Gene 582 (2016) 69–76

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Genome-wide transcriptome analysis in the ovaries of two goats identifies differentially expressed genes related to fecundity

Xiangyang Miao *, Qingmiao Luo, Xiaoyu Qin

Institute of Animal Sciences, Chinese Academy of Agricultural Sciences, Beijing 100193, China

ARTICLE INFO

Article history:
Received 11 September 2015
Received in revised form 6 January 2016
Accepted 28 January 2016
Available online 4 February 2016

Keywords:
Goat
High prolificacy
mRNA
RNA-Seq

ABSTRACT

The goats are widely kept as livestock throughout the world. Two excellent domestic breeds in China, the Laiwu Black and Jining Grey goats, have different fecundities and prolificacies. Although the goat genome sequences have been resolved recently, little is known about the gene regulations at the transcriptional level in goat. To understand the molecular and genetic mechanisms related to the fecundities and prolificacies, we performed genome-wide sequencing of the mRNAs from two breeds of goat using the next-generation RNA-Seq technology and used functional annotation to identify pathways of interest. Digital gene expression analysis showed 338 genes were up-regulated in the Jining Grey goats and 404 were up-regulated in the Laiwu Black goats. Quantitative real-time PCR verified the reliability of the RNA-Seq data. This study suggests that multiple genes responsible for various biological functions and signaling pathways are differentially expressed in the two different goat breeds, and these genes might be involved in the regulation of goat fecundity and prolificacy. Taken together, our study provides insight into the transcriptional regulation in the ovaries of 2 species of goats that might serve as a key resource for understanding goat fecundity, prolificacy and genetic diversity between species.

1. Introduction

Previously, the most widely used method of transcriptome analysis was the microarray (Eklund et al. 2006). The data obtained from microarray analyses showed that transcriptome analysis is an effective way to analyze biological processes at the molecular level. Thus, research efforts aimed at improving the technology arose. Recently, a new whole transcriptome shotgun sequencing method was developed. This new technology, the RNA-Seq approach, uses next-generation deep-sequencing and is described as a powerful tool for RNA analysis (Wang et al. 2009). The resulting sequence reads are individually mapped to the source genome and counted to obtain the number and density of reads corresponding to the RNA from each known exon, splice event or new candidate gene (Mortazavi et al. 2008). This technique has been successfully used for genome-wide analysis of mRNAs in multiple organisms, such as yeast (Nagalakshmi et al. 2008; Wilhelm et al. 2008), sheep (Miao et al. 2015a, 2015b, 2015c, 2015d) and human (Peng et al. 2012).

The domestic goat (*Capra hircus*) is one of the oldest species and is raised all over the world for meat, milk, skin and fiber. Over 90% of the world's goat population is kept in small herds by farmers in developing countries (Mak 2013). Recently, goats have been recognized for their value as animal models in biomedical research and the genome sequence of a female Yunnan Black goat was recently published (Ko et al. 2000; Dong et al. 2013). These types of molecular analyses have been performed in other domestic animals to examine fecundity. For example, previous studies indicated that the bone morphogenetic protein receptor-1B (BMPR1B) gene regulates the fecundity and ovulation rate of sheep (Wilson et al. 2001; Davis 2005; Mulsant et al. 2001). In addition, a variety of molecular studies regarding fecundity have been conducted in sheep. Sheep fecundity is dependent on multiple genes, including the bone morphogenetic protein 15 gene (BMP15) (Shimasaki et al. 1999) and the BMP15 receptor gene in the Booroola breed (BMPR1B) (Paradis et al. 2009). The Booroola (FecB) phenotype, which is associated with a mutation in BMPR1B (Souza et al. 2001), is highly related to sheep fecundity and prolificacy. However, little is understood about the genetic determinants for fecundity in goat. Goat fecundity, which is important for agriculture, is said to vary between species of different genetic backgrounds.

