BMC Genomics

RESEARCH ARTICLE

Open Access

CrossMark

Genome-wide analysis of long non-coding RNAs at early stage of skin pigmentation in goats (Capra hircus)

Hangxing Ren[1,2†] ID, Gaofu Wang[1,2†], Lei Chen[1], Jing Jiang[1,2], Liangjia Liu[1,2], Nianfu Li[1,3], Jinhong Zhao[1,2], Xiaoyan Sun[1,2] and Peng Zhou[1,2*]

Abstract

Background: Long noncoding RNAs (lncRNAs) play roles in almost all biological processes; however, their function and profile in skin development and pigmentation is less understood. In addition, because lncRNAs are species-specific, their function in goats has not been established.

Result: We systematically identified lncRNAs in 100-day-old fetal skin by deep RNA-sequencing using the Youzhou dark goat (dark skin) and Yudong white goat (white skin) as a model of skin pigmentation. A total of 841,895,634 clean reads were obtained from six libraries (samples). We identified 1336 specific lncRNAs in fetal skin that belonged to three subtypes, including 999 intergenic lncRNAs (lincRNAs), 218 anti-sense lncRNAs, and 119 intronic lncRNAs. Our results demonstrated significant differences in gene architecture and expression among the three lncRNA subtypes, particularly in terms of density and position bias of transpose elements near the transcription start site. We also investigated the impact of lncRNAs on its target genes in *cis* and *trans*, indicating that these lncRNAs have a strict tissue specificity and functional conservation during skin development and pigmentation.

Conclusion: The present study provides a resource for lncRNA studies in diseases involved in pigmentation and skin development. It expands our knowledge about lncRNA biology as well as contributes to the annotation of the goat genome.

Keywords: Skin, lncRNA, Goat, Pigmentation, Transcriptome

Background

As a species of ubiquitous noncoding RNAs, long non-coding RNAs (lncRNAs) are unambiguously distinguished from mRNAs in terms of sequence structure, positional characteristics, expression level, and evolutional conservation [1–5]. Moreover, subspecies of lncRNAs have also been categorized and characterized in human [1, 4], zebrafish [5], and *Caenorhabditis elegans* [6]. Recent reports have shown that similar to mRNA, lncRNA is functional and spatiotemporally expressed in tissues [7–9]. Researchers have identified several functional lncRNAs associated with skin biology such as *ANCR*, *TINCR*, *U1 RNA*, *PRINS*, *BANCR*, and *SPRY4-IT1* [10]. In addition, it has also been shown that a few well-known oncogenes, including *H19*, *HOTTIP*, *Nespas*, *Kcnq1ot1*, *lincRNA-p21*, *mHOTAIR*, *Malat1*, *SRA*, *Foxn2-as*, *Gtl2-as*, and *H19-as*, are involved in vitamin D receptor protection against skin cancer formation by maintaining the balance of oncogenic to tumor-suppressing lncRNAs [11]. In current lncRNA databases, most of the identified lncRNAs were mainly derived from human and mouse [12–14]. Several recent studies in bovine [15–17], chicken [2], and pig [3, 18] have enriched the animal lncRNA datasets; however, our understanding of goat lncRNAs is limited. Despite the abundance of lncRNAs in the genome, only a few have been fully characterized. Currently, there are only two reports on the identification of the skin lncRNAs in mammals. RNA sequencing (RNA-seq) analysis conducted by

* Correspondence: cqzp@163.com
†Equal contributors
[1]Chongqing Academy of Animal Sciences, Chongqing 402460, China
[2]Chongqing Engineering Research Center for Goats, Chongqing 402460, China
Full list of author information is available at the end of the article

Ren *et al. BMC Genomics* (2016) 17:67

Page 2 of 12

Weikard et al. (2013) identified 4365 potential intergenic lncRNAs in cow with a piebald phenotype [17], which differs from that of Youzhou dark goat (as described in the next section). Another skin lncRNA catalog was derived from human skin cancer [11]. To our knowledge, only a few reports have described the involvement of skin lncRNAs in prenatal pigmentation and development. During embryonic development, fetal skin undergoes growth at a relatively high rate for 100 gestational days in goat [19]. Therefore, it is significant and necessary to investigate skin pigmentation during this specific developmental stage.

RNA-seq is a powerful approach that unravels the differential expression profiles underlying phenotypic differences, as well as deciphers non-annotated transcriptional activity by identifying various novel transcripts (protein-coding and noncoding) and additional alternative splice variants of known annotated transcripts [20–22]. In the current study, we elucidated the lncRNA profiles of two different phenotypes involved in skin pigmentation in goats using deep RNA-sequencing. Our study subject featured dark skin in its entire body, including its visible mucous membranes, and this phenotypic feature has not been reported in other mammals to date. Our study provides a valuable resource for studying lncRNAs that are involved in diseases, as well as contributes to better understanding the biology of skin pigmentation and development.

Results

Identification of lncRNAs in goat fetal skin

A total of 923,013,870 raw reads were produced from the Illumina HiSeq 2000 platform. After discarding adaptor sequences and low-quality sequences, we obtained 841,895,634 clean reads (accounting for 84.2 Gb), and the percentage of clean reads among raw tags in each library ranged from 88.39–93.02 % (Additional file 1). Subsequently, we mapped the clean reads based on the latest goat reference genome (http://goat.kiz.ac.cn). Considering the characteristics of lncRNA sequences (≥200 nt, exon count ≥ 2) and its differences from other classes of RNA (mRNA, tRNA, rRNA, snRNA, snoRNA, pre-miRNA, and pseudogenes), we classified the transcripts into different subtypes using both Scripture (beta2) and Cufflinks (v2.1.1). Our results showed that 93.6 % of the 46,933 identified transcripts were known reference transcripts, whereas 6.29 % (2952) were the presumed lncRNAs. To further confirm these 2952 lncRNAs, we performed coding potential analysis using the software CNCI, CPC, Pfam-scan, and PhyloCSF. After screening using rigorous criteria and four analytic tools, a total of 1336 lncRNAs from the skin of fetal goats were identified and subjected to further analysis

