Gene 676 (2018) 146–155

Contents lists available at ScienceDirect

Gene

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

Research paper

Transcriptome analyses reveal genes of alternative splicing associated with muscle development in chickens

Zhixiong Li*, Yaou Xu, Yaqiu Lin

College of Life Science and Technology, Southwest Minzu University, Key Laboratory of Qinghai-Tibetan Plateau Animal Genetic Resource Reservation and Utilization, Chengdu, Sichuan 610041, PR China

ARTICLE INFO

Keywords:
Transcriptomics
Alternative splicing
Muscle
Breed
Age

ABSTRACT

Alternative splicing (AS) of pre-mRNA is a central mode of genetic regulation in higher eukaryotes. High-throughput experimental verification of alternative splice forms, functional characterization, and regulation of alternative splicing are key directions for research. However, little information is available on the transcriptome-wide changes during different ages in different chicken breeds. In this study, the sequencing reads of chicken muscle tissues collected from White feather broiler (day 42) and Luning Chicken (day 70, 120, 150) were mapped to the chicken genome. Results showed that a total of 16,958 genes were annotated, with 2230 differentially expressed genes (DEGs) when comparing White feather broiler and Luning Chicken, and an average of 700 DEGs when comparing different ages in Luning Chicken. Functional classification by Gene Ontology (GO) and pathways analysis for selecting the genes showed most DEGs were related to muscle development and immune response. Of the 16,958 genes, a total of 6249 genes (36.85%) underwent AS events, and over 40% were specifically expressed in each library. Additionally, 6 DEGs (*SRPK3*, *ENSGALG00000022884*, *CCL4*, *GATM*, *SESN1*, *PTTG1IP*) involved in muscle development and immunity response were found to be alternatively spliced among all the four muscle tissues. In conclusion, we present a complete dataset involving the spatial and temporal transcriptome of chicken muscle tissue using RNA -seq. These data will facilitate the understanding of the effects of breed and age on the development of muscle and uncover that AS events of candidate genes may have important functional roles in regulation of muscle development in chicken.

1. Introduction

Luning Chicken mainly lives in the region of Lunning District, Mianning County, Sichuan, China. Luning Chicken is a well-known native breed, which grows fast in the early stage and has high suitability and resistance to disease. The market demand for Luning chicken is getting bigger and bigger because of its good flavor and excellent quality, and the price of Luning Chicken is 10 times of ordinary broiler. But the growth of Luning Chicken that usually takes 120 to 150 days which is more than ordinary broiler.

Alternative splicing (AS) of pre–messenger RNA (pre-mRNA) was considered to be a key factor underlying increased cellular and functional complexity in higher eukaryotes (Matlin et al., 2005; Blencowe, 2006; Ben-Dov et al., 2008). Variability in splicing patterns is a major source of protein diversity from the genome (Black, 2003). It has been estimated that two-thirds of human genes contain one or more alternatively spliced exon (Modrek and Lee, 2002; Johnson et al., 2003). AS plays an important role in different tissues and stages of development, physiology, and pathophysiological processes (Lopez, 1998), and it has been estimated that up to 15% of point mutations that causing human genetic disease affect splicing (Krawczak et al., 1992). Many recent studies have pointed to the importance of detection and measurement of AS. However, the extent of human alternative splicing is not known because of the limited depth of coverage and sensitivity afforded by conventional sequencing and microarray profiling methods (Sorek et al., 2006).

High-throughput sequencing technologies offer the potential to address this question (Calarco et al., 2007), and many studies have

Abbreviations: AS, alternative splicing; pre-mRNA, pre–messenger RNA; DEGs, differentially expressed genes; DETs, differentially expressed transcripts; WF42, 42-day-old White feather broiler; LN70, 70-day-old Luning Chicken; LN120, 120-day-old Luning Chicken; LN150, 150-day-old Luning Chicken; FPKM, Fragments per kilobase of exon per million mapped reads; FDR, false discovery rate; GO, Gene ontology; SE, skipped exon; RI, retained intron; A5SS, alternative 5′ splice site; A3SS, alternative 3′ splice site; 5′UTR, alternative 5′UTR splice site; 3′UTR, alternative 3′UTR splice site

* Corresponding author.
E-mail address: lzx4113@126.com (Z. Li).

https://doi.org/10.1016/j.gene.2018.07.027
Received 15 April 2018; Received in revised form 26 June 2018; Accepted 10 July 2018

applied analyses of mRNA sequencing (mRNA-Seq) to survey alternative splicing in higher eukaryotes. Both known and new AS events were investigated by a massive-scale RNA sequencing for understanding of transcriptional complexity in undifferentiated mouse embryonic stem cells and embryoid bodies (Cloonan et al., 2008), and AS were observed prominently, with 3500 different genes expressing one or more alternate internal splices in mouse brain, liver and skeletal muscle tissues (Mortazavi et al., 2008). AS events were detected in 85% of multiexon genes, and there are approximately 100,000 such events in major human tissues (Pan et al., 2008). RNA-Seq in human embryonic kidney and a B cell line identified 94,241 splice junctions and showed that exon skipping is the most prevalent form of AS (Sultan et al., 2008). 4567 bovine AS genes were analyzed compared to 16,715 human and 16,491 mouse AS genes, along with Gene Ontology (GO) analysis for the bovine genome annotation (Chacko and Ranganathan, 2009a).

