Mol Biol Rep (2011) 38:3161–3171
DOI 10.1007/s11033-010-9987-3

Characterization of microRNAs from sheep (*Ovis aries*) using computational and experimental analyses

Xihui Sheng · Xuemei Song · Yan Yu · Lili Niu · Shangang Li · Hongbin Li · Caihong Wei · Tao Liu · Li Zhang · Lixin Du

Received: 23 October 2009 / Accepted: 20 January 2010 / Published online: 7 February 2010

Abstract *Ovis aries* is one of the most important agricultural livestock for meat production, and also is an ideal model organism for biological and comparative genomics studies. Many miRNAs have been reported for their important roles in developmental processes in various animals, but there is limited information about *O. aries* miRNAs. In this study, combining a computational method based on expressed sequence tag (EST) analysis with experimental identification based on small RNA cDNA library, we identified 31 miRNAs belong to 24 families in sheep, 2 of which were novel miRNAs which had never been previously identified in any species. Especially, we cloned 12 miRNAs from the sheep skeletal muscle, which were good candidate miRNAs to be studied about the miRNA-dependant regulated process of muscle development, and we identified four pairs of miRNA/miRNA* and one pair of miRNA-3p/miRNA-5p from sheep EST sequences. Expression analysis indicated that some miRNAs were expressed in a specific tissue, and the pair of miRNA-3p/miRNA-5p and one pair of miRNA/miRNA* had a similar relative expression pattern in some tissues, respectively. Further, we predicted 120 potential target genes of 31 oar-miRNAs on the 3′UTR of *O. aries* genes. Gene ontology analysis showed that most of these genes took part in the cellular process and metabolic process. Our results enriched the *O. aries* miRNA database and provided useful information for investigating biological functions of miRNAs and miRNA* in sheep.

Electronic supplementary material The online version of this article (doi:10.1007/s11033-010-9987-3) contains supplementary material, which is available to authorized users.

X. Sheng · C. Wei · T. Liu · L. Zhang · L. Du (✉)
National Center for Molecular Genetics and Breeding of Animal, Institute of Animal Sciences, Chinese Academy of Agricultural Sciences, No. 2 Yuanmingyuan West Road, Haidian District, Beijing 100193, People's Republic of China
e-mail: lxdu@263.net; xihuisheng@gmail.com

X. Sheng
College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA

X. Song
The Institute of Animal Husbandry and Veterinary Science, ZheJiang Academy Of Agricultural Sciences, Shiqiao Rd., Hangzhou, Zhejiang 310021, People's Republic of China

Y. Yu
Reproductive Medical Center, The Second Hospital Affiliated to Shandong University of Traditional Chinese Medicine, Jinan 250001, People's Republic of China

L. Niu
Division of Applied Life Science, Graduate School of Gyeongsang National University, Jinju 660-701, Korea

S. Li
Department of Laboratory Animal Sciences, School of Medicine, Shanghai Jiao Tong University, 200025 Shanghai, People's Republic of China

H. Li
Institute of Zoology, Chinese Academy of Sciences, 1 Beichen West Road, Chaoyang District, Beijing 100101, People's Republic of China

Keywords Sheep · MicroRNA · MicroRNA* · Expressed sequence tag (EST) · Small cDNA library · Target gene

Introduction

MicroRNAs (miRNAs) are non-coding RNAs, which are approximately 22 nucleotides, and can regulate the expression of target genes by binding to complementary sites. It is more prevalent for miRNAs to down-regulate the expression of target genes by binding to the complementary sites in transcripts and cause transcript degradation or translational repression [1, 2]. However, recent studies found that miRNAs could increase protein translation by binding to the complementary promoter sequences [3, 4]. Numerous cellular and developmental processes in various organisms have been reported to be regulated by miRNAs, such as neural development [5, 6], cell division [7], insulin secretion [8], fat metabolism [9], and even as tumour suppressor [10, 11] and oncogenes [12, 13].

