenchymal signaling is Sonic hedgehog (SHH) [10, 11]. SHH depends on WNT signaling, and this protein is required for the proliferation of follicular epithelium and the development of the dermal condensate into the dermal papilla [4]. In addition, many key genes related to skin appendage development have been identified, such as TGF- α receptor (epidermal growth factor receptor) and the transcription factor ETS2, which regulate the shape of hair follicles and are responsible for hair follicle architecture and wavy hair [12, 13], and HOXC13 and KRT75, which control the differentiation of the hair shaft [14–17]. Further, EDA and EDAR interact with members of the bone morphogenetic protein (BMP) family, some of which inhibit follicle development to establish follicle patterning [2, 18–20].

The abovementioned studies primarily focused on mammalian hairs and scales and little is known about changes in gene expression during the development of spine appendages in mammals. Therefore, here, we aimed to perform a detailed analysis of the transcriptome of the skin of abdomen hair and dorsum spines of hedgehogs (*Atelerix albiventris*) by comparing groups of transcripts differentially expressed at different hedgehog developmental stages. We identified key candidate genes related to the development and differentiation of skin appendages. Their importance should be verified in future studies. Nonetheless, the presented novel information will be widely applicable in many fields, e.g., skin disease, skin immune genes, also providing insight into the molecular mechanisms of skin appendage development in general.

Methods

Biological samples

Ten hedgehogs at six developmental stages (Stage 0: Emb, the embryo; Stage I: HH1N/HS1N, within 2 h of birth; Stage II: HH1Y/HS1Y, over 2 h on the first day of birth; Stage III: HH5/HS5, day 5 after birth; Stage IV: HH14/HS14, day 14 after birth; Stage V: HH43/HS43, day 43 after birth) were collected from a commercial animal farm (Dongguan City, China) (Supplementary Table 1). For the skin appendage collection, the animals were first anaesthetized by using diethyl ether and then killed via cervical dislocation. Two embryos and 20 skin tissues representing the two types of appendages (abdomen spine-type and dorsal hair-type) were rapidly excised, immediately snap-frozen on dry ice, and stored at -80°C until RNA extraction.

RNA extraction, library preparation, and Illumina sequencing

Total RNA from each tissue sample was extracted using the RNeasy Kit (Qiagen, Germany). RNA purity was determined using a NanoPhotometer® spectrophotometer (IMPLEN, CA, USA). RNA integrity was assessed using the RNANano 6000 assay kit and the Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA). RNA concentration (ng/ml) was determined using the Qubit® RNA assay kit and Qubit® 2.0 Fluorometer (Life Technologies, CA, USA). Qualified RNAs were used for cDNA library construction and sequencing. All the sequencing stages were conducted at Novogene sequencing company (Beijing, China).

***De novo* transcriptome assembly and functional annotation**

Sequencing reads from 22 samples were first evaluated using FastQC to assess the initial read quality. Clean reads were obtained by removing reads containing adapters or poly-N, and low-quality reads from raw data. Simultaneously, Q20, Q30, GC content, and sequence duplication level of the clean data were calculated. All downstream analyses employed clean data of high quality. The trimmed paired-end reads were resynchronized (<https://github.com/enormandeu/Scripts>) and assembled *de novo* using the Trinity assembler [21].

All unigenes were annotated in the course of analysis with the basic local alignment search tool (BLASTX), considering hits with e-values of $1E-5$ against seven databases: NCBI non-redundant protein sequences (Nr); NCBI non-redundant nucleotide sequences (Nt); Protein family (Pfam) database; Clusters of Orthologous Groups of Proteins (KOG/COG) database; Swiss-Prot (a manually annotated and reviewed protein sequence database); KEGG Ortholog (KO) database; Gene Ontology (GO) database. The best-aligned results were used to determine the sequence direction of the transcripts. In case of discrepancies of analyses using different databases, the sequence direction of the transcript was determined according to the prioritization order of NR > Swiss-Prot > KEGG > COG.

Differential expression analysis

Differentially expressed genes (DEGs) were identified by comparing gene expression levels between samples (or sample groups). Bowtie2 [22] was used to align the clean reads with all unigenes, and RSEM [23] was used to calculate the gene expression levels for each sample. Differential gene expression was analyzed using DESeq2 V1.16.1 in RStudio V1.0.143 running R72 V3.4.1.

For samples with biological replicates, differential expression analysis of two conditions or groups was performed using the DESeq R package (1.10.1). DESeq [24] provides statistical approaches for determining differential expression in digital gene expression datasets using a model based on negative binomial distribution. The resulting P-values were adjusted using the Benjamini and Hochberg's approach to control the false discovery rate. Genes with an adjusted P-value < 0.05 identified by DESeq [25] were deemed to be differentially expressed. For samples without biological replicates, prior to differential gene expression analysis, for each sequenced library, the read counts were adjusted by using edgeR program package [26] and a single scaling normalized factor. Differential expression analysis of two samples was

performed using the DEGseq R package. The P-values were adjusted using the q-value [27]; q-value < 0.005 and $|\log_2(\text{foldchange})| > 1$ were set as the threshold for significantly differential expression.

Further, KOBAS software [28] was used to test the statistical enrichment of DEGs in KEGG pathways. GO enrichment analysis of DEGs was implemented by the GO seq R package, based on Wallenius non-central hyper-geometric distribution [29], adjusting for gene length bias.

Quantitative real-time reverse-transcription PCR (RT-qPCR)

RNA was extracted from 22 skin appendage tissues using TRIzol reagent. Then, cDNA was synthesized by using the Toyobo reverse-transcription kit. Eleven genes related to hair and/or spine development were selected for analysis based on the functional annotation data, and fluorescent quantitative PCR primers were designed accordingly (Supplementary Table 2). RT-qPCR was performed in a 20-ml reaction volume, with four technical replicates for each sample, using the TransStart® Top Green qPCR SuperMix kit (TransGen Biotech, Beijing, China). The relative gene expression levels were analyzed by using the $2^{-\Delta\Delta CT}$ method (Bustin *et al.* 2010). The PCR conditions were as follows: pre-denaturation at 95 °C for 10 min; 40 cycles of 15 s at 95 °C (denaturation), 30 s at 58 °C (annealing), and 20 s at 72 °C (extension); and a final melting curve stage from 60 °C to 95 °C, to verify the specificity of the amplicons.

Results

Transcriptomic atlas for developing hedgehog skin appendages

To identify the transcriptome and molecular mechanisms governing the development and differentiation of skin appendages, we analyzed the temporal changes in the transcript abundance of *A. albiventris* (Fig. 1a). RNA sequencing generated 81.4 ± 6.6 G (mean \pm SD) read pairs for 22 skin tissues (Supplementary Table 3). After quality trimming, $96.8\% \pm 0.6\%$ of reads were retained, indicating a high-quality dataset (>90% reads with $\geq Q30$). The *de novo* assembly using Trinity revealed 32,8576 unigenes, which were used for subsequent analysis. The main length distribution of 94.15% of unigenes was in the range of 0.2–2 kb (Fig. 1b).

A neural network graph based on SOM analysis revealed dynamic transcriptional changes occurring at the individual developmental stages of *A. albiventris* (Fig. 1c). Overall, 85.6% of unigenes were successfully annotated with significant hits (e-value < $1E-10$). Absolute unigenes were annotated using seven major public protein databases (Table 1).

Among the 56 annotated GO terms, cellular process (38,648 unigenes) was the most common annotation in the biological process (BP) category and metabolic process (31,933 unigenes) and single-organism process (28,291 unigenes) were the next most abundant GO terms. In addition, a low proportion of unigenes were annotated as terms related to skin appendage development, such as cellular component organization (9844 unigenes), developmental process (2127 unigenes), cell killing (79 unigenes), and cell aggregation (1 unigene) (Fig. 2a). In the molecular function (MF) and cellular component (CC) categories,

binding (31,320 unigenes) and cell (25,455 unigenes) were the most common annotation terms. Details of GO classification of the *A. albiventris* unigenes are provided in Supplementary Table 4.

In addition, we annotated all unigenes using KEGG pathway analysis to understand their top-level function and role in the hedgehog biological system. Overall, 36,283 unigenes were mapped to 237 biological pathways clustered in five KO hierarchies in the functional ortholog system. Among these KO hierarchies, 15 unique KEGG pathways represented cellular processes and 30, 22, 98, and 72 pathways represented environmental information processing, genetic information process, metabolism, and organismal systems, respectively. The largest KO group was translation (5743 unigenes, 13%), followed by signal transduction (5551 unigenes, 12.6%), and endocrine system (3223 unigenes, 7.3%). Further, a small proportion of unigenes were annotated into categories related to skin appendage development, such as SHH (95 unigenes), p53 (161 unigenes), apoptosis (189 unigenes), and TGF- β (206 unigenes) signaling pathway (Fig. 2b).

Characterization of the developmental cycle of skin appendages in hedgehog

To systematically assess the key regulatory genes related to the development and differentiation of spine and hair in *A. albiventris*, we compared the gene expression profiles of spine-type and hair-type tissues in *A. albiventris* at six developmental stages (Stages 0–V). Highly expressed DEGs and those with type-specific expression in the skin appendage were screened using the following seven comparison combinations: EMH1 (Emb vs. HH1N), EMS1 (Emb vs. HS1N), 1N (HH1N vs. HS1N), 1Y (HH1Y vs. HS1Y), 5D (HH5 vs. HS5), 14D (HH14 vs. HS14), and 43D (HH43 vs. HS43).