(Fig. 1). The 1336 lncRNAs consisted of 999 large intergenic noncoding RNAs (lincRNAs), 218 intronic_lncRNAs, and 119 anti-sense_lncRNAs. A preliminary analysis revealed major differences in gene architecture and expression levels among the three subtypes of lncRNAs. For example, the length of intronic lncRNAs was longer than that of lincRNAs (Kolmogorov-Smirnov test, $P = 0.000$) and anti-sense lncRNAs (Kolmogorov-Smirnov test, $P = 0.005$), with a median length of 1.831 kb *vs.* 0.842 kb and 1.194 kb, respectively. Significant differences in transcript length between lincRNAs and intronic lncRNAs were also observed (Kolmogorov-Smirnov test, $P = 0.001$; Fig. 2a). On the other hand, clear differences in the number of exons were also observed among the three lncRNA subtypes (Fig. 2b). In particular, the anti-sense lncRNAs showed a higher number of exons and wider size distribution than that observed in the lincRNAs (Kolmogorov-Smirnov test, $P = 0.222$) and intronic lncRNAs (Kolmogorov-Smirnov test, $P = 0.001$). We also detected significant differences in exon distribution between lincRNAs and intronic lncRNAs (Kolmogorov-Smirnov test, $P = 0.016$). In terms of expression level based on fragments per kb for a million reads (FPKM) values, the intronic lncRNAs showed a higher expression level than that of the lincRNAs (Kolmogorov-Smirnov test, $P = 0.000$) and anti-sense lncRNAs (Kolmogorov-Smirnov test, $P = 0.037$), with a median of 1.229 *vs.* 0.8607 and 1.035, respectively (Fig. 2c). The diversity in gene architecture and expression levels among various types of lncRNAs may
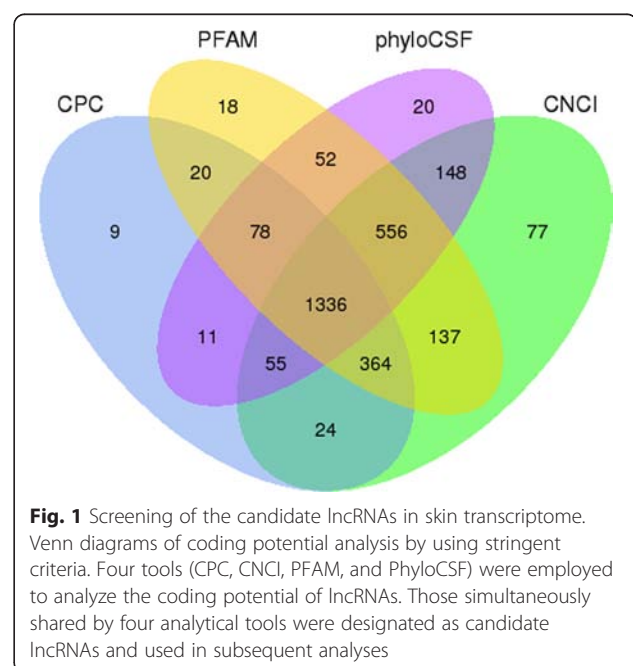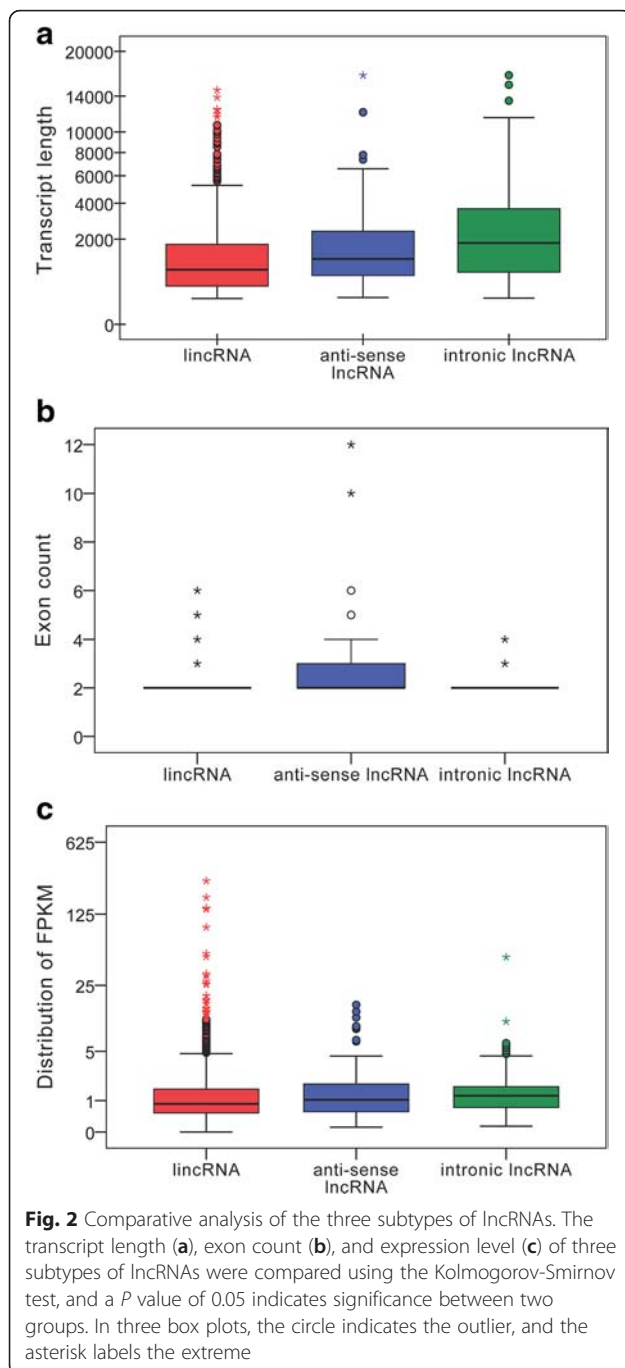

Fig. 1 Screening of the candidate lncRNAs in skin transcriptome. Venn diagrams of coding potential analysis by using stringent criteria. Four tools (CPC, CNCI, PFAM, and PhyloCSF) were employed to analyze the coding potential of lncRNAs. Those simultaneously shared by four analytical tools were designated as candidate lncRNAs and used in subsequent analyses

Ren *et al. BMC Genomics* (2016) 17:67

Page 3 of 12

Fig. 2 Comparative analysis of the three subtypes of lncRNAs. The transcript length (**a**), exon count (**b**), and expression level (**c**) of three subtypes of lncRNAs were compared using the Kolmogorov-Smirnov test, and a *P* value of 0.05 indicates significance between two groups. In three box plots, the circle indicates the outlier, and the asterisk labels the extreme

have implications in its specific function in the goat genome.

Transposable elements characterize various subtypes of lncRNAs

Transposable elements (TEs) are mobile genetic elements that are capable of movement and proliferation within the genome, and its remnants account for one to two thirds of mammalian genomes [23, 24]. TEs are also considered as one of three evolutionary scenarios

involved in the origin of lncRNAs [25]. We were thus prompted to identify differences in TE components between lncRNAs and mRNAs, as well as among the three lncRNA subtypes. Our analysis revealed TE component characteristics that distinguished the three lncRNA subtypes (Fig. 3a). At the global level, significant differences were observed between mRNAs and the individual subtype of lncRNAs (Additional file 2). Among the three subspecies of lncRNAs, the intronic lncRNAs showed a lower TE density than that observed in the lincRNAs and anti-sense lncRNAs (23.75 % *vs.* 36.44 % and 34.91 %). In particular, the lincRNAs and anti-sense lncRNAs have a relatively higher proportion of LINEs/L1s (Fisher's Exact, $P_{\text{linc vs. intronic}} = 0.0034$; $P_{\text{anti-sense vs. intronic}} = 0.0011$) and LINEs/RTE-BovBs (Fisher's Exact, $P_{\text{linc vs. intronic}} = 5.55\text{E-}05$; $P_{\text{anti-sense vs. intronic}} = 0.0024$) than that observed in intronic lncRNAs. In contrast to the lincRNAs and anti-sense lncRNAs, the intronic lncRNAs showed a deletion of LTRs/ERVLs (Fisher's Exact, $P_{\text{linc vs. intronic}} = 5.79\text{E-}08$; $P_{\text{anti-sense vs. intronic}} = 2.11\text{E-}05$) and SINEs/Core-RTEs (Fisher's Exact, $P_{\text{linc vs. intronic}} = 2.38\text{E-}05$; $P_{\text{anti-sense vs. intronic}} = 0.0005$). Differences in the density of other TE subspecies were observed among the three lncRNA subtypes (Additional file 3). These structural characteristics of TE components may underlie the differences in the evolution of the three lncRNA subtypes. Long terminal repeats (LTRs) harbor promoter signals that modulate gene expression in genomes [26, 27], and a recent study has indicated that endogenous retroviruses (ERVs), which is a class of LTRs, exhibit position and orientation biases, often preferring the 5′ end of lincRNA transcripts and sense orientation within the transcript, and avoiding the mRNA transcription start sites (TSSs) [28]. We were thus prompted to identify position bias for LTRs relative to TSSs among the three lncRNA subtypes. The LTR/ERV1 showed a large coverage peak right at the TSS of lincRNAs, whereas a deletion of the LTR/ERV1 was observed in the anti-sense lncRNAs and intronic lncRNAs (Fig. 3b). Furthermore, the anti-sense lncRNAs and intronic lncRNAs also exhibited a relatively higher coverage of LINEs/L1s at its TSSs. These findings were suggestive of differential mechanisms of transcription regulation among the lincRNAs and the other two lncRNA subtypes.

