

1 Title

2 Comparative transcriptome analysis by deep RNA sequencing at early stage of skin pigmentation
3 in goats (*Capra hircus*)

4 Authors

5 Hangxing Ren^{*,†,1}, Gaofu Wang^{*,†,1}, Jing Jiang^{*,†}, Liangjia Liu^{*,†}, Nianfu Li[‡], Jie Li^{*,†}, Lin Fu^{*,†},
6 Haiyan Zhang[‡], Risu Na[§], Yongfu Huang[§], Li Zhang^{*}, Lei Chen^{*}, Yong Huang^{*,2}, and Peng
7 Zhou^{*,†,2}

8 Affiliation

9 ^{*} Chongqing Academy of Animal Sciences, Chongqing, 402460, China

10 [†] Chongqing Engineering Research Center for Goats, Chongqing, 402460, China

11 [‡] Youyang Animal Husbandry Bureau, Chongqing, 409800, China

12 [§] Department of Animal Science, Southwest University, Chongqing, 400715, China

13 ¹ The authors contribute equally to this work

14 ² Corresponding author

15

16

17 **Running title**

18 Dermal hyperpigmentation in goat

19

20 **Key words**

21 dermal pigmentation; transcriptome; melanin; goat

22

23 **Corresponding author**

24 Yong Huang, Changlong Street 51#, Rongchang District, Chongqing, China 406420; Tel:

25 +86-023-46792355; huangyongcqbb@126.com

26 Peng Zhou, Changlong Street 51#, Rongchang District, Chongqing, China 406420; Tel:

27 +86-023-46792015; cqzp2006@163.com

28

29 **ABSTRACT**

30 Although specific genes have been found to be associated with skin pigmentation, the global
31 gene expression profile for the early stage of skin pigmentation and development in mammals is
32 still not well understood. Here we reported a rare natural group of goat (Youzhou dark goat)
33 featuring the dark skin of body including the visible mucous membranes, which may be an
34 exclusive kind of large mammalian species with this special phenotype so far. In the present study,
35 we characterized the 100-day-old fetal skin transcriptome in hyperpigmented (dark-skinned) and
36 wild-type (white-skinned) goats using deep RNA-sequencing. A total of 923,013,870 raw reads
37 from 6 libraries were obtained, and a large number of alternative splicing events were identified
38 in the transcriptome of fetal skin, including the well-known melanogenic genes *ASIP*, *TYRP1*,
39 and *DCT*, which were differentially expressed in the skin between the dark-skinned and
40 white-skinned goats. Further analysis demonstrated that differential genes including *ASIP*, *TYRP1*,
41 *DCT*, *WNT2*, *RAB27A*, *FZD4*, and *CREB3L1* were significantly overrepresented in the
42 *melanogenesis* pathway and several biological process associated with pigmentation. On the
43 other hand, we identified 1616 novel transcripts in goat skin based on the characteristics of their
44 expression level and gene composition. These novel transcripts may represent two distinct groups
45 of nucleic acid molecules. Our findings contribute to the understanding of the characteristics of
46 global gene expression at early stages of skin pigmentation and development, as well as describe
47 an animal model for human diseases associated with pigmentation.

48 **INTRODUCTION**

49 Skin pigmentation is a complex process that includes melanin biosynthesis in melanocytes, which

50 is transferred to keratinocytes (Wolff 1973; Seiberg 2001). Unlike the numerous variations of
51 coat color, there are only a few natural variants of skin color in other mammals compared to the
52 phenotypes of human skin color. As a result, most studies on the genetics of skin pigmentation
53 are conducted in humans (Baxter and Pavan ; Sturm 2009; Quillen and Shriver 2011; Meng *et al.*
54 2012) and mice (Bennett and Lamoreux 2003; Garcia *et al.* 2008), in which 378 color
55 genes (171 cloned genes and 207 uncloned genes) have been identified
56 (<http://www.espcr.org/micemut>) to date. However, only a small portion of these color genes have
57 been identified as candidate causative genes for skin pigmentation in different populations of
58 humans, mice/rats, and sheep, including *ASIP*, *TYRP1*, *DCT*, *TYR*, *HERC2*, *OCA2*, *MC1R*,
59 *SLC24A5*, *SLC45A2*, *IRF4*, *KITLG* (Slominski *et al.* 2004; Yamaguchi *et al.* 2007; Ebanks *et al.*
60 2009; Yamaguchi and Hearing 2009; Garcia-Gamez *et al.* 2011; Kondo and Hearing 2011; Liu *et*
61 *al.* 2013; Raadsma *et al.* 2013; Han *et al.* 2015). In addition, few studies have characterized the
62 global gene expression profile of skin pigmentation and development. Although investigations of
63 the skin transcriptome have recently been conducted on coat color (Fan *et al.* 2013) and hair
64 follicles (Xu *et al.* 2013a; Xu *et al.* 2013b; Wang *et al.* 2015; Yue *et al.* 2015; Gao *et al.* 2016) in
65 sheep and goats, and skin color in the common carp (Jiang and Bikle 2014b; Wang *et al.* 2014),
66 red tilapia (Zhu *et al.* 2016), and chickens (Zhang *et al.* 2015), the genetic basis for skin
67 pigmentation is not well understood compared to that of hair or coat color.

68 Here, we report a rare indigenous goat breed (Youzhou dark goat) that features dark skin of the
69 body including the visible mucous membranes (Fig.1A), which is a rare fibromelanosis that has
70 previously only been reported in the Silky fowl (Nozaki and Makita 1998; Muroya *et al.* 2000;
71 Dorshorst *et al.* 2011; Shinomiya *et al.* 2012). The pigmentation phenotype of the Youzhou dark
72 goat significantly differs from the piebald phenotype of the bovine (Weikard *et al.* 2013),

73 whereas it is similar to a recently reported case of dermal melanocytosis in humans (Lee *et al.*
74 2010). However, the underlying mechanism of this hyperpigmentation is yet to be explored.
75 Therefore, the Youzhou dark goat can be used as a medical model to study human diseases
76 associated with pigmentation, such as skin melanopathy, melanosis coli, and mucosal melanosis.
77 Interestingly, based on our long-term observations (unpublished), skin pigmentation parallels
78 skin development during the pregnant and postnatal period in the Youzhou dark goat.
79 Consequently, to understand fibromelanosis in mammals, it is necessary to investigate the
80 biology of skin pigmentation at early developmental stages in goats. In developmental
81 embryology, the growth of the fetal skin peaks at approximately 100 days of gestation in sheep
82 and goats (Wang *et al.* 1996; Qin 2001). In the present study, we used the Youzhou dark goat
83 (hyperpigmented or dark-skinned) and Yudong white goat (wild-type or white-skinned) as
84 models of skin pigmentation to characterize the skin transcriptome in 100-day-old fetal goats
85 using deep RNA-sequencing. Our study not only contributes to the understanding of the biology
86 of skin pigmentation and development but also provides valuable information towards
87 understanding human melanocytosis.

88 **MATERIALS AND METHODS**

89 **Ethical statement**

90 All surgical procedures in goats were performed according to the Regulations for the
91 Administration of Affairs Concerning Experimental Animals (Ministry of Science and
92 Technology, China; revised in June 2004) and adhered to the Reporting Guidelines for
93 Randomized Controlled Trials in Livestock and Food Safety (REFLECT).

94 **Animals**

95 Two goat groups with different skin pigmentation phenotypes were investigated in this study. The
96 Yudong white goat (*Capra hircus*) is distributed in Southwest China (located at 31°14'-32°12' N
97 and 108°15'-109°58' E) and features white color in the coat and skin. The Youzhou dark goat
98 (*Capra hircus*) is an indigenous breed uniquely distributed in Youyang County in Chongqing,
99 China (located at 26°54' N and 108°57' E) and features dark skin of the body including the visible
100 mucous membranes but a generally white coat color. Briefly, three pregnant ewes from each
101 breed were subjected to a caesarean section to collect the fetuses (n=3) at 100 days of gestation,
102 and then the dorsal and ventral skins were collected from each fetus. The first sample (3 grams)
103 was dissected and rapidly frozen whole in isopentane chilled over liquid nitrogen for histological
104 examination. The second sample (3 grams) was snap-frozen in liquid nitrogen for
105 RNA-sequencing and qPCR analysis.