Overall, to identify the core transcriptome, we screened 22,587 shared unigenes expressed at each developmental stage. GO enrichment analysis of these shared unigenes revealed no molecular function change in the most enriched terms during the six sequential developmental stages of hair-type tissue (Fig. 3a). The following were always identified: BP: GO:0042254 Ribosome biogenesis, GO:0006412 Translation; MF: GO:0003735 Structural constituent of ribosome; CC: GO:0005840 Ribosome. However, in the spine-type tissue, except for BP, the most enriched terms in the MF and CC categories changed during stages III–IV (Fig. 3b). We also screened and annotated the top 100 genes based on the FPKM values in tissues at each stage. In general, protein-coding genes related to skin appendage development (e.g., *RPL13A* and *RPLP1*) and keratin (e.g. *KRT1* and *KRT2*) were highly dynamic and their expression was time-dependent throughout a developmental stage. This suggested their widespread roles in organ development. Considering the transformation of biological functions during stages III–IV, the corresponding regulatory genes with highest frequency changed from ribosomal protein-coding genes (*L13a* and lateral stalk subunit P1) to keratin-encoding genes (types I and II) (Fig. 3c).

We also identified a set of 15,158 DEGs based on comparison combinations of 1N, 1Y, 5D, 14D, and 43D, and one intersection with 163 DEGs (Fig. 4a–f). According to the KEGG enrichment analysis, we screened 27 KEGG pathways related to skin appendage development and then analyzed DEG clustering patterns in these 27 KEGG pathways for different comparison combinations (Fig. 5). As shown in the heatmap, three

clusters of gene expression patterns were apparent: Cluster 1, including the groups EMH1 and EMS1; Cluster 2, including the groups 1N, 1Y, and 5D; and Cluster 3, including the groups 14D and 43D.

Establishing gene modules underlying the key developmental DEGs by WGCNA analysis

We first imported hair-type tissue, spine-type tissue, and embryo sample data to R and executed WGCNA analysis with an emphasis on genes that underwent spine-differentiation changes. WGCNA identified 23 modules of co-expressed genes (Fig. 6a). The brown, tan, magenta, green, and purple modules were the top five modules related to the spine-type trait, and the red, light cyan, cyan, light green, and pink modules were the top five modules related to the hair-type trait (Fig. 6b).

GO analysis of the top eight modules revealed several key biological processes that took place during spine and hair development (Fig. 6c). Analysis of the GO terms in each module revealed signaling activation and transmission-related genes (magenta module, C1), including genes for the regulation of cellular catabolism (*ZNF365*, *TPM2*, *TPM1*, and *KRT2*), cell activation (*MUSTN1*, *TGFB3*, *GNDF*, and *PDGFC*), and cell proliferation (*MYH* and *DES*), highly expressed at Stage I and actively preparing the follicular cells for hair growth. Genes enriched in the metabolic and biosynthetic processes (tan module, C2) exhibited increased expression at Stage II compared with that at C1 and included genes related to the primary metabolic process, cellular biosynthetic and organic substance process, and cell. During Stages III–IV of development, the expressed genes were associated with further development (hardening, thickening, etc.) of the spine, including organic substance transport, macromolecule/cellular/protein localization, and cell cycle. During that time, most of the genes in the green module (C3) were related to keratin (*KRT1*, *KRT2*, and keratin-associated proteins). GO terms in the brown module (C4) were involved in the primary metabolic and biosynthetic process, including cellular and organic substance metabolism (*USP*, *TOP2*, *TGFB2*, *SSH*, and *TAF*), location, gene expression, and transport (*CEBPD*, *MYH*, *TCN2*, and *THBS2S*), and cellular biosynthesis (*MT1*, *FCN*, *ISYNA1*, *KBTBD5*, and *GALNT*).

We also performed KEGG analysis of modules related to the two skin appendage traits. We screened five key signaling pathways involved in spine development and differentiation, namely, the HIPPO, TGFB, MAPK, cell cycle, and Wnt signaling pathways. Some key genes that promote cell reproduction, differentiation, and anti-apoptosis were significantly enriched in these pathways, e.g., *FGF*, *EGF*, *PAK1*, *TEAD*, *CDK6*, and *TCF* (Fig. 7a, Table 2). Further, five key signaling pathways were identified as related to hair development, i.e., TGFB, p53, apoptosis, cell cycle, and tight junction signaling pathways (Fig. 7b). Genes that regulate apoptosis, arrest cell development, and reduce cell numbers were enriched in these pathways, e.g., *CCND3*, *CCNB*, *CD1*, *ACTB_G1*, and *TUBA*.

Verification of important DEGs related to hair and/or spine development and differentiation-related molecules using RT-qPCR

To validate the expression patterns indicated by the transcriptome data, 11 differentiation-related DEGs were selected for RT-qPCR analysis. These were *KRT2*, *LEF1*, *RSPO2*, *ARRB*, *TGFB2*, *GRB2*, *SFN*, *TCF7*, *CTNNA1*, *HOXC13*, and *WIF1*. Indeed, the expression trends determined by RT-qPCR significantly

correlated with the RNA-Seq data (Fig. 8a). The expression of these 11 genes at five developmental stages was also analyzed, revealing expression patterns similar to those determined by RNA-Seq experiments (Fig. 8b). The expression of *LEF1*, *TGFB2*, *SFN*, and *WIF1* at stages I–III significantly increased. However, the expression of *KRT1* and *RSPO2* gradually decreased. Overall, both approaches confirmed the observed DEG trend patterns, indicating the accuracy of the transcriptome data and *de novo* RNA-Seq data.

Immune-related genes in skin epidermis of hedgehog and pangolin

In the KEGG classification, we also identified 4,687 unigenes were involved in immunity related pathways. The 16 immune system related pathways were extracted, Platelet activation, Leukocyte transendothelial migration and Chemokine signaling pathway were the top three categories with higher proportions (Fig. 9). Some key immune molecules were identified from transcripts of *A. albiventris*, such as ACTB_G1, AKT, COL1A5, GNAS, MHC2, P38, MYL6F, MAPK1_3 etc. The details of immune-related unigenes were listed in Supplementary Table 6. Among these annotated immune genes, 16 genes were also found in *Manis javanica*, such as ARPC2, SHC1, GNB3, CD5, MHC2, PLG etc.

Discussion

In mammals, the development of skin appendages, such as hair, tooth, and scale, involves complex interactions between the epidermis and the underlying mesenchyme as part of an established hierarchical morphogenetic process [30]. Specifically, mammals develop a coat containing many distinct types of hair. Such diversity is associated with molecular and signaling pathways that drive the formation and induction in a specific spatial and temporal manner [31]. Reciprocal interactions between the epithelial and mesenchymal tissues constitute a central mechanism that determines the location, size, and shapes of organs [32].

In the current study, we aimed to explore the genetic basis for hedgehog differentiation and development and the resulting expression of the spine-trait. First, according to the comparison of the number of up- and down-regulated genes in 5 developmental stages, we found that spines development requires more genes to participate in expression and regulation. Especially from stage I to stage II, the number of DEGs increased by 30%.

Second, according to the core gene analysis, the six developmental stage in the current study indeed covered the main skin development process throughout the hedgehog life cycle. We found some key nodes in the development of hedgehog skin appendages. (1) Spines developed earlier than hairs. At stage II, the hedgehog began to grow spines on its back, but no hair on its abdomen, indicating that the 2-h time point after birth might be a critical time for hair and spine differentiation in *A. albiventris*, contributing to developmental changes in spine activity. (2) 5-14 d after birth, the dorsum hair follicle could be in the anagen stage. The GO enrichment and clustering analysis (Figs. 3, 5) of DEGs showed that during this period, the soft spine gradually became thicker and harder, and some genes involved in such processes as cell transport and protein localization, e.g., *RSPO2*, *KRT1*, and *KRT2*, were significantly upregulated.

This was consistent with the findings of Cadieu *et al.* (2009) who reported that a combination of RSPO2 and keratin genes caused hair to become coarse and stiff. (3) 43 d after birth, the dorsal and abdominal hair follicles reached the catagen or telogen stage. At that time, the development of skin appendages was complete, and some genes involved in basic metabolism and the immune system, e.g., *MHC* and *APOE*, became significantly upregulated. Some keratin-related genes began to be significantly down-regulated.

Third, WGCNA analysis identified several key signaling pathways related to the development of the appendage trait (Fig. 7). The first was Wnt signaling pathway which was involved in the establishment of the first dermal signal [33]. Our DEGs associated with two phenotypes both were annotated in the Wnt pathway, however, genes associated with the spine-type tissue were significantly enriched in this pathway, while hair-type related genes were not. This indicated that the activation of Wnt signaling pathway was required at the early stage of hair and/or spine development in hedgehog, similar to other mammals. As well as three b-catenin genes (*CTNNB1*, *ICAT*, and *JUP*), genes for FZD, LEF/TCF, and Wnt family members were upregulated in the spine-type tissue (Fig. 8). Sustained b-catenin activity leads to increased dermal thickness and proliferation of fibroblasts, that could lead to increased hair follicle substrate and dermal condensate area during development, accelerating hair follicle differentiation [4,5]. Otherwise, a null mutation in a gene encoding the Wnt effector protein LEF1 in mouse leads to a failure of formation of the vibrissae and two-thirds of the body hair follicles [34,35]. This is consistent with the findings for *A. albiventris*. In addition, we identified *FOLR*, *PML*, *APC*, *WNT10*, *WNT5*, *WNT4*, *WNT11*, *WNT6*, and *WNT7* as the *A. albiventris* DEGs related to the activated Wnt/b-catenin signaling cascade, and inducing the placode formation required for normal embryonic and cell development and proliferation. The Wnt signal transduction cascade controls multiple biological processes throughout the development and adult life of all animals. At the same time, aberrant Wnt signaling is associated with a wide range of pathologies in human. It has been reported that some diseases related to human skin appendages are regulated by the Wnt signaling pathway, e.g., tooth agenesis [36], odontoonychodermal dysplasia [37], palmoplantar hyperkeratosis with squamous cell carcinoma of skin and sex reversal [38], acne inversa [39], focal dermal hypoplasia [40], and nonsyndromic congenital nail disorder [41].

