



Sequencing and characterization of lncRNAs in the breast muscle of Gushi and Arbor Acres chicken

Journal:	<i>Genome</i>
Manuscript ID	gen-2017-0114.R2
Manuscript Type:	Article
Date Submitted by the Author:	09-Dec-2017
Complete List of Authors:	<p>Ren, Tuanhui; Henan Agricultural University, College of Animal Science and Veterinary Medicine</p> <p>Li, Zhuanjian; Henan Agricultural University; Henan Innovative Engineering Research Center of Poultry Germplasm Resource</p> <p>Zhou, Yu; Henan Agricultural University, College of Animal Science and Veterinary Medicine</p> <p>Liu, Xuelian; Henan Agricultural University, College of Animal Science and Veterinary Medicine</p> <p>Han, Ruili; Henan Agricultural University, College of Animal Science and Veterinary Medicine; Henan Innovative Engineering Research Center of Poultry Germplasm Resource</p> <p>Wang, Yongcai; Henan Agricultural University, College of Animal Science and Veterinary Medicine</p> <p>Yan, Fengbin; Henan Agricultural University, College of Animal Science and Veterinary Medicine; Henan Agricultural University, Henan Innovative Engineering Research Center of Poultry Germplasm Resource</p> <p>Sun, Guirong; Henan Agricultural University, College of Animal Science and Veterinary Medicine</p> <p>Li, Hong; Henan Agricultural University, College of Animal Science and Veterinary Medicine</p> <p>Kang, Xiangtao; Henan Agricultural University, College of Animal Science and Veterinary Medicine; Henan Innovative Engineering Research Center of Poultry Germplasm Resource</p>
Is the invited manuscript for consideration in a Special Issue? :	N/A
Keyword:	lncRNA, RNA-Seq, chicken, muscle development

Sequencing and characterization of lncRNAs in the breast muscle of Gushi and Arbor Acres chicken

Tuanhui Ren^{1*}, Zhuanjian Li^{1,2*}, Yu Zhou¹, Xuelian Liu¹, Ruili Han^{1,2}, Yongcai Wang¹, FengBinYan^{1,2},
GuiRong Sun^{1,2}, Hong Li^{1,2}, Xiangtao Kang^{1,2,*}

¹College of Animal Science and Veterinary Medicine, Henan Agricultural University, Zhengzhou,
450002, China

²Henan Innovative Engineering Research Center of Poultry Germplasm Resource, Zhengzhou, 450002,
China

*These authors contributed equally to this work.

***Corresponding author: Xiangtao Kang**

Tel: +86-371 63558209 E-mail: xtkang2001@263.net

Address: College of Animal Science and Veterinary Medicine, Henan Agricultural University,
Zhengzhou, 450002, China

The current e-mail addresses of all author are as follows:

Tuanhui Ren, E-mail: 1490026834@qq.com; Zhuanjian Li, E-mail: lizhuanjian@163.com; Yu Zhou,

E-mail: 1538476837@qq.com; Xuelian Liu, E-mail: 305419093@qq.com; Ruili Han, E-mail:

rlhan@126.com; Yongcai Wang, E-mail: wyc0002@126.com; FengBinYan, E-mail:

yanfb790811@sina.com; GuiRong Sun, E-mail: grsun2000@126.com; Hong Li, E-mail:

lihong19871202@163.com; Xiangtao Kang, E-mail: xtkang2001@263.net

Abstract

Chicken muscle quality is one of the most important factors determining the economic value of poultry, and muscle development and growth are affected by genetics, environment, and nutrition. However, little is known about the molecular regulatory mechanisms of long non-coding RNAs (lncRNAs) in chicken skeletal muscle development. Our study aimed to better understand muscle development in chickens and thereby improve meat quality. In this study, Ribo-Zero RNA-Seq was used to investigate differences in the expression profiles of muscle development-related genes and associated pathways between Gushi (GS) and Arbor Acres (AA) chickens. We identified two muscle tissue-specific expression lncRNAs. In addition, the target genes of these lncRNAs were significantly enriched in certain biological processes and molecular functions, as demonstrated by Gene Ontology (GO) analysis, and these target genes participate in five signaling pathways, as revealed by an analysis of the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. Taken together, these data suggest that different lncRNAs might be involved in regulating chicken muscle development and growth and provide new insight into the molecular mechanisms of lncRNAs.

Keywords: lncRNA; RNA-Seq; chicken; muscle development

Introduction

In recent decades, poultry research has made significant progress in genetic improvement and feedstuff optimization. Due to its high protein, low calorie, and low cholesterol content, chicken meat has become a source of high-quality protein. Because broiler chickens have a short rearing period and high feed conversion rate compared with those of pig and cattle, chicken is the second most consumed meat product after pork in China. However, as living standards have increased, people are paying greater attention to the color, tenderness, and flavor of chicken. Therefore, chicken meat quality has been increasingly studied in animal husbandry science (Zhang et al. 2017b).

Geneticists have established new broiler varieties through breeding and modern biotechnology, including molecular breeding and transgenes, and the focus has been on the selection of growth-rate and meat-quality traits to improve production performance. However, biological properties are not only controlled at the DNA level but are also regulated by pre-transcriptional and post-transcriptional modifications at the mRNA level; regulation by mRNA is more comprehensive and precise. In recent years, the study of miRNAs has gradually advanced with great achievements, but the precise functional mechanism of long non-coding RNAs (lncRNAs) remains unknown. Thus, more in-depth study of lncRNAs is needed. Compared with other coding transcripts, there are many types and quantities of lncRNAs with different modes of action. lncRNAs are generally less conserved than coding genes and exhibit apparent low-level tissue-specific expression that creates many problems in the study of their biological function. lncRNAs are transcripts greater than 200 nt in length with decreased coding potential and are widespread in all types of biological organisms (Hung and Chang 2010). lncRNAs participate in the multilevel regulation of transcriptional, post-transcriptional, and epigenetic modifications (Lee 2012; Mattick 2009). lncRNAs play important roles in many biological processes,

such as cell differentiation and proliferation, growth and development, organogenesis, immune response, and tumorigenesis (Cesana et al. 2011; Guttman et al. 2009; Kretz et al. 2013; Necsulea et al. 2014; Schmitt and Chang 2013). The discovery of lncRNAs has improved our understanding of non-coding RNAs and revealed a new pathway for the regulation of gene expression in animals; lncRNAs more effectively complement target mRNA, rapidly regulate expression at the RNA level, and add a level of complexity to the comprehensive network of cellular gene expression regulation. Some functions of lncRNAs in muscle development and adipogenesis have been confirmed (Cabianca et al. 2012; Sun et al. 2013a). Compared with native Chinese chickens, the broiler chicken is a commercial variety that has been bred for a long time; this chicken grows rapidly and has a high feed conversion rate, which leads to poor meat quality and flavor (NariNç et al. 2015; Wang et al. 2006). Arbor Acres (AA) broilers have greater breast weight and myofiber diameter and lower myofiber density than Chinese native chickens (Chen et al. 2007). Therefore, determining the molecular mechanisms of myogenesis and differentiation in broiler breeding is important. lncRNA identification and functional annotation have provided new insights into the molecular mechanisms underlying chicken muscle development.

The regulation of gene expression by small ncRNAs is an active research area in the field of life science. miRNAs are a class of small ncRNAs that often negatively regulate gene expression, and some play essential roles in chicken skeletal muscle, e.g., miR-125b, miR-221, and miR-206, miRNA-133a, and miR-1a (Andreote et al. 2014; Wang et al. 2012). The function of lncRNAs is generally less conserved than that of miRNAs, and lncRNAs act via different mechanisms to play important roles during muscle differentiation. lncRNA H19 acts as a molecular sponge for the major let-7 family of miRNAs, establishing a role for the H19/let-7 axis in myogenic differentiation. H19 depletion causes

precocious muscle differentiation, a phenotype that is recapitulated by *let-7* overexpression (Kallen et al. 2013). In addition, H19 has a trans-regulatory function in promoting skeletal muscle differentiation and regeneration that is mediated by embedded microRNAs (Dey et al. 2014). *linc-MD1* RNA is also expressed in human muscle cells, where it modulates miR-133 and miR-135 targets and plays a relevant role in affecting the timing of myoblast differentiation (Cesana et al. 2011). A muscle-specific lncRNA, *lncMyoD*, is a key downstream target of MyoD and an important regulator of myoblast cell cycle exit and myogenesis (Gong et al. 2015). lncRNA *gga-lnc-0181* is highly expressed in chicken skeletal muscle, which suggests an important role in muscle development (Li et al. 2012). In a recent study, *Linc-Yyl* was demonstrated to be an important regulator of mouse skeletal myoblast differentiation by interaction with the *Yyl* transcription factor (Zhou et al. 2015). This study emphasized muscle development and the identification of lncRNAs involved in muscle development and growth.

The central dogma of molecular biology is that genetic information travels from DNA to RNA to proteins (Kashi et al. 2016). However, this central dogma has faced severe challenges in response to the results of high-throughput transcriptome analyses. High-throughput sequencing, also called RNA sequencing (RNA-Seq), describes gene expression at the global level. In general, the use of oligo (dT) primers for the total RNA enrichment of mRNA transcripts containing a poly(A) tail, i.e., non-ribosomal RNA, ignores mRNA transcripts without a poly(A) tail and partially degraded mRNA. An increasing number of studies have revealed that poly(A)- transcripts play important roles independent of translation. RNA sequencing to remove ribosomal RNA (Ribo-Zero RNA-Seq) can be used to simultaneously capture RNA species with poly(A)- and poly(A)+ tails.