106 **RNA isolation, library preparation and sequencing**

107 In the present study, a total of 6 libraries were generated for sequencing according to the sample
108 size (n=6) in two breeds of goats. For each of the 6 fetal goats, total RNA was isolated using
109 TRIzol reagent (Invitrogen, USA) according to the manufacturer's instructions. RNA degradation
110 and contamination was monitored on 1% agarose gels. RNA purity was checked using the
111 NanoPhotometer spectrophotometer (Implen, USA). RNA concentration was measured using a
112 Qubit RNA Assay Kit with a Qubit 2.0 Fluorometer (Life Technologies, USA). RNA integrity was
113 assessed using the RNA Nano 6000 Assay Kit with the Agilent 2100 Bioanalyzer System
114 (Agilent Technologies, USA). A total of 3 µg of RNA per sample was used as the input material
115 for the RNA sample preparations. First, ribosomal RNA was removed using the Epicentre
116 Ribo-zero™ rRNA Removal Kit (Epicentre, USA), and rRNA free residue was cleaned up using
117 ethanol precipitation. Subsequently, the highly strand-specific libraries were generated using the

118 rRNA-depleted RNA using the NEBNext Ultra™ Directional RNA Library Prep Kit for Illumina
119 (NEB, USA) according to the manufacturer's recommendations. Briefly, fragmentation was
120 carried out using divalent cations under elevated temperature in NEBNext. First strand cDNA
121 was synthesized using random hexamer primers and M-MuLV Reverse Transcriptase (RNaseH-).
122 Second strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase
123 H. In the reaction buffer, dNTPs with dTTP were replaced by dUTP. The remaining overhangs
124 were converted into blunt ends via exonuclease/polymerase activities. After adenylation of the 3'
125 ends of the DNA fragments, NEBNext Adaptor with hairpin loop structure were ligated to
126 prepare for hybridization. To select cDNA fragments of preferentially 150-200 bp in length, the
127 library fragments were purified using the AMPure XP system (Beckman Coulter, USA). Then, 3
128 µl of USER Enzyme (NEB, USA) was used with size-selected, adaptor-ligated cDNA at 37°C for
129 15 min followed by 5 min at 95°C before PCR. PCR was performed with Phusion High-Fidelity
130 DNA polymerase, Universal PCR primers and Index (X) Primer. Finally, the products were
131 purified (AMPure XP system) and the library quality was assessed on the Agilent 2100
132 Bioanalyzer System. The clustering of the index-coded samples was performed on a cBot Cluster
133 Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumina) according to the
134 manufacturer's instructions. After cluster generation, the libraries were sequenced on an Illumina
135 HiSeq 2000 platform and 100 bp paired-end reads were generated.

136 **Quality control**

137 The raw data were firstly processed through in-house perl scripts. In this step, clean data were
138 obtained by removing the reads containing adapters, reads containing over 10% of ploy-N
139 sequences, and low quality reads (more than 50% of bases with Phred scores less than 5%) from
140 the raw data. The Phred score (Q20, Q30) and GC content of the clean data were calculated. All

141 the subsequent analyses were based on the high quality data. The sequencing data were submitted
142 to the Genome Expression Omnibus (Accession Numbers GSE69812) in NCBI.

143 **Mapping to the reference genome**

144 The reference genome and gene model annotation files were downloaded directly from the
145 genome website (<http://goat.kiz.ac.cn>). The index of the reference genome was built using
146 Bowtie v2.0.6 (Langmead *et al.* 2009; Langmead and Salzberg 2012) and paired-end clean reads
147 were aligned to the reference genome using TopHat v2.0.9 (Trapnell *et al.* 2012; Kim *et al.*
148 2013).

149 **Transcriptome assembly**

150 The goat reference genome and gene model annotation files were downloaded directly from the
151 genome website (<http://goat.kiz.ac.cn>). The index of the reference genome was built using
152 Bowtie v2.0.6 (Langmead *et al.* 2009; Langmead and Salzberg 2012) and paired-end clean reads
153 were aligned to the reference genome using TopHat v2.0.9 (Trapnell *et al.* 2012; Kim *et al.* 2013).
154 The mapped reads of each sample were assembled using both Scripture (beta2) (Guttman *et al.*
155 2010) and Cufflinks (v2.1.1) (Trapnell *et al.* 2010) in a reference-based approach. Scripture was
156 run with default parameters. Cufflinks was run with ‘min-frags-per-transfrag=0’ and
157 ‘--library-type fr-firststrand’, with all other parameters set as default.

158 **Quantification of gene expression level**

159 Cuffdiff (v2.1.1) was used to calculate the FPKM (fragments per kb for a million reads) of both
160 lncRNAs and coding genes in each sample (Trapnell *et al.* 2010). For biological replicates (n=3),
161 transcripts or genes with a P -adjust<0.05 were considered differentially expressed between the
162 two groups of goats (dark-skinned and white-skinned).

163 **Alternative splicing analysis**

164 Alternative splicing (AS) events were classified into 12 basic types using the software Asprofile
165 v1.0 (Florea *et al.* 2013). The number of AS events in each of the 6 samples was estimated
166 separately. We used DEXSeq software(Anders *et al.* 2012) for the differential exon usage analysis
167 of the AS transcripts, in which a general linear model was employed for the differential analysis
168 of exon expression and a P -adjust<0.05 indicated a significant result.

169 **Identification of novel transcripts**

170 To identify the novel transcripts from clean data, we first distinguished the novel transcripts and
171 the candidate long noncoding RNAs using four tools including CNCI (v2) (Sun *et al.* 2013), CPC
172 (0.9-r2) (Kong *et al.* 2007), Pfam-scan (v1.3) (Punta *et al.* 2012), and PhyloCSF (v20121028)
173 (Lin *et al.* 2011). Transcripts predicted to be without coding potential by all of four tools above
174 were filtered away (which are considered as the candidate long noncoding RNAs), and those with
175 coding potential (which are selected by either of four tools above) but lacking of any known
176 annotation were kept and considered as the novel transcripts in the present analysis.
177 Quantification of gene expression level were estimated by calculating FPKMs of the novel
178 transcripts.

179 RepeatMasker (http://www.repeatmasker.org/cgi-bin/WEB_RepeatMasker) was used with the
180 default parameters to identify various TE components in goat. To identify the position bias of TEs
181 in the novel transcript, we searched the TEs in the 2, 000 bp upstream of TSS (Transcription Start
182 Site) of each transcript identified in the goat genome (<http://goat.kiz.ac.cn>) and plotted read
183 coverage at TSSs with the ggplot2 package in R(Wickham 2009).

184 **Differential expression and functional enrichment analysis**

185 Cuffdiff provides statistical routines to determine differential expression in digital transcript or
186 gene expression data using a model based on the negative binomial distribution (Trapnell *et al.*
187 2010). For biological replicates, transcripts or genes with a P -adjust <0.05 were considered
188 differentially expressed. To identify the molecular events or cascades involved, the differentially
189 expressed genes or lncRNA target genes were analyzed using the DAVID platform (Huang Da *et*
190 *al.* 2009b; Huang Da *et al.* 2009a). Significance was expressed as a P -value, which was
191 calculated using the EASE score (a P -value of 0.05 was considered significant).