Further, DEGs related to the two phenotypes were also significantly enriched in the TGF- β signaling pathway, affecting follicle fate direction and distribution pattern. *TGFB*, *E2F4*, *DCN*, *PAK1*, and *ID4* were here identified as DEGs upregulated in the spine-type tissue. We hypothesize that these genes are involved in skin appendage development and follicle size. Jackowska *et al.* (2012) observed increased TGFB1, TGFB2, and TGFB3 protein levels in oocytes in large follicles compared with those in small follicles [42]. This is consistent with the hedgehog data from the current study. By contrast, members of the BMP family of secreted signaling molecules appear to act as inhibitors of follicle formation. In the current study, several BMP and Noggin genes were identified as DEGs, with opposite trends of expression in the hair-type and spine-type tissues. Noggin is a secretory molecule that may inhibit BMP activity and is involved in the development of hair follicle. In support of this hypothesis, mouse lacking Noggin has fewer hair follicles than wild type and retarded follicular development [43]. Further, some diseases related to human skin appendages are regulated by the TGF- β signaling pathway, e.g., the Loeys-Dietz syndrome [44] and Axenfeld-Rieger syndrome [45].

In addition, we identified some DEGs enriched in the MAPK and cell cycle signaling pathways. These pathways often act together by forming signaling loops during organogenesis [46]. They induce the most fundamental biological processes, such as the formation of periodic patterns. In the current study, *FGF*, *EGF*, *CDK6*, and *SFN* were upregulated in the spine-type tissue. Hebert *et al.* (1994) demonstrated that the *FGF* gene plays an important role in the regulation of hair cycle growth, and functions as an inhibitor of hair elongation by promoting progression from the anagen stage, the growth phase of the hair follicle, to the catagen stage, the apoptosis-induced regression phase [16]. Hence, we believe that these genes may be important regulators that affect the development and differentiation of spine.

We identified *UCP1* as a DEG in the Wnt signaling pathway. It encodes a mitochondrial protein responsible for thermogenic respiration that participates in non-shivering adaptive thermogenesis in response to temperature variation and, more generally, in the regulation of energy balance. *UCP1* was significantly upregulated in the hair-type tissue, suggesting that the spiny back of hedgehog is more sensitive to cold than abdomen with hair. We also compared the *KRT1* gene sequences of some mammals with those of the European hedgehog, and identified an approximately 40-bp deletion in the *KRT1* gene in the latter. Keratin gene is highly conserved in mammals. Whether this deletion has a special significance for the development of hedgehog skin appendages requires further study and analysis.

At last, we found some immune-related genes in the skin epidermis transcriptome of hedgehog. By observing hedgehog captive population, we found that there was no obvious immune deficiency in this species, which was consistent with the normal expression of the detected immune genes. Choo *et al.* (2016) found that interferon epsilon (IFNE) which expressed in epithelial cell and mucosal immunity on pangolins skin has been pseudogenized, so they propose that scale development was an innovation that provided protection against injuries or stress and reduced pangolin vulnerability to infection, the protection of pangolin scales on the body can compensate for the low immunity of this species to a certain extent. Interestingly, 16 of the immune genes we found on hedgehog were also annotated on pangolin. However, neither IFNE gene nor the 16 genes we detected showed pseudogenization in hedgehogs, which indicated that the compensatory effect of skin appendages on immunity did not occur in hedgehogs. So, hedgehog's ratchet may have evolved independently of the immune system, and its function is more physical defense.

Conclusions

In the current study, we conducted a comprehensive transcriptome analysis of *A. albiventris*, to explore the genetic basis for the growth of skin appendages. The transcriptome analysis provided a rich list of unigenes expressed in hair-type and spine-type tissues in *A. albiventris* at six different developmental stages. Overall, 32,8576 unigenes and 15,158 DEGs were screened from seven comparison combinations from the transcriptome datasets. Candidate genes related to five key signaling pathways were identified, which are likely involved in the regulation of the growth and differentiation of the hair and spine. The knowledge of molecular and signaling pathways greatly contributes to the current genetic resources for

hedgehog and certain mammalian species that harbor shaggy appendages as well as traits that could potentially be exploited commercially for curing skin diseases of other animals, even human.

Abbreviations

Atelerix albiventris, *A. albiventris*; Hiseq, High-throughput sequencing; FGF, fibroblast growth factor; EGF, Pro-epidermal growth factor; TEAD, Transcriptional enhancer factor; TCF, Transcription factor; SHH, Sonic Hedgehog; EDAR, Tumor necrosis factor receptor superfamily member; BMP, Bone morphogenetic protein; WGCNA, weighted gene correlation network analysis; DEGs, Differentially expressed genes; RT-qPCR, Quantitative real time polymerase chain reaction; SOM, Self-Organizing Feature Map

Declarations

Acknowledgments

We gratefully acknowledge Professor Gang Li for instructing us designed the experiments. We thank Wiley publishing (<http://wileyediting.com>), for editing the English text of a draft of this manuscript.

Author contributions

Li and J. Chen designed research; H. Li analysed data, and wrote the report; B. Yang developed software necessary to perform and record experiments; L. Li, X. Zhang and H. Jiang provided expertise and advice on computational analysis; and all authors edited the report.

Funding

This study was supported by the GDAS' Project of Science and Technology Development (2019GDASYL-0103062), China Postdoctoral Science Foundation (2019M652832), and the GDAS Special of Science and Technology Development (2018GDASCX-0107).

Data availability

The data analyzed in the current study are included within the article and its additional files. All unigene sequences from *A. albiventris* have been deposited in the GenBank Sequence Read Archive (SRA) under the accession number PRJNA561241 for SUB6195278.

Ethics approval and consent to participate

All animal procedures were approved by the ethics committee for animal experiments at the Guangdong Institute of Applied Biological Resources (reference number G2ABR20170523) and followed basic principles.

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Author details

1 Guangdong Key Laboratory of Animal Conservation and Resource Utilization, Guangdong Public Laboratory of Wild Animal Conservation and Utilization, Guangdong Institute of Applied Biological Resources, Guangzhou 510260, China

References

1. Gandolfi B, Outerbridge CA, Beresford LG, Myers JA, Pimentel M, Alhaddad H, et al. The naked truth: Sphynx and Devon Rex cat breed mutations in KRT71. *Mamm Genome*. 2010;21:509–15.
2. Di-Poi N, Milinkovitch MC. The anatomical placode in reptile scale morphogenesis indicates shared ancestry among skin appendages in amniotes. *Sci Adv*. 2016;2:1–9.
3. Zeller M, Reusch TBHTBH, Lampert W. A comparative population genetic study on calanoid freshwater copepods: investigation of isolation-by-distance in two Eudiaptomus species with a different potential for dispersal. *Limnol Oceanogr*. 2006;51:117–24.
4. Millar SE. Molecular mechanisms regulating hair follicle development. *J Invest Dermatol [Internet]*. 2002; 118: 216 – 25.
5. Paus R, Mu S, Veen C Van Der, Maurer M, Eichmu S, Ling G. A Comprehensive Guide for the Recognition and Classification of Distinct Stages of Hair Follicle. *J Invest Dermatol*. 1999;113:523–32.
6. Won HI, Schulze TT, Clement EJ, Watson GF, Watson SM, Rosalie C, et al. De novo Assembly of the Burying Beetle *Nicrophorus orbicollis* (Coleoptera : Silphidae) Transcriptome Across Developmental Stages with Identification of Key Immune Transcripts. *J Genomics*. 2018;6:41–52.
7. Noramly S, Freeman A, Morgan BA. Beta-Catenin Signaling Can Initiate Feather Bud Development. *Development [Internet]*. 1999;126:3509–21.
8. Andl T, Reddy ST, Gaddapara T, Millar SE. WNT signals are required for the initiation of hair follicle development. *Dev Cell*. 2002;2:643–53.
9. Tsai SY, Sennett R, Rezza A, Clavel C, Grisanti L, Zemla R, et al. Wnt/ β -catenin signaling in dermal condensates is required for hair follicle formation. *Dev Biol [Internet]*. Elsevier; 2014;385:179–88.
10. Bitgood MJ, McMahan AP. Hedgehog and Bmp genes are coexpressed at many diverse sites of cell-cell interaction in the mouse embryo. *Dev Biol*. 1995;172:126–38.
11. Iseki S, Araga A, Ohuchi H, Nohno T, Yoshioka H, Hayashi F, et al. Sonic hedgehog is expressed in epithelial cells during development of whisker, hair, and tooth. *Biochem Biophys Res Commun*. 1996;218:688–93.