In the present study, breast muscle tissues from 6-week-old Gushi (GS) chickens, which represent a

slow-growing native Chinese breed, and Arbor Acres (AA) chickens, which are fast-growing broilers, were studied using Ribo-Zero RNA-Seq. We analyzed protein-coding genes, measured their total expression, functionally annotated the differences in gene expression between breeds, and studied alternative splicing and genetic variation at the transcriptional level. We then analyzed lncRNAs by describing the systematic global characteristics of lncRNAs in chicken breast muscle tissue, searched for differentially expressed lncRNAs and analyzed the tissue expression profiles. The results identified two muscle tissue-specific expression lncRNAs (TCONS_00064133 and TCONS_00069348). The target genes of the differentially expressed lncRNAs were significantly enriched in certain biological processes and molecular functions of Gene Ontology (GO) terms and were found to participate in five signaling pathways through an analysis of the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. Taken together, these data suggest that different lncRNAs might be involved in regulating chicken muscle development and growth and provide new insights into the molecular mechanisms of lncRNAs.

Materials and Methods

Animals

All animal experiments were performed in accordance with the Regulation for the Chinese National Research Council (1994) and were approved by the Henan Agricultural University Institutional Animal Care and Use Committee (Permit Number 11-0085) (Han et al. 2011). AA and GS chickens were obtained from the Henan Innovative Engineering Research Center of Poultry Germplasm Resource (Zhengzhou, China). The chickens were raised in stair-step cages under identical recommended environments, with feed and drinking water provided ad libitum. Six individuals were selected from each of the set of AA and GS chickens at three developmental states (1 day, 6 weeks, and 16 weeks),

and their heart, liver, lung, kidney, leg muscle, breast muscle, abdominal fat, sebum cutaneum, and hypothalamus tissues were collected, rinsed with diethyl pyrocarbonate (DEPC) water, immediately frozen in liquid nitrogen, and stored at -80 °C until further analysis. During sample collection, breast muscle samples from 6-week-old AA and GS chickens were snap frozen in liquid nitrogen and sent to Shanghai Biotechnology Corporation (Bohao, Shanghai, China) for sequencing.

RNA extraction, library construction, and sequencing

Total RNA from AA and GS breast muscle was extracted using the RNeasy Micro kit (QIAGEN, Valencia, CA, USA) and treated with DNase according to the manufacturer's instructions to deplete any contaminating genomic DNA. The RNA quality was measured using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA) and an Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA). RNA samples with an integrity number greater than 7.0 and an optical density 260/280 nm ranging from 1.8 to 2.0 were selected for follow-up experiments.

Ribo-Zero RNA-Seq libraries were prepared using an Illumina TruSeq RNA Sample Prep Kit (Illumina, Inc., San Diego, CA, USA) according to the manufacturer's instructions, and three RNA samples were pooled to obtain an "averaged" transcriptome for each breed. During library construction, the Ribo-Zero rRNA Removal Kit (Epicenter, Madison, WI, USA) was used to remove rRNA; the remaining RNA was then purified and fragmented, the first-strand and second-strand cDNA were synthesized, the end of the double strand was repaired, and an A-tailed link connector was added to the 3' end. The concentration and size of the two libraries were confirmed using a Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA) and an Agilent 2100 Bioanalyzer. The libraries were then analyzed using one lane of 2 × 100 nucleotide paired-end Illumina HiSeq 2500 with a TruSeq SBS v3-HS Kit (Illumina, Inc., USA). The quality of RNA-Seq reads was assessed using FastQC (Babraham

Bioinformatics, UK, Version 0.10.1) (Green et al. 2017).

Mapping, assembly, and transcript abundance estimation

Reads that passed quality control were aligned to the *Gallus gallus* reference genome (gal5) from NCBI using TopHat (Version 2.0.9); the *Gallus gallus* reference genome download link is ftp://ftp.ensembl.org/pub/release-87/fasta/gallus_gallus/Dna/Gallus_gallus_gallus-5.0.dna.toplevel.fa.gz. According to the alignment results from TopHat, the transcriptome data for each library were assembled separately using Cufflinks (Version 2.1.1) (Trapnell et al. 2012). The transcript abundances are shown in fragments per kilobase of exon per million mapped reads (FPKM) and measured using Cufflinks. A Cuffcompare program in the Cufflinks package was used to map the assembled transcripts to the reference annotation to acquire a unique set of novel genes. All Ribo-Zero RNA-Seq datasets were submitted to the NCBI Sequence Read Archive (SRA) database, and the files can be found under GenBank accession number SRP111503.

lncRNA prediction and identification

To identify a novel reliable lncRNA database, we extracted three classes of transcripts called “i”, “u”, and “x” and filtered them using the following highly stringent criteria. 1) Size selection of the fragment: transcripts longer than 200 nt and exon ≥ 2 were kept. 2) Read coverage threshold: transcripts with ≤ 3 reads were removed from the database. 3) Open reading frame (ORF) filter: transcripts with a predicted ORF longer than 300 nt were removed. 4) Known protein domain filter: transcripts were aligned to the Pfam database to remove transcripts with significant homology or known protein domains (Finn et al. 2014). 5) Protein-coding-ability prediction: protein-coding transcripts were predicted using the Coding-Non-Coding Index (CNCI) and Coding Potential Calculator (CPC) (Kong et al. 2007; Sun et al. 2013b). 6) Eliminate housekeeping lncRNAs: snRNAs,

tRNAs, and snoRNAs, among others, were removed. 7) Known lncRNAs: The candidate lncRNAs were compared with the lncRNA database, and known lncRNAs were removed (Li et al. 2016).

lncRNA re-annotation and expression analysis

There are known lncRNAs in the NONCODE v4 database genome, and novel lncRNAs have emerged. We next quantified and analyzed the differences between lncRNAs using Cufflinks (Version 2.1.1). Known genes were named beginning with ENS and NON, and novel genes were named beginning with TCONS. A Perl script was used to find lncRNAs in chromosomes corresponding to genes; intronic lncRNAs, sense lncRNAs, and antisense lncRNAs were located near corresponding genes, and intergenic lncRNAs were located between two genes. Differentially expressed lncRNAs identified from an analysis of RNA-Seq data between two analyzed samples were identified according to the FPKM value using DEGseq packages (Wang et al. 2010). Fold change (FC) and Fisher's exact test were used to screen for differentially expressed genes; transcripts with a false discovery rate (FDR) ≤ 0.05 and $FC \geq 2$ were considered significant.

Quantitative real-time PCR (qRT-PCR)

The expression levels of randomly selected lncRNAs were validated by qRT-PCR. Total RNA samples from the heart, liver, lung, kidney, breast muscle, leg muscle, abdominal fat, sebum cutaneum, and hypothalamus tissues of GS chickens were extracted using RNAiso Plus (TaKaRa, Dalian, China) using three biological replicates for each developmental stage, and the quality was assessed using electrophoresis and NanoDrop 1000 spectrophotometry (Thermo Scientific, USA). All cDNAs were synthesized using an RNA reverse transcription kit (TaKaRa, Dalian, China) according to the manufacturer's instructions. The relative expression levels of the lncRNAs were quantified using the SYBR® Premix Ex Taq™ II kit (Takara, Dalian, China) with a Bio-Rad CFX96 Real-Time PCR

instrument. The 25- μ L PCR reaction mixture included 12.5 μ L of SYBR Premix Ex Taq II (TaKaRa, Dalian, China), 1.25 μ L (10 pM/ μ L) of specific forward primer, 1.25 μ L (10 pM/ μ L) of reverse primer, 2.0 μ L (10 ng/ μ L) of diluted cDNA, and 8 μ L of RNase-free water. The amplification conditions included initial denaturation at 95 °C for 3 min, 40 cycles of denaturation at 95 °C for 30 s, annealing at 60 °C for 30 s, and 72 °C for 30 s, and a final 10-min extension at 72 °C. All reactions were run in triplicate. *GAPDH* was used as a housekeeping gene, and the relative expression of genes in different tissues was analyzed using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen 2001). Each column represents the means \pm SE from three biological replicates, and each measurement was repeated three times. The lncRNA primer details are presented in Table S1. The qRT-PCR results were statistically analyzed using SPSS software (Version 21.0; Statistical Product and Service Solutions, IBM Corporation, Armonk, NY, USA). *P*-values < 0.05 were considered statistically significant.

Prediction and annotation of lncRNA targets

lncRNAs function by acting on protein-coding genes via the cis-regulatory elements and trans-regulatory factory. The closest coding genes 10 kb upstream and downstream of lncRNAs were selected as cis-regulatory targets (Quinlan and Hall 2010). In the present study, lncRNA targets were predicted based on the prediction of cis function. The next step was to assess the functional enrichment of target genes using the DAVID database (Huang et al. 2009).

Gene Ontology and Kyoto Encyclopedia of Genes and Genomes analyses

GO (<http://www.geneontology.org>) enrichment and KEGG (<http://www.kegg.jp>) pathway analyses of differentially expressed genes were performed using DAVID software (Huang et al. 2009). In GO significant enrichment analyses, all differentially expressed lncRNAs from RNA-Seq of GS and AA were used to map each term of the GO database. The number of genes for each term was counted, and

a hypergeometric test was then applied to identify the GO terms that were significantly enriched in differentially expressed genes. In the KEGG annotation of differentially expressed genes, the KEGG pathway acted as a unit; each identified group was compared with differentially expressed genes mapped to the KEGG database, and the pathway graph was observed. P-values of GO terms and pathways were corrected using Bonferroni correction, and functional terms with corrected P-values ≤ 0.05 were considered significantly enriched.

Results

RNA quality control

The isolated RNA quality of GS and AA chicken breast muscle tissues was measured using a NanoDrop ND-1000 spectrophotometer and an Agilent Bioanalyzer 2100. Agarose gel electrophoresis showed clear and repeating 18s and 28s bands; better-quality RNA was used for sequencing (Fig 1). The RNA sample baseline was relatively flat according to the Bioanalyzer results; the 18s and 28s genes possessed higher kurtosis and better integrity, and large amounts could be used for library construction and sequencing (Table 1 and Fig S1).