There are basically two kinds of approaches to identify miRNAs. One is to sequence size-fractionated cDNA libraries. Many known miRNAs have been identified by this method [14, 15]. This method allows the identification of both conserved and unconserved miRNAs, but a limitation of which is that some miRNAs are expressed at low levels, and expressed at very specific stages or in rare cell types. In contrast, computational strategies, which may especially miss those that are not phylogenetically conserved, provide an efficient way to predict miRNAs and their targets by surveying genomic sequences or other databases like expressed sequence tags (ESTs), which are based on the secondary structure characteristics, phylogenetic conservation of both sequence and structure, and thermodynamic stability of hairpins. Up to now, computational approaches have been successfully applied in vertebrates [16–19], insect [17, 20, 21], and plant [22–24].

Ovis aries is one of the most important agricultural livestock for meat production, but there is limited information, with only four reported (miRBase 14.0, September 2009) [25–27], about the *O. aries* miRNAs. In our study, *O. aries* miRNAs were identified in two ways. The first method consisted of exploiting miRNAs in the sheep EST database using a computational approach based on the phylogenetic conservation of both miRNA sequence and the secondary structure of pre-miRNAs (precursor-microRNA). In the second, four small RNA cDNA libraries from sheep skeletal muscle, brain, spleen and liver tissues were constructed and sequenced. A total of 31 miRNAs belong to 24 families were identified. By now, several miRNAs were reported to regulate the process of skeletal muscle development, for example, miR-1 and miR-206 were found to contribute to the muscular hypertrophy of Texel sheep [28]. In our study, 12 miRNAs were isolated from the sheep skeletal muscle, which might be researched as the good candidates which could take part in the process of skeletal muscle development. In addition, we identified four pairs of miRNA/miRNA* and one pair of miRNA-3p/miRNA-5p from sheep EST sequences. Expression analysis for several miRNAs indicated that some were expressed in a specific tissue, and the pair of miRNA-3p/miRNA-5p and one pair of miRNA/miRNA* had a similar relative expression pattern in some tissues, respectively. Further, we predicted 120 potential target genes of 31 oar-miRNAs on the 3′UTR of *O. aries* genes. Most of the target genes were found to encode transcription factors which regulated animal cellular process and metabolic process.

Materials and methods

Computational identification of *O. aries* miRNAs

Animal miRNAs and their precursor sequences (total 6785 miRNAs from 49 species) were obtained from miRBase (Release 12:0 April 2008 at http://microrna.sanger.ac.uk/sequences/) [25–27]. The *O. aries* ESTs (total 209808 sequences) were downloaded from NCBI database (http://www.ncbi.nlm.nih.gov/). A five-step prediction method was used to identify *O. aries* miRNAs. First, alignment of known animal miRNAs was conducted by BioEdit software package [29] to remove redundant sequences. Second, we used remaining miRNAs as query sequences for BLAST searched against the *O. aries* ESTs with BLASTN, the parameter settings were as follows: E-value cut-off was 10, the number of descriptions and alignments were 1000. Third, the ESTs with no more than two mismatched nucleotides were chosen to blast with the known pre-miRNA sequences. Fourth, the ESTs which have >90% similarity with the corresponding known precursor sequences were selected to remove the repeated sequences and the protein-coding sequences by BLASTN and BLASTX program. The last step was to apply *mfold* program [30, 31] to further identify the pre-miRNAs. Four criteria were used: (1) the A + U content of the precursor sequences should range from 30 to 70%; (2) the mature miRNA should locate on one arm of the hairpin structure, and the corresponding positions of the mature miRNAs in their pre-miRNAs were nearly identical, which was calculated by δ-len(A,B) [27]. We used δ-len cut-off 10 as the default parameter; (3) the minimum free energy (MFE) of the secondary-structure for each potential pre-miRNA was less than −20 kcal/mol; (4) the hairpin must include at

least 16 bp within the first 22 nt of the miRNA, and should not contain large internal loops or bulges, particularly not large asymmetric bulges, as described in reference [28].