192 **Validation of gene expression in RNA-seq using quantitative PCR**

193 To validate the gene expression in RNA-seq, the total RNA from the RNA-seq analysis was used
194 for qPCR (Fig. 4). Briefly, first strain cDNA was obtained using a One Step cDNA Synthesis Kit
195 (Bio-Rad, USA), and the mRNA was then quantified using a standard SYBR with *GAPDH* as an
196 endogenous control. Quantitative PCR was performed under the following conditions: 95°C for
197 30 sec, 40 cycles of 95°C for 5 sec, and the optimized annealing temperature for 30 sec. The
198 primers and annealing temperatures for the 8 genes are listed in Table S1. All reactions were
199 performed in triplicate for each sample. Gene expression was quantified relative to *GAPDH*
200 expression using the $2^{(-\Delta Ct)}$ method. Corrections for multiple comparisons were performed using
201 the Holm-Sidak method.

202 **Statistical analysis**

203 Data analyses were performed using the R statistical package.

204 **Data availability**

205 Figure S1 contains classification of the raw reads by RNA-seq for each of the libraries (samples)
206 in the hyperpigmented and normal fetal goats. Figure S2 contains a heatmap of the cluster

207 analysis for the differentially expressed genes. Table S1 contains statistics of the mapping reads
208 to the reference genome. Table S2 contains annotation of novel transcripts in Swiss-Prot database.
209 Table S3 contains differential exon usage (DEU) analysis of the mRNAs in goat fetal skin. Table
210 S4 contains the differentially expressed genes between the wild-type and hyperpigmented goat
211 skins. Table S5 contains results of GO and KEGG analysis of the up-regulated DE genes using
212 DAVID. Table S6 contains results of functional annotation clustering of the DE genes using
213 DAVID.

214

215 **RESULTS**

216 **RNA sequencing of skin in fetal goats**

217 As for phenotype, there are obvious differences in skin tissue pigmentation between the two
218 breeds of fetal goat (Fig. 1A and B). Then we examined the phenotypic differences at the
219 histological level using H & E staining. The results demonstrated that there were more melanin
220 granules in the 100-day fetal epidermal and dermal layers in the Youzhou dark goat than in the
221 Yudong white goat (Fig. 1C and D). To further explore the mechanism underpinning the
222 differences in skin pigmentation, we characterized the skin transcriptome using RNA-seq
223 performed with the Illumina HiSeq 2000 platform. In the present study, we obtained a total of
224 923,013,870 raw reads from 6 libraries (samples), of which 841,895,634 clean reads remained for
225 further analysis after discarding the raw reads that contained adapter sequences, N sequences and
226 low quality sequences. The percentage of clean reads among the raw tags in each library ranged
227 from 88.39% to 93.02% (Fig. S1). We then mapped the clean reads to the goat reference genome

228 (<http://goat.kiz.ac.cn>). Of the total reads from each library, more than 83% matched to a unique
229 genomic location, whereas only 3% matched to multiple genomic locations (Table S1). The
230 Phred score (Q20) for each library was greater than 95.4%, which indicates our RNA-seq data
231 were high quality and suitable for the subsequent analyses.

232 **Novel transcripts in goat genome**

233 We identified a total of 27,947 mRNAs and 1616 novel transcripts in fetal goat skin in the
234 present study. To examine their homology with proteins in state-of-the-art database, we made a
235 BLAST alignment for these novel transcripts in Swiss-Prot database (<http://www.uniprot.org/>).
236 Results showed that 660 of them had a similarity in various degree, but with a relative lower
237 coverage to the known proteins (Table S2). Another 956 novel transcripts were found without
238 annotation in the state-of-the-art database presently. Then we examined the differences in gene
239 length and expression levels (FPKMs) between the two groups of novel transcripts, and found
240 that there were significant differences in gene length (Kolmogorov-Smirnov test, $P = 0.011$; Fig
241 2A) and expression levels (Kolmogorov-Smirnov test, $P = 0.000$; Fig 2B) between them. To
242 further characterize the two groups of novel transcripts, we examined the composition of
243 transposable elements (TEs) harboring in their gene sequences. Our results demonstrated that
244 there were considerable differences in density of various TE components between the two groups
245 of novel transcripts, as well as between the novel transcripts and mRNAs (Fig 2C). Further
246 investigation of position bias for TEs relative to TSS (transcription start site) revealed that the
247 LINE/RTE-BovB were higher enriched in novel transcripts without annotation than the other two
248 groups (Fig 2D). These above findings suggest that the identified transcripts may represent
249 different types of novel transcripts in goat genome.

250 **Alternative splicing of the fetal skin transcriptome**

251 To ascertain the alternative splicing (AS) events of the skin transcriptome, we examined the data
252 from six samples using ASprofile software. Our results showed that TSSs (alternative 5' first
253 exon, transcription start sites), TTSs (alternative 3' last exon, transcription terminal sites), and
254 SKIPs (skipped exons) were the three most frequently observed AS events among the 12 AS
255 subtypes in each sample (Fig. 3A). Among the over 66,000 AS events, we observed AS events in
256 two well-known genes involved in pigmentation, *ASIP* and *TYRP1*, which were found in both
257 goat breeds. To further examine the AS events of the differentially expressed genes between the
258 dark- and white-skinned goat breeds, we subsequently investigated differential exon usage of the
259 AS events using DEXSeq software. The results demonstrated that 360 AS exons belonging to 253
260 known and 28 unknown genes were significantly differentially expressed between the dark- and
261 white-skinned groups ($P\text{-adjust}<0.05$) (Table S3). Importantly, we found that some AS exons
262 belonging to well-known genes involved in pigmentation, such as *ASIP*, *DCT*, and *RAB27A*, were
263 significantly differentially expressed between the two different skin colors (Fig. 3B). This
264 indicates that alternative splicing may be an important mechanism by which expression and
265 function of *ASIP*, *DCT*, and *RAB27A* is regulated in melanocytes.

266 **Identification and functional clustering analysis of the differentially** 267 **expressed genes**

268 In our data, we obtained 448 differentially expressed (DE) genes ($P\text{-adjusted}<0.05$ and fold
269 change ≥ 1.5) using Cuffdiff between the dark- and white-skinned goats. Of the 448 DE genes, 101
270 genes were up-regulated and 347 were down-regulated in the dark-skinned goats compared with
271 the white-skinned goats (Fig. 4A, Table S4, Fig. S2). Here, to elucidate the biology of skin
272 pigmentation in goats, we specially focused on the up-regulated DE genes in the dark-skinned

273 goats. Gene functional classification using DAVID demonstrated that a family of keratin (KRT)
274 was significantly enriched in the up-regulated DE genes, including *KRT-I*, *KRT81*, *KRT83*,
275 *KRTAP13-1*, *KRTAP11-1*, and *KRTAP7-1*, which are constitutive components of skin produced
276 by keratinocytes. Gene ontology analysis showed that the top six biological processes were
277 *response to inorganic substance*, *response to extracellular stimulus*, *response to reactive oxygen*
278 *species*, *pigmentation during development*, *pigment biosynthetic process*, and *response to cAMP*
279 (Table S5). KEGG analysis showed that genes involved in *melanogenesis* and the *MAPK*
280 *signaling pathway* were significantly enriched. These results suggest that the up-regulated genes
281 play important roles in the dark skin phenotype of the Youzhou dark goats.

282 To gain a comprehensive insight of the gene expression between the two breeds, we performed
283 Gene ontology and pathway analyses of all DE genes between the normal and dark-skinned goats
284 using DAVID. Functional clustering analysis of all DE genes showed that 65 annotation clusters
285 were significantly enriched, particularly *cluster 41*, which comprises 7 terms related to
286 hyperpigmentation including *pigment biosynthetic process*, *melanin biosynthetic process*,
287 *pigment metabolic process*, *melanin metabolic process*, *secondary metabolic process*,
288 *melanogenesis*, and *pigmentation during development*. Regarding the pathways involved, we
289 identified 14 significantly overrepresented pathways (Table S6) including the *melanogenesis*
290 pathway (Fig. 4B). Selected DE genes from the RNA-seq analysis were validated using
291 quantitative PCR (Fig. 5). These findings provide additional evidence of the underlying diversity
292 of dermal colors in goats on the transcriptome level. Moreover, we explored the enrichment of
293 genes involved in human pigmentation diseases using DAVID to determine whether there are
294 diseases with similar variations in dermal colors in goats as are observed in humans. As we
295 expected, *ASIP* and *TYRP1* were significantly enriched in one type of OMIM disease (Table 1).