Sequencing results and quality control

Quality control of RNA-Seq reads was performed using FastQC. In the quality assessment of sequencing data, all base quality scoring Q-values were >30 ; the error rate E-value (the relationship between the Q-value and sequencing error E-value is expressed as $Q = -10 \log_{10} E$) was 0.1% when the Q-value was equal to 30, and the sequencing data distributed to a green area that represents high-quality scores in the Q-value boxplot (Fig S2). These sequencing data met the criteria necessary to conduct subsequent analysis.

A total of 70,151,126 and 113,719,518 raw reads were produced from the AA and GS libraries, respectively. After removing contaminant reads with fastx (version 0.0.13) software, we obtained 70,137,370 (AA) and 113,689,278 (GS) clean reads that were used for the subsequent analyses (Bolger et al. 2014). The percentages of clean reads among raw reads were 94.5% and 95.3% in AA and GS libraries, respectively. Among the clean reads, 64,426,854 sequences from the AA libraries and 102,625,185 sequences from the GS libraries mapped perfectly to the chicken genome sequence. The mapping ratio of GS chickens was inferior to that of AA broilers.

Identification and characteristic distribution of lncRNAs in AA and GS breast muscle tissue

To identify unknown chicken lncRNAs, we used highly stringent filtering criteria to remove transcripts with positive protein-coding potential and known lncRNAs. In total, 4,815 lncRNAs were obtained from the two breast muscle tissues samples (Table S2). According to the genomic position of the lncRNAs, these 4,815 lncRNAs were divided into 1,030 sense lncRNAs, 2,614 lincRNAs (large intergenic ncRNAs), 525 anti-sense lncRNAs, and 646 bidirectional lncRNAs. Subsequently, we analyzed the characteristics of these lncRNAs, including sequence length, coding capacity, transcript abundance, and exon number. A total of 703 novel lncRNAs and 4,112 known lncRNAs were analyzed. The coding potential of the 703 novel lncRNAs was determined using CPC, CNCI, and Pfam software (Fig 2a). Subsequently, we analyzed the transcript length of lncRNAs and mRNAs in Ribo-Zero RNA-Seq data (Fig 2b). The FPKM method was used to measure expression, as shown in Fig 2c. Moreover, novel lncRNAs showed fewer exons per transcript compared with average mRNA (Fig 2d). Reports indicate that lncRNAs are shorter and have significantly reduced expression compared with the mRNA of protein-coding genes (Cabili et al. 2011; Gloss and Dinger 2016; Guttman et al. 2010), as we observed for chicken lncRNAs.

Differentially expressed lncRNAs of AA and GS chicken breast tissue

In total, 4,815 candidate lncRNAs were obtained from the two samples. As shown in Table S1, a total of 4,592 and 4,751 lncRNAs were found in AA and GS breast muscle tissue, respectively; 64 lncRNAs existed only in AA chickens, and 223 lncRNAs existed only in GS chickens. In this study, RNA-Seq data from the two breeds were useful for detecting differentially expressed lncRNAs. Thus, to identify lncRNAs associated with muscle development or meat quality, we analyzed significantly differentially expressed lncRNAs ($FC \geq 2$ and $FDR < 0.05$) between GS and AA tissues and identified 147 differentially expressed lncRNAs, including 132 up-regulated genes and 15 down-regulated genes (Table S3 and Fig S3).

Validating tissue-specific lncRNAs by qRT-PCR

To validate tissue-specific lncRNAs identified from the RNA-Seq sequencing data, we conducted qRT-PCR. Fifteen differentially expressed lncRNAs, including 10 up-regulated (TCONS_00005187, TCONS_00064133, TCONS_00086249, TCONS_00081971, TCONS_00069348, TCONS_00082375, TCONS_00000618, TCONS_00016748, TCONS_00094682, TCONS_00073392) and five down-regulated (TCONS_00023521, TCONS_000298 27, TCONS_00103822, TCONS_00058184, TCONS_00096145) lncRNAs (Table S3), were randomly selected. The expression patterns of the selected lncRNAs were roughly consistent with the RNA-Seq results (Fig 3 and Fig S4). The up-regulated differentially expressed genes TCONS_00064133 and TCONS_00069348 exhibited apparent tissue specificity. TCONS_00064133 was relatively highly abundant in leg muscle and breast tissue; expressed the lowest in the sebum cutaneum, lung, kidney, and hypothalamus tissue; and not expressed in the heart, liver, or abdominal fat tissue. The expression of TCONS_00069348 in leg muscle and pectoralis was similar to that of TCONS_00064133, and this lncRNA was not expressed in

the lung and sebum cutaneum tissue (Fig 3). TCONS_00086249 was relatively highly expressed in the sebum cutaneum, leg muscle, and breast tissue, showed the lowest expression in the lung, and was not expressed in other tissues (Fig S4).

Predicted and functional analyses of lncRNA target genes

Recent studies have suggested that lncRNAs can serve as cis-regulators and affect the closest coding genes 10 kb upstream and downstream (Ørom et al. 2010). To investigate the relationship between lncRNAs and their neighboring coding genes and to further explore the function of lncRNAs in chicken muscle tissue, we predicted the candidate target genes of lncRNAs by predicting the cis function between lncRNAs and mRNAs. In total, 2,280 target genes were predicted by candidate lncRNA cis-regulation (Table S4).

GO enrichment revealed the functions of genes in stage-specific modules, and KEGG analysis revealed signaling pathways and metabolic networks vital for breast muscle development (Zhang et al. 2017a). In the current study, the GO terms for the target lncRNA genes were related primarily to biological processes and molecular functions (Fig 4). The biological processes included response to endoplasmic reticulum stress, protein glycosylation, glycosylation, glycoprotein metabolic process, glycoprotein biosynthetic process, and carbohydrate metabolic process. The molecular functions included transferase activity, transferring hexosyl groups, transferase activity, transferring glycosyl groups, transferase activity, and cofactor binding. A KEGG pathway analysis revealed that the enriched pathways included the MAPK signaling pathway, insulin signaling pathway, calcium signaling pathway, arginine biosynthesis pathway, and alanine, aspartate, and glutamate metabolism pathway (Fig 5).

Discussion

Chicken muscle is an important source of animal protein for human health, and meat quality affects

the economic value of poultry. The processes regulating the muscle growth of chickens are affected by multiple factors, including genetics, nutrition, and environment (Ouyang et al. 2015). Domestic chickens have high carcass yields; the breast muscle contributes up to approximately 30% of the chicken's total carcass weight, and the total muscle weight constitutes approximately 40% (Schreurs 2000). To date, researchers have used RNA-Seq to identify a large number of lncRNAs (Billerey et al. 2014; Washietl et al. 2014; Weikard et al. 2013). In chickens, hundreds of lncRNAs have been identified using RNA-Seq (Billerey et al. 2014; Koufariotis et al. 2015). A total of 281 new intergenic lncRNAs have been identified in chicken skeletal muscle by RNA-Seq (Li et al. 2012). However, previous studies have focused on poly(A)⁺ lncRNAs and ignored ncRNA poly(A)⁻ transcripts, which are a crucial component. poly(A)⁻ lncRNAs play a vital role in myogenesis (Djebali et al. 2012; White et al. 2014). In the current study, we investigated the integral expression of chicken lncRNAs in 42-day-old GS and AA chickens using the Ribo-Zero RNA-Seq method. Although this study indicated that several lncRNAs act as single exon transcripts, because there are few effective ways to discriminate them from the thousands of assembled artifacts, many single exonic transcripts were often discarded (Sun et al. 2014). To gather trustworthy data, we focused on lncRNAs that included no fewer than two exons. Compared with polyA-Seq, Ribo-Zero RNA-Seq revealed coequal coverage uniformity and rRNA removal efficiency but exhibited superiority to polyA-Seq in that it simultaneously captured poly(A)⁺ and poly(A)⁻ transcripts; this method will doubtlessly become the best sequencing method for transcriptome investigations.

A total of 147 differentially expressed lncRNAs were obtained by pairwise comparisons (6-week-old GS vs. AA) of samples collected from breast muscle tissue. Interestingly, vestigial like family member 2 (Vgll2) is the common target gene of TCONS_00064133 and TCONS_00069348 by cis function, and

Vgll2 expression is associated with all sites of skeletal myogenesis. RNA-Seq revealed that Vgll2 is specifically expressed in the limb tissue of mouse. According to an analysis of Vgll2 protein-protein interactions and co-expression, Vgll2 interacts with a cluster of Wnt proteins, Mef2 (myocyte enhancer factor-2) family proteins, Notch1, Notch2, and Tead (TEA domain) family proteins and is associated with proteins such as Myod1, MyOG, Pitx3 (paired like homeodomain 3), Tbx15 (T-box 15), and Myf5. These results indicated a possible functional connection of Vgll2 with mouse skeletal muscle development (Sun et al. 2017). Similar to MyoD, Vgll2 is positively regulated by the myogenic regulatory factors associated with skeletal differentiation and is negatively regulated by Notch signaling. This result demonstrates that Vgll2 is related to skeletal muscle differentiation in myogenesis downstream of MyoD activation (Bonnet et al. 2010). Based on the qRT-PCR results, we speculate that TCONS_00064133 and TCONS_00069348 likely play similar roles in chicken muscle. Therefore, further study is needed to determine the possible roles of these two genes.