Comparison of features of mRNAs and lncRNAs

In the present study, we obtained a total of 27,947 mRNAs and 1336 lncRNAs from goat fetal skin. To comprehensively examine the differences between the two transcript species, comparative analysis of gene structure, expression, and sequence conservation was performed. Our results showed that 1) most of lncRNAs contained two or three exons, which differs from that of
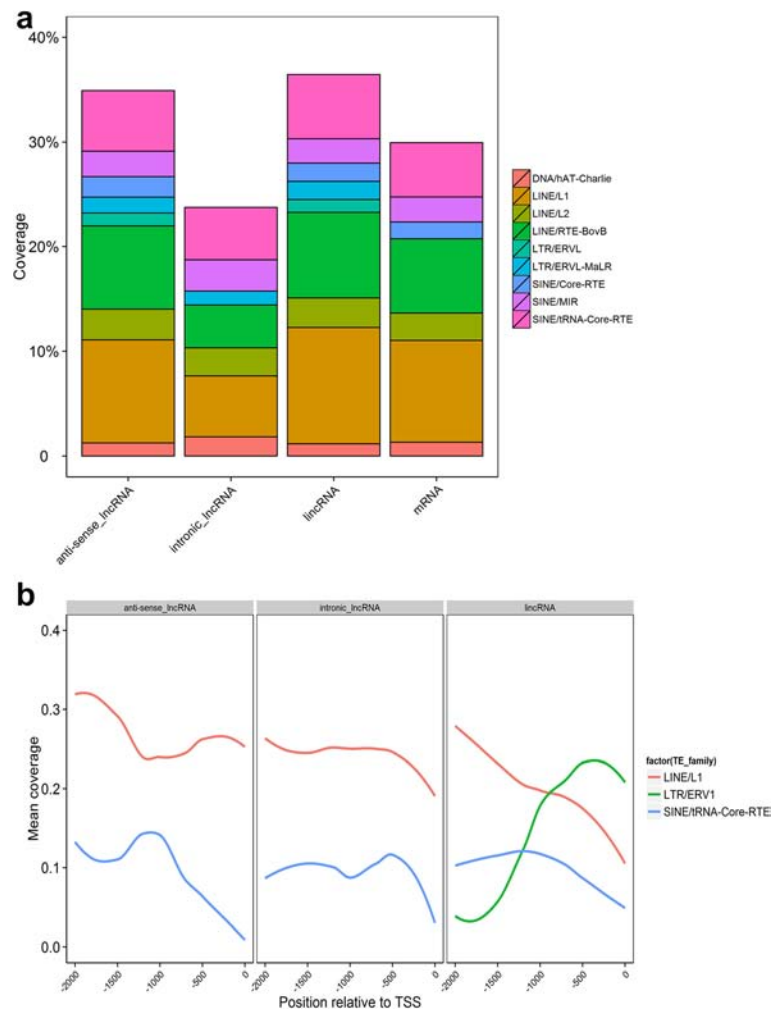
Ren *et al. BMC Genomics* (2016) 17:67

Page 4 of 12

Fig. 3 TE components and position bias in three subtypes of goat lncRNAs. The main TE families were identified using RepeatMasker in the 999 lincRNAs, 218 intronic_lncRNAs, 119 anti-sense_lncRNAs, and 27,947 mRNAs, respectively. Differences in TE components between mRNAs and individual subtype of lncRNAs were measured by using the Fisher Exact test (**a**). To ascertain the position bias of TE components in three subtypes of lncRNAs, we identified TEs approaching TSSs in three classes of lncRNAs in the goat genome (http://goat.kiz.ac.cn) and plotted the coverage of various TE families (**b**)

mRNAs (Fig. 4a); 2) there was a distinct divergence in the distribution of transcript length between mRNAs and lncRNAs (Fig. 4b); 3) most of the lncRNAs contained a relatively shorter ORF, compared to that of mRNAs (Fig. 4c); 4) lncRNAs generally showed a lower level of expression compared to that observed in mRNAs (Fig. 4d); 5) lncRNAs often generate a lower number of alternatively spliced transcripts, in contrast to that in mRNAs (Fig. 4e); and 6) most lncRNAs are slightly less conserved, although not statistically significant (Fig. 4f).

The *cis* and *trans* role of lncRNAs in target genes
To investigate the function of lncRNAs, we predicted the potential targets of lncRNAs in *cis* and *trans*. For the *cis* action of lncRNAs, we searched for protein-coding genes

10 and 100 kb upstream and downstream of the lncRNAs, respectively. Our results included 641 lncRNAs that corresponded to 868 protein-coding genes within a range of 10 kb, as well as 964 lncRNAs that represented 3468 protein-coding genes within a range of 100 kb (Additional file 4). Interestingly, we detected melanogenic genes such as *ASIP*, *Mitf*, *Sox10*, *Wnt7b*, and *Wnt3a*, which were respectively located near the *XLOC_005274*, *XLOC_013722*, *XLOC_020482*, *XLOC_020548*, and *XLOC_022579* loci, thereby suggesting that skin melanogenesis is regulated by the action of five lncRNAs on neighboring protein-coding genes. Gene Ontology (GO) analysis of *cis* lncRNA targets demonstrated that 25 significantly overrepresented terms were mainly involved in the regulation of gene expression. For example, the top five terms were sequence-specific DNA binding, nucleic acid binding transcription factor
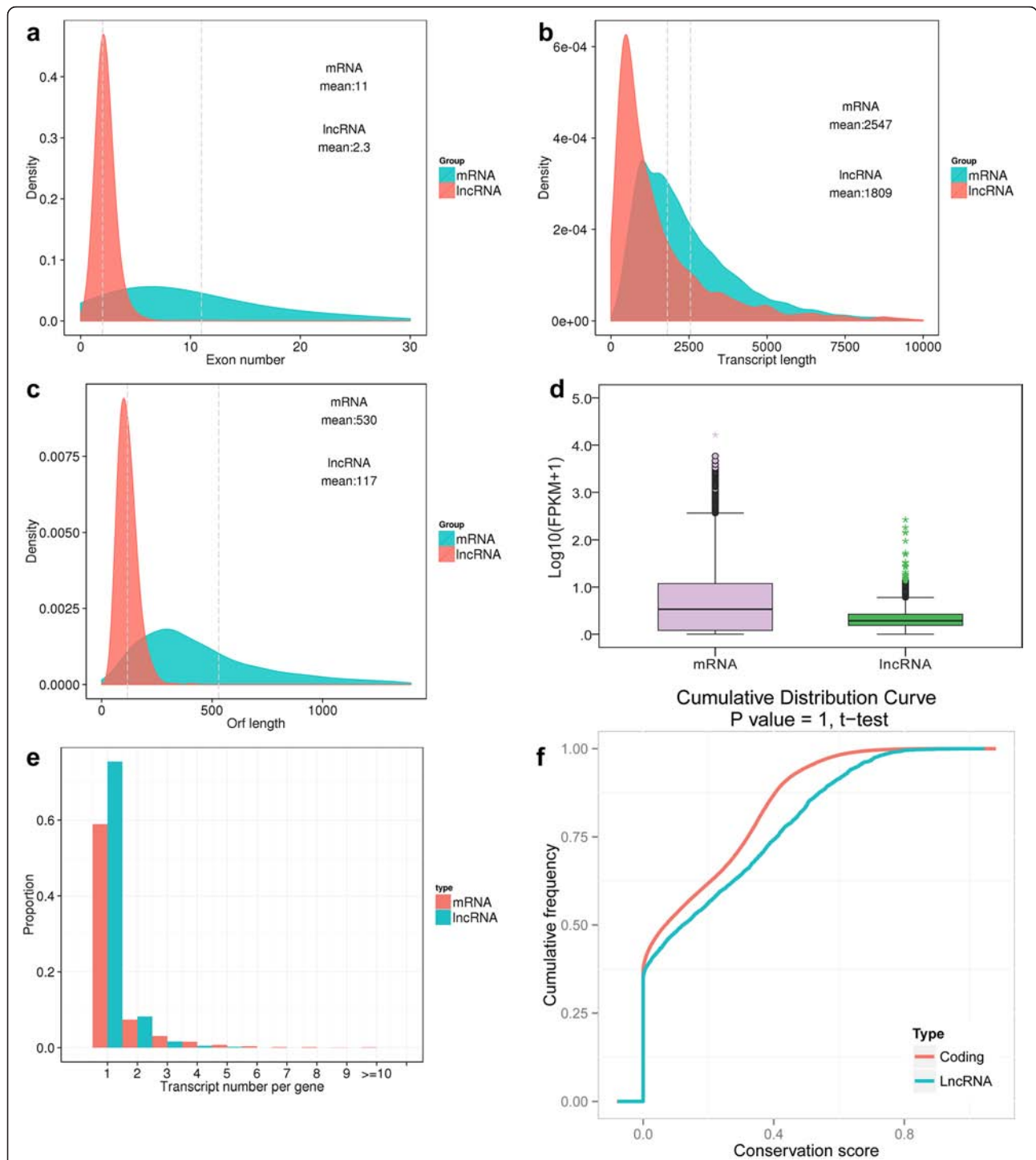
Ren *et al. BMC Genomics* (2016) 17:67

