Step of microRNA sequencing for porcine skeletal muscle

Dongting He

Abstract

Citation: Dongting He Step of microRNA sequencing for porcine skeletal muscle. **protocols.io** dx.doi.org/10.17504/protocols.io.idsca6e
Published: 13 Jun 2017

Protocol

Sample collection

Step 1.

The *longissimus dorsi* (LD) muscle tissues were isolated from pigs, and immediately frozen in liquid nitrogen and stored at -80℃ until RNA isolation. Here,taking two treatment groups as an example. Each treatnment was prepared at least three repetitions.

Small RNA library preparation and sequencing

Step 2.

For small RNA library construction, the muscle total RNA samples were isolated using mirVana™ miRNA isolation kit (Ambion, Austin, USA). The total RNA isolated from pigs generated small RNA librarie. The total RNA quality and purity were measured using a Bioanalyzer 2100 and RNA 6000 Nano LabChip Kit (Agilent, CA, USA) with RIN > 7.0. Approximately 1 μg total RNA per sample was used to prepare small RNA library according to protocol of TruSeq Small RNA Sample Prep Kits (Illumina, San Diego, USA). In general, the processing consisted of the following successive steps: 1040 nt RNA fragment were excised, purified from a polyacrylamide gel electrophoresis (PAGE), and ligated with 5’ and 3’ adaptors using T4 RNA ligase. Then the modified small RNA was reverse transcripted and amplificated by RT-PCR. Subsequently, the amplified cDNA constructs were purified from agarose gel. Finally the enriched cDNA was quantified in an Bioanalyzer 2100 (Agilent, CA, USA) and sequenced in a HiSeq 2500 sequencing system (Illumina, San Diego, USA) at the LC-BIO (Hangzhou, China).

Sequencing data analysis

Step 3.

The raw reads were subjected to the Illunina pipeline filter (Solexa 0.3), and then the dataset was further processed with an in-house program, ACGT101-miR (LC Science, Houston, Texas, USA) to remove sequencing adapters, junk reads and fragments < 18 nt and > 26 nt. Subsequently, the remaining 1826 nt reads were searched against the Rfam, NCBI and Repbase database to remove non-miRNAs (i.e. rRNA, tRNA, snRNA, snoRNA, mRNA and repeats). Those sequenced reads survived from above strict filter rules were deemed to ‘mappable reads’ or high quality reads and used for further analysis.

Identification of porcine miRNA

Step 4.

The mappable reads were mapped to the pig genome (Sscrofa 10.2) using NCBI Local BLAST, including three steps: (1) the mappable reads were aligned to porcine pre-miRNAs/miRNAs , and then to pre-miRNAs/miRNAs from 25 other mammals in miRBase 21.0; (2) the mapped pre-miRNAs /miRNAs in step1 were searched against the pig genome to determine their genomic locations and annotations in Ensemble (Sscrofa 10.2); (3)the unmapped sequences were further blasted against the pig genomes, and the hairpin RNA structures containing sequences were predicated from the flank 80 nt sequences using RNAfold software. After the above analysis, three kinds of porcine miRNAs were identified. First, reads map to specific miRNAs/pre-miRNAs in miRbase and the pre-miRNAs further map to the pig genome & EST, were defined as porcine known miRNAs, marked with “ssc-miR-”. Second, the reads map to other mammalian miRNAs/pre-miRNAs in miRbase and the pre-miRNAs further map to the pig genome & EST, were defined as porcine conserved miRNAs, marked with “PC-”. Third, there are two kinds of reads: (1) the reads can map to selected miRNAs/pre-miRNAs in miRbase. The mapped pre-miRNAs do not map to the genome, but the reads (and of course the miRNAs of the pre-miRNAs) map to genome. The extended genome sequences from the genome loci may form hairpins. (2) The reads do not map to selected pre-miRNAs in miRbase. But the reads map to genome & the extended genome sequences from genome may form hairpins. The above two were defined as porcine candidate miRNAs, they were also predicted novel miRNAs for pig, so marked as “PN-”. These identified porcine miRNAs were assembled into porcine unique miRNAs according to the unique miRNA sequence. In the above tables, the expression of miRNAs in six libraries were normalized by total mappable reads, and used for further analysis.

Differentially expressed miRNA analysis

Step 5.

The differentially expressed miRNAs were identified based on the normalized most abundant sequence reads. The normalized miRNA reads, > 1000 reads counts in either of treatment 1 and treatment 2, from the two treatments were compared using Student *t*-test. The significance threshold was set to be 0.05 in *t*-test. After that the top 50 most abundant DE miRNAs (differentially expressed miRNAs), according to the total reads counts of treatment 1 and treatment 2, were choosed to perform the Go and KEGG pathway analysis.

Functional analysis of miRNAs

Step 6.

To predict the genes targeted by most abundant miRNAs and differentially expressed miRNAs, two computational target prediction algorithms (TargetScan 6.2 and miRanda) [1, 2]were used to identify miRNA biding site. Finally, the data predicted by both algorithms were combined and the overlaps were calculated. As porcine genes were not included in the current version of the above-mentioned algorithms, prediction was performed using human miRNAs. The Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway of these miRNA targets were annotated using DAVID bioinformatics resources (http://david.abcc.ncifcrf.gov/) [3]. The small RNA sequence data and processed files should be submitted to NCBI Gene Expression Omnibus (GEO) to get an accession No..

Validation of miRNA expression by RT-qPCR and statistics analysis

Step 7.

The miRNAs levels were quantified using the S-poly(T) miRNA qPCR-assay method as described [4] using the reagent kit (Geneups, China). SnoRNA-202, a commonly used mouse internal reference, was used as the internal control for normalization. Mus snoRNA-202 and sus scrofa snoRNA-68 shared high sequence similarity when subjected to NCBI BlAST analysis, and Ct values of snoRNA-202 was consistent and had little variation from sample to sample. All of the data are expressed as the mean ± SD, or error bars depict SD. The $2^{-\Delta\Delta Ct}$ method was used to analyze real-time PCR data. Expression of miRNA was presented as the fold of the mean of treatment 1 group. The DE miRNAs (a, b, c...) among six samples (the muscle of treatment 1 and treatment 2 , each has three replicates) was analyzed through two-way ANOVA method, in which breeds and time points were factors. For the validation of sequenced result, the Pearson correlation coefficient (r) and corresponding significance value (*P*) was also calculated by SPSS software. Each fetus was considered as an experimental unit. Statistical significance was defined as $P < 0.05$.