A GO analysis showed that target genes are significantly enriched in certain molecular functions and biological processes. Intriguingly, a few target genes, such as PTPN1, PYGL, GOT2, and ERBB4, were found to be enriched in KEGG pathways. Ubiquitin-specific peptidase 25 (USP25), tropomodulin 4 (TMOD4), and glutathione S-transferase kappa 1 (GSTK1) participate in myogenesis and muscle differentiation, as revealed by a GO term analysis. Three muscle-specific cytoskeletal and sarcomeric proteins specifically interact with USP25m; this enzyme plays an important role in the regulation of muscular differentiation and function (Bosch-Comas et al. 2006). The expression of TMOD4 is activated during adipogenesis but inhibited during myogenesis (Zhao et al. 2013). The level of GSTK1 expression has been linked to obesity, and a mutation in the hGSTK1 promoter area has been

associated with insulin secretion and fat deposition (Fei et al. 2009; Shield et al. 2010).

The vast majority of research suggests that lncRNAs can regulate the expression of neighboring protein-coding genes (Ren et al. 2016; Wang et al. 2016). To investigate lncRNA function, we predicted the potential targets of lncRNAs by a cis method and conducted GO and pathway analyses. After performing a KEGG pathway analysis, we found a significantly enriched MAPK signal pathway that not only participates in cell growth, differentiation, and feed conversion but also has a positive effect on meat tenderness (Taye et al. 2017). As one of the major intracellular signaling pathways influencing myogenesis, MAPK signaling is associated with postmortem meat quality (Ponsuksili et al. 2009). MAP2K3 is related to porcine loin muscle area and fat traits, implying roles in muscle differentiation and growth (Wu et al. 2010). The MAPK pathway in skeletal muscle has been investigated; p38 MAPK activity is necessary for myogenic cell differentiation and plays an essential role in energy expenditure and glucose metabolism (Niu et al. 2003; Puigserver et al. 2001; Zetser et al. 1999). p38 MAPK is a major kinase in the MAPK family; adiponectin inhibits chicken preadipocyte differentiation via the activation of p38 MAPK, and the muscle-specific activation of the p38 MAPK pathway in transgenic mice is sufficient to enhance PGC-1 α (a promoter of slow fiber formation in skeletal muscle) expression and promote mitochondrial biogenesis (Akimoto et al. 2005; Yan et al. 2013).

MAPK pathway genes, including PPM1B, NFATC1, and FGF13, are highly expressed target genes of lncRNAs. NFATc1 represses the slow MyHC2 promoter activity in both fast and fast/slow primary muscle fibers. Research does not support the role of NFAT as a direct regulator of primary muscle fiber type differences. Instead, the results reflect intrinsic differences in the modes of regulation of the slow MyHC2 gene in primary muscle fiber types (Theobald and DiMario 2011). Numerous studies have also indicated that the expression of intracellular FGF13 protein is higher in leg muscle, skeletal muscle,

and single muscle fibers in mice (Chakkalakal et al. 2012; Gecz et al. 1999). On the first day of myogenic differentiation, FGF13 knockdown caused the mRNA and protein levels of MyOG and MyHC to significantly increase in C2C12 myoblasts. In contrast, FGF13 overexpression reduced MyOG and MyHC expression. These data showed that FGF13 inhibits C2C12 cell proliferation and differentiation by down-regulating *Spry1* and indicated that FGF13 plays a negative regulatory role in skeletal muscle development (Lu et al. 2015). Therefore, the expression of these genes might be extremely relevant to muscle development in chicken.

Two target genes, protein tyrosine phosphatase non-receptor type 1 (PTPN1) and glycogen phosphorylase L (PYGL), participate in insulin signaling. Interestingly, PTPN1, the gene coding for PTP-1 β , is located in a genomic region that has been identified in numerous linkage studies as a QTL for obesity and diabetes (Ghosh et al. 1999; Lee et al. 1999; Soro et al. 2002). PTP-1 β affects plasma lipid levels and might result in obesity and hypertension in Japanese and Chinese populations. Given similar associations with insulin resistance and diabetes found in other populations, this gene might play a vital role in the development of the characteristic metabolic changes observed in patients with metabolic syndrome (Olivier et al. 2004). In addition, the mRNA expression of PTPN1 and the activation of PTPN1 and NF- κ B in muscle were significantly decreased in animals fed a DHA-enriched (DE) diet ($P < 0.05$), which suggests that feeding a DE diet increases feeding-induced muscle protein synthesis in growing pigs and that muscle IGF-1 expression and insulin activity are involved (Wei et al. 2013).

Ryanodine receptor 1-like (RyR1), peptidylprolyl isomerase F (PPIF), and erb-b2 receptor tyrosine kinase 4 (ERBB4) are all involved in the calcium signaling pathway. RyR1 inhibition increases slow MyHC2 promoter activity in innervated pectoralis major (PM) muscle fibers, enhances the

transcriptional activity of nuclear factor of activated T cells and myocyte enhancer factor 2, and enhances their interactions with their respective binding sites on the slow MyHC2 promoter; these results reveal that RyR1 activity in innervated fast PM muscle fibers contributes to cell type-specific repression (Jordan et al. 2004). ERBB4 is a receptor for neuregulins that belongs to a family of closely related proteins implicated as regulators of neural and muscle development and the differentiation and oncogenic transformation of mammary epithelia (Carraway and Burden 1995; Plowman et al. 1993). In addition, ERBB4 functions as a neuregulin receptor in the heart; however, differences in the hindbrain phenotype of these mutants are consistent with the activity of a new ERBB4 ligand in the central nervous system (Gassmann et al. 1995).

Intriguingly, glutaminase (GLS) and glutamic-oxaloacetic transaminase 2 (GOT2) simultaneously participate in the arginine biosynthesis pathway and alanine, aspartate, and glutamate metabolism. Glutamine is synthesized mainly in skeletal muscle, lungs, and adipose tissue. The catabolism of glutamine is activated by either of two isoforms of the mitochondrial GLS (Curthoys and Watford 1995). Glutaminase converts glutamine to glutamate, which is further catabolized through the tricarboxylic acid cycle for the production of ATP or serves as a substrate for glutathione synthesis. A previous study revealed a pathway through which the Myc suppression of miR-23a/b, which targets GLS, enhances glutamine catabolism through elevated mitochondrial glutaminase expression in human cancers (Gao et al. 2009). A previous study using an omission and addition test found that glutamic acid (Glu) significantly affects umami and saltiness perception (Fujimoto et al. 1996). Taken together, these two pathways might be associated with the flavor of chicken meat.

Compared with the few pathways and target genes that we investigated in our research, previous research has shown that myogenesis is regulated by a few transcription factors, including Pax3, Pax7

and Gli, and four myogenic regulatory factors, MyoD, Myf-5, myogenin, and MRF-4 (Relaix et al. 2005). The activation of Wnt signaling stabilizes β -catenin, which regulates the expression of Pax3 and Gli (Borycki et al. 2000; Capdevila et al. 1998). Pax3 is necessary for skeletal myogenesis and acts upstream of MyoD during skeletal muscle development; Gli factors could induce the expression of Myf-5, which is a direct target of Wnt/ β -catenin (Borello et al. 2006; Gustafsson et al. 2002; Ridgeway and Skerjanc 2001). The Wnt signaling pathway could enhance myogenesis and inhibit adipogenesis in cultured mesenchymal stem cells (Du et al. 2011; Shang et al. 2007). Livestock carcasses are composed of muscle, fat, connective tissue, and bone; muscle fibers form lean meat, whereas fat cells and fibroblasts are embedded in connective tissue (Du et al. 2010). Thus, skeletal muscle development might be associated with myogenesis, adipogenesis, and fibrogenesis.

Conclusion

In summary, in the present study, Ribo-Zero RNA-Seq technology and bioinformatics tools were used to analyze the primary differentially expressed genes and pathways between GS and AA chicken breast muscle. The annotation of the predicted targets indicate that differentially expressed lncRNAs might be associated with myogenesis and myogenic differentiation in chickens, and this information will help expand our understanding of the molecular repertoire of muscle development- and growth-related genes in GS and AA chickens. Furthermore, we identified two muscle tissue-specific lncRNAs: TCONS_00064133 and TCONS_00069348. Further investigation is required to better understand the molecular mechanisms of lncRNAs in muscle development and growth in GS chickens.

Acknowledgments

This work was supported by grants received from the China Postdoctoral Science Foundation (No. 2014 M552004), the Program for Innovation Research Team of Ministry of Education (No. IRT16R23),

the National Natural Science Foundation of China (No. 31372330), and the Key Science and Technology Research Project of Henan Province (No. 151100110800).

Conflicts of interest

The authors declare that no conflicts of interest exist.

