

OPEN

Transcriptome profiling reveals transcriptional and alternative splicing regulation in the early embryonic development of hair follicles in the cashmere goat

YanJun Zhang^{1,7}, Lele Wang^{2,7}, Zhen Li^{3,7}, Dong Chen³, Wenjing Han¹, Zhihong Wu¹, Fangzheng Shang¹, Erhan Hai¹, Yaxun Wei³, Rui Su¹, Zhihong Liu¹, Ruijun Wang¹, Zhiying Wang¹, Yanhong Zhao¹, Zhixin Wang¹, Yi Zhang^{3*} & Jinquan Li^{1,4,5,6*}

The undercoat fiber of the cashmere goat, from the secondary hair follicle (HF), possesses commercial value. However, very few studies have focused on the molecular details of primary and secondary HF initiation and development in goat embryos. In this study, skin samples at embryonic day 45, 55, and 65 (E45, E55, and E65) were collected and prepared for RNA sequencing (RNA-seq). We found that the HF probably initiated from E55 to E65 by analyzing the functional pathways of differentially expressed genes (DEGs). Most key genes in canonical signaling pathways, including WNT, TGF- β , FGF, Hedgehog, NOTCH, and other factors showed clear expression changes from E55 to E65. We, for the first time, explored alternative splicing (AS) alterations, which showed distinct patterns among these three stages. Functional pathways of AS-regulated genes showed connections to HF development. By comparing the published RNA-seq samples from the E60, E120, and newborn (NB) stages, we found the majority of WNT/ β -catenin signaling genes were important in the initiation of HF development, while other factors including *FOXN1*, *GATA3*, and *DLX3* may have a consistent influence on HF development. Our investigation supported the time points of embryonic HF initiation and identified genes that have potential functions of embryonic HF initiation and development. We further explored the potential regulatory roles of AS in HF initiation, which extended our knowledge about the molecular mechanisms of HF development.

The cashmere goat is well-known for producing high quality cashmere wool. The Inner Mongolia cashmere goat is a traditional cashmere goat breed in China. Previous research has shown that the fleece of cashmere goats consists of two distinct fibers: guard hair and cashmere undercoat, which are produced by primary hair follicles (PHF) and secondary hair follicles (SHF), respectively¹. The undercoat is fine and soft, while guard hair is straighter and much coarser than the undercoat. The undercoat of the cashmere goat is the source of cashmere products such as pashmina or sweaters, implying its high commercial value¹⁻⁴. Hair follicle (HF) organogenesis is generated during fetal skin development and relies on ectodermal-mesodermal interactions^{5,6}. During cashmere goat embryonic development, the hair placodes of PHF are initiated at 45–55 days gestation. The formation of SHF is initiated at 65–75 days gestation, and all PHFs and some SHFs mature at 135 days gestation⁷. These two

¹College of Animal Science, Inner Mongolia Agricultural University, 010018, Hohhot, Inner Mongolia Autonomous Region, China. ²Ulanqab Medical College, 010020, Ulanqab, Inner Mongolia Autonomous Region, China. ³Center for Genome Analysis, ABLife Inc., Wuhan, Hubei, 430072, China. ⁴Key Laboratory of Mutton Sheep Genetics and Breeding, Ministry of Agriculture, 010018, Hohhot, Inner Mongolia Autonomous Region, China. ⁵Key Laboratory of Animal Genetics, Breeding and Reproduction in Inner Mongolia Autonomous Region, 010018, Hohhot, Inner Mongolia Autonomous Region, China. ⁶Engineering Research Center for Goat Genetics and Breeding, Inner Mongolia Autonomous Region, 010018, Hohhot, Inner Mongolia Autonomous Region, China. ⁷These authors contributed equally: Yanjun Zhang, Lele Wang and Zhen Li. *email: yizhang@ablife.cc; lijinquan_nd@126.com

kinds of HFs present different structures⁷. As a consequence, we propose that understanding the molecular features and underlying regulatory mechanisms that define the embryonic precursors in nascent cashmere follicles may help produce high yielding cashmere goats.

There have been several studies focusing on the critical molecular mechanisms of HF development in the past decades^{8–10}. In the past 20 years, many studies have focused on the molecular mechanisms of HF development and cycling control in humans and mice^{5,11–13}. Increasing evidence shows that HF transformation during cycling is caused by alterations in the local signaling milieu^{5,6,13,14}. Previous studies in humans and mice also demonstrated that several pathways like WNT/beta-catenin, TGF-beta/BMP, FGF, Hedgehog, NOTCH, and other factors (SOX9, LHX2, DLX3, GATA3, RUNX1, TP63, MSX2, FOXN1, NEATC1, and TBX1) play important roles in HF development⁸. However, among these studies on HFs, few have been carried out in cashmere goats, whose skin is comprised of two types of HF, and studies are also lacking on the early stages of HF development^{15,16}. The molecular mechanisms of PHF initiation that occur at early stages in embryonic development, and the transcriptional gene expression and alternative splicing related to HF morphogenesis in cashmere goats, still remain to be elucidated. A better understanding of the biological characteristics and regulation of HF morphogenesis may provide approaches to enhance the formation of fleece and achieve higher fleece production.

Nowadays, high-throughput RNA sequencing (RNA-seq) allows a rapid and comprehensive examination of the transcribed genomic elements in the genome and helps researchers discover novel protein-coding and non-coding transcripts, showing more powerful capabilities compared with traditional methods based on public sequence information¹⁷. Several studies have been performed to investigate transcriptome profiling changes during HF development using RNA-seq technology^{18–21}. Seasonal HF cycling differences between cashmere and milk goats showed several key factors for HF cycling²². Embryonic samples taken from day 60 (E60), 120 (E120), and newborn (NB), which correspond to SHF initiation, PHF, and SHF maturation, respectively, were sequenced to explore HF development in cashmere goat embryos²³. Many publications have explored fetal HF development and individual cycling in mice^{24–26}, while the molecular mechanisms of HF initiation and development in cashmere goats needs to be further illuminated.

In order to enhance our understanding of the process of cashmere goat fetal HF initiation, we used RNA-seq and bioinformatics analysis to study the PHF and SHF development in cashmere goat embryos. Skin epidermis samples were taken at early stages of cashmere goat fetal HF development. We assayed the gene expression profiles at these initial stages of HF morphogenesis and explored the potential molecular mechanisms by which signaling cascades in embryonic skin epidermis regulate HF formation. Meanwhile, the impact of alternative splicing (AS) on HF development was investigated for the first time. We found that the AS profile was distinct among the three different embryonic stages, implying potential regulatory roles in HF development.

Results

Transcriptome profiling of gene expression dynamics during early hair follicle development. To determine the transcriptional profiles at the initial stages of embryonic HF development in cashmere goats, we selected skin samples at E45, E55, and E65. According to our previous observation, the formation PHF and SHF initiate at E55–E65 and E65–E75, respectively⁷. A total of nine RNA-seq libraries (three biological replicates at each stage) were constructed. We obtained 66,882,470 to 95,111,216 raw reads for each sequencing sample (Table S1). After aligning the quality filtered reads to the goat (*Capra hircus*) genome²⁷, 64.19% to 81.8% of the total reads from each sample were successfully aligned. Between 39,319,085 and 71,694,863 reads per sample were uniquely aligned to the genome. Overall, 22,987 of the 28,921 annotated genes (79.48%) were detected in these samples. We used RNASeqPowerCalculator²⁸ to calculate the statistical power (Fig. S1) and other parameters (Table S2) for our RNA-seq data.

Principal component analysis (PCA) was performed to cluster the nine samples from the three stages based on gene expression level. PC1 showed that the distance between the E45 and E55 samples is less than that between E55 and E65. The three E65 samples were close to each other but distinct from the E45 and E55 samples. However, the E45 and E55 samples were more discrete than the E65 samples (Fig. 1A). Sample correlation analysis of the nine samples confirmed the results from PCA (Fig. 1B). To explore the transcriptional differences among these stages, we identified a total of 4,962 differentially expressed genes (DEGs) using the edgeR package²⁹ with a false discovery rate (FDR) < 0.05 and $|\log_2 \text{fold change}| \geq 1$. Among them, 1,083, 1,783, and 3,119 were downregulated and 321, 670, and 1,182 DEGs were upregulated in the E55 vs. E45, E65 vs. E55, and E65 vs. E45 groups, respectively (Fig. 1C). These results showed that more downregulated genes were detected than upregulated genes at the initial stages of HF development, and the transcriptional difference between E55 and E65, during which time PHF is initiated, was more pronounced than that between E45 and E55.

Transcriptional activation of genes in a number of HF developmental pathways occur before and during PHF formation.