296 This suggests that *ASIP* and *TYRP1* are the two likely candidate genes responsible for dermal
297 hyperpigmentation in Youzhou dark goats.

298 **DISCUSSION**

299 Although transcriptomic investigations on variations in skin color have been conducted in the
300 common carp (Jiang and Bikle 2014b; Wang *et al.* 2014), red tilapia (Zhu *et al.* 2016), and
301 chickens (Zhang *et al.* 2015), few studies have investigated this topic in mammals to date. We
302 first characterized the skin transcriptome in two breeds of fetal goats that feature different skin
303 colors using deep RNA-seq methods. The present study produced a large amount of data (over
304 84.2 Gb of clean reads for six samples) and more than 86% of the total reads were mapped to the
305 goat reference genome (<http://goat.kiz.ac.cn>), which indicated that the RNA-seq data were of
306 high quality and valuable for further analysis. First, we identified a significant number of AS
307 events (Fig. 3A) in the skin transcriptome of fetal goats, which suggests that gene transcription
308 and regulation is complex in the skin during this developmental stage in goats. AS events are
309 tissue specific and widespread in the genome (Xu *et al.* 2002). AS produces variety in the
310 proteins translated from a limited number of genes through significant alterations in protein
311 conformations that modulate cell functions (Yura *et al.* 2006). Our findings (Fig. 3; Table S3)
312 indicate that alternative splicing may be an important mechanism by which the function of *ASIP*,
313 *TYRP1*, *DCT*, and *RAB27A* are regulated in melanocytes. Therefore, it is necessary to identify the
314 isoforms of these genes and their roles in pigmentation in the future. Interestingly, in the present
315 study, we did not observe alternative splicing of the *MITF* gene in the skin, whereas multiple
316 splice variants of *MITF* were previously found in other species such as sheep (Saravanaperumal
317 *et al.* 2014), mice (Bismuth *et al.* 2005), and humans (Kuiper *et al.* 2004). There might be

318 differences in the regulatory mechanisms of skin pigmentation between goats and the species
319 mentioned above. Another valuable finding of the present study is the DE genes identified in the
320 fetal skin between the two goat groups. In particular, members of the keratin family were highly
321 differentially expressed in the fetal skin between the two goat groups, including *KRT1*, *KRT81*,
322 *KRT83*, *KRTAP13-1*, *KRTAP11-1*, and *KRTAP7-1* (Table S5). Keratins make up the largest
323 subgroup of intermediate filament proteins and represent the most abundant proteins in epithelial
324 cells. These proteins are essential to sustain normal epidermal function and play a role in
325 signaling (Porter and Lane 2003). Mutations of a member of the KRT family resulted in an
326 unusual skin pigmentation in humans (Irvine *et al.* 1997; Horiguchi *et al.* 2005; Pascucci *et al.*
327 2006; Geller *et al.* 2013). Therefore, these DE keratins might participate in the signaling
328 transduction involved in melanin synthesis in melanocytes or the transportation of melanin from
329 melanocytes to keratinocytes through cellular interactions (Nakazawa *et al.* 1995; Seiberg 2001;
330 Joshi *et al.* 2007).

331 Further functional clustering analysis of the DE genes showed that the terms involved in
332 pigmentation and melanogenesis were highly overrepresented in our findings (Table S5 and S6),
333 which is similar to the findings from studies on coat color in sheep (Fan *et al.* 2013) and skin
334 color in the common carp (Jiang and Bikle 2014a). However, the specific DE genes of the skin
335 differed in this study compared with the studies above. Specifically, the melanogenic *ASIP* (fold
336 change >31) and *TYRPI* (fold change >87) were not only among the most significantly DE genes
337 between the dark- and white-skinned goats, but were also associated with human pigmentation
338 (Table 1). Previous studies have demonstrated that polymorphisms of *ASIP* are associated with
339 darker skin color in African Americans (Bonilla *et al.* 2005) and fair skin color in Caucasians
340 (Nan *et al.* 2009). *TYRPI*, a key regulator of melanin biosynthesis in melanocytes, is reported to

341 be involved in the diversity of skin color in Europeans (Lao *et al.* 2007) and is a candidate gene
342 for skin color (such as mouth, nose, and ear) in sheep (Raadsma *et al.* 2013). In addition, *DCT* is
343 also associated with variations of skin pigmentation in Asians (Myles *et al.* 2007). However, the
344 results of the present study suggest that *ASIP* is more likely the candidate gene for skin color in
345 goats than *TYRP1* and *DCT*. We believe that *TYRP1* and *DCT* are two downstream genes whose
346 expression levels are affected by *ASIP* in the KEGG pathway *melanogenesis* (Fig. 4B). However,
347 efforts still be taken to ascertain the candidate genes underlying variations of skin color in goats
348 using gene mapping techniques such as genome wide association studies (GWAS), next
349 generation sequencing (NGS), or quantitative trait loci (QTL) analyses.

350 Our another valuable finding is the novel transcripts identified by RNA-seq (Table S1). Cabili, *et*
351 *al* (2011) characterized a class of novel transcripts that were excluded by their coding potential
352 criteria (a Pfam domain, a positive PhyloCSF score, or previously annotated as pseudogenes),
353 and firstly termed them as TUCP (transcripts of uncertain coding potential) (Cabili *et al.* 2011).
354 Since they merely focused the lincRNAs in human genome under a certain classification strategy,
355 other subtypes of lincRNAs such as intronic lincRNAs and antisense lincRNAs were retained and
356 thus grouped into the catalog of TUCP in their findings. However, in the present study, the
357 intronic lincRNAs and antisense lincRNAs were absolutely excluded from the novel transcripts
358 based on our recent study (Ren *et al.* 2016). Thus there are certain differences in classification of
359 novel transcripts between the study from Cabili *et al* and ours. In view of subsequent analyses,
360 we are certain that the identified transcripts with annotation and without annotation are two
361 different groups of novel transcripts. Especially for those without annotation, the characteristics
362 of a relatively low expression level and a highly enrichment of TE component (LINE/RTE-BovB)
363 around TSS (transcription start site) (Fig 2) which are similar with that of the lincRNAs in recent

364 studies (Cabili *et al.* 2011; Derrien *et al.* 2012; Kelley and Rinn 2012; Li *et al.* 2012; Nam and
365 Bartel 2012; Billerey *et al.* 2014; Ren *et al.* 2016), strongly suggest this group of novel
366 transcripts could be long nocoding RNAs. However, we are not certain that the novel transcripts
367 with some annotation should be protein coding genes or long nocoding RNAs based on the
368 current knowledge. It still need further efforts to identify their identity. Generally, this study
369 provides a valuable resource for the genetic mechanisms involved in pigmentation diseases and
370 contributes to the understanding of the biology of skin pigmentation and development in
371 mammals.

372

373 **ACKNOWLEDGEMENTS**

374 We thank Dr. Sun Wu, Sun Xiaowei, and their colleagues from Southwest University for assistance
375 in the sampling and experiments. We also thank Ms Yuan Jingxian and her colleagues for data
376 processing in Novogene LTD. Co (Beijing). This work was supported by grants from the
377 Chongqing Fund of application and development (cstc2013yykfC80003), the Chongqing
378 Fundamental Research Funds (2013cstc-jbky-00106-zj; No.15435), and the Chongqing Fund of
379 Agriculture Development (No.14412, No.15404).