Small RNA cDNA library construction

Total RNA was isolated from four tissues (skeletal muscle, brain, liver, and spleen) of 6-month-old Texel sheep using the TRIzol reagent (Invitrogen) according to the manufacturer's instructions. Small RNAs cloning was performed by constructing a cDNA library for small RNAs as previously described [32] with some modification. Briefly, total RNA (~500 μg) was separated on a 15% denaturing polyacrylamide gel, and small RNAs of 18–26 nt in length were recovered, purified, and ligated to 3′ adapter (5′pCTGTAGGCACCATCAAx (x: DMT-O-C3-CPG)′3). The ligation products were gel-purified and ligated to 5′ adapter (5′ ATCGTaggcaccugaaa 3′, lowercase RNA). Then reverse transcription was performed (RT primer: 5′A TTGATGGTGCCTAC′3), followed by PCR amplification (5′primer: 5′ATCGTAGGCACCTGAAA′3, 3′primer: 5′ATTGATGGTGCCTACAG′3). The PCR products were gel-purified, cloned into pGEM-T Vector (Promega, USA) and transformed into competent cells. The recombinant plasmids carrying fragments were isolated from individual colonies and subjected to DNA sequencing analysis.

Sequence analysis

All cloned sequences longer than 18 nt were searched in the miRBase database (http://microrna.sanger.ac.uk/sequences/index.shtml) [25–27]. If it was identical or nearly identical with a registered miRNA from other animal species but not previously identified in sheep, it was considered as a new *O. aries* miRNA. The remaining sequences were searched against the bovine genome in UCSC genome sequence database (http://genome.ucsc.edu/cgi-bin/hgBlat) and NCBI (http://www.ncbi.nlm.nih.gov), because the sheep genome sequencing have not completed yet, and there is a high genome identity between sheep and cattle. If the sequence had a perfect match or fewer than two mismatches with the bovine genome sequences, then it was compared with the RefSeq protein database, and the RefSeq RNA database in Genbank. If a sequence was identical with a known rRNA, tRNA, mRNA, mtRNA (mitochondrial RNA), snRNA (small nuclear RNA) or piRNA (Piwi protein interacting small RNA), it was excluded from further evaluation. Further, the genomic sequence containing the candidate miRNA (~80 nt) was used to analysis the secondary structure of pre-miRNA by the program Mfold 3.2 [30]. The secondary structure should fulfill the criteria as described above.

Northern blot

Total RNA was isolated from seven tissues (heart, liver, spleen, lung, kidney, brain, and skeletal muscle) of 6-month-old sheep using TRIzol (Invitrogen, USA). 30 μg of total RNAs were fractionated on a denaturing 15% polyacryamide–8 mol/l urea gel, and then were electrophoretically transferred to nylon membrane by using a Mini Trans-Blot Electrophoretic Transfer Cell (Bio-Rad, USA). After transfer, the RNAs were fixed to the membrane by baking in a vacuum oven at 80°C for 2 h. The 5′ end of the DNA probes were labeled with DIG using a DIG Oligonucleotide 3′-End Labeling Kit (Roche, Switzerland). Pre-hybridization, hybridization and washing were performed as manual described. Signal from the membrane was detected using a DIG Luminescent Detection Kit (Roche, Switzerland) and the ChemiDoc XRS system (Bio-Rad, USA). The tRNA and 5S rRNA bands were visualized by ethidium bromide staining of polyacrylamide gels and used as loading controls.