To further understand how the functions of DEGs in E65 vs. E55 and E55 vs. E45 are related to embryonic HF development, we performed Reactome analysis. For the upregulated DEGs in E55 vs. E45, the top 10 enriched pathways contained well-known pathways involved in HF development, including cornified envelope, keratinization, apoptotic cleavage of cell adhesion proteins³⁰, collagen degradation³¹, degradation of the extracellular matrix³², and pathways related to TP53³³ (Fig. 1D). Two muscle contraction functional terms were also included. These results probably suggested that the gene expression patterns were shifted to support HF development from E45 to E55 when the epidermal structure developed and the phenotypic appearance of the HF structure emerged. For the upregulated DEGs in E65 vs. E55, the top 10 enriched pathways revealed the presence of two HF pathways (Fig. 1E), including formation of the cornified envelope and keratinization, which were also activated during development from E45 to E55. This finding is consistent with the hypothesis that genes in these functional pathways were further activated during PHF formation. Additionally, a number of new functional pathways regulating HF formation were also activated, including the negative regulation of the TFAP2 (AP-2) family of transcription factors²⁴, WNT ligand biogenesis and trafficking⁸, GLI protein binding to

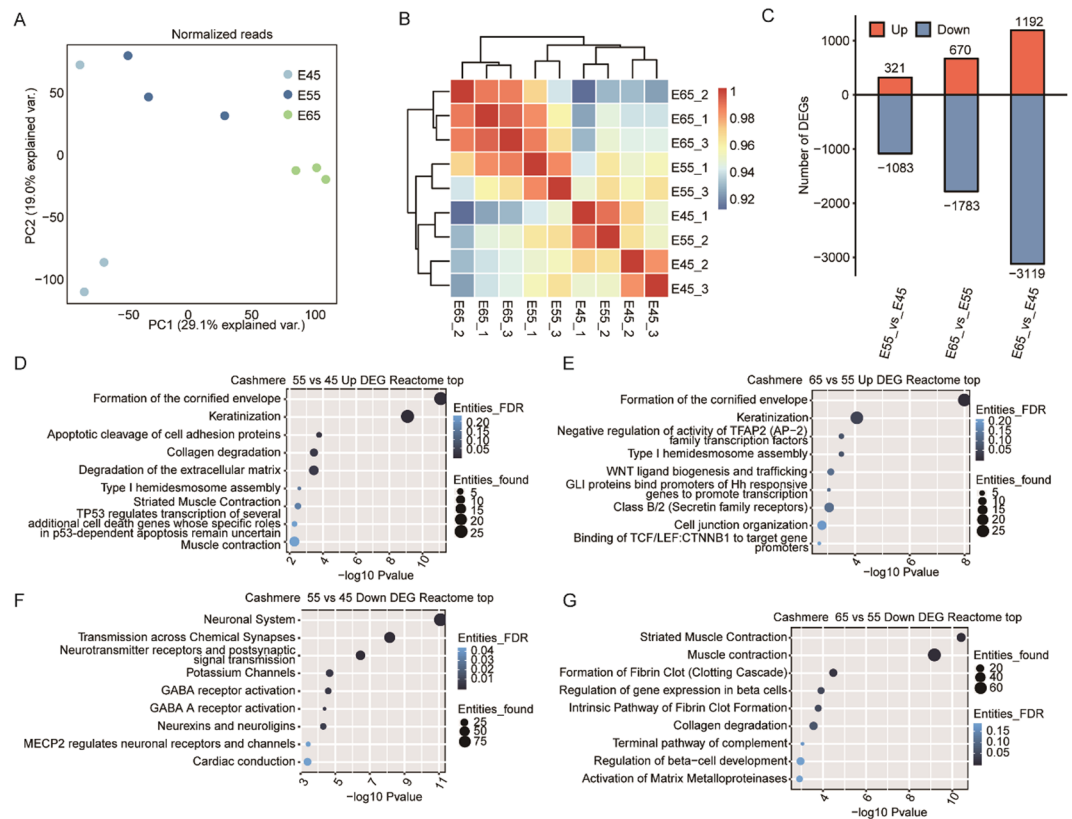


Figure 1. The global transcriptional profiles and the enriched functional analysis of DEGs. **(A)** Principal component analysis (PCA) of nine samples based on normalized mRNA expression level. The samples were grouped by different embryonic days. **(B)** Heat map showing the clustering of the nine embryonic samples by calculating the Pearson's correlation coefficients using expression values. **(C)** Bar plot showing the number of differentially expressed genes in three comparison groups (E55 vs. E45, E65 vs. E55, E65 vs. E45). Red bar represents the number of upregulated genes while blue represents the number of downregulated genes. **(D–G)** Bubble plot showing the top enriched Reactome pathways of the DEGs.

promoters of Hh responsive genes to promote transcription³⁴, and binding of TCF/LEF:CTNNB1 to target gene promoters³⁵ (Fig. 1E).

Among the top 10 enriched pathways among the downregulated DEGs in E55 vs. E45, a number of neuronal development related pathways were identified, including neuronal system, neurexins³⁶, and neuroligins³⁷, GABA receptor activation and GABA A receptor activation³⁸, and pathways of MECP2³⁹ (Fig. 1F). Muscle contraction related pathways were enriched in the downregulated DEGs in the E65 vs. E55 group (Fig. 1G). We also performed gene ontology (GO) enrichment analysis to explore the enriched functional terms for the DEGs (Fig. S2). Skin barrier and HF development related terms were also enriched in the upregulated DEG sets (Fig. S2).

To further explore the expression dynamics of the DEGs during the initial stages of HF development, we performed hierarchical clustering analysis for all identified DEGs. Four distinct expression clusters were identified, including one E65 up-regulated and three E45/E55 up-regulated gene clusters (Fig. 2A). We then performed Reactome analysis for the DEGs from these four clusters. The activated genes at E65 were highly enriched for HF development related terms, including formation of the cornified envelope, keratinization, degradation of the extracellular matrix³², extracellular matrix organization³², collagen degradation⁴⁰, and collagen chain trimerization^{40,41} (Fig. 2B). Interestingly, the E45/E55 upregulated gene clusters included a muscle cluster that was enriched for muscle related functional terms (muscle contraction, striated muscle contraction, and smooth muscle contraction), and HF development terms (assembly of collagen fibrils and other multimeric structures⁴⁰ (Fig. 2C). Genes in the muscle cluster were highly expressed in both E45 and E55, and then repressed at E65. Another group of E45/E55 up-regulated genes was classified into a neuronal cluster, including neuronal system, transmission across chemical synapses, neurexins and neuroligins, MECP2 regulates neuronal receptors and channels⁴². This was a large cluster showing a high level of expression, specifically at E45 (Fig. 2D). Taken together, the DEG results provided some candidate genes that perhaps have key functions during HF initiation and development.

In addition, we conducted an RT-qPCR experiment to verify the RNA-seq results. Six DEGs from the RNA-seq analysis results were tested: *SMAD1*, *FGFR2*, *LEF1*, *WNT11*, *FGFR10*, and *LRP6*, which were involved in different signaling pathways discussed in this paper. We observed a significant increase from E55 to E65 for *SMAD1*, *FGFR2*, *LEF1*, and *WNT11*, and a significant decrease for *FGFR10* and *LRP6*, showing a high agreement with the RNA-seq results (Fig. 2E).