380 **LITERATURE CITED**

381 Anders, S., A. Reyes and W. Huber, 2012 Detecting differential usage of exons from RNA-seq data. *Genome Res* 22:
382 2008-2017.
383 Baxter, L. L., and W. J. Pavan, The etiology and molecular genetics of human pigmentation disorders. *Wiley*
384 *Interdiscip Rev Dev Biol* 2: 379-392.
385 Bennett, D. C., and M. L. Lamoreux, 2003 The color loci of mice--a genetic century. *Pigment Cell Res* 16: 333-344.
386 Billerey, C., M. Boussaha, D. Esquerre, E. Rebours, A. Djari *et al.*, 2014 Identification of large intergenic non-coding
387 RNAs in bovine muscle using next-generation transcriptomic sequencing. *BMC Genomics* 15: 499.
388 Bismuth, K., D. Maric and H. Arnheiter, 2005 MITF and cell proliferation: the role of alternative splice forms. *Pigment*
389 *Cell Res* 18: 349-359.

- 390 Bonilla, C., L. A. Boxill, S. A. Donald, T. Williams, N. Sylvester *et al.*, 2005 The 8818G allele of the agouti signaling
391 protein (ASIP) gene is ancestral and is associated with darker skin color in African Americans. *Hum Genet*
392 116: 402-406.
- 393 Cabili, M. N., C. Trapnell, L. Goff, M. Koziol, B. Tazon-Vega *et al.*, 2011 Integrative annotation of human large
394 intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev* 25: 1915-1927.
- 395 Derrien, T., R. Johnson, G. Bussotti, A. Tanzer, S. Djebali *et al.*, 2012 The GENCODE v7 catalog of human long
396 noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res* 22: 1775-1789.
- 397 Dorshorst, B., A. M. Molin, C. J. Rubin, A. M. Johansson, L. Stromstedt *et al.*, 2011 A complex genomic
398 rearrangement involving the endothelin 3 locus causes dermal hyperpigmentation in the chicken. *PLoS*
399 *Genet* 7: e1002412.
- 400 Ebanks, J. P., R. R. Wickett and R. E. Boissy, 2009 Mechanisms regulating skin pigmentation: the rise and fall of
401 complex coloration. *Int J Mol Sci* 10: 4066-4087.
- 402 Fan, R., J. Xie, J. Bai, H. Wang, X. Tian *et al.*, 2013 Skin transcriptome profiles associated with coat color in sheep.
403 *BMC Genomics* 14: 389.
- 404 Florea, L., L. Song and S. L. Salzberg, 2013 Thousands of exon skipping events differentiate among splicing patterns
405 in sixteen human tissues. *F1000Res* 2: 188.
- 406 Gao, Y., X. Wang, H. Yan, J. Zeng, S. Ma *et al.*, 2016 Comparative Transcriptome Analysis of Fetal Skin Reveals Key
407 Genes Related to Hair Follicle Morphogenesis in Cashmere Goats. *PLoS One* 11: e0151118.
- 408 Garcia-Gamez, E., A. Reverter, V. Whan, S. M. McWilliam, J. J. Arranz *et al.*, 2011 Using regulatory and epistatic
409 networks to extend the findings of a genome scan: identifying the gene drivers of pigmentation in merino
410 sheep. *PLoS One* 6: e21158.
- 411 Garcia, S., F. Diaz and C. T. Moraes, 2008 A 3' UTR modification of the mitochondrial rieske iron sulfur protein in
412 mice produces a specific skin pigmentation phenotype. *J Invest Dermatol* 128: 2343-2345.
- 413 Geller, L., L. Kristal and K. D. Morel, 2013 Epidermolysis bullosa simplex with mottled pigmentation due to a rare
414 keratin 5 mutation: cutaneous findings in infancy. *Pediatr Dermatol* 30: 631-632.
- 415 Guttman, M., M. Garber, J. Z. Levin, J. Donaghey, J. Robinson *et al.*, 2010 Ab initio reconstruction of cell type-specific
416 transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nat Biotechnol* 28:
417 503-510.
- 418 Han, J. L., M. Yang, Y. J. Yue, T. T. Guo, J. B. Liu *et al.*, 2015 Analysis of agouti signaling protein (ASIP) gene
419 polymorphisms and association with coat color in Tibetan sheep (*Ovis aries*). *Genet Mol Res* 14:
420 1200-1209.
- 421 Horiguchi, Y., D. Sawamura, R. Mori, H. Nakamura, K. Takahashi *et al.*, 2005 Clinical heterogeneity of 1649delG
422 mutation in the tail domain of keratin 5: a Japanese family with epidermolysis bullosa simplex with mottled
423 pigmentation. *J Invest Dermatol* 125: 83-85.
- 424 Huang da, W., B. T. Sherman and R. A. Lempicki, 2009a Bioinformatics enrichment tools: paths toward the
425 comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 37: 1-13.
- 426 Huang da, W., B. T. Sherman and R. A. Lempicki, 2009b Systematic and integrative analysis of large gene lists using
427 DAVID bioinformatics resources. *Nat Protoc* 4: 44-57.
- 428 Irvine, A. D., K. E. McKenna, H. Jenkinson and A. E. Hughes, 1997 A mutation in the V1 domain of keratin 5 causes
429 epidermolysis bullosa simplex with mottled pigmentation. *J Invest Dermatol* 108: 809-810.
- 430 Jiang, Y. J., and D. D. Bikle, 2014a LncRNA profiling reveals new mechanism for VDR protection against skin cancer
431 formation. *J Steroid Biochem Mol Biol* 144 Pt A: 87-90.
- 432 Jiang, Y. J., and D. D. Bikle, 2014b LncRNA: a new player in 1 α , 25(OH) $_2$ vitamin D $_3$ /VDR protection against
433 skin cancer formation. *Exp Dermatol* 23: 147-150.
- 434 Joshi, P. G., N. Nair, G. Begum, N. B. Joshi, V. P. Sinkar *et al.*, 2007 Melanocyte-keratinocyte interaction induces
435 calcium signalling and melanin transfer to keratinocytes. *Pigment Cell Res* 20: 380-384.
- 436 Kelley, D., and J. Rinn, 2012 Transposable elements reveal a stem cell-specific class of long noncoding RNAs.
437 *Genome Biol* 13: R107.
- 438 Kim, D., G. Pertea, C. Trapnell, H. Pimentel, R. Kelley *et al.*, 2013 TopHat2: accurate alignment of transcriptomes in
439 the presence of insertions, deletions and gene fusions. *Genome Biol* 14: R36.
- 440 Kondo, T., and V. J. Hearing, 2011 Update on the regulation of mammalian melanocyte function and skin
441 pigmentation. *Expert Rev Dermatol* 6: 97-108.
- 442 Kong, L., Y. Zhang, Z. Q. Ye, X. Q. Liu, S. Q. Zhao *et al.*, 2007 CPC: assess the protein-coding potential of transcripts

