Gene 647 (2018) 164–173

Contents lists available at ScienceDirect

Gene

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

Research paper

Comprehensive analysis of lncRNAs and mRNAs with associated co-expression and ceRNA networks in C2C12 myoblasts and myotubes

Rui Chen[a,*,1], Ting Jiang[b,1], Yanling She[a,1], Sidong Xie[b], Shanyao Zhou[a], Cheng Li[a], Jun Ou[c], Yulin Liu[c]

[a] Guangdong Traditional Medical and Sports Injury Rehabilitation Research Institute, Guangdong Second Provincial General Hospital, 466 Xin Gang Zhong Road, Guangzhou 510317, China
[b] Department of Radiology, The Third Affiliated Hospital, Sun Yat-sen University, 600 Tian He Road, Guangzhou 510630, China
[c] Guangzhou FitGene Biotechnology CO., LTD, Building D, 3 Ju Quan Road, Guangzhou 510663, China

ARTICLE INFO

Keywords:
Long non-coding RNA
Microarray analysis
Muscle differentiation
Bioinformatics analysis

ABSTRACT

Long non-coding RNAs (lncRNAs) are emerging as important regulators in the modulation of muscle development and muscle-related diseases. To explore potential regulators of muscle differentiation, we determined the expression profiles of lncRNAs and mRNAs in C2C12 mouse myoblast cell line using microarray analysis. Gene ontology and Kyoto Encyclopedia of Genes and Genomes pathway enrichment analyses were performed to explore their function. We also constructed co-expression, cis/trans-regulation, and competing endogenous RNA (ceRNA) networks with bioinformatics methods. We found that 3067 lncRNAs and 3235 mRNAs were differentially regulated (fold change ≥2.0). Bioinformatics analysis indicated that the principal functions of the transcripts were related to muscle structure development and morphogenesis. Co-expression analysis showed 261 co-expression relationships between 233 lncRNAs and 10 mRNAs, and nine lncRNAs interacted with myog and MEF2C collectively. Cis/trans-regulation prediction revealed that lncRNA Myh6 could be a valuable gene via cis-regulation, and lncRNAs such as 2310043L19Ris, V00821, and AK139352 may participate in particular pathways regulated by transcription factors, including myog, myod1, and foxo1. The myog-specific ceRNA network covered 10 lncRNAs, 378 miRNAs, and 1960 edges. The upregulated lncRNAs Filip1, Myl1, and 2310043L19Rik may promote myog expression by acting as ceRNAs. Our results offer a new perspective on the modulation of lncRNAs in muscle differentiation.

1. Introduction

Long non-coding RNAs (lncRNAs) are a class of non-coding RNAs longer than 200 nucleotides that make up the majority of the transcriptome (Bolta a et al., 2016). LncRNAs were initially thought to be transcriptional noise; however, recent studies have indicated that they are involved in various physiological processes through interactions with chromatin modifiers, genomic imprinting, transcriptional activation, transcriptional interference, and nuclear transport (L. Wang et al., 2015c). The discovery and characterization of lncRNAs have helped elucidate some issues of genome complexity, providing a better understanding of biology from the viewpoint of gene regulatory networks.

Approximately 40% of total body weight is skeletal muscle, an important part of the locomotion system (Grumati and Bonaldo, 2012; Luo et al., 2013). Muscle differentiation is a multistep process requiring myoblast migration, proliferation, determination, cell cycle withdrawal, fusion into multinucleated and highly specialized myotubes, and finally maturation (or differentiation) of myotubes into myofibers with contractile properties (Luo et al., 2013; Millay et al., 2013; Wang et al., 2012). This process is dependent on a series of specific transcription factors (TFs) such as the myogenic regulatory factors, including myogenic differentiation antigen (MyoD), myogenic factor (Myf)5, Myf6, myogenin (Myog) (Braun and Gautel, 2011), and the myocyte enhancer factor 2A–D (MEF2A-D) family (Buckingham and Rigby, 2014). Myog is essential for muscle metabolism and development and is activated early on in development when it requires a