Real-time PCR

Total RNA was extracted from six tissues of three 6-month-old Texel sheep, which included heart, liver, spleen, kidney, brain and skeletal muscle. The quantification method similar to that descried by Chen et al. [33] was applied to validate the predicted miRNAs. The miRNA cDNA was prepared by the specific stem-loop primer. Real-time PCR was performed using Power SYBR-Green PCR master mix kit (Applied Biosystems) on the BIO-RAD iQ5 Real-time PCR Detection System. The 15 μl PCR included 1 μl RT product, 2× SYBR-Green PCR master mix, 2 μM forward primer and 2 μM reverse primer. The reactions were incubated at 95°C for 5 min followed by 40 cycles of 95°C for 25 s, 60°C for 30 s and 72°C for 30 s. The standard curve was made by a consecutive dilution of 10^{-2}–10^{-8} copies. All reactions were run in triplicate and included no template controls. The 5S ribosomal RNA was used as an internal control. The level of miRNAs was determined as the relative ratio (RR), which was calculated by dividing the level of miRNAs by the level of the 5S ribosomal RNA in the same samples. The significant level was set to 0.05.

Prediction of *O. aries* miRNAs targets

252 conserved 3′UTR of *O. aries* genes were extracted by using UTR Sequences from 5-way alignments downloaded from the targetscan database (http://www.targetscan.org/) [16, 34] searched against 3779 *O. aries* mRNA sequences downloaded from NCBI database. The *RNAhybrid* program [35] was applied to predict the potential target sites of

Table 1 Information of 14 identified *O. aries* miRNAs by computational method, including the name, sequence, EST ID, size, number of mismatches, number of predicted target genes, and δ_len of the mature miRNAs, and the percent of A/U, MFE of the pre-miRNAs

miRNA	Sequence	EST ID	Size (nt)	NM[a] (nt)	SP[b] (nt)	δ_len	A + U (%)	MFE (kcal/mol)	NT[c]
oar-miR-374b	AUAUAAUACAACCUGCUAAGUG	EE839959.1	22	0	72	0	59.72	−34.8	0
*oar-miR-374b**	CUUAUCAGGUUGUAUUAUCAUU	EE839959.1	22	1	72	0	59.72	−34.8	0
oar-miR-421	AUCAACAGACAUUAAUUGGGCGC	EE839959.1	23	0	85	2	61.18	−35.2	4
*oar-miR-7-1**	CAACAAAUCACAGUCUGCCAUA	EE795723.1	22	0	109	1	57.80	−45.2	2
oar-miR-7	UGGAAGACUAGUGAUUUUGUUGU	EE795723.1	23	0	109	0	57.80	−45.2	8
*oar-miR-144**	GGAUAUCAUCAUAUACUGUAAG	EE764430.1	22	0	84	0	53.57	−34.5	0
oar-miR-144	UACAGUAUAGAUGAUGUACU	EE764430.1	20	0	84	0	53.57	−34.5	0
oar-miR-425-5p	AUGACACGAUCACUCCCGUUGA	EE826429.1	22	0	88	0	43.18	−34.7	2
oar-miR-425-3p	UCGGGGAAUGUCGUGUCCGCCC	EE826429.1	22	3	88	0	43.18	−34.7	10
oar-miR-147b	GUGUGCGGAAAUGCUUCUGCUA	DY476730.1	22	0	80	9	61.25	−22.6	2
*oar-miR-24-1**	GUGCCUACUGAGCUGAUAUCAGU	EE839419.1	23	0	68	0	50.00	−26.3	7
oar-miR-24	UGGCUCAGUUCAGCAGGAACAGG	EE839419.1	23	0	68	0	50.00	−26.3	16
oar-miR-1826	AUUGAUCAUCGACACUUCGAACGCA	EE827923.1	25	2	86	10	39.53	−36.7	0
oar-miR-1244	AAGUAGUUGGUUUGUAUGAGAUGGUU	CF116113.1	26	0	86	0	69.41	−16.5	11

O. aries miRNAs identified by computational method

[a] The mismatch number of cloned miRNA sequences compared to those of known miRNAs in related species

[b] Size of precursor-microRNAs

[c] Number of predicted target genes

identified *O. aries* miRNAs. A series of rules were used: (1) a perfect Watson–Crick match between miRNA and target at 2–7 positions in the 5′end of miRNA; (2) one G:U pair in the seed was allowable; (3) the size of max internal loop and bulge loop was 2; (4) the threshold for the minimum free energy (MFE) of the hybridization was −20 kcal/mol. The Gene Ontology analysis was conducted on AmiGO website.