The Jining Grey goat is an excellent local breed in China that possesses the characteristics of high prolificacy and year-round estrus (Huang et al. 2012). The mean litter size of the Jining Grey has been reported to be 2.94 (Tu 1989). The fecundity of the Laiwu Black goat is

Abbreviations: BMPR1B, Gene; BMPR1B, Bone morphogenetic protein receptor-1B; BMP15, Bone morphogenetic protein 15 gene; BMPR2, Bone morphogenetic protein receptor, type II; COGs, Clusters of orthologous groups; FDR, False discovery rate; FecB, Fecundity gene, Boorla, of sheep; FPKM, Fragments per kilobase of transcript per million mapped reads; GO, Gene Ontology; KEGG, Kyoto encyclopedia of genes and genomes; IGF1R, Insulin-like growth factor1 receptor; MAPK, Mitogen-activated protein kinase; PMSG, Pregnant mare serum gonadotropin; RPKM, Reads per kilobase of the exon model per million mapped reads; SMAD3, SMAD family member 3; TGFBR1, Transforming growth factor, beta receptor 1.

* Corresponding author.
E-mail addresses: mxy32@sohu.com, miaoxy32@163.com (X. Miao).

http://dx.doi.org/10.1016/j.gene.2016.01.047
0378-1119/

much lower than that of the Jining Grey goat. Limited studies have identified genes that might regulate fecundity in goats. The BMPR1B gene might be a major gene that influences prolificacy of Black Bengal goats (Polley et al. 2009) and also associated with the reproductive differences between the Lezhi Black and Tibetan goats (Yang et al. 2012). Previous studies indicated that bone morphogenetic protein receptor-1B (BMPR1B) gene regulates the fecundity and ovulation rate of sheep (Wilson et al. 2001; Davis 2005; Mulsant et al. 2001). These preliminary data, together with what has been shown in sheep, suggest that through genome-wide analyses of mRNAs, these two goat breeds with different prolificacies should be examined to uncover potential molecular mechanisms of fecundity in goats.

In this study, using RNA-Seq technology, we performed genome-wide sequencing of the mRNAs prepared from the ovaries of the Jining Grey goat and Laiwu Black goat. In total, 3493 candidates were predicted as new genes, and 742 genes were differentially expressed between these two goats. The functional annotation of the differentially regulated mRNAs points to multiple cellular functions and pathways that might be involved in regulating the different fecundities of these two goat species.

2. Materials and methods

2.1. Ethics statement

All of the experiments were performed in accordance with the relevant guidelines and regulations and approved by the Institutional Animal Care and Use Committee of Institute of Animal Sciences, Chinese Academy of Agricultural Sciences.

2.2. Goat sample preparation

The goats used in this study were obtained from the Qingdao Aote Farm (Shandong, China). Five female Laiwu Black goats and five female Jining Grey goats of similar ages and good status were selected. All of the 10 goats were treated with intravaginal sponges followed by injection of pregnant mare serum gonadotropin (PMSG) intramuscularly to synchronize estrus, as previously described (Miao and Luo 2013). The estrus status was verified afterwards and the goats were killed 4–5 h after estrus. Whole ovaries were excised and were immediately snap-frozen in liquid nitrogen and stored at −70 °C for total RNA extraction (Miao and Luo 2013).