12. Luetkeke NC, Qiu TH, Peiffer RL, Oliver P, Smithies O, Lee DC. TGF α deficiency results in hair follicle and eye abnormalities in targeted and waved-1 mice. *Cell*. 1993;73:263–78.
13. Yamamoto H, Flannery ML, Kupriyanov S, Pearce J, McKercher SR, Henkel GW, et al. Defective trophoblast function in mice with a targeted mutation of *Ets2*. *Genes Dev*. 1998;12:1315–26.
14. Godwin AR, Capecchi MR. *Hoxc13* mutant mice lack external hair. *Genes Dev*. 1998;12:11–20.
15. Potter CS, Pruett ND, Kern MJ, Baybo MA, Godwin AR, Potter KA, et al. The nude mutant gene *Foxn1* is a *HOXC13* regulatory target during hair follicle and nail differentiation. *J Invest Dermatol*. 2011;131:828–37.
16. Cadieu E, Neff MW, Quignon P, Walsh K, Chase K, Parker HG, et al. Coat variation in the domestic dog is governed by variants in three genes. *Science* (80-). 2009;326:150–3.
17. Koch PJ, Sundberg JP, Roop DR, Jaeger K, Den Z, Chen J. Mice Expressing a Mutant *Krt75* (*K6hf*) Allele Develop Hair and Nail Defects Resembling *Pachyonychia Congenita*. *J Invest Dermatol* [Internet]. Elsevier Masson SAS; 2007;128:270–9.
18. Mou C, Jackson B, Overbeek PA, Schneider P, Headon DJ. Generation of the primary hair follicle pattern. *Proc Natl Acad Sci*. 2006;103:9075–80.
19. Pummila M, James MJ, Jaatinen R, Schneider P, Mikkola ML, Thesleff I, et al. Ectodysplasin has a dual role in ectodermal organogenesis: inhibition of *Bmp* activity and induction of *Shh* expression. *Development*. 2006;134:117–25.
20. Kamberov YG, Wang S, Tan J, Gerbault P, Wark A, Tan L, et al. Modeling recent human evolution in mice by expression of a selected *EDAR* variant. *Cell* [Internet]. Elsevier Inc.; 2013;152:691–702.
21. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 2011;29:644–52.
22. Langmead B, Salzberg S. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9:357–9.
23. Li B, Dewey CN. RSEM : accurate transcript quantification from RNA-Seq data with or without a reference genome. *Bioinformatics*. 2011;12:323–323.
24. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol*. 2010;11:106–7.
25. Wang L, Feng Z, Wang X, Wang X, Degseq ZX, Wang L, et al. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics*. 2010;26:136–48.
26. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26:139–40.
27. Storey JD, Tibshirani R. Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A*. 2003;100:9440–5.
28. Mao X, Cai T, Olyarchuk JG, Wei L. Databases and ontologies Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics*. 2005;21:3787–93.

29. Young MD, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 2010;11:1–12.
30. Duverger O, Morasso MI. Epidermal Patterning and Induction of Different Hair Types during Mouse Embryonic Development. *Birth Defects Res C Embryo.* 2009;87:263–72.
31. Hardy MH. The secret life of the hair follicle. *Trends Genet.* 1992;8:55–61.
32. Laurikkala J, Pispä J, Jung H-S, Nieminen P, Mikkola M, Wang X, et al. Regulation of hair follicle development by the TNF signal ectodysplasin and its receptor Edar. *Development.* 2002;129:2541–53.
33. Wodarz A, Nusse R. Mechanisms of Wnt signaling in development. *Annu Rev Cell Dev Biol.* 1998;14:59–88.
34. Van Genderen C, Okamura RM, Farinas I, Quo RG, Parslow TG, Bruhn L, et al. Development of several organs that require inductive epithelial- mesenchymal interactions is impaired in LEF-1-deficient mice. *Genes Dev.* 1994;8:2691–703.
35. Kratochwil K, Dull M, Farinas I, Galceran J, Grosschedl R. Lef1 expression is activated by BMP-4 and regulates inductive tissue interactions in tooth and hair development. *Genes Dev.* 1996;10:1382–94.
36. Lammi L, Arte S, Somer M, Järvinen H, Lahermo P, Thesleff I, et al. Mutations in AXIN2 Cause Familial Tooth Agenesis and Predispose to Colorectal Cancer. *Am J Hum Genet.* 2004;74:1043–50.
37. Adaimy L, Chouery E, Mégarbané H, Mroueh S, Delague V, Nicolas E, et al. Mutation in WNT10A Is Associated with an Autosomal Recessive Ectodermal Dysplasia: The Odonto-onycho-dermal Dysplasia. *Am J Hum Genet.* 2007;81:821–8.
38. Parma P, Radi O, Vidal V, Chaboissier MC, Dellambra E, Valentini S, et al. R-spondin1 is essential in sex determination, skin differentiation and malignancy. *Nat Genet.* 2006;38:304–1309.
39. Dessinioti C, Katsambas A, Antoniou C. Hidradenitis suppurativa (acne inversa) as a systemic disease. *Clin Dermatol.* 2014;32:397–408.
40. Wang X, Reid Sutton V, Omar Peraza-Llanes J, Yu Z, Rosetta R, Kou YC, et al. Mutations in X-linked PORCN, a putative regulator of Wnt signaling, cause focal dermal hypoplasia. *Nat Genet.* 2007;39:836–8.
41. Mohammadi-asl J, Pourreza MR, Mohammadi A, Eskandari A, Mozafar-Jalali S, Tabatabaiefar MA. A novel pathogenic variant in the FZD6 gene causes recessive nail dysplasia in a large Iranian kindred. *J Dermatol Sci.* 2017;88:134–8.
42. Jackowska M, Kempisty B, Woźna M, Piotrowska H, Antosik P, Zawierucha P, et al. Differential expression of GDF9, TGFB1, TGFB2 and TGFB3 in porcine oocytes isolated from follicles of different size before and after culture in vitro. *Acta Vet Hung.* 2012;61:99–115.
43. Botchkarev VA, Botchkareva N V, Roth W, Nakamura M, Chen L, Herzog W, et al. Noggin is a mesenchymally derived stimulator of hair-follicle induction. *Nat Cell Biol.* 1999;1:158–64.
44. Loeys BL, Schwarze U, Holm T, Callewaert BL, Thomas GH, Pannu H, et al. Aneurysm syndromes caused by mutations in the TGF- β receptor. *N Engl J Med.* 2006;355:788–98.

45. Ittner LM, Wurdak H, Schwerdtfeger K, Kunz T, Ille F, Leveen P, et al. Compound developmental eye disorders following inactivation of TGF β signaling in neural-crest stem cells. *J Biol.* 2005;4:<https://doi.org/10.1186/jbiol29>.
46. Jung HS, Francis-West PH, Widelitz RB, Jiang TX, Ting-Berreth S, Tickle C, et al. Local inhibitory action of BMPs and their relationships with activators in feather formation: Implications for periodic patterning. *Dev Biol.* 1998;196:11–23.

Tables

Table 1 Summary of the functional annotation of assembled hedgehog unigenes with public protein databases using BlastX cut-off E-value of 1E-5

	Number of unigenes	% Annotated unigenes
Annotated in NR	54658	16.63
Annotated in NT	224390	68.29
Annotated in KO	36283	11.04
Annotated in SwissProt	67656	20.59
Annotated in PFAM	70833	21.55
Annotated in GO	70833	21.55
Annotated in KOG	24944	7.59
Annotated in all Databases	12570	3.82
Annotated in at least one Database	243133	73.99
Total Unigenes	328576	100

Table 2 The candidate pathway and genes related to spine and hair development and differentiation

ID	Pathway	Candidate genes
ko04110	Cell cycle	ABL1, APC10, APC11, APC3, CCNB1, CCND2, CCN3, CCNH, CDC14, CDC20, CDC7, CDH1, CDK1, CDK6, CUL1, E2F3, SKP1, SMAD2_4, SMC1, SMC3, STAG1, TFDP1, TGFB2, TGFB3, YWHAB_Q_Z, YWHAE
ko04010	MAPK	AKT, ANGPT1, ARRB, BRAF, CACNA1B, CACNA1C, CACNA1H, CACNA2D, CACNG5_7, CDC42, DRK, DUSP, EFNA, EGFR, MAPK1_3, FGF, FGFR3, FLNA, GRB2, HSPA1S, HSPB1, JUN, JUND, KRAS, LAMTOR3, MAP2K1, MAP2K7, MAX, MKNK, NF1, NFKB1, NGFR, NLK, NTF3, PAK1, PDGFC, PPM1A, PPP3C, PRKCA, PAC, RAP1A, RRAS, SRF, TAO, TGFB2, TGFB3, TGFBR1, VEGFB, ZAK
ko04350	TGF-beta	SKP1, ACVR1, BMP5, BMP8, BMPR1B, CUL1, DCN, EP300, ERK, FST, GDF5, ID4, MADHIP, PP2C, RGMB, RHOA, RPS6KB, SKP1, SMAD2_4, SMAD9, SMURF, TFDP1, TGFB2, TGFB3, TGFBR1
ko04350	Hippo	ACTB_G1, AMOT, BMP5, BMP8, BMPR1B, CCN2, CRB, CSNK1E, CTTNA, CTTNB1, DLG1, DVL, FZD, GDF5, GSK3B, LAST1, LEF1, LLGL, MOB1, PARD3, SCRIB, SMAD2, SNAI2, SOX2, STK3, TCF1, TEAD, TGFB2, TGFB3, WNT16, WNT3, WNT4, WNT5, YWHAB_Q_Z, WWC1, YAP1, YWHAE
ko04310	Wnt	CSNK2A, CSNK2B, CTNNB1, DVL, FZD4, CD344, GSK3B, PLCB, PPP3C, PPKCA, TCF7
ko04210	Apoptosis	ACTB_G1, AKT, BIRC5, CTSL, CTSB, CTSK, CTSS, CYC, GZMB, LMNA, NGFB, PARP, TUBA
ko04115	p53	CCND3, CCNB, CYC, GFBP3, CCNG1, TP53I3, IGFBP3, RRM2
ko04530	Tight junction	ACTB_G1, AKT, AMOTL1, RRAS, CLDN, CSDA, CTNNB1, MYH, MYL10, PPP2R2, PRKCZ, RRAS2