- 443 using sequence features and support vector machine. *Nucleic Acids Res* 35: W345-349.
- 444 Kuiper, R. P., M. Schepens, J. Thijssen, E. F. Schoenmakers and A. G. van Kessel, 2004 Regulation of the MITF/TFE
445 bHLH-LZ transcription factors through restricted spatial expression and alternative splicing of functional
446 domains. *Nucleic Acids Res* 32: 2315-2322.
- 447 Langmead, B., and S. L. Salzberg, 2012 Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9: 357-359.
- 448 Langmead, B., C. Trapnell, M. Pop and S. L. Salzberg, 2009 Ultrafast and memory-efficient alignment of short DNA
449 sequences to the human genome. *Genome Biol* 10: R25.
- 450 Lao, O., J. M. de Gruijter, K. van Duijn, A. Navarro and M. Kayser, 2007 Signatures of positive selection in genes
451 associated with human skin pigmentation as revealed from analyses of single nucleotide polymorphisms.
452 *Ann Hum Genet* 71: 354-369.
- 453 Lee, S., D. H. Kim, G. Lee, K. U. Whang, J. S. Lee *et al.*, 2010 An unusual case of congenital dermal melanocytosis.
454 *Ann Dermatol* 22: 460-462.
- 455 Li, T., S. Wang, R. Wu, X. Zhou, D. Zhu *et al.*, 2012 Identification of long non-protein coding RNAs in chicken skeletal
456 muscle using next generation sequencing. *Genomics* 99: 292-298.
- 457 Lin, M. F., I. Jungreis and M. Kellis, 2011 PhyloCSF: a comparative genomics method to distinguish protein coding
458 and non-coding regions. *Bioinformatics* 27: i275-282.
- 459 Liu, F., B. Wen and M. Kayser, 2013 Colorful DNA polymorphisms in humans. *Semin Cell Dev Biol* 24: 562-575.
- 460 Meng, S., M. Zhang, L. Liang and J. Han, 2012 Current opportunities and challenges: genome-wide association
461 studies on pigmentation and skin cancer. *Pigment Cell Melanoma Res* 25: 612-617.
- 462 Muroya, S., R. Tanabe, I. Nakajima and K. Chikuni, 2000 Molecular characteristics and site specific distribution of the
463 pigment of the silky fowl. *J Vet Med Sci* 62: 391-395.
- 464 Myles, S., M. Somel, K. Tang, J. Kelso and M. Stoneking, 2007 Identifying genes underlying skin pigmentation
465 differences among human populations. *Hum Genet* 120: 613-621.
- 466 Nakazawa, K., H. Nakazawa, C. Collombel and O. Damour, 1995 Keratinocyte extracellular matrix-mediated
467 regulation of normal human melanocyte functions. *Pigment Cell Res* 8: 10-18.
- 468 Nam, J. W., and D. P. Bartel, 2012 Long noncoding RNAs in *C. elegans*. *Genome Res* 22: 2529-2540.
- 469 Nan, H., P. Kraft, D. J. Hunter and J. Han, 2009 Genetic variants in pigmentation genes, pigmentary phenotypes, and
470 risk of skin cancer in Caucasians. *Int J Cancer* 125: 909-917.
- 471 Nozaki, A., and T. Makita, 1998 The surface color measurement of major tissues of silky fowls and White Leghorns. *J*
472 *Vet Med Sci* 60: 489-493.
- 473 Pascucci, M., P. Posteraro, C. Pedicelli, A. Provini, L. Auricchio *et al.*, 2006 Epidermolysis bullosa simplex with
474 mottled pigmentation due to de novo P25L mutation in keratin 5 in an Italian patient. *Eur J Dermatol* 16:
475 620-622.
- 476 Porter, R. M., and E. B. Lane, 2003 Phenotypes, genotypes and their contribution to understanding keratin function.
477 *Trends Genet* 19: 278-285.
- 478 Punta, M., P. C. Coghill, R. Y. Eberhardt, J. Mistry, J. Tate *et al.*, 2012 The Pfam protein families database. *Nucleic*
479 *Acids Res* 40: D290-301.
- 480 Qin, P., 2001 *Embryology in Mammals*. Science Press, Beijing.
- 481 Quillen, E. E., and M. D. Shriver, 2011 Unpacking human evolution to find the genetic determinants of human skin
482 pigmentation. *J Invest Dermatol* 131: E5-7.
- 483 Raadsma, H. W., E. Jonas, M. R. Fleet, K. Fullard, J. Gongora *et al.*, 2013 QTL and association analysis for skin and
484 fibre pigmentation in sheep provides evidence of a major causative mutation and epistatic effects. *Anim*
485 *Genet* 44: 547-559.
- 486 Ren, H., G. Wang, L. Chen, J. Jiang, L. Liu *et al.*, 2016 Genome-wide analysis of long non-coding RNAs at early stage
487 of skin pigmentation in goats (*Capra hircus*). *BMC Genomics* 17: 67.
- 488 Saravanaperumal, S. A., D. Pediconi, C. Renieri and A. La Terza, 2014 Alternative splicing of the sheep MITF gene:
489 novel transcripts detectable in skin. *Gene* 552: 165-175.
- 490 Seiberg, M., 2001 Keratinocyte-melanocyte interactions during melanosome transfer. *Pigment Cell Res* 14: 236-242.
- 491 Shinomiya, A., Y. Kayashima, K. Kinoshita, M. Mizutani, T. Namikawa *et al.*, 2012 Gene duplication of endothelin 3 is
492 closely correlated with the hyperpigmentation of the internal organs (Fibromelanosis) in silky chickens.
493 *Genetics* 190: 627-638.
- 494 Slominski, A., D. J. Tobin, S. Shibahara and J. Wortsman, 2004 Melanin pigmentation in mammalian skin and its
495 hormonal regulation. *Physiol Rev* 84: 1155-1228.

- 496 Sturm, R. A., 2009 Molecular genetics of human pigmentation diversity. *Hum Mol Genet* 18: R9-17.
- 497 Sun, L., H. Luo, D. Bu, G. Zhao, K. Yu *et al.*, 2013 Utilizing sequence intrinsic composition to classify protein-coding
498 and long non-coding transcripts. *Nucleic Acids Res* 41: e166.
- 499 Trapnell, C., A. Roberts, L. Goff, G. Pertea, D. Kim *et al.*, 2012 Differential gene and transcript expression analysis of
500 RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* 7: 562-578.
- 501 Trapnell, C., B. A. Williams, G. Pertea, A. Mortazavi, G. Kwan *et al.*, 2010 Transcript assembly and quantification by
502 RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*
503 28: 511-515.
- 504 Wang, C., M. Wachholtz, J. Wang, X. Liao and G. Lu, 2014 Analysis of the skin transcriptome in two oujiang color
505 varieties of common carp. *PLoS One* 9: e90074.
- 506 Wang, L., L. Peng, W. Zhang, Z. Zhang, W. Yang *et al.*, 1996 Initiation and development of skin follicles in the inner
507 Mongolian Cashmere goat. *Acta Veterinaria et Zootechnica Sinica* 27: 524-530.
- 508 Wang, L., Y. Zhang, M. Zhao, R. Wang, R. Su *et al.*, 2015 SNP Discovery from Transcriptome of Cashmere Goat Skin.
509 *Asian-Australas J Anim Sci* 28: 1235-1243.
- 510 Weikard, R., F. Hadlich and C. Kuehn, 2013 Identification of novel transcripts and noncoding RNAs in bovine skin by
511 deep next generation sequencing. *BMC Genomics* 14: 789.
- 512 Wickham, H., 2009 *Elegant graphics for data analysis*. Springer, New York.
- 513 Wolff, K., 1973 Melanocyte-keratinocyte interactions in vivo: the fate of melanosomes. *Yale J Biol Med* 46: 384-396.
- 514 Xu, Q., B. Modrek and C. Lee, 2002 Genome-wide detection of tissue-specific alternative splicing in the human
515 transcriptome. *Nucleic Acids Res* 30: 3754-3766.
- 516 Xu, T., X. Guo, H. Wang, X. Du, X. Gao *et al.*, 2013a De Novo Transcriptome Assembly and Differential Gene
517 Expression Profiling of Three *Capra hircus* Skin Types during Anagen of the Hair Growth Cycle. *Int J*
518 *Genomics* 2013: 269191.
- 519 Xu, T., X. Guo, H. Wang, F. Hao, X. Du *et al.*, 2013b Differential gene expression analysis between anagen and telogen
520 of *Capra hircus* skin based on the de novo assembled transcriptome sequence. *Gene* 520: 30-38.
- 521 Yamaguchi, Y., M. Brenner and V. J. Hearing, 2007 The regulation of skin pigmentation. *J Biol Chem* 282:
522 27557-27561.
- 523 Yamaguchi, Y., and V. J. Hearing, 2009 Physiological factors that regulate skin pigmentation. *Biofactors* 35: 193-199.
- 524 Yue, Y. J., J. B. Liu, M. Yang, J. L. Han, T. T. Guo *et al.*, 2015 De novo assembly and characterization of skin
525 transcriptome using RNAseq in sheep (*Ovis aries*). *Genet Mol Res* 14: 1371-1384.
- 526 Yura, K., M. Shionyu, K. Hagino, A. Hijikata, Y. Hirashima *et al.*, 2006 Alternative splicing in human transcriptome:
527 functional and structural influence on proteins. *Gene* 380: 63-71.
- 528 Zhang, J., F. Liu, J. Cao and X. Liu, 2015 Skin transcriptome profiles associated with skin color in chickens. *PLoS One*
529 10: e0127301.
- 530 Zhu, W., L. Wang, Z. Dong, X. Chen, F. Song *et al.*, 2016 Comparative Transcriptome Analysis Identifies Candidate
531 Genes Related to Skin Color Differentiation in Red Tilapia. *Sci Rep* 6: 31347.
- 532
- 533