References

- Akimoto, T., Pohnert, S.C., Li, P., Zhang, M., Gumbs, C., Rosenberg, P.B., Williams, R.S., and Yan, Z. 2005. Exercise stimulates Pgc-1 α transcription in skeletal muscle through activation of the p38 MAPK pathway. *Journal of Biological Chemistry* **280**(20): 19587-19593.
- Andreote, A.P.D., Rosario, M., Ledur, M., Jorge, E., Sonstegard, T., Matukumalli, L., and Coutinho, L. 2014. Identification and characterization of microRNAs expressed in chicken skeletal muscle. *Genetics and Molecular Research* **13**(1): 1465-1479.
- Billerey, C., Boussaha, M., Esquerré, D., Rebours, E., Djari, A., Meersseman, C., Klopp, C., Gautheret, D., and Rocha, D. 2014. Identification of large intergenic non-coding RNAs in bovine muscle using next-generation transcriptomic sequencing. *BMC genomics* **15**(1): 499.
- Bolger, A., Lohse, M., and Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**(15): 2114-2120.
- Bonnet, A., Dai, F., Brand-Saber, B., and Duprez, D. 2010. Vestigial-like 2 acts downstream of MyoD activation and is associated with skeletal muscle differentiation in chick myogenesis. *Mechanisms of development* **127**(1): 120-136.
- Borello, U., Berarducci, B., Murphy, P., Bajard, L., Buffa, V., Piccolo, S., Buckingham, M., and Cossu, G. 2006. The Wnt/ β -catenin pathway regulates Gli-mediated Myf5 expression during somitogenesis. *Development* **133**(18): 3723-3732.
- Borycki, A., Brown, A., and Emerson, C. 2000. Shh and Wnt signaling pathways converge to control Gli gene activation in avian somites. *Development* **127**(10): 2075-2087.
- Bosch-Comas, A., Lindsten, K., Gonzalez-Duarte, R., Masucci, M., and Marfany, G. 2006. The ubiquitin-specific protease USP25 interacts with three sarcomeric proteins. *Cellular and molecular life sciences* **63**(6): 723-734.
- Cabianca, D.S., Casa, V., Bodega, B., Xynos, A., Ginelli, E., Tanaka, Y., and Gabellini, D. 2012. A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. *Cell* **149**(4): 819-831.
- Cabili, M.N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., and Rinn, J.L. 2011. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes & development* **25**(18): 1915-1927.
- Capdevila, J., Tabin, C., and Johnson, R.L. 1998. Control of dorsoventral somite patterning by Wnt-1 and β -catenin. *Developmental biology* **193**(2): 182-194.
- Carraway, K.L., and Burden, S.J. 1995. Neuregulins and their receptors. *Current opinion in neurobiology* **5**(5): 606-612.
- Cesana, M., Cacchiarelli, D., Legnini, I., Santini, T., Sthandier, O., Chinappi, M., Tramontano, A., and Bozzoni, I. 2011. A long noncoding RNA controls muscle differentiation by functioning as a competing

- endogenous RNA. *Cell* **147**(2): 358-369.
- Chakkalakal, J.V., Jones, K.M., Basson, M.A., and Brack, A.S. 2012. The aged niche disrupts muscle stem cell quiescence. *Nature* **490**(7420): 355-360.
- Chen, X., Ma, Q., Tang, M., and Ji, C. 2007. Development of breast muscle and meat quality in Arbor Acres broilers, Jingxing 100 crossbred chickens and Beijing fatty chickens. *Meat science* **77**(2):220-227.
- Curthoys, N.P., and Watford, M. 1995. Regulation of glutaminase activity and glutamine metabolism. *Annual review of nutrition* **15**(1): 133-159.
- Dey, B.K., Pfeifer, K., and Dutta, A. 2014. The H19 long noncoding RNA gives rise to microRNAs miR-675-3p and miR-675-5p to promote skeletal muscle differentiation and regeneration. *Genes & development* **28**(5): 491-501.
- Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., and Schlesinger, F. 2012. Landscape of transcription in human cells. *Nature* **489**(7414): 101-108.
- Du, M., Yan, X., Tong, J.F., Zhao, J., and Zhu, M.J. 2010. Maternal obesity, inflammation, and fetal skeletal muscle development. *Biology of reproduction* **82**(1): 4-12.
- Du, M., Zhao, J.X., Yan, X., Huang, Y., Nicodemus, L.V., Yue, W., McCormick, R., and Zhu, M. 2011. Fetal muscle development, mesenchymal multipotent cell differentiation, and associated signaling pathways. *Journal of animal science* **89**(2): 583-590.
- Gao, F., Fang, Q., Zhang, R., Lu, J., Lu, H., Wang, C., Ma, X., XU, J., Jia, W., and Xiang, K. 2009. Polymorphism of DsbA-L gene associates with insulin secretion and body fat distribution in Chinese population. *Endocrine journal* **56**(3): 487-494.
- Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., and Mistry, J. 2014. Pfam: the protein families database. *Nucleic acids research* **42**(D1): D222-D230.
- Fujimoto, S., Koga, H., Takeda, H., Tone, N., Kadowaki, M., and Ishibashi, T. 1996. Role of taste-active components, glutamic acid, 5'-inosinic acid and potassium ion in taste of chicken meat extract. *Animal Science and Technology (Japan)* **67**(5): 423-429.
- Gao, P., Tchernyshyov, I., Chang, T.C., Lee, Y.S., Kita, K., Ochi, T., Zeller, K.I., De Marzo, A.M., Van Eyk, J.E., and Mendell, J.T. 2009. c-Myc suppression of miR-23a/b enhances mitochondrial glutaminase expression and glutamine metabolism. *Nature* **458**(7239): 762-765.
- Gassmann, M., Casagrande, F., Orioli, D., and Simon, H. 1995. Aberrant neural and cardiac development in mice lacking the ErbB4 neuregulin receptor. *Nature* **378**(6555): 390-394.
- Gecz, J., Baker, E., Donnelly, A., Ming, J.E., McDonald-McGinn, D.M., Spinner, N.B., Zackai, E.H., Sutherland, G.R., and Mulley, J.C. 1999. Fibroblast growth factor homologous factor 2 (FHF2): gene structure, expression and mapping to the Börjeson-Forssman-Lehmann syndrome region in Xq26 delineated by a duplication breakpoint in a BFLS-like patient. *Human genetics* **104**(1): 56-63.
- Ghosh, S., Watanabe, R.M., Hauser, E.R., Valle, T., Magnuson, V.L., Erdos, M.R., Langefeld, C.D., Balow, J., Ally, D.S., and Kohtamaki, K. 1999. Type 2 diabetes: evidence for linkage on chromosome 20 in 716 Finnish affected sib pairs. *Proceedings of the National Academy of Sciences* **96**(5): 2198-2203.
- Gloss, B.S., and Dinger, M.E. 2016. The specificity of long noncoding RNA expression. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* **1859**(1): 16-22.
- Gong, C., Li, Z., Ramanujan, K., Clay, I., Zhang, Y., Lemire-Brachat, S., and Glass, D.J. 2015. A long non-coding RNA, LncMyoD, regulates skeletal muscle differentiation by blocking IMP2-mediated mRNA translation. *Developmental cell* **34**(2): 181-191.
- Green, R., Wilkins, C., Ferris, M. T., and Gale, M. 2017. RNA-Seq in the Collaborative Cross. *Systems*

- Genetics: Methods and Protocols, 251-263.
- Gustafsson, M.K., Pan, H., Pinney, D.F., Liu, Y., Lewandowski, A., Epstein, D.J., and Emerson, C.P. 2002. Myf5 is a direct target of long-range Shh signaling and Gli regulation for muscle specification. *Genes & development* **16**(1): 114-126.
- Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., and Cassady, J.P. 2009. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**(7235): 223-227.
- Guttman, M., Garber, M., Levin, J.Z., Donaghey, J., Robinson, J., Adiconis, X., Fan, L., Koziol, M.J., Gnirke, A., and Nusbaum, C. 2010. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nature biotechnology* **28**(5): 503-510.
- Han, R., Li, Z., Li, M., Li, J., Lan, X., Sun, G., Kang, X., and Chen, H. 2011. Novel 9-bp indel in visfatin gene and its associations with chicken growth. *British poultry science* **52**(1): 52-57.
- Huang, D.W., Sherman, B.T., and Lempicki, R.A. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* **4**(1): 44-57.
- Hung, T., and Chang, H.Y. 2010. Long noncoding RNA in genome regulation: prospects and mechanisms. *RNA biology* **7**(5): 582-585.
- Jordan, T., Jiang, H., Li, H., and DiMario, J.X. 2004. Inhibition of ryanodine receptor 1 in fast skeletal muscle fibers induces a fast-to-slow muscle fiber type transition. *Journal of cell science* **117**(25): 6175-6183.
- Kallen, A.N., Zhou, X.B., Xu, J., Qiao, C., Ma, J., Yan, L., Lu, L., Liu, C., Yi, J.-S., and Zhang, H. 2013. The imprinted H19 lincRNA antagonizes let-7 microRNAs. *Molecular cell* **52**(1): 101-112.
- Kashi, K., Henderson, L., Bonetti, A., and Carninci, P. 2016. Discovery and functional analysis of lincRNAs: Methodologies to investigate an uncharacterized transcriptome. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* **1859**(1): 3-15.
- Kong, L., Zhang, Y., Ye, Z.Q., Liu, X.Q., Zhao, S.Q., Wei, L., and Gao, G. 2007. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic acids research* **35**(suppl 2): W345-W349.
- Koufariotis, L.T., Chen, Y.P.P., Chamberlain, A., Vander Jagt, C., and Hayes, B.J. 2015. A catalogue of novel bovine long noncoding RNA across 18 tissues. *PloS one* **10**(10): e0141225.
- Kretz, M., Siprashvili, Z., Chu, C., Webster, D.E., Zehnder, A., Qu, K., Lee, C.S., Flockhart, R.J., Groff, A.F., and Chow, J. 2013. Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature* **493**(7431): 231-235.
- Lee, J.H., Reed, D.R., Li, W.D., Xu, W., Joo, E.J., Kilker, R.L., Nanthakumar, E., North, M., Sakul, H., and Bell, C. 1999. Genome scan for human obesity and linkage to markers in 20q13. *The American Journal of Human Genetics* **64**(1): 196-209.
- Lee, J.T. 2012. Epigenetic regulation by long noncoding RNAs. *Science* **338**(6113): 1435-1439.
- Li, M., Sun, X., Cai, H., Sun, Y., Plath, M., Li, C., Lan, X., Lei, C., Lin, F., and Bai, Y. 2016. Long non-coding RNA ADNCR suppresses adipogenic differentiation by targeting miR-204. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* **1859**(7): 871-882.
- Li, T., Wang, S., Wu, R., Zhou, X., Zhu, D., and Zhang, Y. 2012. Identification of long non-protein coding RNAs in chicken skeletal muscle using next generation sequencing. *Genomics* **99**(5), 292-298.
- Livak, K.J., and Schmittgen, T.D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods* **25**(4): 402-408.
- Lu, H., Shi, X., Wu, G., Zhu, J., Song, C., Zhang, Q., and Yang, G. 2015. FGF13 regulates proliferation and