Figures

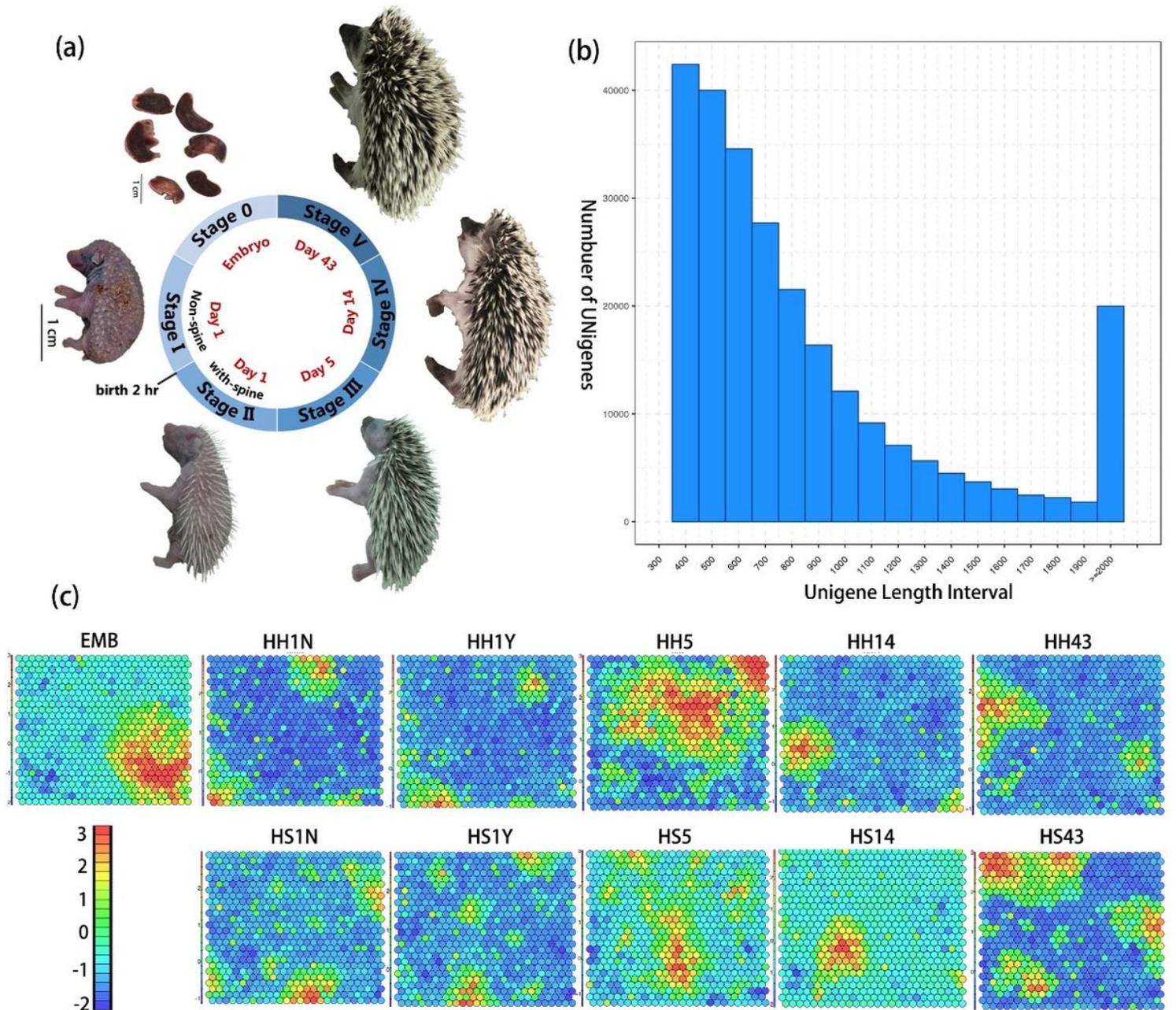


Figure 2

Aterix albiventris skin developmental life cycle transcriptome. (a) Different phenotype of 6 different developmental stages of *A. albiventris* in this study. Stage 0 = Embryo; Stage I = Day 1, < 2 hours; Stage II = Day 1, > 2 hours; Stage III = Day 5; Stage IV = Day 14; Stage V = Day 43; (b) Length distribution of unigenes; (c) Gene expression-specific and phenotype-specific gene-trait correlation analysis based on SOM module analysis.

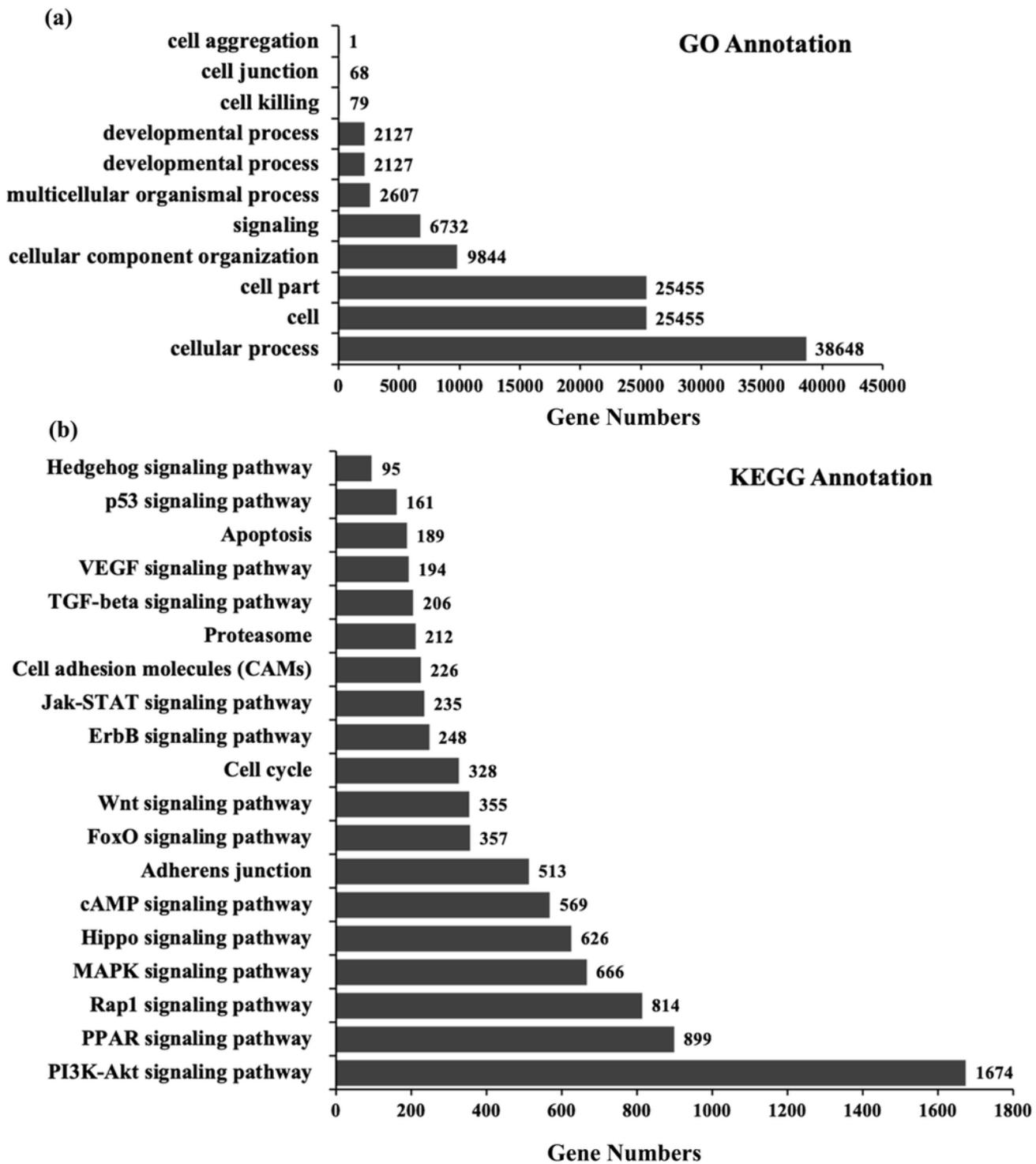


Figure 4

GO and KEGG categories of 36,283 unigenes involved in the skin appendages. (a) Terms related to skin appendage development in Cell development process (GO annotation); (b) categories related to cell development pathway (KEGG annotation).