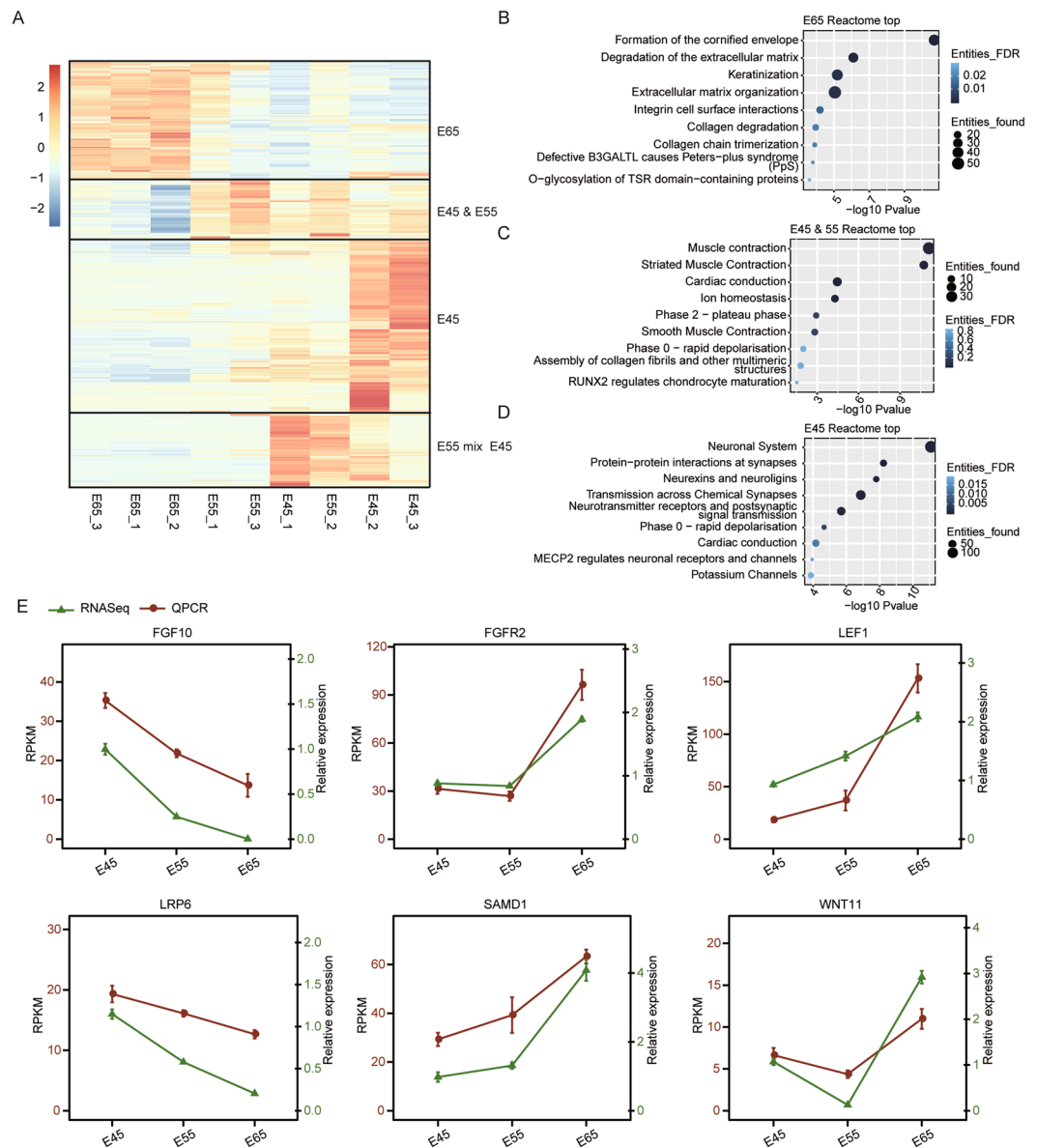


Figure 2. Pathway analysis of four groups clustered by heat map based on E45, E55, and E65 samples. **(A)** Heat map showing the expression pattern of all DEGs from the nine sequenced samples. The DEGs were divided into four clusters according to their distinct expression profiles. **(B–D)** Bubble plot showing the enriched Reactome pathways of the DEGs from three clusters. **(E)** Coupled line plot showing the consistency between RNA-seq and RT-qPCR quantification for the selected DEGs.

The expression dynamics of HF development related genes. Accumulating studies have suggested that canonical signaling cascades, such as WNT/ β -catenin, TGF- β /BMP, FGF, Hedgehog, NOTCH, and other factors (SOX9, LHX2, DLX3, GATA3, RUNX1, TP63, MSX2, FOXN1, NFATC1, and TBX1) play important roles in HF development^{8,18}. We picked these genes from the published literature and related Reactome pathways, and then examined their expression patterns from the RNA-seq data in order to gain an updated view of the molecular regulatory network that governs the initial stage of prenatal HF formation in the goat (Fig. 3). The reported genes from each canonical signaling pathway and genes we obtained from canonical signaling pathways in the E45 vs. E45 and E65 vs. E65 Reactome analysis results were used as query genes. Genes in each signaling pathway can be arbitrarily classified into two groups: increased or decreased expression patterns from E45 to E65. A majority of genes in the WNT/ β -catenin, TGF- β /BMP, and other categories showed increased expression patterns in E65, while genes from the other three signaling pathways showed the opposite pattern (Fig. 3). These results suggested different functional roles for these pathways during the initiation of HF development.

The transcriptional shift between E60 and E120 shows significant similarity between E55–E65. A transcriptome study of the embryonic primary and secondary HF development in cashmere goat has been recently published²³, including samples from cashmere goat embryos at E60 and E120, and newborns

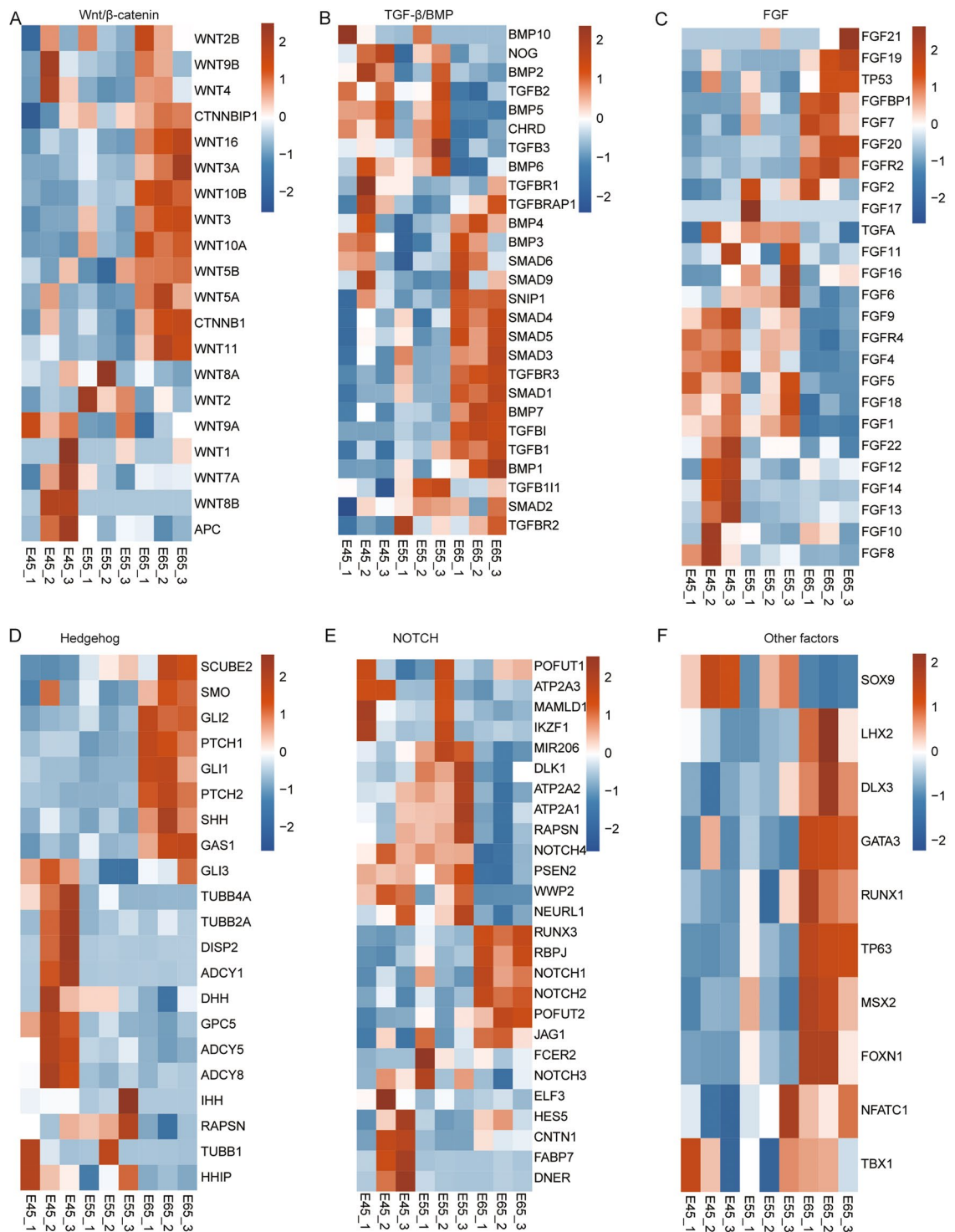


Figure 3. Heat map of key canonical pathway genes in hair follicle development in the E45, E55, and E65 samples (A–F). Genes from defined signaling pathways are developmentally regulated in embryonic skin. Published key genes for specific signaling pathways during HF development and DEGs (E55 vs. E45 and E65 vs. E55) in these canonical signaling pathways were used as query genes. Goat genes involved in WNT/ β -catenin (A), TGF- β /BMP (B), FGF (C), Hedgehog (D), NOTCH (E) signaling, and other factors (F) were represented in the heat map if expressed (RPKM > 0) in embryonic skin samples. In the heat map, FPKM values of each gene were replaced by \log_2 (FPKM), then all values in each row were normalized by a scale factor S so that the sum of the squares of the values in each row was 1.0. Colors ranged from blue to red, corresponding to low to high expression levels.