- differentiation of skeletal muscle by down-regulating *Spry1*. *Cell proliferation* **48**(5): 550-560.
- Mattick, J.S. 2009. The genetic signatures of noncoding RNAs. *PLoS Genet* **5**(4): e1000459.
- NariNç, D., Aksoy, T., Öneç, A., and ÇÜrek, D. 2015. The influence of body weight on carcass and carcass part yields, and some meat quality traits in fast- and slow-growing broiler chickens. *Kafkas Üniversitesi Veteriner Fakültesi Dergisi*, **21**(4).
- Necsulea, A., Soumillon, M., Warnefors, M., Liechti, A., Daish, T., Zeller, U., Baker, J.C., Grützner, F., and Kaessmann, H. 2014. The evolution of lncRNA repertoires and expression patterns in tetrapods. *Nature* **505**(7485): 635-640.
- Niu, W., Huang, C., Nawaz, Z., Levy, M., Somwar, R., Li, D., Bilan, P.J., and Klip, A. 2003. Maturation of the regulation of GLUT4 activity by p38 MAPK during L6 cell myogenesis. *Journal of Biological Chemistry* **278**(20): 17953-17962.
- Olivier, M., Hsiung, C.A., Chuang, L.M., Ho, L.T., Ting, C.T., Bustos, V.I., Lee, T.M., de Witte, A., Chen, Y.-D.I., and Olshen, R. 2004. Single nucleotide polymorphisms in protein tyrosine phosphatase 1 β (*PTPN1*) are associated with essential hypertension and obesity. *Human molecular genetics* **13**(17): 1885-1892.
- Ørom, U.A., Derrien, T., Beringer, M., Gumireddy, K., Gardini, A., Bussotti, G., Lai, F., Zytnicki, M., Notredame, C., and Huang, Q. 2010. Long noncoding RNAs with enhancer-like function in human cells. *Cell* **143**(1): 46-58.
- Ouyang, H., He, X., Li, G., Xu, H., Jia, X., Nie, Q., and Zhang, X. 2015. Deep sequencing analysis of miRNA Expression in breast muscle of fast-growing and slow-growing broilers. *International journal of molecular sciences* **16**(7): 16242-16262.
- Plowman, G.D., Green, J.M., Culouscou, J.M., Carlton, G.W., Rothwell, V.M., and Buckley, S. 1993. Heregulin induces tyrosine phosphorylation of HER4/p180erbB4. *Nature* **366**(6454): 473.
- Ponsuksili, S., Murani, E., Phatsara, C., Schwerin, M., Schellander, K., and Wimmers, K. 2009. Porcine muscle sensory attributes associate with major changes in gene networks involving *CAPZB*, *ANKRD1*, and *CTBP2*. *Functional & integrative genomics* **9**(4): 455-471.
- Puigserver, P., Rhee, J., Lin, J., Wu, Z., Yoon, J.C., Zhang, C.Y., Krauss, S., Mootha, V.K., Lowell, B.B., and Spiegelman, B.M. 2001. Cytokine stimulation of energy expenditure through p38 MAP kinase activation of PPAR γ coactivator-1. *Molecular cell* **8**(5): 971-982.
- Quinlan, A.R., and Hall, I.M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**(6): 841-842.
- Relaix, F., Rocancourt, D., Mansouri, A., and Buckingham, M. 2005. A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. *Nature* **435**(7044): 948-953.
- Ren, H., Wang, G., Chen, L., Jiang, J., Liu, L., Li, N., Zhao, J., Sun, X., and Zhou, P. 2016. Genome-wide analysis of long non-coding RNAs at early stage of skin pigmentation in goats (*Capra hircus*). *BMC genomics* **17**(1): 67.
- Ridgeway, A.G., and Skerjanc, I.S. 2001. Pax3 is essential for skeletal myogenesis and the expression of *Six1* and *Eya2*. *Journal of Biological Chemistry* **276**(22): 19033-19039.
- Schmitt, A.M., and Chang, H.Y. 2013. Gene regulation: Long RNAs wire up cancer growth. *Nature* **500**(7464): 536-537.
- Schreurs, F. 2000. Post-mortem changes in chicken muscle. *World's poultry science journal* **56**(04): 319-346.
- Shang, Y., Zhang, C., Wang, S., Xiong, F., Zhao, C., Peng, F., Feng, S., Yu, M., Li, M., and Zhang, Y. 2007. Activated β -catenin induces myogenesis and inhibits adipogenesis in BM-derived mesenchymal

- stromal cells. *Cytotherapy* **9**(7): 667-681.
- Shield, A.J., Murray, T.P., Cappello, J.Y., Coggan, M., and Board, P.G. 2010. Polymorphisms in the human glutathione transferase Kappa (GSTK1) promoter alter gene expression. *Genomics* **95**(5): 299-305.
- Soro, A., Pajukanta, P., Lilja, H.E., Ylitalo, K., Hiekkalinna, T., Perola, M., Cantor, R.M., Viikari, J.S., Taskinen, M.R., and Peltonen, L. 2002. Genome scans provide evidence for low-HDL-C loci on chromosomes 8q23, 16q24.1-24.2, and 20q13.11 in Finnish families. *The American journal of human genetics* **70**(5): 1333-1340.
- Sun, K., Wang, H., and Sun, H. 2017. mTFkb: a knowledgebase for fundamental annotation of mouse transcription factors. *Scientific Reports* **7**(1).
- Sun, K., Zhao, Y., Wang, H., and Sun, H. 2014. Sebnif: an integrated bioinformatics pipeline for the identification of novel large intergenic noncoding rnas (lincrnas)-application in human skeletal muscle cells. *PloS one* **9**(1): e84500.
- Sun, L., Goff, L.A., Trapnell, C., Alexander, R., Lo, K.A., Haciosuleyman, E., Sauvageau, M., Tazon-Vega, B., Kelley, D.R., and Hendrickson, D.G. 2013a. Long noncoding RNAs regulate adipogenesis. *Proceedings of the National Academy of Sciences* **110**(9): 3387-3392.
- Sun, L., Luo, H., Bu, D., Zhao, G., Yu, K., Zhang, C., Liu, Y., Chen, R., and Zhao, Y. 2013b. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic acids research* **41**(17): e166-e166.
- Taye, M., Kim, J., Yoon, S.H., Lee, W., Hanotte, O., Dessie, T., Kemp, S., Mwai, O.A., Caetano-Anolles, K., and Cho, S. 2017. Whole genome scan reveals the genetic signature of African Ankole cattle breed and potential for higher quality beef. *BMC genetics* **18**(1): 11.
- Theobald, J., and DiMario, J.X. 2011. Lineage-based primary muscle fiber type diversification independent of MEF2 and NFAT in chick embryos. *Journal of muscle research and cell motility* **31**(5-6): 369-381.
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols* **7**(3): 562-578.
- Wang, L., Feng, Z., Wang, X., Wang, X., and Zhang, X. 2010. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* **26**(1): 136-138.
- Wang, X., Yu, J., Zhang, Y., Gong, D., and Gu, Z. 2012. Identification and characterization of microRNA from chicken adipose tissue and skeletal muscle. *Poultry science* **91**(1): 139-149.
- Wang, Y., Xue, S., Liu, X., Liu, H., Hu, T., Qiu, X., Zhang, J., and Lei, M. 2016. Analyses of long non-coding RNA and mRNA profiling using RNA sequencing during the pre-implantation phases in pig endometrium. *Scientific reports* **6**:20238.
- Wang, Z., Zhang, J., Chen, W., and Ji, C. 2006. Study on growth, nutrient deposition and meat quality among Gushi chicks, Roman Layers and Avain Broilers. *Chinese Journal of Animal Nutrition* (02):117-121. (in Chinese)
- Washietl, S., Kellis, M., and Garber, M. 2014. Evolutionary dynamics and tissue specificity of human long noncoding RNAs in six mammals. *Genome research* **24**(4): 616-628.
- Wei, H.K., Zhou, Y., Jiang, S., Tao, Y.X., Sun, H., Peng, J., and Jiang, S. 2013. Feeding a DHA-enriched diet increases skeletal muscle protein synthesis in growing pigs: association with increased skeletal muscle insulin action and local mRNA expression of insulin-like growth factor 1. *British Journal of Nutrition* **110**(04): 671-680.
- Weikard, R., Hadlich, F., and Kuehn, C. 2013. Identification of novel transcripts and noncoding RNAs in

- bovine skin by deep next generation sequencing. *BMC genomics* **14**(1): 789.
- White, N.M., Cabanski, C.R., Silva-Fisher, J.M., Dang, H.X., Govindan, R., and Maher, C.A. 2014. Transcriptome sequencing reveals altered long intergenic non-coding RNAs in lung cancer. *Genome biology* **15**(8): 429.
- Wu, H., Zhao, S., and Fan, B. 2010. Investigation of effects of the MKK3 and MKK6 genes on meat production traits in the pig (Brief Report). *Archiv Fur Tierzucht* **53**(2): 242-245.
- Yan, J., Gan, L., Chen, D., and Sun, C. 2013. Adiponectin impairs chicken preadipocytes differentiation through p38 MAPK/ATF-2 and TOR/p70 S6 kinase pathways. *PLoS one* **8**(10): e77716.
- Zetser, A., Gredinger, E., and Bengal, E. 1999. p38 mitogen-activated protein kinase pathway promotes skeletal muscle differentiation participation of the MEF2C transcription factor. *Journal of Biological Chemistry* **274**(8): 5193-5200.
- Zhang, T., Zhang, X., Han, K., Zhang, G., Wang, J., Xie, K., and Xue, Q. 2017a. Genome-wide analysis of lncRNA and mRNA expression during differentiation of abdominal preadipocytes in the chicken. *G3: Genes, Genomes, Genetics* **7**(3): 953-966.
- Zhang, T., Zhang, X., Han, K., Zhang, G., Wang, J., Xie, K., Xue, Q., and Fan, X. 2017b. Analysis of long noncoding RNA and mRNA using RNA sequencing during the differentiation of intramuscular preadipocytes in chicken. *PLoS one* **12**(2): e0172389.
- Zhao, X., Huang, Z., Liu, X., Chen, Y., Gong, W., Yu, K., Qin, L., Chen, H., and Mo, D. 2013. The switch role of the Tmod4 in the regulation of balanced development between myogenesis and adipogenesis. *Gene* **532**(2): 263-271.
- Zhou, L., Sun, K., Zhao, Y., Zhang, S., Wang, X., Li, Y., Lu, L., Chen, X., Chen, F., Bao, X., Zhu, X., Wang, L., Tang, L., Esteban, M., Wang, C., Jauch, R., Sun, H., and Wang, H. 2015. Linc-YY1 promotes myogenic differentiation and muscle regeneration through an interaction with the transcription factor YY1. *Nature communications* **6**.