533

534 Figure legends

535 **Fig. 1 Histomorphological examination of the hyperpigmented and normal skin in fetal goats.**

536 To explore the histomorphological differences between the two breeds of goats at the early stage
537 of skin pigmentation and development, the 100-day-old fetal skin from the Youzhou dark goat (A

538 and C) and Yudong white goat (B and D) were examined using H & E staining. The scale bar for
539 the images in (C) and (D) is 10 μ m.

540 **Fig.2 Characteristic differences between two groups of novel transcripts in goat genome**

541 The transcript length (A) and expression level (B) of two classes of novel transcripts were
542 compared using the Kolmogorov-Smirnov test, and a P value of 0.05 indicates significance
543 between two groups. In two box plots, the circle indicates the outlier, and the asterisk labels the
544 extreme. (C) The proportion of the main TE families in two groups of novel transcripts and
545 mRNAs in the goat genome (<http://goat.kiz.ac.cn>). Differences in TE components between them
546 were measured by using the Fisher Exact test. (D) The position bias of TE components in the
547 2,000 bp upstream of TSS in transcripts above mentioned.

548 **Fig. 3 Alternative splicing (AS) events in the goat genome.**

549 Following the analysis using Cufflinks, the classification of AS events of transcripts were
550 performed for each library (sample) using the ASprofile tool. The 12 main AS events (TSS, TTS,
551 SKIP, XSKIP, MSKIP, XMSKIP, IR, XIR, MIR, XMIR, AE, and XAE) in the goat genome are
552 summarized in (A). Y1, Y2, and Y3 indicate the dark skin goats, whereas B1, B2, and B3 indicate
553 the white skin goats. (B) Differential exon usage of the AS events in the *ASIP* and *DCT* genes
554 between the dark-skinned (Y, blue) and white-skinned (B, red) goats.

555 **Fig. 4 The differentially expressed genes between the dark- and white-skinned goats.**

556 (A) The differential expression of genes was identified based on the FPKM values (corrected
557 P -value \leq 0.05 and fold change \geq 1.5) in the two groups. The volcano plot was generated based on
558 the values above. Y indicates the dark-skinned goats, and B refers to the white-skinned goats. (B)
559 Enrichment of the DE genes was conducted under KEGG pathways. Red indicates the

560 up-regulated DE genes and green refers to the down-regulated genes in the dark-skinned goats
561 compared to the control goats (white-skinned).

562 **Fig. 5 Validation of gene expression in the dark- and white-skinned goats using qPCR.**

563 Some of the identified melanogenic genes were examined in the dark- and white-skinned fetal
564 goats using qPCR. Gene expression was quantified relative to *GAPDH* expression using the $2^{(-\Delta Ct)}$
565 method. Corrections for multiple comparisons were performed using the Holm-Sidak method.
566 The data are shown as the mean \pm 1 SE (n = 3). * $P < 0.05$, ** $P < 0.01$.

567 **Additional files**

568 Fig. S1. Classification of the raw reads by RNA-seq for each of the libraries (samples) in the
569 hyperpigmented and normal fetal goats.

570 Fig. S2. Heatmap of the cluster analysis for the differentially expressed genes.

571 Table S1. Statistics of the mapping reads to the reference genome.

572 Table S2. Annotation of novel transcripts in Swiss-Prot database.

573 Table S3. Differential exon usage (DEU) analysis of the mRNAs in goat fetal skin.

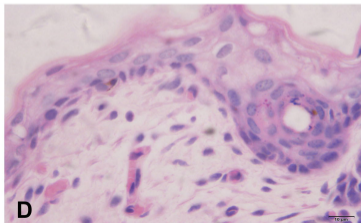
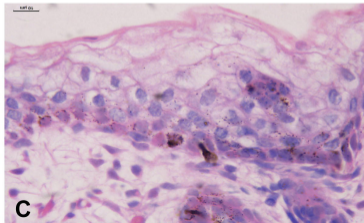
574 Table S4. Differentially expressed genes between the wild-type and hyperpigmented goat skins.

575 Table S5. GO and KEGG analysis of the up-regulated DE genes using DAVID.

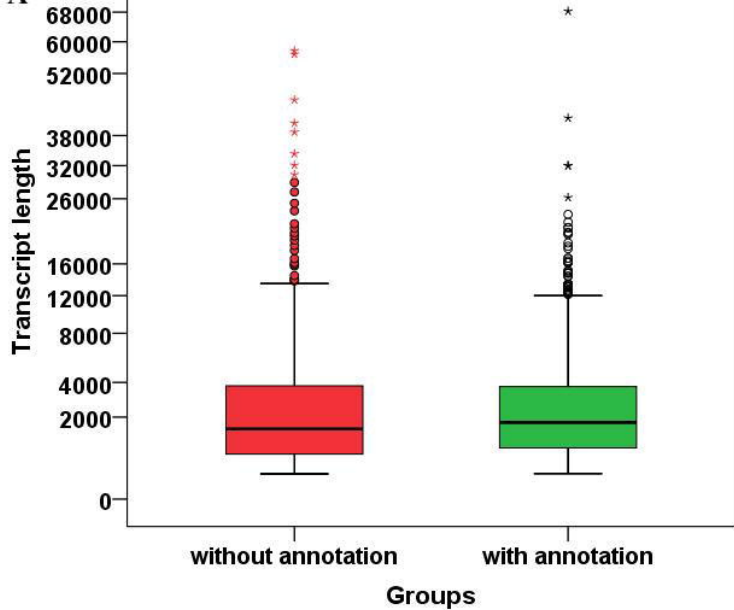
576 Table S6. Functional annotation clustering of the DE genes using DAVID.