Page 5 of 12

Fig. 4 Comparison of genomic architecture and expression level between mRNAs and lncRNAs. The mRNAs and lncRNAs identified in the present study were used for comparison of primary differences in two classes of transcripts. **a** Distribution of transcript lengths in the mRNAs and lncRNAs in skin. The horizontal axis of indicates the length of transcripts, and the vertical axis represents density. **b** Distribution of the number of exons in the mRNAs and lncRNAs. In the present study, single-exon lncRNAs were filtered out from the goat genome due to the limitations of the algorithm. **c** Distribution of the number of open reading frames (Orfs) in the mRNAs and lncRNAs. The Orf was identified using Estscan in the present study. **d** Expression level indicated by $\log_{10}(FPKM + 1)$ in the mRNAs and lncRNAs. **e** Proportional distribution of alternative splicing transcripts in mRNAs and lncRNAs. **f** Conservation of the sequence in mRNAs and lncRNAs were evaluated using phastCons (http://compgen.bscb.cornell.edu/phast/)

Ren *et al. BMC Genomics* (2016) 17:67

Page 6 of 12

activity, sequence-specific DNA binding transcription factor activity, regulation of transcription, and DNA-dependent regulation of RNA metabolic processes. These findings clearly demonstrated one of the roles of lncRNAs in the genome, namely, regulation of gene expression. Pathway analysis showed that these *cis* target genes of lncRNAs were enriched in 266 KEGG pathways, in which several pathways were related to pigmentation such as tyrosine metabolism, cAMP signaling pathway, MAPK signaling pathway, Wnt signaling pathway, melanogenesis, and melanoma (Additional file 5). These findings suggested that lncRNAs act on its neighboring protein-coding genes in *cis* to regulate skin pigmentation during dermal development.

On the other hand, the *trans* role of 1336 lncRNAs in protein-coding genes was examined based on its expression correlation coefficient (Pearson correlation ≥ 0.95 or ≤ −0.95). A total of 123,969 interaction relationships were detected in *trans* between 1150 lncRNAs and the protein-coding genes in the goat genome (Additional file 6). Functional analysis illustrated that the *trans* target genes were enriched in 2643 GO terms, encompassing a variety of biological processes. Importantly, we observed a few of melanogenic terms, including pigment biosynthetic process, tyrosine 3-monooxygenase activity, melanin-concentrating hormone activity, pigment metabolic process, nitrogen compound metabolic process, and others. Of the 256 KEGG pathways identified, five were associated with pigmentation such as melanogenesis, melanoma, Wnt signaling pathway, cAMP signaling pathway, and tyrosine metabolism (Additional file 7). These findings indicated that lncRNAs act on the protein-coding genes associated with skin pigmentation in *trans*.

To further ascertain lncRNA-protein-coding gene pairs that belong to both co-localization (*cis* action) and expression correlation (*trans* action) relationships, detailed examination was conducted, which identified 26 lncRNA-protein coding gene pairs that fulfilled to these criteria (Additional file 8). This finding suggested that lncRNAs act on its neighboring protein-coding genes to regulate gene expression. We also noticed that one lncRNA, *XLOC_020022*, which was significantly differentially expressed between goats with dark skin and white skin, interacted with an early development-related gene, *HOXC11*.

Tissue and functional specificities of lncRNAs

Expression correlation analysis revealed an interesting phenomenon wherein an lncRNA in *trans* acted on two protein-coding genes that were specifically expressed in a particular type of cell or belonged to a certain functional cluster. For example, *XLOC_013372* targets *ASIP* and *MITF*, yet with opposite correlations.

A group of lncRNAs, including *XLOC_023806*, *XLOC_019686*, *XLOC_008226*, *XLOC_013939*, *XLOC_015399*, *XLOC_017870*, *XLOC_000404*, and *XLOC_002582* simultaneously act on both *TYRP1* and *DCT*; *XLOC_010430*, *XLOC_000995*, *XLOC_019547*, *XLOC_009688*, *XLOC_005961*, and *XLOC_006605* target both *WNT2* and *CREB3L1*; and lncRNAs target *WNT2* and *FZD4*, respectively. On the other hand, *ASIP*, *MITF*, *TYRP1*, and *DCT* are members of melanogenic pathways and are expressed specifically in melanocytes. These findings indicate that the lncRNAs are tissue- or function-specific. Furthermore, the three unique differentially expressed lncRNAs (*XLOC_010430*, *XLOC_004341*, and *XLOC_015448*) between the normal and dark skin in goats require further investigation because their targets were also differentially expressed, except for *MITF* (the target of *XLOC_015448*). We suspect that these lncRNAs most probably participated in the regulation of melanogenesis, although its underlying mechanisms require additional investigations. Selected lncRNAs and target genes related to pigmentation were validated by quantitative PCR analysis (Fig. 5, Table 1).

Discussion

In the present study, we identified a total of 1336 multiple-exon lncRNAs in 100-day-old fetal goat skin. In contrast to the number of protein-coding genes identified in the present study (27,947 mRNAs), the expression of lncRNAs was tissue-specific [1]. Comparative analysis of lncRNAs and mRNAs revealed characteristics that were similar to those of recent studies [1–5]. In addition to the preliminary examination of the three lncRNA subtypes, our extensive characterization revealed major differences in TE components (LINEs/L1s, LINEs/RTE-BovBs, LTRs/ERVLs, and SINEs/Core-RTEs) among lincRNAs, intronic lncRNAs, and anti-sense lncRNAs (Fig. 3a), which may in turn be responsible for the observed differences in their evolution and function. Because about half of mammalian genomes consist of lncRNAs [25], our results might provide insights into the scenario of genome evolution via lncRNA evolution. LTRs are known to harbor promoter signals that modulate gene expression in genomes [26, 27]. Our findings have demonstrated that lincRNAs are highly enriched with LTRs/ERV1 at TSSs, but absent in anti-sense lncRNAs and intronic lncRNAs (Fig. 3b), which suggest that the regulatory mechanism of expression of lincRNAs differs from that of the other two subtypes. A recent study showed distinct differences in TE density and position bias between the lincRNAs and mRNAs [28], whereas the present study improves our understanding of lncRNA biology.
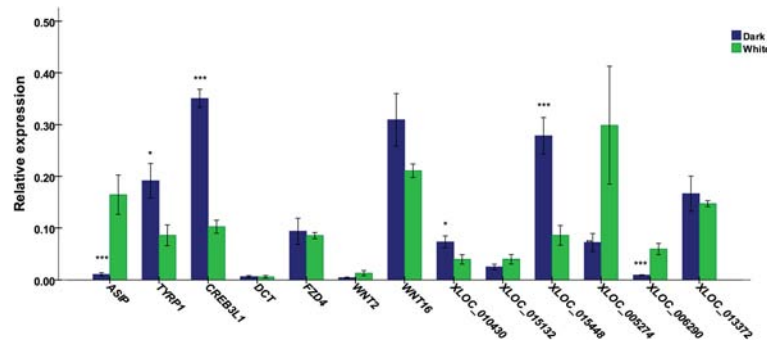
Ren *et al. BMC Genomics* (2016) 17:67

Page 7 of 12

Fig. 5 Validation of gene expression in dark and white skin by quantitative PCR. Some identified melanogenic genes and lncRNAs were examined in dark and white skin of fetal goats using quantitative PCR. Gene expression was quantified relative to the expression level of *β-actin* using the comparative cycle threshold (ΔCT) method. Correction for multiple comparisons was performed using the Holm-Sidak method. The data are expressed as the mean ± 1 SE ($n = 3$). * $P < 0.05$, **$P < 0.01$, *** $P < 0.001$