Results

Computational and experimental identification of new *O. aries* miRNAs

First, we used a computational approach to exploit *O. aries* miRNAs in sheep EST database. Following a set of strict filtering criteria, we finally identified 14 *O. aries* miRNAs (Table 1). Among which, 13 identified miRNAs sequences were identical or highly similar (1 or 2 nt mismatches) with those of known miRNAs in related species, which agreed with that most miRNAs are conserved among mammalian species [36, 37]. Only *oar-miR-425-3p* had three mismatched nucleotides with corresponding known miRNA, but we also considered the oar-miR-425-3p as a miRNA candidate, because it had a homolog in *Bos Taurus*, which had the highest conserved genomic information with *O. aries*. In the end, it was validated by Real-Time PCR. The MFE of all precursor sequences was lower than −20 kcal/mol, except *oar-miR-1244*. However, the MFE of its homologs in other species also has a higher value, for example, −16.7 kcal/mol in *Homo sapiens* and −15.2 kcal/mol in *Pan troglodytes*. In addition, we found four pairs of miRNA/miRNA*, oar-miR-374b/oar-miR-374b*, oar-miR-7/oar-miR-7-1*, oar-miR-144/oar-miR-144*, oar-miR-24/oar-miR-24-1*, and a pair of oar-miR-425-3p/oar-miR-425-5p. A miRNA cluster was also identified, including oar-miR-374b, oar-miR-374b* and oar-miR-421.

By constructing four cDNA libraries for small RNAs in the size range of 18–26 nt from the tissues of skeletal

Table 2 Information of 17 identified *O. aries* miRNAs by experimental method, including the name, sequence, number of clones, size, number of mismatches, number of predicted target genes of the mature miRNAs, and the tissues

miRNA	Sequence[a]	No. of clones	Size (nt)	NM[b] (nt)	NT[c] (nt)	Tissue
oar-let-7b	UGAGGUAGUAGGUUGUGUGGU	3	21	1	29	Brain
oar-let-7c	UGAGGUAGUAGGUUGUAUGGUU	2	22	0	25	Brain
	UGAUGUAGUAGGUUGUAUGGU	1	21	2	–[d]	Skeletal muscle
oar-miR-1	UGGAAUGUAAAGAAGUAUGUAU	2	22	0	4	Skeletal muscle
	GGAAUGUAAAGAAGUAUGUAA	1	21	2	–	Skeletal muscle
	UGGAAUGUAAAGAAGUAUGUAA	2	22	1	–	Skeletal muscle
oar-miR-21	UAGCUUAUCAGACUGAUGUUG	1	21	1	1	Skeletal muscle
	UAGCUUAUCAGACUGAUGUUGAC	1	23	1	–	Skeletal muscle
oar-miR-26a	UUCAAGUGAUCCAGGAUAGGCU	1	22	1	9	Skeletal muscle
oar-miR-29a	UAGCACCAUCUGAAAUCGGUU	1	21	1	3	Skeletal muscle
oar-miR-29c	UAGCACCAUUUGAAAUCGGUUA	1	22	0	3	Skeletal muscle
oar-miR-122	UGGAGUGUGACAAUGGUGUUUG	6	22	0	15	Liver
oar-miR-124	UAAGGCACGCGGUGAAUGCCA	4	21	0	6	Brain
oar-miR-125b	UCCCUGAGACCCUAACUUGUG	1	21	1	10	Brain
oar-miR-133	UUGGUCCCCUUCAACCAGCUGU	1	22	0	14	Skeletal muscle
oar-miR-181a	AACAUUCAACGCUGUCGGUGAG	1	22	0	3	Brain, Skeletal muscle
oar-miR-206	UGGAAUGUAAGGAAGUGUGUGG	2	22	1	16	Brain, skeletal muscle
oar-miR-378	ACUGGACUUGGAGUCAGAAGGC	1	22	0	14	Skeletal muscle
oar-miR-451	AAACCGUUACCAUUACCGAGUUU	1	23	1	1	Skeletal muscle
oar-NEW-1	CCACUCGGCCACCUCGUC	1	18	Unknown	2	Spleen
oar-NEW-2	UUGUCCCUUCCUUCCCACCAGA	1	22	Unknown	9	Skeletal muscle