Abbreviations: lncRNAs, long non-coding RNAs; ceRNAs, competing endogenous RNAs; TFs, specific transcription factors; MyoD, myogenic differentiation antigen; Myf, myogenic factor; myog, myogenin; MEF, myocyte enhancer factor; bHLH, basic helix loop helix; GO, Gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PCC, Pearson correlation coefficient

* Corresponding author.
E-mail address: rui.c.med@163.com (R. Chen).
[1] Rui Chen, Ting Jiang and Yanling She contributed equally to this work.

https://doi.org/10.1016/j.gene.2018.01.039
Received 14 November 2017; Received in revised form 29 December 2017; Accepted 9 January 2018
Available online 10 January 2018
0378-1119/

change in DNA methylation status (Antoniou et al., 2014). MyoD and Myf5 are unable to substitute myog's functions during differentiation. The myogenic program can be strictly controlled by two essential motifs of myog transcription: consensus binding sites for basic helix–loop–helix proteins and serum response factor proteins (Yee and Rigby, 1993). Myog expression plays a direct role in cell cycle withdrawal by downregulating cell cycle genes, inhibiting proliferation (Liu et al., 2012). Furthermore, mice that lack myog gene expression die after birth because of severe skeletal muscle deficiency, as myoblasts are unable to fuse into multinucleated myofibres (Hasty et al., 1993). Myog is also involved in regenerating adult myofibers; its expression is induced 4–5 days after muscle damage (Novitch et al., 1996). Our previous study found that myog is repressed and exhibits less spindly ring-shaped myotube formation under hypoxia (Chen et al., 2017).

There is increasing evidence that lncRNAs are also part of the muscle regulatory network (Sohi and Dilworth, 2014). To date, microRNAs (miRNAs) have been the most widely studied and characterized. Muscle-specific miR-1, miR-206, and miR-133 are necessary for muscle development and function. Overexpression of miR-1 or miR-206 accelerates myogenic differentiation, whereas overexpression of miR-133 promotes myoblast proliferation but inhibits the differentiation process (Townley-Tilson et al., 2010). In the past few years, lncRNAs have emerged as critical regulators of muscle development. The first lncRNA identified as playing a role in myogenesis was linc-MD1, which governs the timing of muscle differentiation by acting as a competing endogenous RNA (ceRNA) in mouse and human myoblasts (Cesana et al., 2011). H19, the first lncRNA described in mammalian cells, is decreased in Wilms tumor and rhabdosarcoma (Hao et al., 1993). More recently, Dey et al. reported that H19 promotes skeletal muscle differentiation, and its trans-regulatory function is mediated by miR-675-3p and miR-675-5p (Dey et al., 2014a, 2014b).

Although several studies on lncRNAs have been published, lncRNAs are only now being discovered in skeletal muscle. The formation of skeletal muscle fibers depends on the fusion of myoblasts to produce multinucleated muscle fibers. Muscle biologists frequently employ the differentiated C2C12 myoblast line as a cell model for mature myocytes (Han et al., 2017). In this study, we performed comprehensive analysis of lncRNAs and mRNAs with associated co-expression and ceRNA networks in C2C12 myoblasts and myotubes. Our findings may provide a new avenue for investigating muscle differentiation.

2. Materials and methods

2.1. Cell culture

The mouse myoblast cell line C2C12 (Stem Cell Bank, Chinese Academy of Sciences, Shanghai, China) was cultured in high glucose Dulbecco's Modified Eagle's Medium (DMEM; Gibco-BRL, Grand Island, NY, USA) supplemented with 10% fetal bovine serum (HyClone, Logan, UT, USA), 100 U/mL penicillin, and 100 μg/mL streptomycin in 5% CO_2 at 37 °C. When the cells reached 80–90% confluency, they were differentiated by incubation in DMEM containing 2% horse serum (HyClone).