In 2013, the goat genome was sequenced and assembled *de novo* using whole-genome mapping technology [29], which endows its high quality of genome assembly and annotation among farm animals, including horse, pig, cattle, yak, and sheep [30–34]. Although some single-exon lncRNAs were filtered out of the goat genome due to the limitations of the algorithm of the present study, authentic multiple-exon lncRNAs were generated, which could then be utilized in future investigations, as well as considerably improve the annotation of the goat genome. On the other hand, unlike protein-coding genes where sequence motifs are usually indicative of function, lncRNA sequences are currently uninformative for predicting function. This particular

limitation hinders the prediction of the function of lncRNAs. Interestingly, several lncRNAs are transcribed with their associated protein-coding transcripts [35], and various examples of recently characterized ncRNAs such as Evf2 [36], HOTAIR [37], Kcnq1ot1 [38], and Air [39] support a functional relationship between lncRNAs and its associated or related-protein coding gene(s). Therefore, functional predictions for mammalian lncRNAs have often been based on guilt-by-association analyses [1, 5, 40–42], although this may not be the most appropriate model to explain the function of lncRNAs.

We predicted the potential function of lncRNAs in goat skin and determined that protein-coding genes can act with lncRNAs in *cis* or *trans*. In particular, *ASIP,*

Table 1 LncRNAs and its potential target genes that are involved in melonagenesis

Protein-coding genes	lncRNAs in *cis*	lncRNAs in *trans*
ASIP[a]	XLOC_005274	*XLOC_013372, XLOC_006290, XLOC_013615, XLOC_024549,*
	XLOC_004858	*XLOC_000129, XLOC_006932, XLOC_020962, -XLOC_020019*[b]*, -XLOC_018830*[b]
MITF	XLOC_013722	*-XLOC_013372*[b]*, -XLOC_022057*[b]*, XLOC_000689, XLOC_018361, XLOC_012968, XLOC_009509, XLOC_023640, -XLOC_013890*[b]*, XLOC_005088, XLOC_015448*[a]
TYRP1[a] *(BROWN)*		*XLOC_023806, XLOC_019686, XLOC_008226, XLOC_013939, XLOC_015399, XLOC_017870, XLOC_000404, XLOC_002582*
DCT[a]		*XLOC_023806, XLOC_019686, XLOC_008226, XLOC_013939, XLOC_015399, XLOC_017870, XLOC_000404, XLOC_002582, -XLOC_005274*[b]
TYR		*-XLOC_010559*[b]*, -XLOC_024478*[b]*, XLOC_021855, XLOC_006064*
CREB3L1[a]		*XLOC_000995, -XLOC_009688*[b]*, XLOC_005961, XLOC_006605, XLOC_004319, XLOC_008730, XLOC_010430*[a]*, XLOC_019547, -XLOC_000912*[b]*, XLOC_024598, XLOC_004263*
FZD4[a]		*XLOC_000995, XLOC_018035, XLOC_005957, XLOC_006451, XLOC_004319, XLOC_020603, XLOC_009285, XLOC_019333, XLOC_004805, XLOC_008730, XLOC_010430*[a]*, XLOC_006605, XLOC_003840, XLOC_002867*[a]*, XLOC_023214, XLOC_023692, XLOC_002389*[a]
WNT2[a]		*XLOC_000995, XLOC_018035, XLOC_005957, XLOC_006451, XLOC_004319, XLOC_020603, XLOC_009285, XLOC_019333, XLOC_004805, XLOC_008730, XLOC_010430*[a]*, XLOC_006605, XLOC_003840, XLOC_019547, -XLOC_009688*[b]*, XLOC_005961, XLOC_004341*[a]*, -XLOC_014182*[b]*, XLOC_013150, XLOC_022462, XLOC_013012, XLOC_025297*[a]*, XLOC_008538, -XLOC_007438*[b]*, XLOC_005975, XLOC_004597*

[a]Differentially expressed in dark and normal skins of goats
[b]Negative correlation between the lncRNAs and their targets in trans. Italic font indicates that one lncRNA acts on at least two different protein-coding genes in trans. For example, *XLOC_013372* regulated *ASIP* and *MITF* in *trans*

Ren *et al. BMC Genomics* (2016) 17:67

Page 8 of 12

which was most differentially expressed in dark and white skin as indicated by RNA-seq analysis, was determined to be regulated by several lncRNAs both in *cis* and in *trans* (Table 1). However, the mechanism by which the lncRNAs act on *ASIP* in *cis* and *trans* remains to be elucidated. An intriguing observation is that *XLOC_013372* acts on both *ASIP* and *MITF* in *trans*, but with reverse correlations (Table 1). This is the first report of such an interesting observation, which is worth investigating further, as well as indicates the functional complexity of lncRNAs. Furthermore, a certain cluster of lncRNAs in *trans* often target protein-coding genes that were specifically expressed in melanocytes (*ASIP*, *MITF*, *TYPR1*, and *DCT*) and/or involved in melanogenesis (*WNT2*, *WNT16*, *FZD4*, and *CREB3L1*) (Table 1). This finding indicates that lncRNAs play a regulatory role in melanogenesis. Moreover, a cluster of eight lncRNAs act on both *TYRP1* and *DCT*, which evolved from a common ancestral tyrosinase gene [43–45]. The observation of highly identical regulatory lncRNAs suggests that these homologous sequences in the tyrosinase family genes are involved in its evolution and functionality. A third interesting observation is that *FZD4* and *WNT2*, which are members of the WNT signaling pathway, share a few of regulatory lncRNAs, including significantly differentially expressed *XLOC_010430*. This again indicates that lncRNAs are highly functionally conserved, similar to their targets, namely, the WNT signaling proteins [46]. Several recent studies also indicate that lncRNAs are conserved in function [41, 47, 48]. Functional conservation, despite variations in sequence, is a characteristic of lncRNAs. The differentially expressed lncRNAs between dark and white skin in goats such as *XLOC_015448*, *XLOC_002867*, *XLOC_002389*, *XLOC_010430*, *XLOC_004341*, and *XLOC_025297* (Table 1), as well as the 26 lncRNA-protein coding gene pairs that belong to both co-localization (*cis* role) and correlation (*trans* role) (Additional file 8) require additional investigations.

As far as we know, only a small portion of the pathways involved in pigmentation have been validated to date, including the protein kinase C pathway [49, 50], cAMP pathway [51], SCF-KIT pathway [52], cGMP pathway [53], phosphatidylinositol 3-kinase-Akt pathway [54], protein kinase A pathway [55], BMP signaling [56], Notch pathway [57], ERK pathway [58], Wnt signal [59], KITLG and the KITLG/c-Kit pathway [60], CXCR3-mediated pathway [61], JAK2-STAT6 signaling pathway [62], nitric oxide/protein kinase G signaling pathway [63], FGF/MAPK/Ets signaling [64], p38MAPK [65], MITF-GPNMB pathway [66], Galphas-SASH1-IQGAP1-E-cadherin pathway [67], CREB/MITF/tyrosinase pathway [68], and necrosis factor receptor-associated factor 2 (TRAF2)-caspases pathway [69]. However, reports on

the role of lncRNA in pigmentation are limited. In the present study, the enriched KEGG pathways associated with pigmentation (Additional files 5 and 7) in the potential lncRNA targets clearly indicated that these lncRNAs play roles in skin pigmentation in goats. However, the predicted targets based on guilt-by-association analyses should be carefully assessed because of the low number of sample examined, and experimental validations are also warranted.

Conclusions

We elucidated the skin lncRNA profiles of fetal goats using deep RNA-seq analysis. The characterization of three lncRNA subtypes casts light on the mechanism underlying the origin and evolution of lncRNAs, as well as its regulation of expression. LncRNAs are tissue-specific and functionally conserved during skin development and pigmentation in goats. Our findings have further expanded our knowledge on lncRNA biology, as well as contributed to the annotation of the goat genome. The present study also provides valuable resources for studying lncRNAs.