O. aries miRNAs identified by experimental method

[a] Some miRNAs are represented by more than one sequence due to variation at the 3′end

[b] The mismatch number of cloned miRNA sequences compared to those of known miRNAs in related species

[c] Number of predicted target genes

[d] No target prediction for miRNAs with variation

muscle, brain, liver and spleen, we cloned 35 miRNA sequences, represent 17 *O. aries* miRNAs. Among these miRNAs, 12 could be isolated from skeletal muscle (Table 2), which nearly included total miRNAs that have been proved to regulate the muscle development, such as miR-1 and miR-133.

Altogether, we identified 31 oar-miRNAs belong to 24 families, all of which were first identified in sheep. Two of them were novel *O. aries* miRNAs that had never been annotated in the miRBase, named oar-NEW-1 and oar-NEW-2. The remaining 29 miRNAs were sheep orthologs of known mammalian miRNAs. We confirmed that the flanking sequences of 29 new miRNAs could form the typical hairpin precursor structure by the program mfold (Supplementary Data 1). Fourteen mature miRNAs were located at the 5′ arm of the precursor, and the other fifteen were located at the 3′ arm. Two miRNAs (oar-miR-26a and oar-miR-451) identified from the small RNA library could not match the bovine genome, likely because of the incomplete genome identity between sheep and cattle. The new miRNAs identified in our study will be subsequently named in accordance with recognized standards (the official name will be assigned via the Rfam Sanger website as soon as this paper is accepted).

Expression profiling of new *O. aries* miRNAs

To further understand the biological function of newly identified miRNAs in sheep, we first detected the expression profile of three miRNAs (oar-miR-1, oar-miR-122 and oar-NEW-1) identified from the small RNA cDNA library in seven sheep tissues (heart, liver, spleen, lung, kidney, brain and skeletal muscle) by northern blot. The results showed that all of these three miRNAs appeared to be extremely tissue specific (Fig. 1).

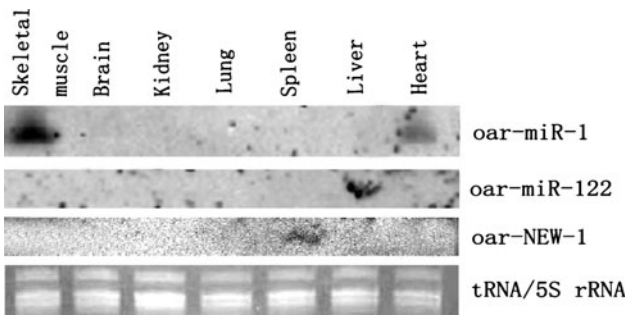

Fig. 1 Expression analysis result of three newly identified *O. aries* miRNAs (oar-miR-1, oar-miR-122, and oar-NEW-1) by northern blot. Total RNA from seven sheep tissues (heart, liver, spleen, lung, kidney, brain, and skeletal muscle) was blotted and probed with a 5′ DIG labeled DNA probes. 30 μg total RNA was loaded on each lane and tRNA/5S rRNA served as loading control