Hyperpigmented

Wild type



A

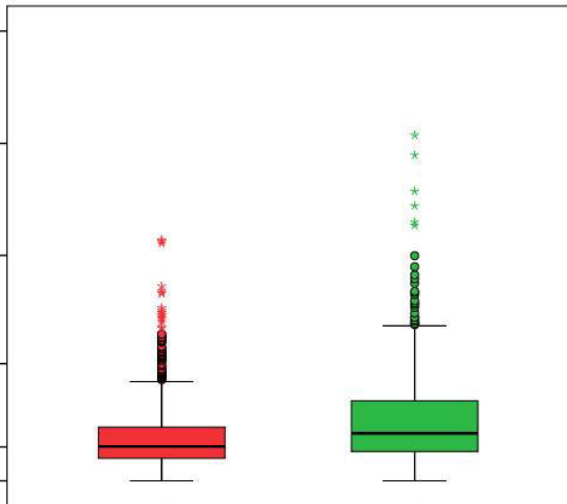


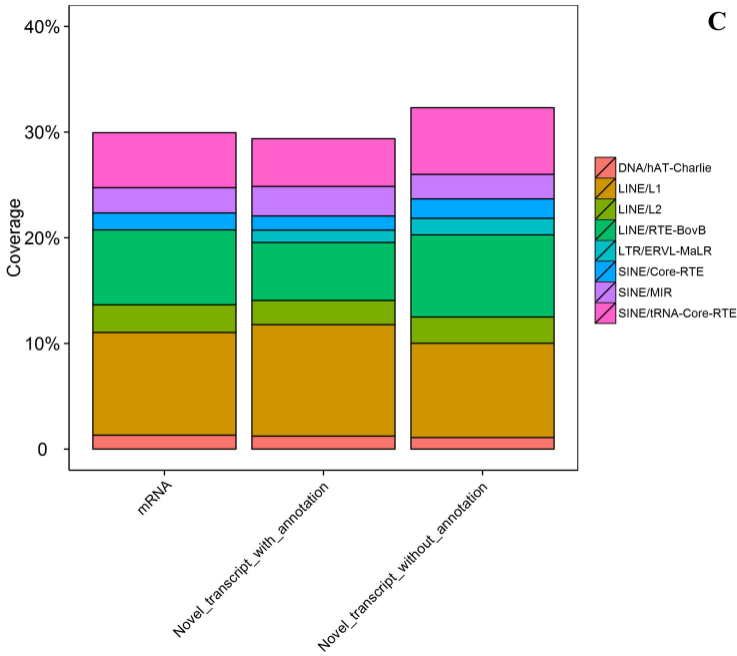
B**Distribution of FPKM**

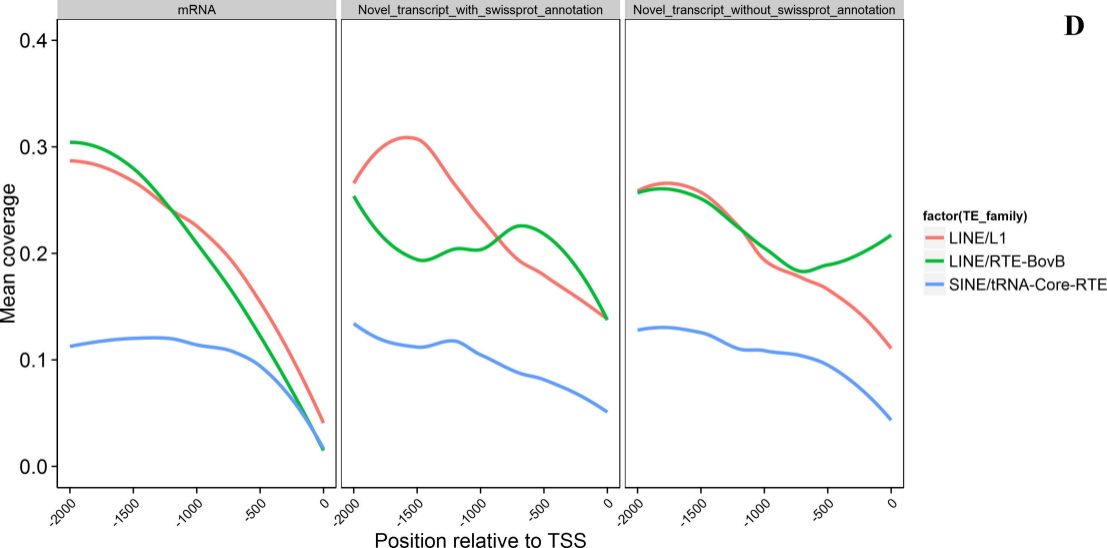
10000.000
1000.000
100.000
10.000
1.000
.000

without annotation

with annotation

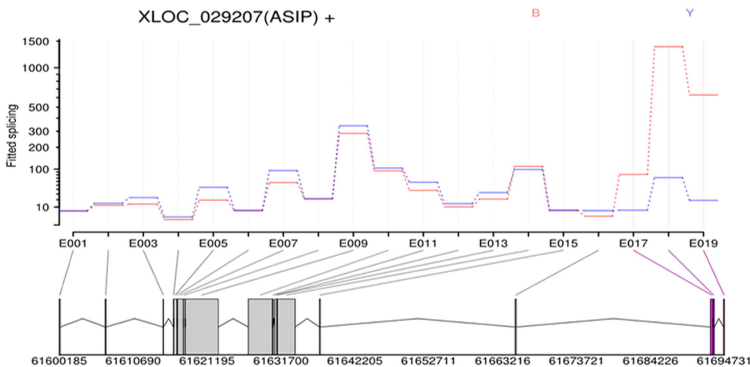
Groups



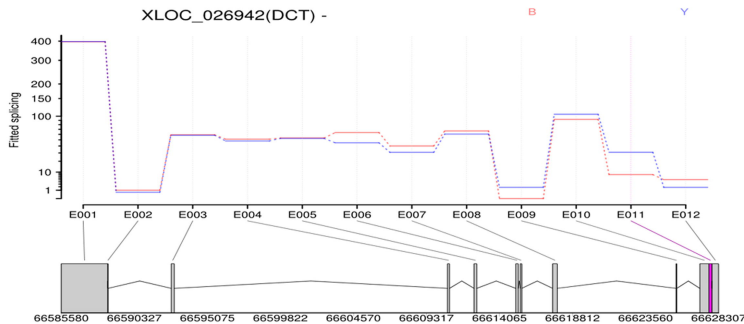
D

B

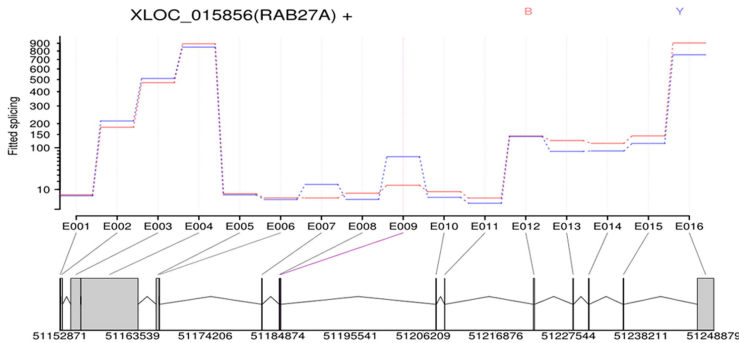
XLOC_029207(ASIP) +

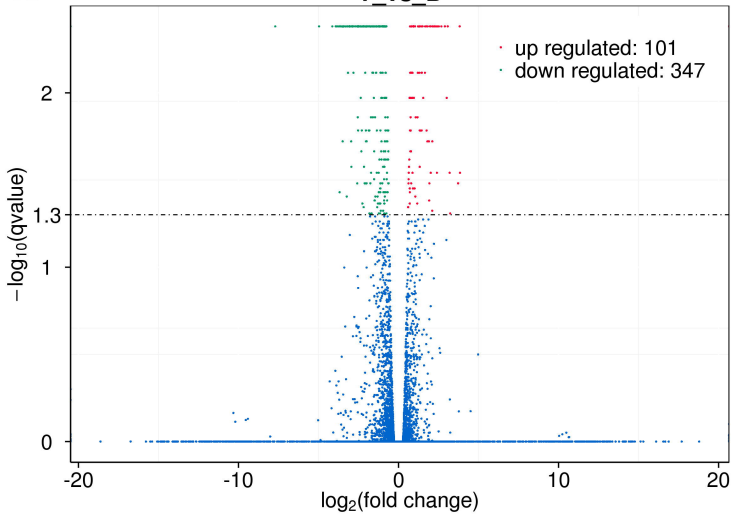


XLOC_026942(DCT) -



XLOC_015856(RAB27A) +



A**Y_vs_B**

B