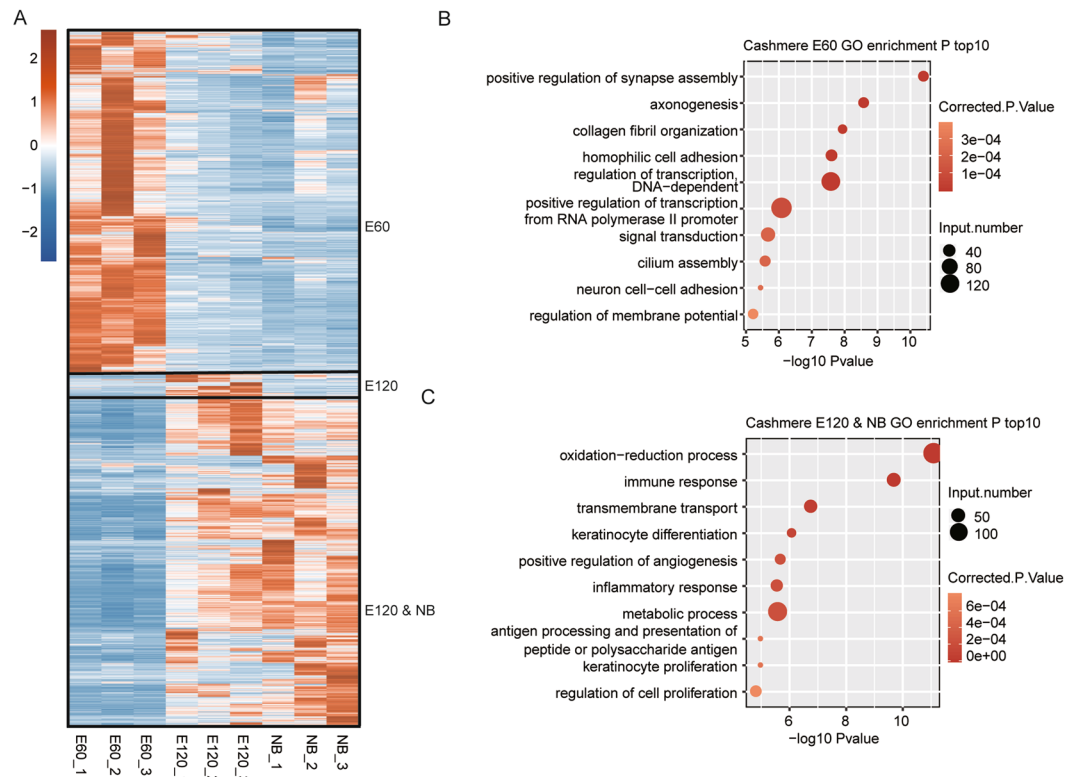


Figure 4. Expression patterns and functional analysis of detected DEGs based on E60, E120, and NB samples. (A) Heat map showing the three clusters by clustering DEGs according to their expression level. (B–C) Bubble plot showing the top 10 enriched GO biological process terms of DEGs in the two dominant groups.

(NB). To comprehensively compare between these two datasets, we reanalyzed the RNA-seq data from that study, showing that the transcriptome profile of the E60 samples was highly distinct from those of the E120 and NB samples (Fig. S3A,B). The above results indicated that the transcriptional activation from E60 to E120 was similar to that between E55 and E65. To validate this hypothesis, we overlapped the DEGs between these two comparison groups and found the overlapped DEGs between these two DEGs were significant (p -value = $4.528515e-94$, hypergeometric test, Fig. S3C), suggesting that thousands of genes were commonly regulated at both the early and late stages of HF development. Heat map clustering of these DEGs revealed two major clusters, the E60 and E120-NB clusters (Fig. 4A). The E120 cluster contained a very small number of genes. To explore the biological functions of genes from these clusters, GO enrichment analysis was performed. This analysis revealed that the genes from the E60 cluster were enriched in collagen fibril organization, cilium assembly, and positive regulation of synapse assembly (Fig. 4B). In contrast, the genes from the E120 and NB clusters were highly enriched in HF development related terms, including keratinocyte differentiation and proliferation, regulation of cell proliferation, transmembrane transport, and oxidation-reduction process (Fig. 4C).

To further explore the similarities and differences between the transcriptional switches during these stages, we plotted the expression dynamics of HF developmental genes (Fig. 5). Different from E45, E55, and E65 related to PHF initiation and development, E60, E120, and NB stages were concerned with PHF development and maturation, SHF initiation, and SHF development and maturation. We found that most genes in the WNT/ β -catenin and NOTCH signaling pathways were repressed during these three stages (Fig. 5A,E), which was contrary to the expression pattern from E45 to E65. The other factors were consistently increased during these three stages (Fig. 5F).

Alternative splicing regulation occurs during early HF development. It is well-known that AS regulation is one of the key post-transcriptional mechanisms that control gene expression and protein diversity⁴³. To explore the potential involvement of AS regulation in HF development, we performed AS analysis for the same samples collected from E45, E55, and E65. Regulated AS events (RASEs) during early HF development were identified (Fig. 6A). The major RASE types included alternative 5' splice site (A5SS), alternative 3' splice site (A3SS) and Cassette Exons. Different from transcriptional regulation, the RASEs from E55 vs E45 were only a little bit similar to those from E65 vs E55 (Fig. 6A).

We next performed heat map clustering analysis of the ratio between the alternative and model splice isoforms of all identified RASEs (Fig. 6B). Different from the dynamics of transcriptional patterns, we identified RASE clusters specific for E45, E55, and E65. Many RASEs were specific either for E55 or E45, suggesting that AS may play potential regulatory roles in HF initiation (Fig. 6B). GO functional enrichment analysis revealed that functional pathways related to protein modifications and mRNA transcription and processing emerged in genes

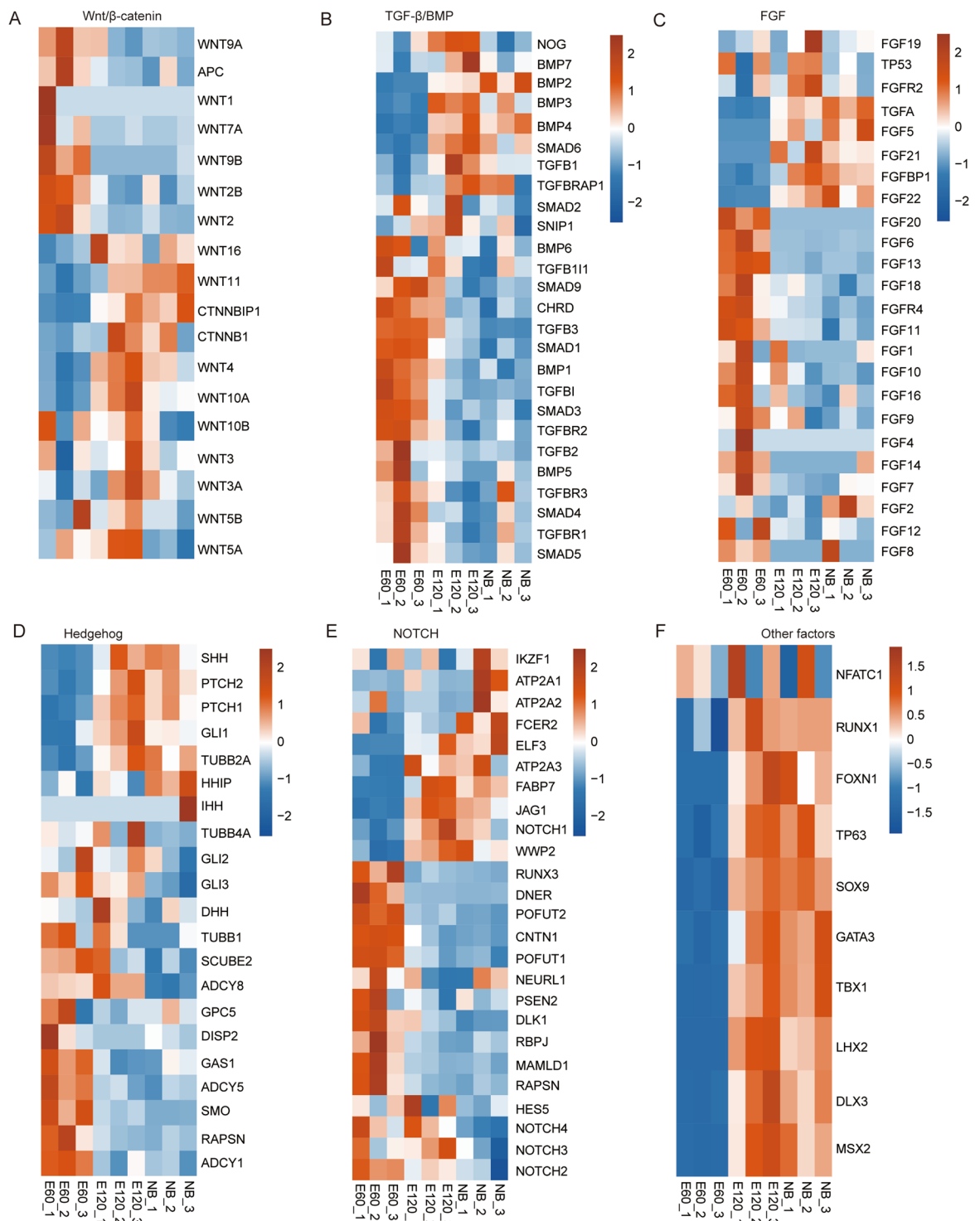


Figure 5. Heat map of known key genes from canonical pathways in HF development (E60, E120, and NB samples). Expressed goat genes involved in WNT/beta-catenin (A), TGF-beta/BMP (B), FGF (C), Hedgehog (D), NOTCH signaling (E), and other factors (F) are presented in the heat map.

containing RASEs from E55 vs E45 (Fig. 6C). Additionally, genes positively regulating GTPase activity were also identified (Fig. 6C). We also found that positive regulation of GTPase activity was the top pathway in the functional terms of genes containing RASEs from E65 vs E55 (Fig. 6D), suggesting that GTPase activity may be related to HF development.