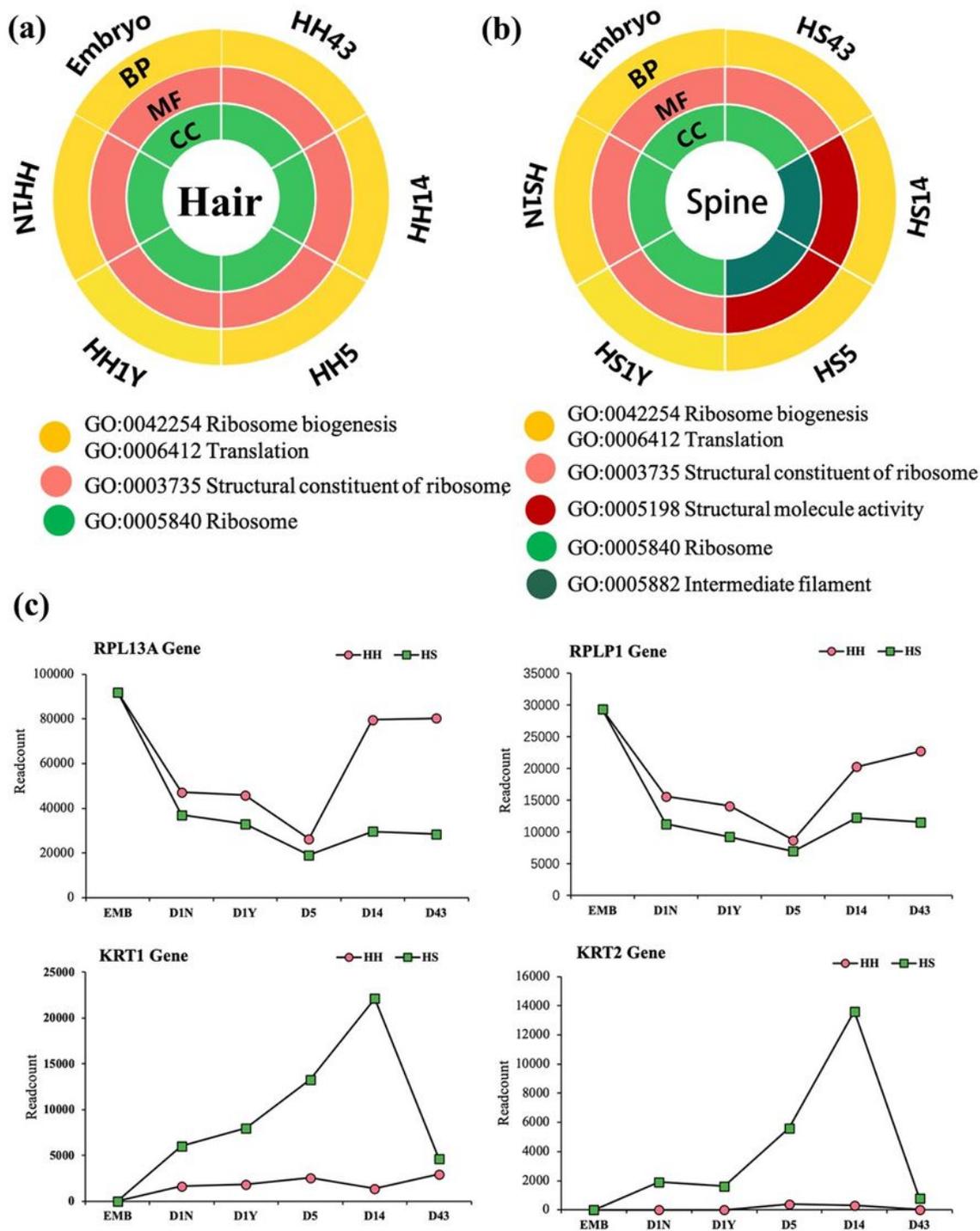


Figure 6

Core top 100 transcripts GO enrichment analysis based on FPKM value, and the top 3 regulated genes in each developmental stage. (a) Hair-type group; (b) Spine-type group; (c) Dynamic changes of FPKM value of RPL13A, RPLP1, KRT1 and KRT2 gene at different development stages.

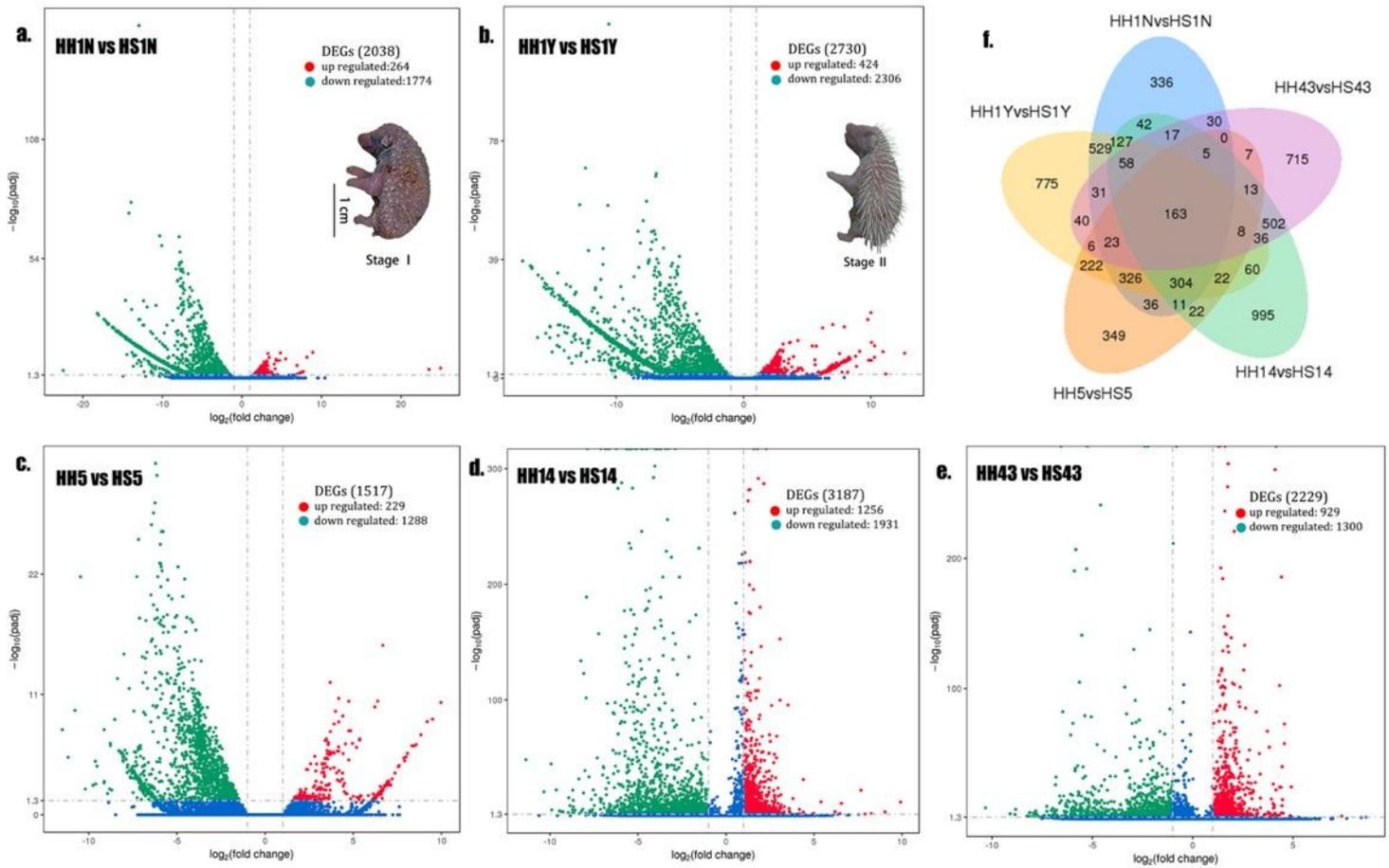


Figure 8

Summary of the differentially expressed genes (DEGs) between the hair and spine appendage in hedgehog skin. MA plot showing the mean (x-axis) and log2 ratios (y-axis) of differential expression (adjusted p-value less than the significance level) during different kinds of appendages, based on normalized RNA-seq read counts. Each data point represents a transcript, and those indicating differential expression are colored in red and green. Gray dashed lines indicate thresholds specified for the fold change.