Methods

Animals

Two goat groups with diverse phenotypes of skin pigmentation were investigated in this study. The Yudong white goat (*Capra hircus*) is distributed in Southwest China (located at 31°14′–32°12′ N and 108°15′–109°58′ E), which features white color coat and skin. The Youzhou dark goat (*Capra hircus*), a indigenous breed uniquely distributed in Youyang county in Chongqing, China (located at 26°54′ N and 108°57′ E), is characterized by dark skin, including the visible mucous membranes, yet is generally white in coat color. Briefly, three pregnant ewes from each breed were subjected to caesarean section to collect the fetuses at 100 days of gestation, and then the dorsal and ventral skins were collected from each fetus. Three grams of skin were dissected and rapidly frozen in liquid nitrogen for RNA extraction.

All surgical procedures involving goats 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 adhered to the Reporting Guidelines for Randomized Controlled Trials in Livestock and Food Safety (REFLECT).

RNA isolation, library preparation, and sequencing

Total RNA was isolated using TRIzol reagent (Invitrogen, Carlsbad, CA, USA), according to the manufacturer's instructions. RNA degradation and contamination was monitored on 1 % agarose gels. RNA purity was checked using the NanoPhotometer spectrophotometer (Implen,

Ren *et al. BMC Genomics* (2016) 17:67

Page 9 of 12

Los Angeles, CA, USA). RNA concentration was measured using a Qubit RNA Assay Kit in a Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA). RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA). Approximately 3 μg RNA per sample was used as input material for the RNA sample preparations. First, ribosomal RNA was removed by using an Epicentre Ribo-zero rRNA Removal Kit (Epicentre, Madison, WI, USA), and rRNA-free residue was removed by ethanol precipitation. Subsequently, the high strand-specificity of the libraries was generated using the rRNA-depleted RNA of the NEBNext Ultra Directional RNA Library Prep Kit for Illumina (NEB, Ipswich, MA, USA), following manufacturer's recommendations. Briefly, fragmentation was conducted using divalent cations under elevated temperature in NEBNext. First-strand cDNA was synthesized using random hexamer primers and M-MuLV Reverse Transcriptase (RNaseH-). Second-strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. In the reaction buffer, dNTPs with dTTP were replaced by dUTP. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of the 3′ ends of the DNA fragments, NEBNext adaptors with a hairpin loop structure were ligated to prepare for hybridization. To preferentially select cDNA fragments of 150–200 bp in length, the library fragments were purified with an AMPure XP system (Beckman Coulter, Brea, CA, USA). Then 3 μL USER Enzyme (NEB, Ipswich, MA, USA) was used with size-selected, adaptor-ligated cDNA at 37 °C for 15 min followed by 5 min at 95 °C before PCR. Then, PCR was performed with Phusion High-Fidelity DNA polymerase, universal PCR primers, and Index (X) Primers. Finally, the PCR products were purified (AMPure XP system), and library quality was assessed on an Agilent Bioanalyzer 2100 system. Clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumina, San Diego, CA, USA), following the manufacturer's instructions. After cluster generation, the libraries were sequenced on an Illumina Hiseq 2000 platform and 100-bp paired-end reads were generated.

Quality control

Raw data were first processed using in-house Perl scripts. In this step, clean data were obtained by removing reads containing adapter, reads containing over 10 % of ploy-N, and low-quality reads (>50 % of bases whose Phred scores were <5 %) from the raw data. The Phred score (Q20, Q30) and GC content of the clean data were calculated. All subsequent analyses were based on the high-quality data.

Transcriptome assembly

Goat reference genome and gene model annotation files were downloaded from the goat genome website (http://goat.kiz.ac.cn) directly. Index of the reference genome was built using Bowtie v2.0.6 [70, 71] and paired-end clean reads were aligned to the reference genome using TopHat v2.0.9 [72, 73]. The mapped reads of each sample were assembled using both Scripture (beta2) [74] and Cufflinks (v2.1.1) [20] in a reference-based approach. Scripture was run with default parameters. Cufflinks was run with min-frags-per-transfrag = 0′ and –library-type fr-firststrand′, and other parameters were set as default. In the present study, single-exon lncRNAs were filtered out from the goat genome due to limitations of the algorithm. This operation that at least two exons are preferred is a purely technical one. To avoid false-positive results as much as possible, the transcripts with a single exon were usually considered as background transcripts and were discarded, whereas multiple-exon lncRNAs were retained [75].

Quantification of gene expression level

Cuffdiff (v2.1.1) was used to calculate fragments per kb for a million reads (FPKM) of both lncRNAs and coding genes in each sample [20]. For biological replicates, transcripts or genes with a *P*-adjust of <0.05 were described as differentially expressed between two groups of goats with the dark and white skin.

Coding potential and conserved analysis of lncRNAs

To achieve high-quality data, we used four analytic tools, including CNCI (v2) [76], CPC (0.9-r2) [77], Pfam-scan (v1.3) [78], and PhyloCSF (v20121028) [79] to identify the candidate lncRNAs. Transcripts predicted with coding potential by any of the four tools earlier described were filtered out, and those without coding potential were retained. Then, we selected those shared by four tools as the final candidate lncRNAs and use for further analysis. Quantification of gene expression level was estimated by calculating the FPKMs of the transcripts. The pipeline used to identify putative lncRNAs from the deep sequencing data is presented in Supplementary Figure S1.

To investigate the sequence conservation of transcripts, we used the phyloFit program in the Phast (v1.3) package [80] to compute phylogenetic models for conserved and non-conserved regions among species. Then, we used phastCons to compute a set of conservation scores of lncRNAs and coding genes.

LncRNA target gene prediction and functional enrichment analysis

To explore the function of lncRNAs, we first predicted the target genes of lncRNAs *in cis* and *trans*. The *cis* role

Ren *et al. BMC Genomics* (2016) 17:67

Page 10 of 12

refers to lncRNAs' action on neighboring target genes. In the present study, we searched coding genes 10/100 k upstream and downstream of an lncRNA. The *trans* role refers to the influence of lncRNAs on other genes at the expression level. Here, we calculated for Pearson's correlation coefficients between expression levels of 1336 lncRNAs and 27,947 mRNAs with custom scripts (r > 0.95 or r < −0.95). Then, we performed functional enrichment analysis of the target genes for lncRNAs by using the DAVID platform [81, 82]. Significance was expressed as a *p*-value, which was calculated using the EASE score (*P* value < 0.05 was considered significant).

Enrichment analysis of TE in goat lncRNAs

RepeatMasker (http://www.repeatmasker.org) was used with default parameters to identify various TE components in goat. To detect position bias of TEs in each class of lncRNAs, we searched for TEs at the 2000-bp upstream region of the TSS of each lncRNA identified in the goat genome (http://goat.kiz.ac.cn) and plotted its read coverage with the ggplot2 package in R [83].

Validation of gene expression in RNA-seq by quantitative PCR analysis

Total RNAs from fetal skin in two groups of goats were used for quantitative PCR analysis. Briefly, the first cDNA strains were obtained using a One Step cDNA Synthesis Kit (Bio-Rad, USA), and were then subjected to quantification of mRNAs or lncRNAs with *β-actin* as an endogenous control using a standard SYBR Green PCR kit (Bio-Rad) on the Bio-Rad CFX96 Touch™ Real-Time PCR Detection System. The quantitative PCR was performed using the following conditions: 95 °C for 30 s, 40 cycles of 95 °C for 5 s, and the optimized annealing temperature for 30 s. The primers and annealing temperatures for 14 genes are listed in Additional file 9. All reactions were performed in triplicate for each sample. Gene expression was quantified relative to *β-actin* expression using the comparative cycle threshold (ΔCT) method. Differences in gene expression between the dark and white skin were detected by using the t-test. Corrections for multiple comparisons were performed using the Holm-Sidak method.

Statistical analysis

Data analyses were performed using the statistical R package.

Data availability

The sequencing data were submitted to the Genome Expression Omnibus (Accession Numbers GSE69812) in NCBI.