To illustrate the reliability of the RASE results, we selected three RASEs and corresponding genes to show their read density and statistical significance, including *APC* (exon skipping, Fig. 7A), *POFUT1* (intron retention, Fig. 7B), and *TGFB3* (cassette exons, Fig. 7C). These three genes were among the HF development related

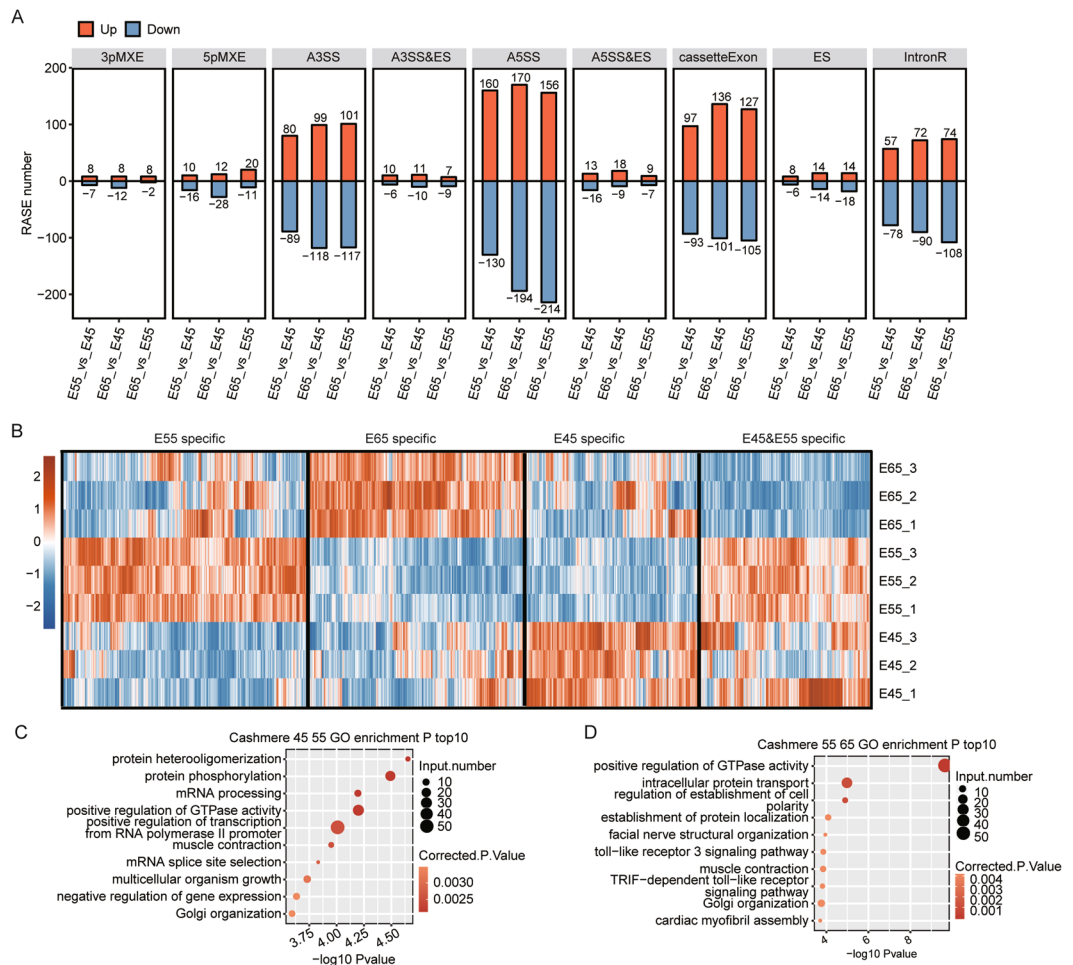


Figure 6. Alternative splicing analysis of E45, E55, and E65 samples. **(A)** Classification of different types of expression changes of alternative splicing events between the three stages. **(B)** Splicing patterns of different samples. A hierarchical clustering dendrogram of different AS events from the three groups. Red: relatively high ratio; blue: relatively low ratio. AS events regulated by different stages in cashmere goat embryos were divided into three groups, and the top 10 GO biological process terms of genes with RASEs in different groups were analyzed **(C–D)**.

genes^{8,44}, which were also regulated by alternative splicing. In summary, these results suggested that posttranscriptional regulation, especially AS, played potential important roles in HF initiation and development.

Discussion

Previous studies on cashmere goat embryonic HF development showed by morphological observation that PHF initiated between E55 and E65⁷. Consistent with previous studies, in our current study, samples from E45 to E55 of cashmere goat showed that most key genes, from essential pathways or important factors that influenced HF development, were upregulated in E65. We also compared the gene expression patterns between E45, E55, and E65 with the published RNA-seq data, including E60, E120, and NB samples²³. By performing AS analysis, we found that the AS patterns of expressed genes were extensively regulated, showing that the AS profile exhibited clear differences between these three stages. Some enriched signaling pathways of AS-regulated genes were also related to HF development, suggesting that AS may participate in the initiation of PHFs.

During cashmere goat embryonic development, the complete structure of the skin epidermis was formed during E45 to E55⁷. At this stage, keratinocytes were evenly distributed in the stratum basale. The onset of PHF development occurred from E55 to E65, which is marked by the formation of a regular array of placodes, and epidermal keratinocytes were clustered in response to signals emanating from the dermis⁷. Consequently, the DEGs in E55 vs. E45 were expected to differ from those in E65 vs. E55. We found that there were 321 upregulated and 1083 downregulated genes (E55 vs. E45), and 670 upregulated and 1783 downregulated genes (E65 vs. E55). This large transcriptional shift gave us a possibility that HF initiation most probably occurred from E55 to E65. The dynamic expression profiles of members of the Wnt/beta-catenin family, the Hedgehog family, members of the TGF-beta/BMP pathway, the Fgf family, the Notch family, and other important factors (SOX9, LHX2, DLX3, GATA3, RUNX1, TP63, MSX2, FOXN1, NFATC1, and TBX1) further supported the above conclusion.