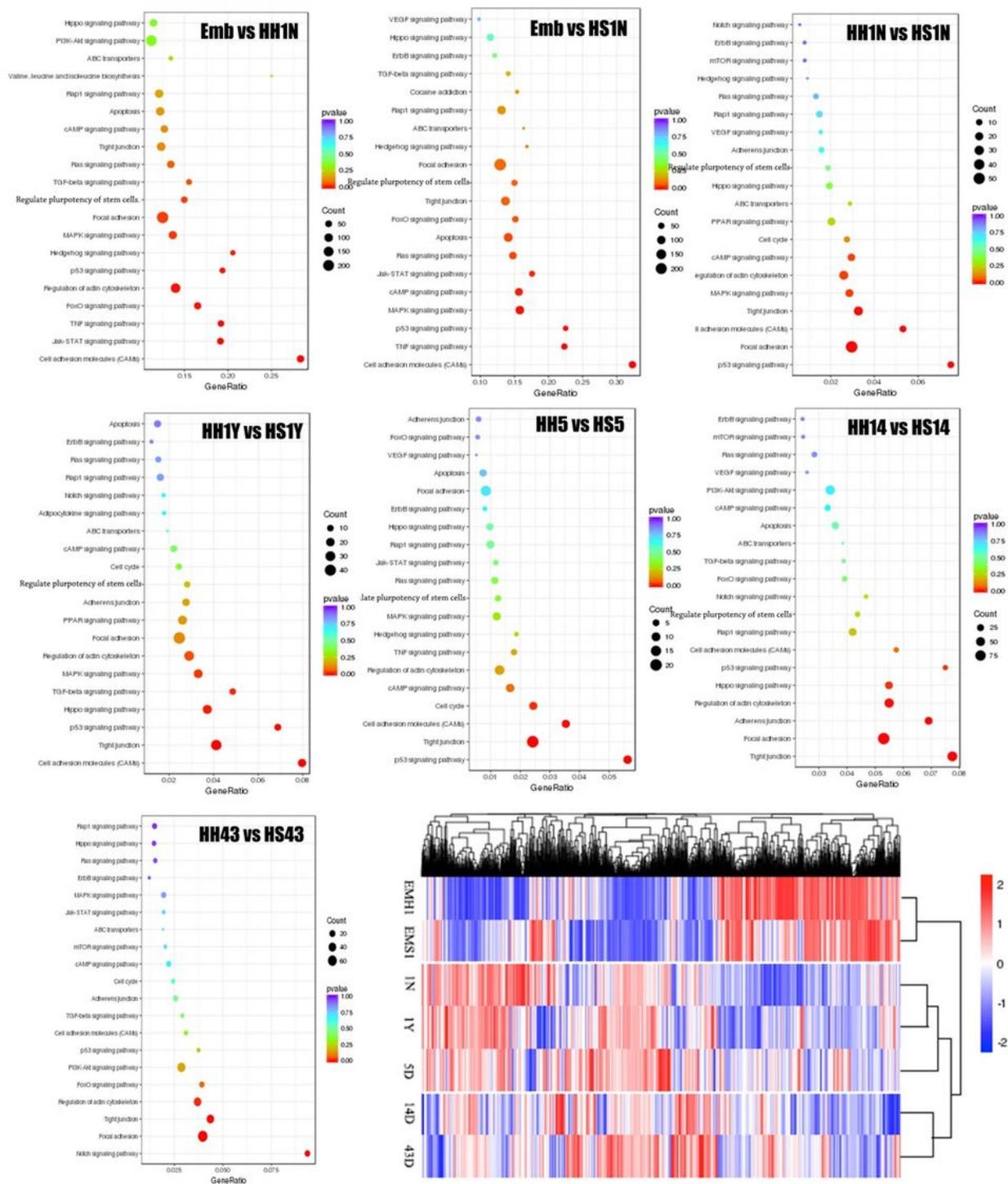


Figure 10

KEGG enrichment analysis based on DEGs of each development stage, and the cluster analysis of heatmap based on 15,158 DEGs

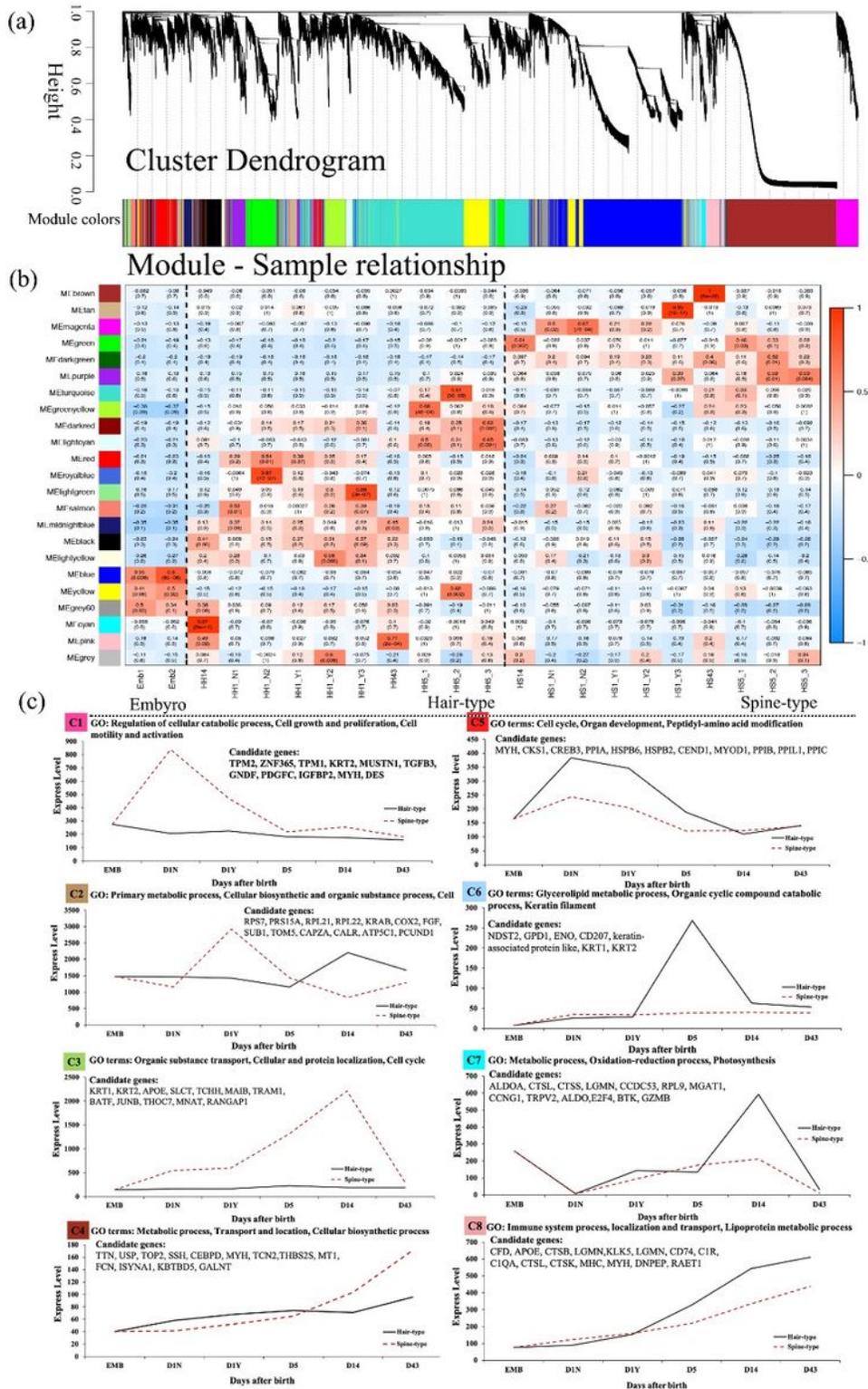


Figure 12

WGCNA of skin appendages samples in hair- and spine-type identified gene modules underlying phenotypic characteristics at different development time. (a) Hierarchical cluster dendrogram of two type skin appendages showing co-expression modules identified using WGCNA on the 15,158 DEGs. Modules corresponding to branches are labeled with colors indicated by the color bands underneath the tree. (b) Heatmap of relevance between samples and modules. The numerical value of each lattice represents the

correlation between the module and the sample. Numbers in parentheses represent significant P value. (c) Top 3 GO terms associated with top 4 modules related to spine phenotype, averaged expression levels of top 100 genes in the corresponding module at each development time point for spine-type samples, and gene symbols of candidate gene within the 4 modules were shown.

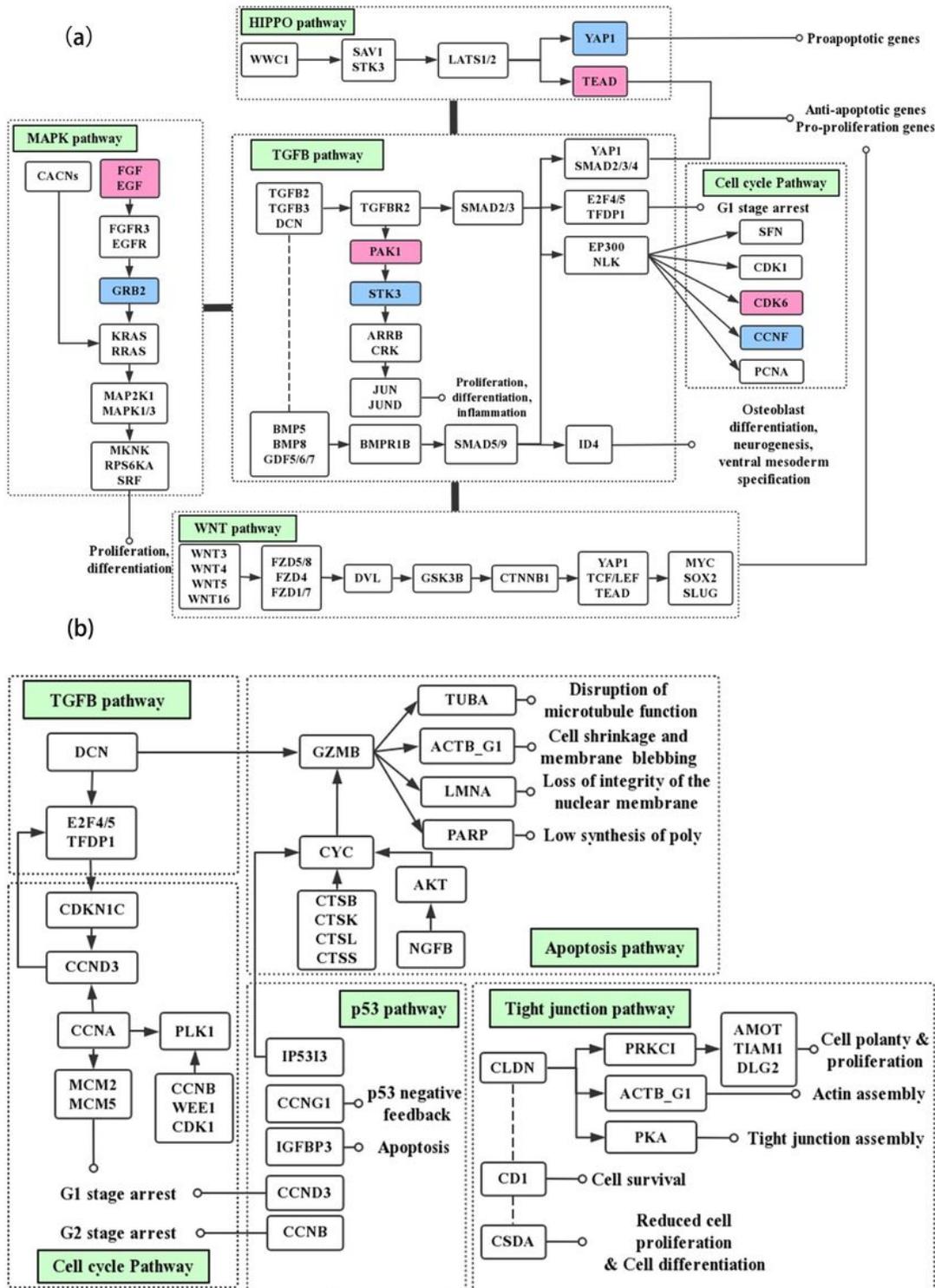


Figure 14

The main regulatory signaling pathways and key genes for the development of hedgehog skin appendages. (a) the key regulatory pathways and genes for spine development; (b) the key regulatory pathways and genes for hair development.