Table 1. Total RNA detection information from breast muscle of GS and AA chicken.

Note: RIN > 7.0 and 28S/18S > 0.7 total RNA quality conformance; G1, G2, G3 represent GS

Sample	Con.($\mu\text{g}/\mu\text{L}$)	Vol.(μL)	Amount.(μg)	A260/A280	2100 Result		Result
					RIN	28s/18s	
G1	0.383	100	38.3	1.87	8	1.4	qualified
G2	0.396	100	39.61	1.89	7.9	1.2	qualified
G3	0.349	100	34.93	1.89	7.6	1	qualified
A1	0.209	100	20.91	1.92	8.2	1.3	qualified
A2	0.204	100	20.41	1.91	8.2	1.3	qualified
A3	0.581	100	58.11	1.91	8.2	1.3	qualified

chicken three biology repetition, A1, A2, A3 represent AA chicken three biology repetition.

Draft

Figure Legends

Fig 1. Detection by electrophoresis of the total RNA isolated from the breast muscle of GS and AA chickens.

Fig 2. Screened and genomic features of lncRNAs. (a) Novel lncRNAs screened by CPC, CNCI, and Pfam software. (b) Distribution of transcript lengths of lncRNAs (red) and mRNA (blue). (c) Distribution of exon numbers of lncRNAs (red) and mRNAs (blue). (d) Comparison of the expression levels of known (red) and novel (green) lncRNAs, plotted as fragments per kilobase of exon per million fragments mapped (FPKM). The dashed line represents the respective mean for different gene categories.

Fig 3. Six screened differentially expressed lncRNAs were randomly selected by qRT-PCR. Gene expression levels were examined in different tissues by qRT-PCR; the data represent mean \pm SE (n= 3). The GAPDH was used as the control. *P < 0.05, **P < 0.01.

Fig 4. GO (Gene Ontology) enrichment analysis of different lncRNA target genes.

The circle represents the biological process, the triangle represents the cellular component, and the square represents the molecular function. The size of the circle, triangle, and square indicates the number of genes enriched for related GO terms, and a larger icon indicates more genes enriched for a specific term. Color indicates significance, and a deeper red color indicates a smaller *P*-value.

Fig 5. Scatter plot of the top 30 KEGG enrichments.

The abscissa represents the richness factor, and the ordinate represents the enriched pathway terms. The size of the dot indicates the number of genes enriched for the pathway, and a larger dot indicates more genes enriched for a specific pathway. Color indicates significance, and a deeper red color indicates a smaller P -value.

Draft

Supplementary Files

gen-2017-0114.R2Suppla: **Fig S1.** Agilent 2100 analysis of total RNA isolated from breast muscle of GS and AA chickens.

gen-2017-0114.R2Supplb: **Fig S2. Quality analyses of sequenced reads.** G and A represent Gushi and Arbor Acres chickens, respectively. In Q-value boxplot, the blue line indicates the mean quality value, the green area is the high quality number, the orange area is the eligible quality number, and the red area is the low quality number.

gen-2017-0114.R2Supplc: **Fig S3. Correlation analysis of significantly differentially expressed lncRNAs in GS and AA chickens.** G and A represent Gushi and Arbor Acres chickens, respectively. Blue dots represent down-regulated lncRNAs, red dots represent up-regulated lncRNAs, and gray dots represent the not-significant lncRNAs.

gen-2017-0114.R2Suppld: **Fig S4. Nine screened differentially expressed lncRNAs were randomly selected by qRT-PCR.** Gene expression levels were examined in different tissues by qRT-PCR; the data represent mean \pm SE (n= 3). The GAPDH was used as the control. *P < 0.05, **P < 0.01.

gen-2017-0114.R2Supple:**Table S1.** Specific primers used for lncRNA qRT-PCR

validation.

gen-2017-0114.R2Supplf:**Table S2.** lncRNA quantitative expression.

Note: IDs beginning with TCONS are novel lncRNAs; IDs beginning with NON and ENS are known lncRNAs.

gen-2017-0114.R2Supplg: **Table S3.** Analysis of lncRNA differential expression.

Note: IDs beginning with TCONS are novel lncRNAs; IDs beginning with NON and ENS are known lncRNAs.

gen-2017-0114.R2Supplh:**Table S4.** Prediction of lncRNA target genes by cis function.

Note: IDs beginning with TCONS are novel lncRNAs; IDs beginning with ENS and NON are known lncRNAs.

Draft

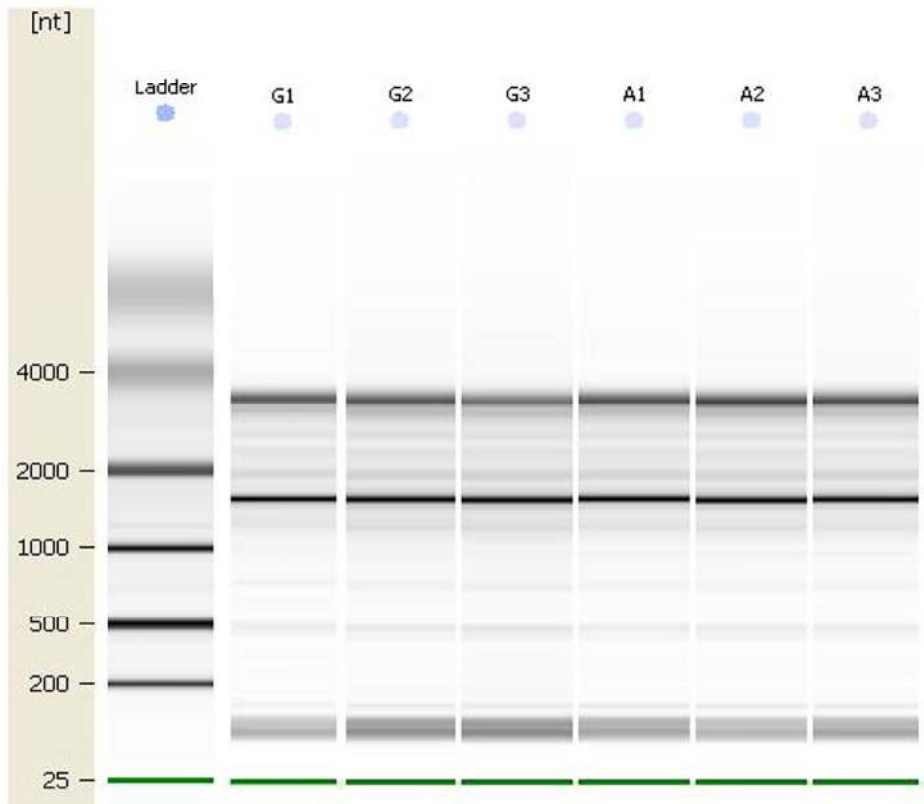


Fig 1. Detection by electrophoresis of the total RNA isolated from the breast muscle of GS and AA chickens.
134x115mm (300 x 300 DPI)

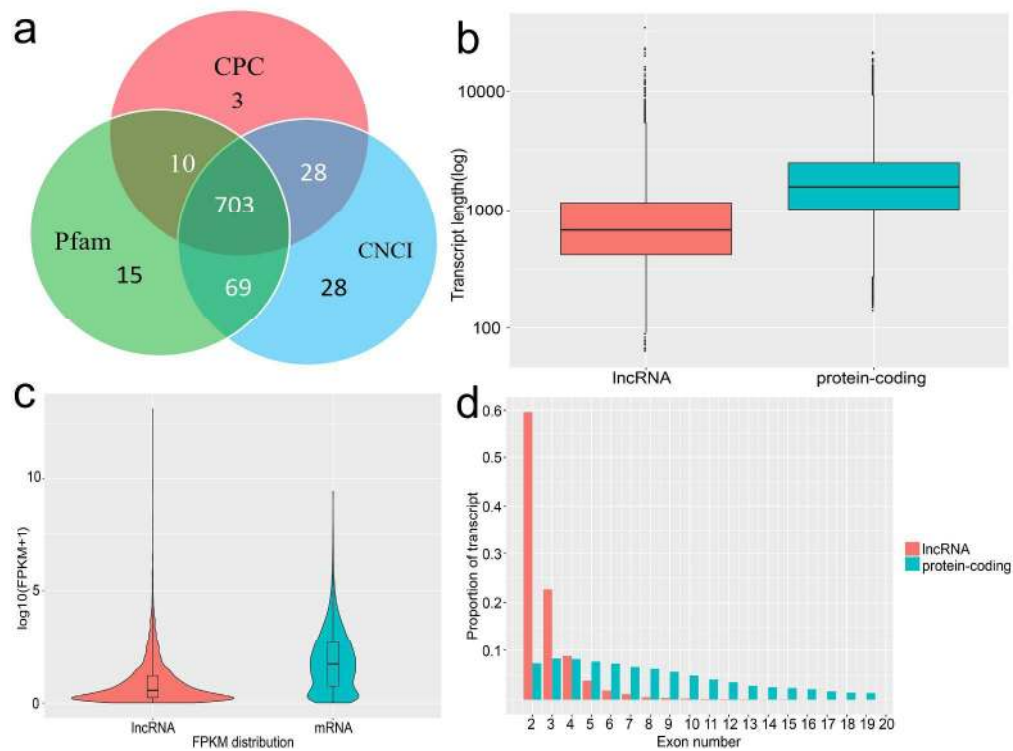


Fig 2. Screened and genomic features of lncRNAs. (a) Novel lncRNAs screened by CPC, CNCI, and Pfam software. (b) Distribution of transcript lengths of lncRNAs (red) and mRNA (blue). (c) Distribution of exon numbers of lncRNAs (red) and mRNAs (blue). (d) Comparison of the expression levels of known (red) and novel (green) lncRNAs, plotted as fragments per kilobase of exon per million fragments mapped (FPKM). The dashed line represents the respective mean for different gene categories.