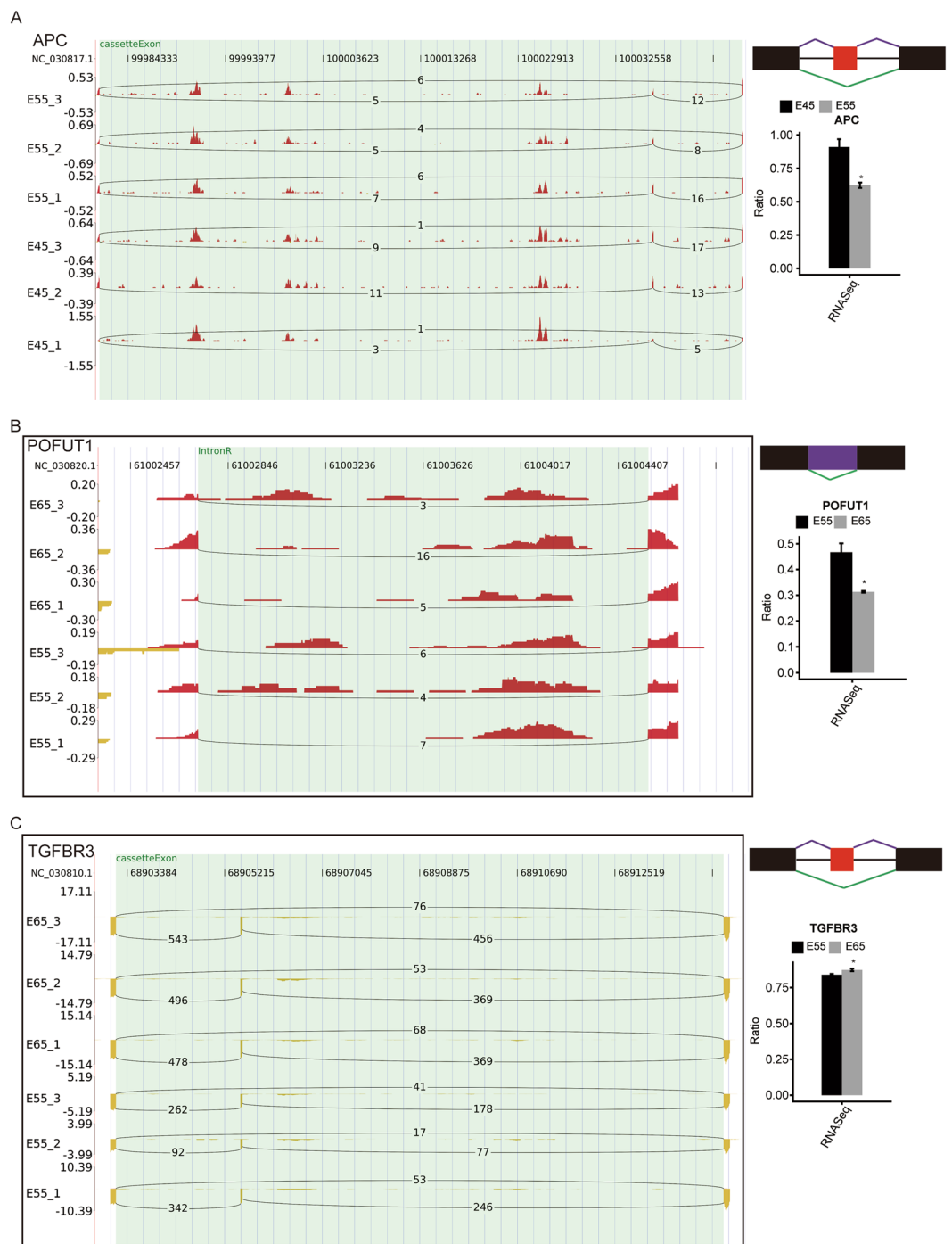


Figure 7. Presentation of HF development-affected RASEs. Three HF development-related genes, *APC* (A), *POFUT1* (B), and *TGFBR3* (C) are shown. IGV-sashimi plots show the alternative splicing changes that occurred in cashmere goat embryonic skin cells in different stages (left panel). The schematic diagrams depict the structures of alternative splicing events (right panel, top), AS1 (shown in purple), and AS2 (shown in green); exon sequences are denoted by boxes, and intron sequences by the horizontal line. RNA-seq quantification and statistical significance of RASEs are shown (right panel, bottom). The altered ratio of RASEs was calculated using the formula: AS1 junction reads/AS1junction reads + AS2 junction reads.

For example, in the Wnt/beta-catenin signaling pathway, several lines of evidence have shown that wnt signaling and beta-catenin, as well as their interactors, are critical for HF induction^{10,25,45–47}. In mice, Wnt10a, Wnt10b, CTNNB1, and Lef1 are highly expressed in hair placode epithelium^{48,49}. Conversely, constitutive activation of CTNNB1 induces precocious placode formation and multiple hair germ (HG) budding out from the original follicle^{25,46,50}. The downstream transcription factors of wnt signaling, including LEF1, TCF3, and TCF4, have been shown to extensively regulate the development of HF^{35,51–53}. From our results, the higher expression of *LEF1* and

TCF3/4 at E65 compared with E45 and E55 suggests that they play a critical role in HF initiation. Twelve genes in the wnt signaling pathway showed an increased expression level, while only seven genes showed a decreased expression level during HF development, confirming their essential roles as master regulators during HF morphogenesis¹⁰. In addition to Wnt/beta-catenin related genes, genes in the TGF-beta/BMPs signaling pathway are also critical for HF morphogenesis, growth, and differentiation^{8,54–56}. Consistent with these previous reports, the expression of TGF-beta/BMPs-related genes showed the highest level at either E55 or E65. *Shh*, *Notch1*, *Notch2*, and *Runx1* exhibited upregulation specifically at E65, supporting the finding that HF initiation might occur from E55 to E65.

Studies on cashmere goat embryonic HF have suggested considerable differences during physiological process at different stages of HF development. In this study, we reanalyzed the RNA-seq data from a previous study²³ to compare the DEGs in E45/E55/E65 with the DEGs in E60/E120/NB (especially E120/NB) to further study the transcriptome dynamics from the early stage to the late stage of cashmere goat HF development. The early stage of HF development was only associated with PHF initiation while the later stage was associated with both PHF and SHF development⁷. Supporting the previous discovery, the global gene expression profile of E55 was different from that of E45 but similar with that of E65. From published data, the profile of E60 was different from that of both E120 and NB, while E120 and NB had similar gene expression profiles. The pathway analysis results of upregulated genes in specific stages were consistent with previous studies. Enriched functional pathways of genes in E60 and E65 were related to HF development. Pathways of genes upregulated in E60 and E65 were also different from those in E45, E120, and NB.

We also analyzed the expression patterns of key genes from canonical pathways in published RNA-seq data to explore their different functions between early and late stages of HF development. We took the Wnt/beta-catenin signaling pathway as an example, and saw that the expression profiles of *Wnt7a*, *APC*, *Wnt10a*, and *FZD8* showed similar patterns in the E45, E55, and E65 groups and E60, E120, and NB groups, while other genes showed different expression profiles in these two groups. From our results, we propose that these canonical pathways play different roles in the initiation or maturation of HF development. High expression of most genes from the WNT/ β -catenin and NOTCH signaling pathways in the E65 and E60 samples implied their important regulatory role in HF initiation. Other factors that showed consistent increased expression levels during the course of embryonic development suggested that they have regulatory roles in both the initiation and development of HF. Other signaling pathways, such as TGF- β /BMP, FGF, and Hedgehog, had varied expression during development. Based on our analysis results, we extended the potential regulatory mechanisms of these signaling pathways during HF development. Further studies are necessary to explore their fundamental molecular mechanisms.

Because different embryonic stages play specific roles in cashmere goat HF development, transcriptional profiles have stage-specific features. This kind of stage-specific feature not only presented in the DEGs, but also in AS. It was clear that the RASEs could be divided into four groups, E55 specific, E65 specific, E45 specific, and E45&E55 specific, suggesting stage-specific features of AS events. Similar with the DEGs, AS-regulated genes between E45 and E55 were enriched for genes involved in the neurexins and neuroligins pathways, which are related to the neuronal system. Meanwhile, SHC-related events triggered by IGF1R and signaling by ERBB2 also emerged. AS-regulated genes in E65 vs E55 were enriched in pathways including cell-extracellular matrix interactions, smooth muscle contraction, and RHO GTPases. Although the functions of enriched pathways of AS genes and DEGs were similar, they may implement their functions through different pathways. From these results, we propose that AS regulation also influenced cashmere goat embryo HF initiation and development through pathways different from DEGs. HF initiation and development may be under the regulation of both transcription and post-transcriptional mechanisms.

In summary, we not only confirmed that the canonical signaling pathways and factors that influence human and mice HF development were highly regulated at the initiation stage of PHF (E45 to E65), but also showed that AS could play potentially important roles in embryonic HF initiation and development. Our study made a further step to understand the dynamic epithelial-mesenchymal interactions during cashmere HF initiation and development.

Materials and Methods

Ethics statement. The animal (including maternal goats and fetal goats) experiments were performed according to the Regulations for the Administration of Affairs Concerning Experimental Animals (Ministry of Science and Technology, China; revised in June 2004) and were approved by the ethics committee of Inner Mongolia Agricultural University. The animals did not experience unnecessary pain or distress at any stage of this experiment.

Animal and skin tissue preparation. The Inner Mongolia cashmere goat is a traditional outstanding breed, which is famous for its excellent cashmere quality and strong adaptation to the semi-desert and desert pastures. The goats were grazed under the same environment with free access to feed and water at the Aerbas White Cashmere Goat Breeding Farm (Inner Mongolia, China). Nine embryos (three samples at each stage) were randomly selected and any lineage was avoided during the sampling process. The skin samples (approximately 1 cm² for each individual) were collected from the mid flank region of embryos at three different embryonic days (E45, E55, E65). Tissue was frozen in liquid nitrogen and stored at -80°C for further experiments.

RNA extraction and RNA sequencing. Total RNA was extracted using TRIzol[®] reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions, and RQ1 DNase (Promega, Madison, WI, USA) was used to remove contaminating genomic DNA. The quality and quantity of the purified RNA were monitored at the ratios of A260/A280 and A260/230 on a SmartSpecPlus Spectrophotometer (BioRad, Philadelphia, PA, USA). RNA integrity was further verified by 1.5% agarose gel electrophoresis.

For RNA-seq, total RNA was extracted from each embryo, and 10 mg of total RNA was used for RNA sequencing (RNA-seq) library preparation. Polyadenylated RNAs were purified and concentrated with Magnetic Beads Oligo (dT) (Invitrogen, Carlsbad, CA, USA) before being used for directional RNA-seq library preparation. Purified RNAs were iron fragmented at 95 °C followed by end repair and 5' adaptor ligation. Then, reverse transcription was performed with RT primer harboring a 3' adaptor sequence and randomized hexamer. The cDNAs were purified and amplified, and PCR products corresponding to 200–500 bps were purified, quantified, and applied to Illumina NextSeq. 500 system151nt pair-end sequencing by ABlife, Inc. (Wuhan, China).