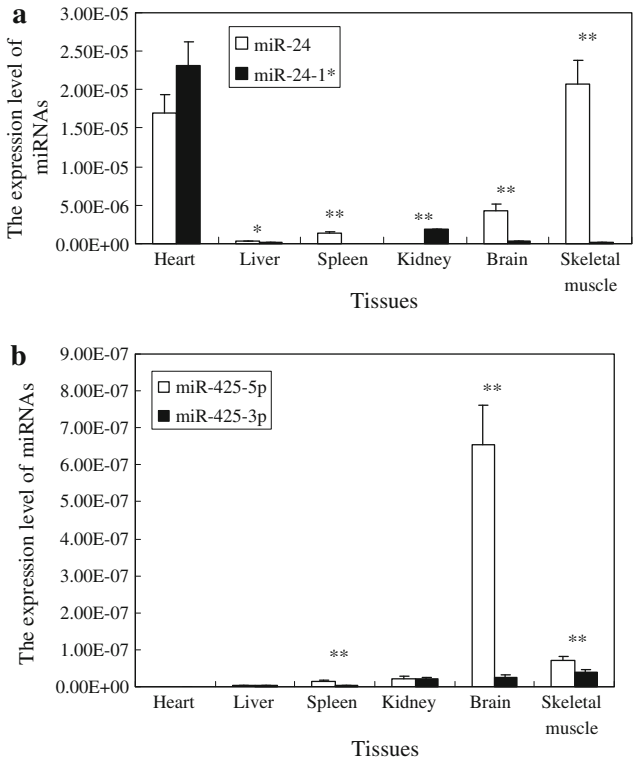

Fig. 2 Comparison of expression level between each pair of *O. aries* miRNAs in six sheep tissues, including heart, liver, spleen, kidney, brain and skeletal muscle. Comparison of expression level between **a** oar-miR-24 and oar-24-1* in six tissues, **b** oar-miR-425-5p and oar-425-3p in six tissues. The expression of four *O. aries* miRNAs was detected by real-time PCR with 5S ribosomal RNA as loading control. The *asterisks* indicate a significant difference between oar-miR-24 and oar-miR-24-1* or oar-miR-425-5p and oar-miR-425-3p (* $P < 0.05$; ** $P < 0.01$)

oar-miR-1 could be only detected in heart and skeletal muscle, which was reported to be specifically expressed in adult cardiac and skeletal muscle tissues and play the important role in regulating muscle development. oar-miR-122 could be only detected in liver, which was reported as the most enriched and the specific miRNA in liver [38]. And oar-NEW-1 could be only detected in spleen, which indicated that oar-NEW-1 might play the regulatory role in spleen. However, further study to discover its regulatory function is needed.

Then, we chose four miRNAs (the pair of oar-miR-24/oar-miR-24-1*, and the pair of oar-miR-425-3p/oar-miR-425-5p) identified by computational method for experimental validation and the expression analysis between miRNA and miRNA* by Real-time PCR using three independent samples from each sheep (the primer sequences are available in Supplementary Data 2). All of four miRNAs could be detected in total tissues, except that oar-miR-425-5p could not be detected in heart. In four

tissues, the expression level of oar-miR-24 was significantly higher than that of oar-miR-24-1*, and the oar-miR-425-5p had a stronger expression than oar-miR-425-3p in three tissues (Fig. 2). The results also showed that two pairs of miRNAs had a similar relative expression pattern in some tissues, respectively. Both oar-miR-24 and oar-miR-24-1* were expressed strongly in heart, and the expression level of oar-miR-425-3p and oar-miR-425-5p were high in brain (Fig. 3).

Prediction of *O. aries* miRNA targets

According to the criteria, described in "Materials and methods", 120 target genes, including 230 target sites were predicted. Of these 31 identified miRNAs, 5 miRNAs (*oar-miR-374b, oar-miR-374b**, *oar-miR-144**, *oar-miR-144 and oar-miR-1826*) had no target genes, whereas *oar*-let-7b had a maximum of 29 targets (Supplementary Data 3). By gene ontology analysis, 85 target genes were assigned to the GO terms of "molecular function" ontology, "binding" and "catalytic activity" of which over-represented these genes; 78 were assigned to the "biological process" ontology, most of which were found to take part in the "cellular process" and "biological regulation process" (Fig. 4). This may expose the potential effect of newly identified *O. aries* miRNAs in the biological process.