254x190mm (300 x 300 DPI)

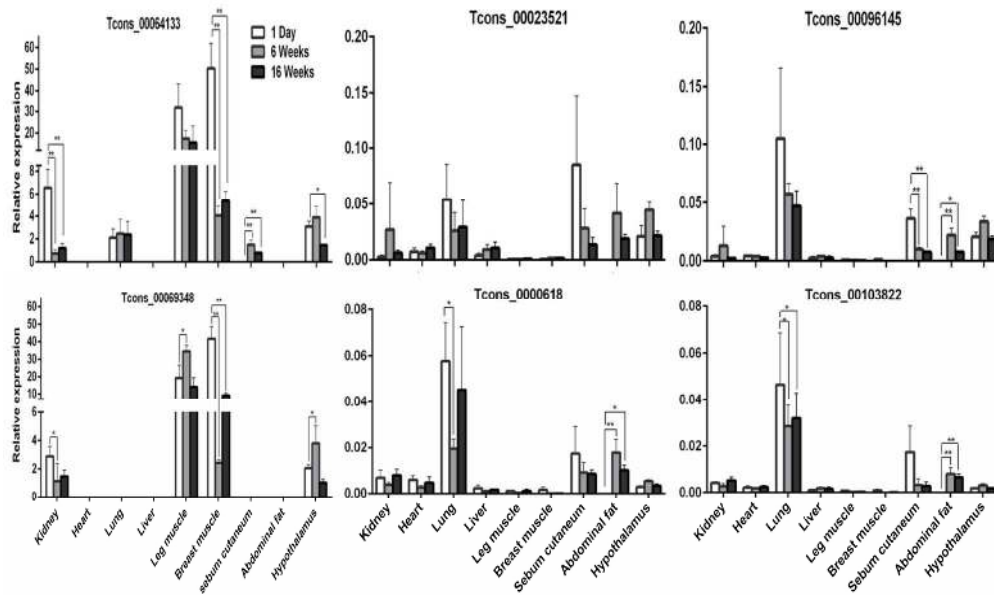


Fig 3. Six screened differentially expressed lncRNAs were randomly selected by qRT-PCR. † † Gene expression levels were examined in different tissues by qRT-PCR; the data represent mean \pm SE (n= 3). The GAPDH was used as the control. *P < 0.05, **P < 0.01.

137x82mm (300 x 300 DPI)

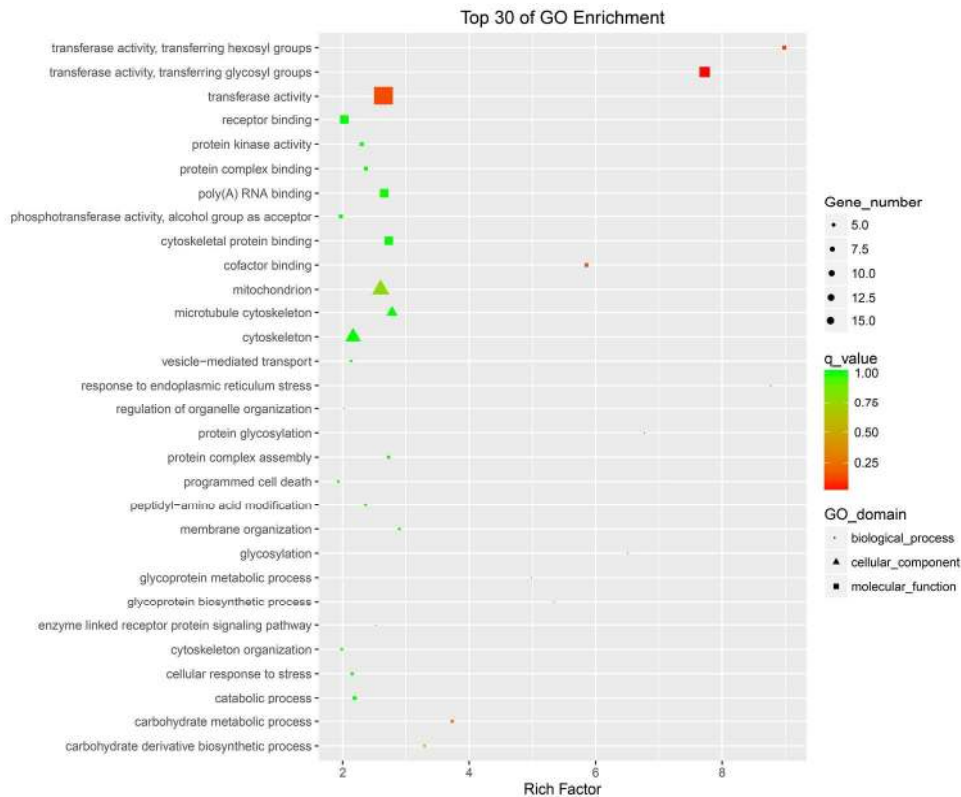


Fig 4. GO (Gene Ontology) enrichment analysis of different lncRNA target genes. The circle represents the biological process, the triangle represents the cellular component, and the square represents the molecular function. The size of the circle, triangle, and square indicates the number of genes enriched for related GO terms, and a larger icon indicates more genes enriched for a specific term. Color indicates significance, and a deeper red color indicates a smaller P-value.

203x162mm (300 x 300 DPI)

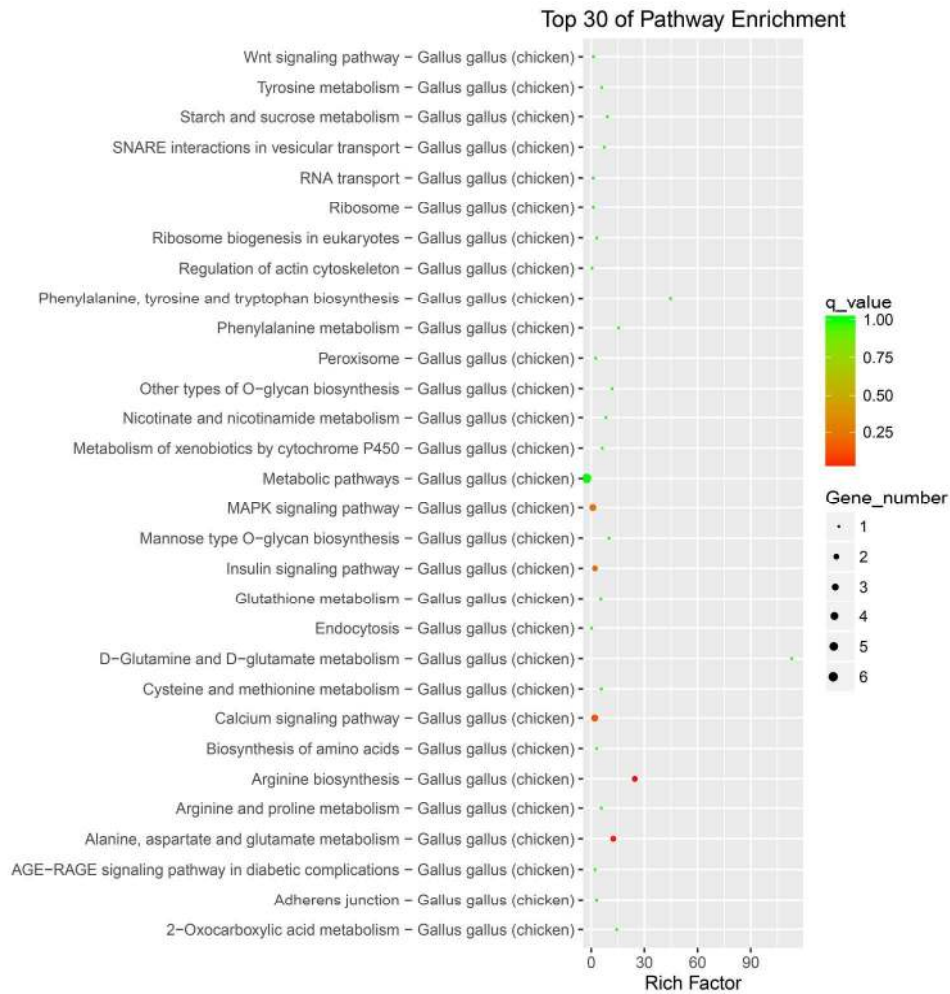


Fig 5. Scatter plot of the top 30 KEGG enrichments.

The abscissa represents the richness factor, and the ordinate represents the enriched pathway terms. The size of the dot indicates the number of genes enriched for the pathway, and a larger dot indicates more genes enriched for a specific pathway. Color indicates significance, and a deeper red color indicates a smaller P-value.

203x203mm (300 x 300 DPI)