RT-qPCR experiments. To confirm the DEGs revealed by RNA-seq, six genes identified to be differentially expressed among three developmental stages were randomly chosen for reverse transcription qPCR (RT-qPCR) validation. Beta-actin was used as the internal control. The qPCR was carried out on an ABI 7300 system (ABI, USA) using FastStart Universal SYBR Green Master (ROX) according to the manufacturer's instructions. The thermal cycling conditions used in the RT-qPCR reaction were 95 °C for 10 min, followed by 40 cycling of 95 °C for 10 s and 60 °C for 30 s. The specificity of the SYBR green PCR signal was confirmed by melting curve analysis. There were three biological and technical replicates. Primers for RT-qPCR are provided in Table S2.

RNA-seq raw data cleaning and alignment statistics. Raw reads with more than 2-N bases were first discarded. Then, reads were processed by clipping adaptors and removing low quality bases, dropping too short reads (less than 16nt). FASTX-Toolkit (Version 0.0.13) was used to obtain the quality filtered reads. After that, the filtered reads were aligned to the goat (*Capra hircus*) genome²⁷ by TopHat⁵⁷. Based on the gene annotation of the genome, aligned reads with more than one genome location were discarded due to their ambiguous location. Uniquely localized reads were used to calculate read number and fragments per kilobase and per million (FPKM) value for each gene, according to the genomic locations of the aligned reads and genes. Other statistical results, such as gene coverage and depth and read distribution around start codons and stop codons, were also obtained.

Differentially expressed genes (DEGs) analysis. We used RNASeqPowerCalculator²⁸ to calculate the statistical power and other parameters for our RNA-seq data. Each sample was considered as a biological replicate. Differentially expressed genes (DEGs) between the three stages, E45, E55, and E65, were analyzed using the edgeR²⁹ package in R software. For each gene, the false discovery rate (FDR) and log₂ fold change (log₂FC) were obtained based on the negative binomial distribution model. 0.05 FDR and 2-fold change were set as the thresholds to define DEGs.

Alternative splicing analysis. The alternative splicing events (ASEs) and regulated alternative splicing events (RASEs) among these three developmental stages were defined and quantified using the ABLas pipeline as described previously⁵⁸. In brief, detection of seven types of ASEs was based on the splice junction reads. The eight types of ASE included cassette exon (CassetteExon), exon skipping (ES), mutual exclusive exon skipping (MXE), A5SS, A3SS, the MXE combined with alternative 5' promoter (5pMXE), and with alternative polyadenylation site (3pMXE). After detecting the ASEs in each RNA-seq sample, Fisher's exact test was chosen to calculate the statistical *p*-value, with the alternative reads and model reads of samples as input data, respectively. The changed ratio of alternatively spliced reads and constitutively spliced reads between compared samples was defined as the RASE ratio. A *p*-value < 0.05 and RASE ratio > 0.2 were set as the thresholds for RASE definition.

Functional enrichment analysis. Because goat genes lack functional annotation, we aligned the goat protein sequence to the sheep protein database by blastp⁵⁹. The sheep protein with highest score was regarded as the homologous protein of the goat. Meanwhile, all differentially expressed genes were mapped to pathways in gene ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Reactome databases. Fisher's exact test was used to calculate the enrichment of each GO term and KEGG pathway. Reactome pathway analysis was performed online using Reactome version V62 (<http://reactome.org/>)⁶⁰.

Other statistical analysis. Hierarchical clustering was used to calculate the clustering of gene sets using Cluster3.0. In the heat map, the FPKM value of each gene was replaced by log₂ (FPKM), the row-wise median was subtracted from the values in each row, then all values in each row were multiplied by a scale factor *S* so that the sum of the squares of the values in each row was 1.0. Other statistical results were obtained by R software.

Data deposition

The accession number for the RNA-seq data reported in this paper is GEO: GSE76538.

Received: 13 May 2019; Accepted: 31 October 2019;

Published online: 28 November 2019

References

1. Ansari-Renani, H. R. *et al.* Cashmere quality of Raeini goats kept by nomads in Iran. *Small Ruminant Research* **104**, 10–16, <https://doi.org/10.1016/j.smallrumres.2011.11.003> (2012).
2. Nixon, A. J., Gurnsey, M. P., Betteridge, K., Mitchell, R. J. & Welch, R. A. S. Seasonal hair follicle activity and fibre growth in some New Zealand cashmere-bearing goats (*Caprus hircus*). *Proceedings of the Zoological Society of London* **224**, 589–598 (1991).
3. L. Ryder, M. *Coat structure and seasonal shedding in goats*. **8**, (1966).
4. Hardy, M. H. The secret life of the hair follicle. *Trends in Genetics* **8**, 55–61 (1992).
5. Schmidt-Ullrich, R. & Paus, R. Molecular principles of hair follicle induction and morphogenesis. *Bioessays* **27**, 247–261, <https://doi.org/10.1002/bies.20184> (2005).
6. Schneider, M. R., Ruth, S. U. & Ralf, P. The hair follicle as a dynamic miniorgan. *Current Biology Cb* **19**, R132–R142 (2009).
7. Zhang, Y. J., Yin, J., Jin-Quan, L. I. & Chang-Qing, L. I. Study on Hair Follicle Structure and Morphogenesis of the Inner Mongolian Arbas Cashmere Goat. *Scientia Agricultura Sinica* **40**, 1017–1023 (2007).