However, there is no available information to present the alteration of AS using RNA-Seq data in chicken. Considering that the growth rate is different between Luning Chicken and ordinary broiler, it is hypothesized that the AS of genes associated with growth and development are differentially expressed during different ages in different breeds. Therefore, we hope to identify differences in transcriptional expression profiles by comparing the muscle tissues of individuals at various ages and breeds. These data would be a significant contribution towards the understanding of AS complexity for the growth and development of muscle tissue in chicken. In the present study, RNA-Seq was used for analysis of AS complexity of muscle tissue from Luning Chicken of 3 ages and 42-day-old White feather broiler. The results determined the differentially expressed genes (DEGs), AS events, differentially expressed transcripts (DETs) and revealed potential candidates with important role in regulation of muscle growth and development.

2. Methods

2.1. Animals and sample collection

Five male chickens were used for each chicken breed in different development stage in this study respectively. Luning chickens used in this study were raised in Mianning Animal Husbandry Co., Ltd. (Sichuan, China). Gastrocnemius was taken from each animal from three key stages of myogenesis and muscle maturation (day 70, 120 and 150, simplified as LN70, LN120 and LN150). Gastrocnemius was also taken from five 42-day-old White feather broiler (simplified as WF42). In each group, all fresh tissue samples were washed briefly with PBS (Phosphate Buffered Saline) and divided into 1.5 mL plastic centrifuge tubes (each sample weighing approximately 100 mg) and then immediately frozen in liquid nitrogen. Procedures involving animals were in compliance with guidelines of care and use of experimental animals established by the Institutional Animal Care and Use Committee, Southwest Minzu University, Chengdu, Sichuan, China.

2.2. RNA isolation, library construction and sequencing

Total RNA was extracted from every sample using the TRIzol reagent (Invitrogen, CA, USA) and DNA contamination was evaluated using DNase (TaKaRa, Dalian, China). Total RNA content of each sample was measured by using NanoDrop 2000c UV–Vis Spectrophotometer (Thermo Fisher Scientific Inc.), and the quality of RNA samples was assessed by agarose gel electrophoresis and Agilent 2100 Bioanalyzer (Agilent technologies). According to the recommendation of TruSeq™ RNA sample preparation Guide, all samples displayed a 260/280 ratio > 2.0 and RNA integrity numbers (RIN) ≥ 8.0.

Equal amount of 4 grouped RNA samples were then respectively pooled for cDNA synthesis and RNA-seq. Before the construction of library, ribosomal and viral RNA were removed and poly(A) + mRNAs were isolated with magnetic Oligo-dT beads (Invitrogen, USA), and then cDNA libraries were constructed and sequenced by Majorbio Biotech Co., Ltd., Shanghai, China. Briefly, 5 μg of total RNA for each group was used for the construction of libraries by using TruseqTM RNA sample prep Kit (Illumina, USA) according to the manufacturer's protocol. The constructed DNA template was enriched by PCR amplification (15 cycles). Amplicons were collected and purified by Certified Low Range Ultra Agarose (Bio-Rad, USA) gel electrophoresis. Before sequencing, the DNA libraries were quantified by using TBS-380 micro fluorometer with Picogreen® reagent (Invitrogen, USA). Clone clusters were generated on Illumina cBot, using Truseq PE Cluster Kit v3-cBot-HS, and high-throughput sequencing was performed on the Illumina Hiseq4000 Truseq SBS Kit v3-HS (200 cycles).

2.3. Sequence assembly and primary analysis

The raw reads were cleaned by removing adapter sequences, empty reads, and low-quality sequences (reads with over 10% unknown base pairs 'N' and nucleotide with q quality score lower than 20). Q20, Q30, GC-content and sequence duplication level of the clean reads were calculated. All of the downstream analyses were based on high quality clean reads.

The clean reads were aligned to the chicken reference genome (ftp://ftp.ensembl.org/pub/release-84/fasta/gallus_gallus/dna/) for assembly using TopHat (http://tophat.cbcb.umd.edu/). Sequence segments were spliced, annotated, and transcript expression was calculated by Cufflinks. Fragments per kilobase of exon per million mapped reads (FPKM) was employed to quantify gene expression, the read counts were further normalized into FPKM values. The FPKM values from the four libraries were pairwise compared. The fold changes ($\log_2$ Ratio) and the corresponding significance threshold of the P-value (decided by controlling the false discovery rate (FDR)) were estimated according to the normalized gene expression level. Based on the expression levels, the significance DEGs were identified with "FDR < 0.001 and $|\log_2$ fold change$|$ > 1" used as the threshold to judge the DEGs in this study.

2.4. Gene ontology and pathway enrichment analysis of DEGs

Gene ontology (GO) is an international standard gene functional classification system. The hypergeometric test was applied to map all DEGs to terms in the GO database. The corrected P-value < 0.05 was used as the threshold to find significantly enriched GO terms in the input list of DEGs compared to their genomic background. The formula of P-value was as follows:

$$P = 1 - \sum_{i=0}^{m-1} \frac{\left(\frac{M}{i}\right)\left(\frac{N-M}{n-i}\right)}{\frac{N}{n}}$$

The N represents the number of GO annotated genes in genome; n represents the number of DEGs in N; M represents the number of particular GO annotated genes in genome; m represents the number of particular GO annotated genes expressed differentially in M. And then P-value < 0.05 was used as the threshold to judge significant enrichment GO term in this study.

KEGG is the major public pathway-related database that helps to further understanding the biological functions of genes with high level functions and the utilities of the biological system from large-scale molecular datasets (http://www.genome.jp/kegg/). Pathway enrichment analysis identifies significantly enriched metabolic pathways or signal transduction pathways10 using the corrected P-value < 0.05 as a threshold to find significantly enriched KEGG terms in the input list of DEGs compared to their genomic background. The calculation formula is the same as that used for the GO analysis.