Additional files

Additional file 1: Pipeline for lncRNA analysis of the present study. (DOCX 13 kb)

Additional file 2: Comparison of TE components between individual subtypes of lncRNAs and mRNAs. (XLSX 17 kb)

Additional file 3: Comparison of TE components among three subtypes of lncRNAs. (XLSX 13 kb)

Additional file 4: The protein-coding genes within 10 k and 100 k upstream and downstream of an lncRNA. (XLSX 55 kb)

Additional file 5: Functional enrichment analysis of the protein-coding genes of lncRNAs in *cis*. (XLS 360 kb)

Additional file 6: Pearson correlations between protein-coding genes and lncRNAs. (LIST 1829 kb)

Additional file 7: Functional enrichment analysis of protein-coding genes of lncRNAs in *trans*. (XLS 7903 kb)

Additional file 8: LncRNA-protein coding gene pairs with both co-localization and correlation relationships. (XLS 21 kb)

Additional file 9: Primers used in quantitative PCR analysis. (XLSX 9 kb)

Author details
[1]Chongqing Academy of Animal Sciences, Chongqing 402460, China. [2]Chongqing Engineering Research Center for Goats, Chongqing 402460, China. [3]Youyang Animal Husbandry Bureau, Chongqing 409800, China.

References
1. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev. 2011;25:1915–27.
2. Li T, Wang S, Wu R, Zhou X, Zhu D, Zhang Y. Identification of long non-protein coding RNAs in chicken skeletal muscle using next generation sequencing. Genomics. 2012;99:292–8.

Ren *et al. BMC Genomics* (2016) 17:67

Page 11 of 12

3. Zhou ZY, Li AM, Adeola AC, Liu YH, Irwin DM, Xie HB, et al. Genome-wide identification of long intergenic noncoding RNA genes and their potential association with domestication in pigs. Genome Biol Evol. 2014;6:1387–92.

4. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res. 2012;22:1775–89.

5. Pauli A, Valen E, Lin MF, Garber M, Vastenhouw NL, Levin JZ, et al. Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. Genome Res. 2012;22:577–91.

6. Nam JW, Bartel DP. Long noncoding RNAs in C. elegans. Genome Res. 2012; 22:2529–40.

7. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. Nat Rev Genet. 2009;10:155–9.

8. Moran VA, Perera RJ, Khalil AM. Emerging functional and mechanistic paradigms of mammalian long non-coding RNAs. Nucleic Acids Res. 2012; 40:6391–400.

9. Wilusz JE, Sunwoo H, Spector DL. Long noncoding RNAs: functional surprises from the RNA world. Genes Dev. 2009;23:1494–504.

10. Wan DC, Wang KC. Long noncoding RNA: significance and potential in skin biology. Cold Spring Harb Perspect Med. 2014;4(5). doi:10.1101/cshperspect. a015404.

11. Jiang YJ, Bikle DD. LncRNA profiling reveals new mechanism for VDR protection against skin cancer formation. J Steroid Biochem Mol Biol. 2014;144 Pt A:87–90.

12. Volders PJ, Verheggen K, Menschaert G, Vandepoele K, Martens L, Vandesompele J, et al. An update on LNCipedia: a database for annotated human lncRNA sequences. Nucleic Acids Res. 2015;43:D174–80.

13. Quek XC, Thomson DW, Maag JL, Bartonicek N, Signal B, Clark MB, et al. lncRNAdb v2.0: expanding the reference database for functional long noncoding RNAs. Nucleic Acids Res. 2015;43:D168–73.

14. Bu D, Yu K, Sun S, Xie C, Skogerbo G, Miao R, et al. NONCODE v3.0: integrative annotation of long noncoding RNAs. Nucleic Acids Res. 2012;40:D210–5.

15. Huang W, Long N, Khatib H. Genome-wide identification and initial characterization of bovine long non-coding RNAs from EST data. Anim Genet. 2012;43:674–82.

16. Billerey C, Boussaha M, Esquerre D, Rebours E, Djari A, Meersseman C, et al. Identification of large intergenic non-coding RNAs in bovine muscle using next-generation transcriptomic sequencing. BMC Genomics. 2014;15:499.

17. Weikard R, Hadlich F, Kuehn C. Identification of novel transcripts and noncoding RNAs in bovine skin by deep next generation sequencing. BMC Genomics. 2013;14:789.

18. Zhao W, Mu Y, Ma L, Wang C, Tang Z, Yang S, et al. Systematic identification and characterization of long intergenic non-coding RNAs in fetal porcine skeletal muscle development. Sci Rep. 2015;5:8957.

19. Wang L, Peng L, Zhang W, Zhang Z, Yang W, Ding L, et al. Initiation and development of skin follicles in the inner Mongolian Cashmere goat. Acta Veterinaria et Zootechnica Sinica. 1996;27:524–30.

20. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 2010;28:511–5.

21. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. 2009;25:1105–11.

22. Griffith M, Griffith OL, Mwenifumbo J, Goya R, Morrissy AS, Morin RD, et al. Alternative expression analysis by RNA sequencing. Nat Methods. 2010;7: 843–7.

23. Gregory TR. Synergy between sequence and size in large-scale genomics. Nat Rev Genet. 2005;6:699–708.

24. de Koning AP, Gu W, Castoe TA, Batzer MA, Pollock DD. Repetitive elements may comprise over two-thirds of the human genome. PLoS Genet. 2011;7, e1002384.

25. Kapusta A, Feschotte C. Volatile evolution of long noncoding RNA repertoires: mechanisms and biological implications. Trends Genet. 2014;30:439–52.

26. Ju G, Cullen BR. The role of avian retroviral LTRs in the regulation of gene expression and viral replication. Adv Virus Res. 1985;30:179–223.

27. Cohen CJ, Lock WM, Mager DL. Endogenous retroviral LTRs as promoters for human genes: a critical assessment. Gene. 2009;448:105–14.

28. Kelley D, Rinn J. Transposable elements reveal a stem cell-specific class of long noncoding RNAs. Genome Biol. 2012;13:R107.

29. Dong Y, Xie M, Jiang Y, Xiao N, Du X, Zhang W, et al. Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (Capra hircus). Nat Biotechnol. 2013;31:135–41.

30. Jiang Y, Xie M, Chen W, Talbot R, Maddox JF, Faraut T, et al. The sheep genome illuminates biology of the rumen and lipid metabolism. Science. 2014;344:1168–73.

31. Qiu Q, Zhang G, Ma T, Qian W, Wang J, Ye Z, et al. The yak genome and adaptation to life at high altitude. Nat Genet. 2012;44:946–9.

32. Groenen MA, Archibald AL, Uenishi H, Tuggle CK, Takeuchi Y, Rothschild MF, et al. Analyses of pig genomes provide insight into porcine demography and evolution. Nature. 2012;491:393–8.

33. Gibbs RA, Taylor JF, Van Tassell CP, Barendse W, Eversole KA, Gill CA, et al. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. Science. 2009;324:528–32.

34. Wade CM, Giulotto E, Sigurdsson S, Zoli M, Gnerre S, Imsland F, et al. Genome sequence, comparative analysis, and population genetics of the domestic horse. Science. 2009;326:865–7.

35. Engstrom PG, Suzuki H, Ninomiya N, Akalin A, Sessa L, Lavorgna G, et al. Complex Loci in human and mouse genomes. PLoS Genet. 2006;2, e47.

36. Feng J, Bi C, Clark BS, Mady R, Shah P, Kohtz JD. The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. Genes Dev. 2006;20:1470–84.

37. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell. 2007;129:1311–23.

38. Thakur N, Tiwari VK, Thomassin H, Pandey RR, Kanduri M, Gondor A, et al. An antisense RNA regulates the bidirectional silencing property of the Kcnq1 imprinting control region. Mol Cell Biol. 2004;24:7855–62.

39. Sleutels F, Zwart R, Barlow DP. The non-coding Air RNA is required for silencing autosomal imprinted genes. Nature. 2002;415:810–3.