8. Lee, J. & Tumber, T. Hairy tale of signaling in hair follicle development and cycling. *Seminars in Cell & Developmental Biology* **23**, 906–916, <https://doi.org/10.1016/j.semcdb.2012.08.003> (2012).
9. Paus, R. & Cotsarelis, G. The biology of hair follicles. *N Engl J Med* **341**, 491–497, <https://doi.org/10.1056/NEJM199908123410706> (1999).
10. Rishikaysh, P. *et al.* Signaling involved in hair follicle morphogenesis and development. *Int J Mol Sci* **15**, 1647–1670, <https://doi.org/10.3390/ijms15011647> (2014).
11. Stenn, K. S. & Paus, R. Controls of hair follicle cycling. *Physiological Reviews* **81**, 449, <https://doi.org/10.1152/physrev.2001.81.1.449> (2001).
12. Bot, N. L. Hair follicles run by clockwork. *Nature Cell Biology* **13**, 1394 (2011).
13. Millar, S. E. Molecular mechanisms regulating hair follicle development. *Journal of Investigative Dermatology* **118**, 216–225, <https://doi.org/10.1046/j.0022-202x.2001.01670.x> (2002).
14. Geyfman, M. & Andersen, B. Clock genes, hair growth and aging. *Aging* **2**, 122, <https://doi.org/10.18632/aging.100130> (2010).
15. Su, R. *et al.* Characterization of BMP2 gene expression in embryonic and adult Inner Mongolia Cashmere goat (*Capra hircus*) hair follicles. *Revue Canadienne De Science Animale* **89**, 457–462 (2009).
16. Jiang-Hong, W. U., Zhang, W. G. & Jin-Quan, L. I. Hoxc13 Expression Pattern in Cashmere Goat Skin During Hair Follicle Development. *Agricultural Sciences in China* **8**, 491–496 (2009).
17. Weikard, R., Goldammer, T., Eberlein, A. & Kuehn, C. Novel transcripts discovered by mining genomic DNA from defined regions of bovine chromosome 6. *BMC genomics* **10**, 186–186, <https://doi.org/10.1186/1471-2164-10-186> (2009).
18. Sennett, R. *et al.* An Integrated Transcriptome Atlas of Embryonic Hair Follicle Progenitors, Their Niche, and the Developing Skin. *Developmental Cell* **34**, S153458071500430X (2015).
19. Genander, M. *et al.* BMP signaling and its pSMAD1/5 target genes differentially regulate hair follicle stem cell lineages. *Cell stem cell* **15**, 619–633, <https://doi.org/10.1016/j.stem.2014.09.009> (2014).
20. Zhu, B., Xu, T., Yuan, J., Guo, X. & Liu, D. Transcriptome Sequencing Reveals Differences between Primary and Secondary Hair Follicle-derived Dermal Papilla Cells of the Cashmere Goat (*Capra hircus*). *PLOS ONE* **8**, e76282, <https://doi.org/10.1371/journal.pone.0076282> (2013).
21. Wu, Z. *et al.* Identification of differentially expressed miRNAs between white and black hair follicles by RNA-sequencing in the goat (*Capra hircus*). *International journal of molecular sciences* **15**, 9531–9545, <https://doi.org/10.3390/ijms15069531> (2014).
22. Chen, S. *et al.* The mRNA Export Receptor NXF1 Coordinates Transcriptional Dynamics, Alternative Polyadenylation, and mRNA Export. *Mol Cell* **74**, 118–131 e117, <https://doi.org/10.1016/j.molcel.2019.01.026> (2019).
23. Gao, Y. *et al.* Comparative Transcriptome Analysis of Fetal Skin Reveals Key Genes Related to Hair Follicle Morphogenesis in Cashmere Goats. *Plos One* **11**, e0151118, <https://doi.org/10.1371/journal.pone.0151118> (2016).
24. Panteleyev, A. A., Mitchell, P. J., Paus, R. & Christiano, A. M. Expression patterns of the transcription factor AP-2alpha during hair follicle morphogenesis and cycling. *The Journal of investigative dermatology* **121**, 13–19, <https://doi.org/10.1046/j.1523-1747.2003.12319.x> (2003).
25. Zhang, Y. *et al.* Activation of beta-catenin signaling programs embryonic epidermis to hair follicle fate. *Development* **135**, 2161–2172, <https://doi.org/10.1242/dev.017459> (2008).
26. Rendl, M., Lewis, L. & Fuchs, E. Molecular dissection of mesenchymal-epithelial interactions in the hair follicle. *PLoS Biol* **3**, e331, <https://doi.org/10.1371/journal.pbio.0030331> (2005).
27. Yang, D. *et al.* Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (*Capra hircus*). *Nature Biotechnology* **31**, 135–141 (2013).
28. Ching, T., Huang, S. & Garmire, L. X. Power analysis and sample size estimation for RNA-Seq differential expression. *RNA* **20**, 1684–1696, <https://doi.org/10.1261/rna.046011.114> (2014).
29. McCarthy, D. J., Smyth, G. K. & Robinson, M. D. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)* **26**, 139–140, <https://doi.org/10.1093/bioinformatics/btp1616> (2009).
30. Kaplan, E. D. & Holbrook, K. A. Dynamic expression patterns of tenascin, proteoglycans, and cell adhesion molecules during human hair follicle morphogenesis. *Dev Dyn* **199**, 141–155 (2010).
31. Messenger, A. G., Elliott, K., Temple, A. & Randall, V. A. Expression of basement membrane proteins and interstitial collagens in dermal papillae of human hair follicles. *Journal of Investigative Dermatology* **96**, 93–97, <https://doi.org/10.1111/1523-1747.ep12515907> (1991).
32. Mcdonagh, A. J., Cawood, L. & Messenger, A. G. Expression of extracellular matrix in hair follicle mesenchyme in alopecia areata. *British Journal of Dermatology* **123**, 717–724 (1991).
33. Botchkarev, V. A. & Flores, E. R. p53/p63/p73 in the epidermis in health and disease. *Cold Spring Harbor Perspectives in Medicine* **4**, a015248–a015248, <https://doi.org/10.1101/cshperspect.a015248> (2014).
34. Dahmane, N. *et al.* Activation of the transcription factor Gli1 and the Sonic hedgehog signalling pathway in skin tumours. *Nature* **389**, 876–881, <https://doi.org/10.1038/39918> (1997).
35. Dasgupta, R. & Fuchs, E. Multiple roles for activated LEF/TCF transcription complexes during hair follicle development and differentiation. *Development* **126**, 4557 (1999).
36. Li, X., Zhang, J., Cao, Z., Wu, J. & Shi, Y. Solution structure of GOPC PDZ domain and its interaction with the C-terminal motif of neuroligin. *Protein science: a publication of the Protein Society* **15**, 2149–2158, <https://doi.org/10.1110/ps.062087506> (2006).
37. Fabrichny, I. P. *et al.* Structural analysis of the synaptic protein neuroligin and its beta-neurexin complex: determinants for folding and cell adhesion. *Neuron* **56**, 979–991, <https://doi.org/10.1016/j.neuron.2007.11.013> (2007).
38. Kuffler, S. W. & Edwards, C. Mechanism of gamma aminobutyric acid (GABA) action and its relation to synaptic inhibition. *Journal of Neurophysiology* **21**, 589–610, <https://doi.org/10.1152/jn.1958.21.6.589> (1958).
39. Amir, R. *et al.* Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2 (see comments). *Nature Genetics* **23**, 185–188 (1999).
40. Parakkal, P. F. Role of macrophages in collagen resorption during hair growth cycle. *J Ultrastruct Res* **29**, 210–217 (1969).
41. Boudko, S. P., Jürgen, E. & Hans Peter, B. C. Trimerization and triple helix stabilization of the collagen XIX NC2 domain. *Journal of Biological Chemistry* **283**, 34345, <https://doi.org/10.1074/jbc.M806352200> (2008).
42. Fasolino, M. & Zhou, Z. The Crucial Role of DNA Methylation and MeCP2 in Neuronal Function. *Genes* **8**, 141–, <https://doi.org/10.3390/genes8050141> (2017).
43. Keren, H., Lev-Maor, G. & Ast, G. Alternative splicing and evolution: diversification, exon definition and function. *Nature reviews. Genetics* **11**, 345–355, <https://doi.org/10.1038/nrg2776> (2010).
44. Glover, J. D. *et al.* Hierarchical patterning modes orchestrate hair follicle morphogenesis. *Plos Biology* **15**, e2002117, <https://doi.org/10.1371/journal.pbio.2002117> (2017).
45. Petiot, A. *et al.* A crucial role for Fgfr2-IIIb signalling in epidermal development and hair follicle patterning. *Development* **130**, 5493, <https://doi.org/10.1242/dev.00788> (2003).
46. Huelsken, J., Vogel, R., Erdmann, B., Cotsarelis, G. & Birchmeier, W. β -Catenin Controls Hair Follicle Morphogenesis and Stem Cell Differentiation in the Skin. *Cell* **105**, 533–545 (2001).
47. Andl, T., Reddy, S. T., Gaddapara, T. & Millar, S. E. WNT signals are required for the initiation of hair follicle development. *Dev Cell* **2**, 643–653 (2002).
48. Hendrik, U., Panteleyev, A. A., David, D. B., Tobin, D. T. & Christiano, A. M. Activation of Notch1 in the hair follicle leads to cell-fate switch and Mohawk alopecia. *Differentiation* **72**, 396–409 (2004).

49. Merrill, B. J. *et al.* Tcf3: a transcriptional regulator of axis induction in the early embryo. *Development* **131**, 263–274, <https://doi.org/10.1242/dev.00935> (2004).
50. Silva-Vargas, V. *et al.* beta-catenin and hedgehog signal strength can specify number and location of hair follicles in adult epidermis without recruitment of bulge stem cells. *Developmental Cell* **9**, 121–131 (2005).
51. 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 & Development* **10**, 1382–1394 (1996).
52. Hoang, N. *et al.* Tcf3 and Tcf4 are essential for long-term homeostasis of skin epithelia. *Nature Genetics* **41**, 1068 (2009).
53. Nguyen, H., Rendl, M. & Fuchs, E. Tcf3 Governs Stem Cell Features and Represses Cell Fate Determination in Skin. *Cell* **127**, 171–183, <https://doi.org/10.1016/j.cell.2006.07.036> (2006).
54. Blessing, M., Nanney, L. B., King, L. E., Jones, C. M. & Hogan, B. L. Transgenic mice as a model to study the role of TGF-beta-related molecules in hair follicles. *Genes Dev* **7**, 204–215, <https://doi.org/10.1101/gad.7.2.204> (1993).
55. Botchkarev, V. A. Bone morphogenetic proteins and their antagonists in skin and hair follicle biology. *The Journal of investigative dermatology* **120**, 36–47, <https://doi.org/10.1046/j.1523-1747.2003.12002.x> (2003).
56. Botchkarev, V. A. *et al.* Modulation of BMP signaling by noggin is required for induction of the secondary (nonyltoch) hair follicles. *The Journal of investigative dermatology* **118**, 3–10, <https://doi.org/10.1046/j.1523-1747.2002.01645.x> (2002).
57. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics (Oxford, England)* **25**, 1105–1111, <https://doi.org/10.1093/bioinformatics/btp120> (2009).
58. Xia, H. *et al.* CELF1 preferentially binds to exon-intron boundary and regulates alternative splicing in HeLa cells. *Biochim Biophys Acta* **1860**, 911–921, <https://doi.org/10.1016/j.bbagr.2017.07.004> (2017).
59. Lobo, I. Basic Local Alignment Search Tool (BLAST). *Journal of Molecular Biology* **215**, 403–410 (2012).
60. Croft, D. *et al.* The Reactome pathway knowledgebase. *Nucleic acids research* **42**, D472–477, <https://doi.org/10.1093/nar/gkt1102> (2014).

Acknowledgements

The reported work was supported by the National Natural Science Foundation of China (31860627, 31560619). This work was also supported by Appreciate the Beauty of Life Incorporation (ABLIFE) (201404001), Inner Mongolia Natural Science Foundation of China (2017MS0309), Plan Project of Science and Technology in Inner Mongolia, Program for Young Talents of Science and Technology in Universities of Inner Mongolia Autonomous Region (NJYT-17-A04). We thank members in ABLIFE for their valuable assistance in data processing and analysis.

Author contributions

Yanjun Z., J.L. and Yi Z. conceived and designed the experiments. L.W., Z.L. participated in the sequence alignment and analyzed the data. W.H. performed the validation experiments. Z.W., F.S., E.H., R.S., Z.L., R.W., Z.W., Yanhong Z., Z.W. and Y.W. contribute to figure/dreagents/materials/analysis tools. L.W., Z.L., and D.C. wrote the paper. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-019-54315-7>.

Correspondence and requests for materials should be addressed to Y.Z. or J.L.

Reprints and permissions information is available at www.nature.com/reprints.

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



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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019