2.2. Western blot analysis

Samples were lysed in radioimmunoprecipitation buffer containing protease inhibitors (Beyotime, Jiangsu, China) and phenylmethylsulfonyl fluoride to extract total protein. An equal amount of protein (20 μg) was separated by 10% sodium dodecyl sulfate polyacrylamide gel electrophoresis and transferred to polyvinylidene fluoride membranes. The membranes were blocked with 5% nonfat milk and incubated with primary antibodies targeting myog (1:500; Millipore, Billerica, MA, USA) and myosin heavy chain (MyHC) (1:1000; R&D Systems, Minneapolis, MN, USA) overnight at 4 °C. Then we incubated the membranes with goat anti-mouse or anti-rabbit secondary antibodies for 1 h at room temperature. Band intensity was detected using a chemiluminescence detector (KeyGEN Biotech, Jiangsu, China) and analyzed with a chemiluminescent imaging system (Tanon, Shanghai, China). Tubulin (1:5000; ABclonal, Woburn, MA, USA) was used as a control for protein quantification.

2.3. RNA extraction and quantity control

Total RNA was extracted from cells using TRIzol reagent (Life Technologies, Grand Island, NY, USA) according to the manufacturer's instructions. RNA quantity and quality were measured with a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). RNA integrity was assessed by standard denaturing agarose gel electrophoresis.

2.4. Microarray analysis

ArrayStar Mouse LncRNA Expression Microarray V3.0 containing approximately 35,923 lncRNAs and 24,881 coding transcripts was used for the global profiling of mouse lncRNAs and protein-coding transcripts. Sample labeling and array hybridization were performed according to the Agilent One-Color Microarray-Based Gene Expression Analysis protocol (Agilent Technology, Santa Clara, CA, USA). Data were collected using Agilent's Feature Extraction software (version 11.0.1.1). The Gene Expression Omnibus (GEO) accession number is GSE101499.

2.5. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses

GO analyses were performed to investigate differentially expressed mRNA attributes in any organism, including molecular function, biological processes, and cellular components. Fisher's exact test was used to determine whether there was more overlap between the differentially expressed (DE) gene list and the GO annotation list than would be expected by chance. p values denote the significance of GO term enrichment in the DE genes. The lower the p value, the more significant the GO term ($p \leq .05$). KEGG mapping pathways were used for pathway analysis. p values denote the significance of the pathway correlation. The lower the p value, the more significant the pathway (the p value cutoff was 0.05).

2.6. Quantitative real-time polymerase chain reaction (qRT-PCR)

The myogenic-specific genes myog and MyHC were detected to confirm the degree of muscle differentiation. Five upregulated (AK156028, AK052193, Gm10510, AK046587, and 4930452A19Rik) and four downregulated (Mboat1, Haus6, Gm19898, and Gm14537) lncRNAs were randomly selected for validation by qRT-PCR. Briefly, total RNA was extracted from cells, and aliquots of 1 μg RNA were reverse-transcribed to cDNA with PrimeScript™ RT Master Mix (Takara Biotechnology, Otsu, Japan). The abundance of lncRNAs was measured with SYBR® Green Mix (Takara Biotechnology). The qPCR procedure was performed with denaturation at 95 °C for 30 s, followed by 40 cycles of denaturation at 95 °C for 5 s, annealing at 60 °C for 30 s, and extension at 60 °C for 1 min. The expression of 18S ribosomal RNA was used as an internal control. The primers used for the examined mRNAs and lncRNAs are shown in Table 1.

2.7. Correlation and co-expression analysis

Pearson correlation coefficients (PCCs) were used as co-expression measures to construct gene co-expression networks based on mRNA and lncRNA expression levels. The absolute value of parameter PCC ≥ 0.99 ($p < .01$) and FDR < 0.01 was selected to construct the network.