2.3. Construction of mRNA libraries and sequencing

Total RNA was extracted with Trizol (Invitrogen Inc., California, USA) from the ovaries of the Laiwu Black and Jining Grey goats according to the manufacturer's instructions. The quality and quantity of the RNA samples were assessed on a Bioanalyzer 2100 system using an RNA 6000 Nano kit (Agilent Technologies, Palo Alto, CA). The RNA from the 5 goats in each group was pooled to generate two libraries. The recently developed deep-sequencing technology was used for RNA-Seq (Wang et al. 2009). RNase-free DNase I (Ambion Inc., Texas, USA) was used to eliminate potential genomic DNA contamination. Approximately 1 mg of RNA was used to generate RNA-Seq cDNA libraries for sequencing using the TruSeq RNA Sample Prep Kit v2 (Illumina Inc., San Diego, CA) according to the manufacturer's instructions. Briefly, the workflow included the isolation of polyadenylated RNA molecules using poly-T oligo-attached magnetic beads, enzymatic RNA fragmentation, cDNA synthesis, ligation of bar-coded adapters, and PCR amplification. DNA size and purity of the cDNA library were checked using a high sensitivity DNA 1000 kit on a Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA). The quantification of the cDNA libraries was performed with Qubit™dsDNA HS kit on Qubit® 2.0 Fluorometer (Life Technologies, Carlsbad, CA). The cDNA libraries were then diluted to 4 nM, and a 120 ml aliquot was used to generate clusters on a paired-end flow cell using the cBOT (Illumina) and sequenced on the Illumina Genome Analyzer IIx (GAIIx) using the SBS 36-cycle Sequencing Kit (v5) at Shanghai Biotechnology Corporation (Shanghai, China) according to manufacturer-recommended cycling parameters. One lane for the samples from each species was sequenced as 100-bp reads. The image analysis and base calling were performed with SCS2.8/RTA1.8 (Illumina). The FASTQ file generation and the removal of failed reads were performed using CASAVA ver.1.8.2 (Illumina).

2.4. Mapping reads to the reference genome and transcript reconstruction

The clean reads were mapped to goat reference gene sequences (http://goat.kiz.ac.cn/GGD/download.htm) and to goat reference genome sequences (http://goat.kiz.ac.cn/GGD/download.htm) set using SOAPaligner/SOAP2 (Li et al. 2009). No more than two mismatches were permitted in the alignment. The alignment of the RNA-Seq reads and the assembly of the alignments into a parsimonious set of transcripts was done using Cufflinks (http://cole-trapnell-lab.github.io/cufflinks/).

2.5. Alternative splicing analysis

To analyze the alternative splicing, the program AS detector was used (ASD can be found at http://www.novelbio.com/asd/ASD.html). The different transcriptomes were reconstructed and were merged to yield comprehensive reannotated transcripts for the subsequent alternative splicing (AS) analysis. After mapping, an 'accepted_hits.bam' file was generated that contained information regarding the chromosome position for exonic reads and exon–exon junction reads. A program AS detector is utilized to: (i) reconstruct exon-clusters so as to identify common modes of AS events for each exon-cluster; (ii) count the number of junction reads that align either to the inclusion or exclusion isoforms in all of the samples and finally calculating a P-value; (iii) calculate the read coverage for the alternative exon and its corresponding gene in all of the samples and calculate a second P-value by a Fisher's exact test based on the alternative exon read coverage relative to its gene read coverage between the samples and (iv) combine the P-values to get an adjusted P-value for assessing the statistical difference of the AS between the two samples.

2.6. Workflow of bioinformatics analysis

The workflow of the bioinformatics analysis of the RNA-Seq results is shown (Fig. 1). The summary of the workflow figure is that the reference genome and the gene model annotation files were downloaded from genome website directly. The index of the reference genome was built using Bowtie v2.0.6 and the clean reads were aligned to the reference genome using TopHat v2.0.9 The Cufflinks v2.1.1 Reference Annotation Based Transcript (RABT) assembly method was used to construct and identify both the known and novel transcripts from TopHat alignment results. We selected TopHat as the mapping tool. Then, HTSeq v0.6.1 was used to count the reads numbers mapped to each, and the RPKM of each gene was calculated based on the length of the gene and the reads count mapped to this gene. In addition, the differential expression analysis of the two conditions/groups was performed using the IDEG6 software. The resulting LOG2FC were adjusted using the EBSeq's approach for controlling the false discovery rate. The genes with an adjusted LOG2FC > 1 or <− 1 (FDR < 0.05) found by IDEG6 software were assigned as differentially expressed.

2.7. Expression level analysis

The original mRNA sequencing reads were filtered and only the unique mapping tags were used for gene expression analysis (Zhou et al. 2010). To enable transparent comparison of transcript levels, the mapped read counts for each gene were normalized for RNA length