40. Dinger ME, Amaral PP, Mercer TR, Pang KC, Bruce SJ, Gardiner BB, et al. Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. Genome Res. 2008;18:1433–45.

41. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature. 2009;458:223–7.

42. Gaiti F, Fernandez-Valverde SL, Nakanishi N, Calcino AD, Yanai I, Tanurdzic M, et al. Dynamic and Widespread lncRNA Expression in a Sponge and the Origin of Animal Complexity. Mol Biol Evol. 2015;32:2367–82.

43. Sturm RA, O'Sullivan BJ, Box NF, Smith AG, Smit SE, Puttick ER, et al. Chromosomal structure of the human TYRP1 and TYRP2 loci and comparison of the tyrosinase-related protein gene family. Genomics. 1995; 29:24–34.

44. Oetting WS, Stine OC, Townsend D, King RA. Evolution of the tyrosinase related gene (TYRL) in primates. Pigment Cell Res. 1993;6:171–7.

45. Jackson IJ. Evolution and expression of tyrosinase-related proteins. Pigment Cell Res. 1994;7:241–2.

46. Croce JC, McClay DR. Evolution of the Wnt pathways. Methods Mol Biol. 2008;469:3–18.

47. Ulitsky I, Shkumatava A, Jan CH, Sive H, Bartel DP. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. Cell. 2011;147:1537–50.

48. Nitsche A, Rose D, Fasold M, Reiche K, Stadler PF. Comparison of splice sites reveals that long noncoding RNAs are evolutionarily well conserved. RNA. 2015;21:801–12.

49. Imokawa G, Yada Y, Okuda M. Allergic contact dermatitis releases soluble factors that stimulate melanogenesis through activation of protein kinase C-related signal-transduction pathway. J Invest Dermatol. 1992;99:482–8.

50. Kishi H, Mishima HK, Yamashita U. Involvement of the protein kinase pathway in melanin synthesis by chick retinal pigment epithelial cells. Cell Biol Int. 2000;24:79–83.

51. Englaro W, Rezzonico R, Durand-Clement M, Lallemand D, Ortonne JP, Ballotti R. Mitogen-activated protein kinase pathway and AP-1 are activated during cAMP-induced melanogenesis in B-16 melanoma cells. J Biol Chem. 1995;270:24315–20.

52. Grichnik JM, Burch JA, Burchette J, Shea CR. The SCF/KIT pathway plays a critical role in the control of normal human melanocyte homeostasis. J Invest Dermatol. 1998;111:233–8.

53. Romero-Graillet C, Aberdam E, Biagoli N, Massabni W, Ortonne JP, Ballotti R. Ultraviolet B radiation acts through the nitric oxide and cGMP signal transduction pathway to stimulate melanogenesis in human melanocytes. J Biol Chem. 1996;271:28052–6.

54. Oka M, Nagai H, Ando H, Fukunaga M, Matsumura M, Araki K, et al. Regulation of melanogenesis through phosphatidylinositol 3-kinase-Akt

Ren *et al. BMC Genomics* (2016) 17:67

Page 12 of 12

pathway in human G361 melanoma cells. J Invest Dermatol. 2000;115: 699–703.

55. Lei TC, Virador V, Yasumoto K, Vieira WD, Toyofuku K, Hearing VJ. Stimulation of melanoblast pigmentation by 8-methoxypsoralen:the involvement of microphthalmia-associated transcription factor, the protein kinase a signal pathway, and proteasome-mediated degradation. J Invest Dermatol. 2002;119:1341–9.

56. Sharov AA, Fessing M, Atoyan R, Sharova TY, Haskell-Luevano C, Weiner L, et al. Bone morphogenetic protein (BMP) signaling controls hair pigmentation by means of cross-talk with the melanocortin receptor-1 pathway. Proc Natl Acad Sci U S A. 2005;102:93–8.

57. Schouwey K, Beermann F. The Notch pathway: hair graying and pigment cell homeostasis. Histol Histopathol. 2008;23:609–19.

58. Li X, Guo L, Sun Y, Zhou J, Gu Y, Li Y. Baicalein inhibits melanogenesis through activation of the ERK signaling pathway. Int J Mol Med. 2010;25: 923–7.

59. Dorsky RI, Raible DW, Moon RT. Direct regulation of nacre, a zebrafish MITF homolog required for pigment cell formation, by the Wnt pathway. Genes Dev. 2000;14:158–62.

60. Picardo M, Cardinali G. The genetic determination of skin pigmentation: KITLG and the KITLG/c-Kit pathway as key players in the onset of human familial pigmentary diseases. J Invest Dermatol. 2011;131:1182–5.

61. Kurata R, Fujita F, Oonishi K, Kuriyama KI, Kawamata S. Inhibition of the CXCR3-mediated pathway suppresses ultraviolet B-induced pigmentation and erythema in skin. Br J Dermatol. 2010;163:593–602.

62. Choi H, Han J, Jin SH, Park JY, Shin DW, Lee TR, et al. IL-4 inhibits the melanogenesis of normal human melanocytes through the JAK2-STAT6 signaling pathway. J Invest Dermatol. 2013;133:528–36.

63. Choi YJ, Uehara Y, Park JY, Chung KW, Ha YM, Kim JM, et al. Suppression of melanogenesis by a newly synthesized compound, MHY966 via the nitric oxide/protein kinase G signaling pathway in murine skin. J Dermatol Sci. 2012;68:164–71.

64. Squarzoni P, Parveen F, Zanetti L, Ristoratore F, Spagnuolo A. FGF/MAPK/Ets signaling renders pigment cell precursors competent to respond to Wnt signal by directly controlling Ci-Tcf transcription. Development. 2011;138: 1421–32.

65. Huang YC, Liu KC, Chiou YL, Yang CH, Chen TH, Li TT, et al. Fenofibrate suppresses melanogenesis in B16-F10 melanoma cells via activation of the p38 mitogen-activated protein kinase pathway. Chem Biol Interact. 2013; 205:157–64.

66. Zhang P, Liu W, Yuan X, Li D, Gu W, Gao T. Endothelin-1 enhances the melanogenesis via MITF-GPNMB pathway. BMB Rep. 2013;46:364–9.

67. Zhou D, Wei Z, Deng S, Wang T, Zai M, Wang H, et al. SASH1 regulates melanocyte transepithelial migration through a novel Galphas-SASH1-IQGAP1-E-Cadherin dependent pathway. Cell Signal. 2013;25:1526–38.

68. Chiang HM, Chien YC, Wu CH, Kuo YH, Wu WC, Pan YY, et al. Hydroalcoholic extract of Rhodiola rosea L. (Crassulaceae) and its hydrolysate inhibit melanogenesis in B16F0 cells by regulating the CREB/ MITF/tyrosinase pathway. Food Chem Toxicol. 2014;65:129–39.

69. Wang Q, Jiang M, Wu J, Ma Y, Li T, Chen Q, et al. Stress-induced RNASET2 overexpression mediates melanocyte apoptosis via the TRAF2 pathway in vitro. Cell Death Dis. 2014;5, e1022.

70. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357–9.

71. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10:R25.

72. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012;7:562–78.

73. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013;14:R36.

74. Guttman M, Garber M, Levin JZ, Donaghey J, Robinson J, Adiconis X, et al. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. Nat Biotechnol. 2010;28:503–10.

75. Prensner JR, Iyer MK, Balbin OA, Dhanasekaran SM, Cao Q, Brenner JC, et al. Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lincRNA implicated in disease progression. Nat Biotechnol. 2011;29:742–9.

76. Sun L, Luo H, Bu D, Zhao G, Yu K, Zhang C, et al. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. Nucleic Acids Res. 2013;41:e166.

77. Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, et al. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. Nucleic Acids Res. 2007;35:W345–9.

78. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, et al. The Pfam protein families database. Nucleic Acids Res. 2012;40:D290–301.

79. Lin MF, Jungreis I, Kellis M. PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. Bioinformatics. 2011;27: i275–82.

80. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. 2005;15:1034–50.

81. da Huang W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4:44–57.

82. da Huang W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009;37:1–13.

83. Wickham H. Elegant graphics for data analysis. New York: Springer; 2009.