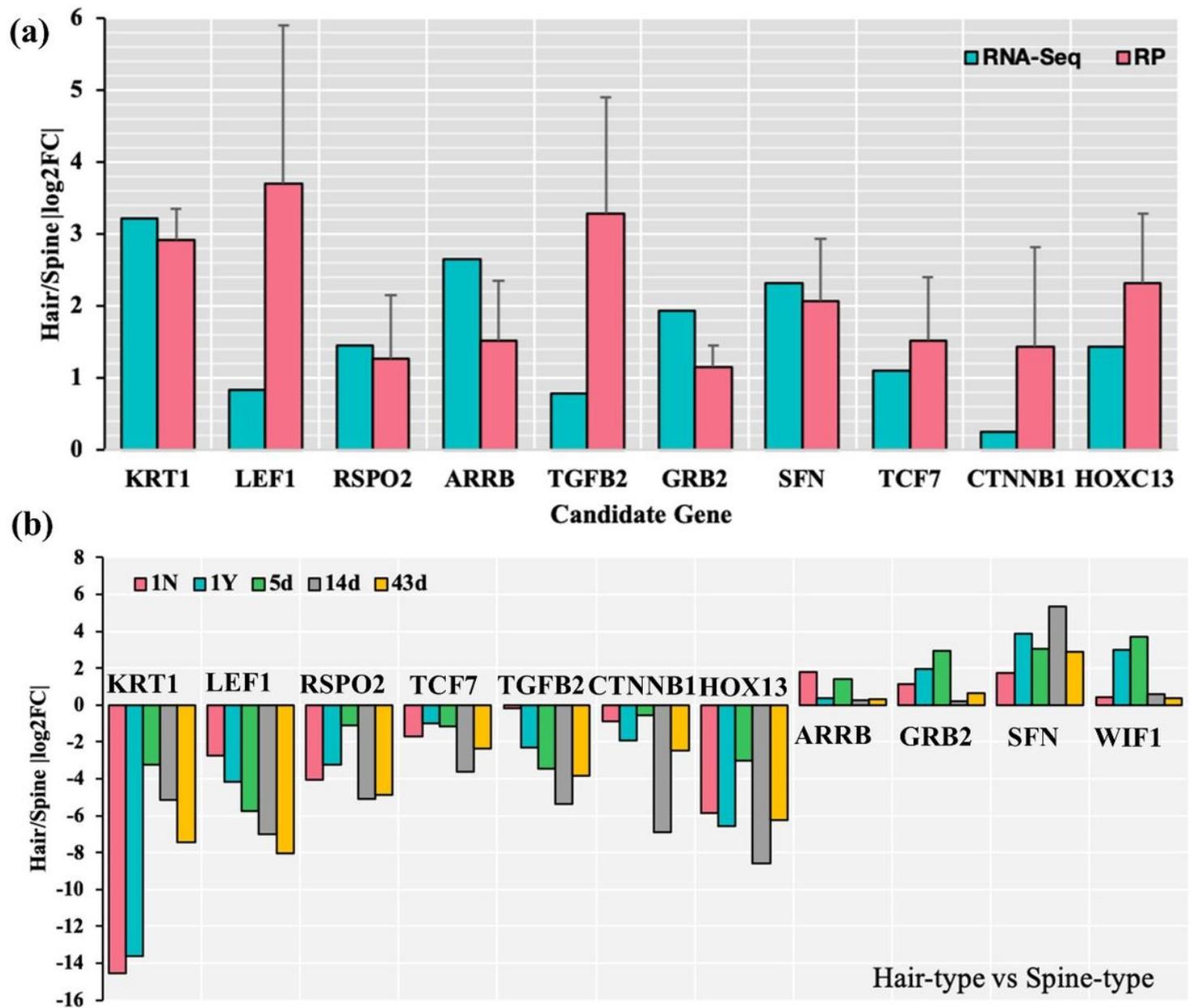


Figure 16

Quantitative real-time PCR analysis. (a) qPCR confirmation of 10 DEGs identified by RNA-seq with stage III tissues; (b) Relative expression levels with 10 DEGs in 5 stages tissues. Bar represent the relative expression levels of unigenes in stage III normalized with respect to the internal control GAPDH. Error bars represent standard error of three biological replicates.

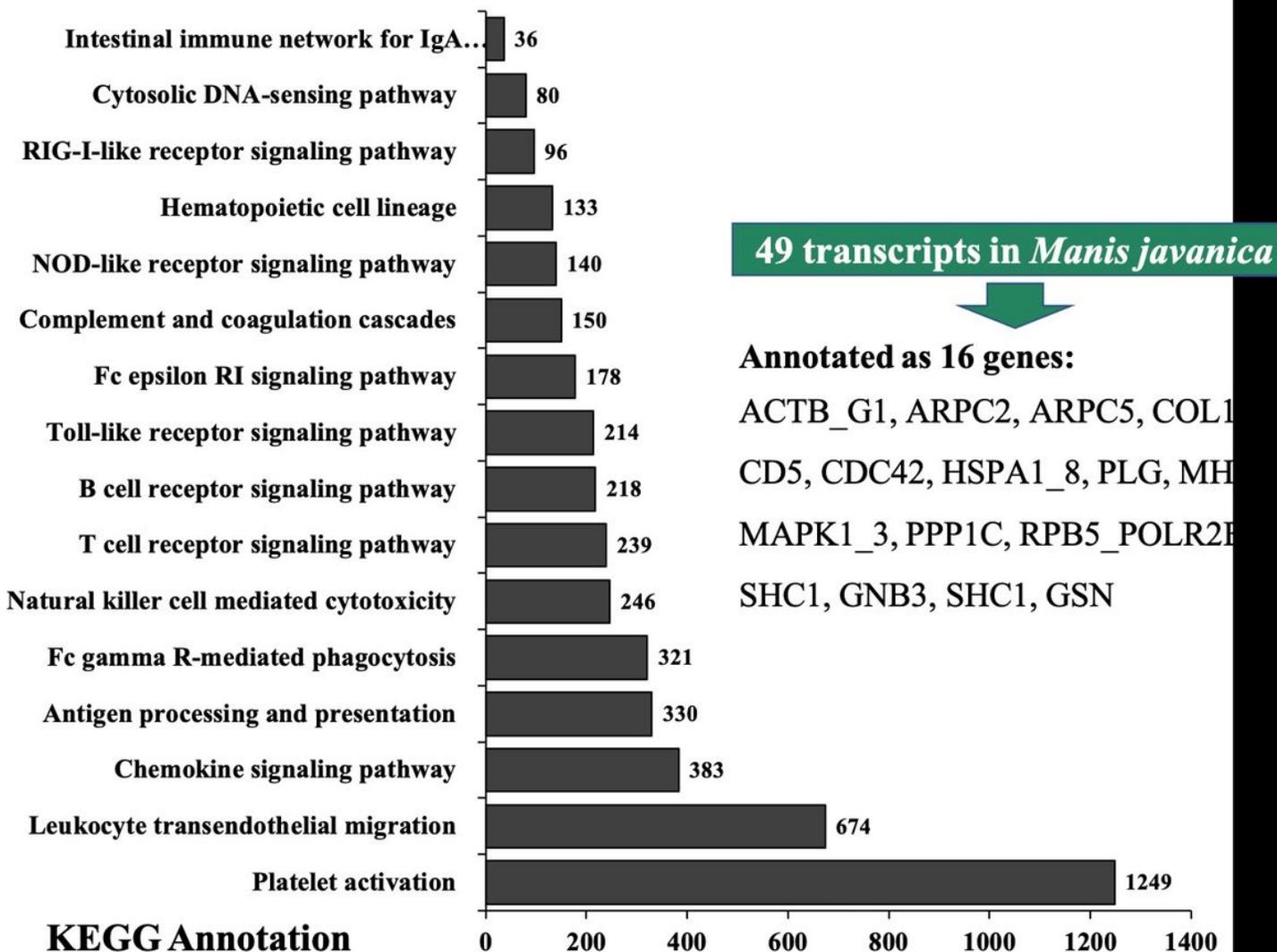


Figure 18

KEGG categories of 4,687 unigenes involved in immunity, of which 16 unigenes have been annotated to pangolin.

Supplementary Files

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

- [SUPPLEMENTARYTABLE16.docx](#)
- [SUPPLEMENTARYTABLE16.docx](#)