Discussion

New miRNAs identified from sheep

Sequence alignment revealed that three newly cloned miRNAs had end variants, and most of them differed at the 3′ end nucleotide(s) (Table 2), which was accordant with previous studies [39, 40]. oar-miR-1 was represented by three variants, whereas the others had two variants. These 3′-end variants may because of imprecise processing of miRNA precursors by Dicer or preferential degradation at the 3′ end [41, 42]. However, we could not exclude the probability that these miRNA variants were artifacts of the cloning procedure.

In addition, we identified one miRNA cluster by computational method: *oar-miR-374b*, *oar-miR-374b** and *oar-miR-421*. The existence of miRNA clusters has been already reported in many species [43–45]. Weber suggested that miRNAs in a cluster should in the same direction, and not be separated by a transcription unit or a miRNA in the opposite direction. The length of the identified cluster in our study was 242 bp, similar to that observed in other species, such as in *H. sapiens*, *P. troglodytes*, *Mus musculus* and *Canis familiaris*.

During the evolution history, some variations entered into the pre-miRNA sequences, so we carried out phylogenetic analysis of pre-miRNA sequences to further

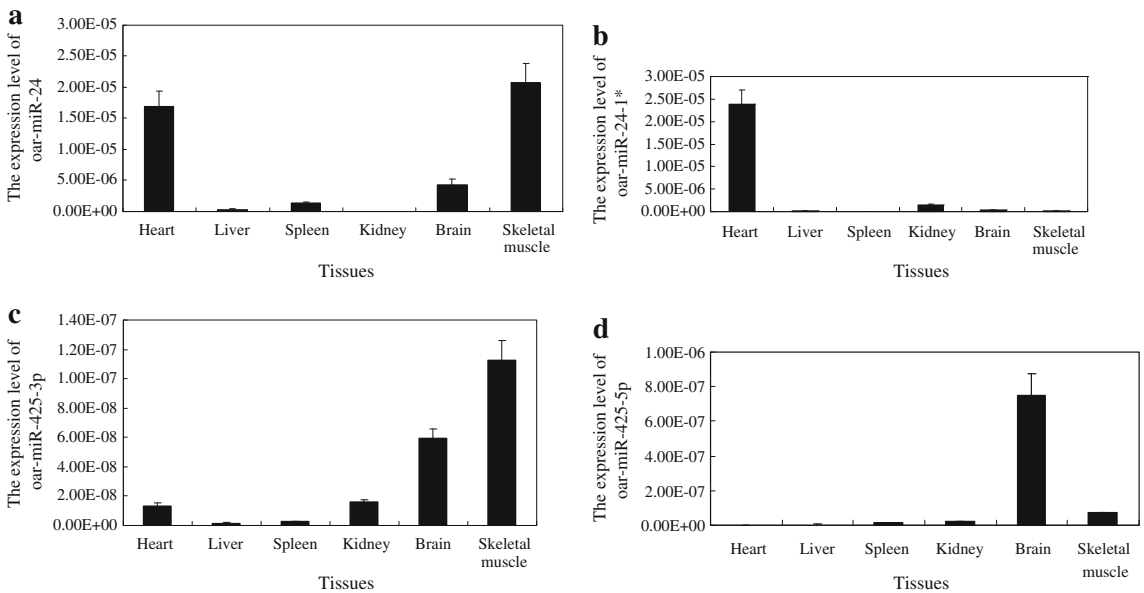

Fig. 3 The expression level of four *O. aries* miRNA in six sheep tissues, including heart, liver, spleen, kidney, brain and skeletal muscle. The expression level of **a** oar-miR-24 between different tissues, **b** oar-miR-24-1* between different tissues, **c** oar-miR-425-3p between different tissues, **d** oar-miR-425-5p between different tissues. The expression of four *O. aries* miRNAs was detected by real-time PCR with 5S ribosome RNA as loading control

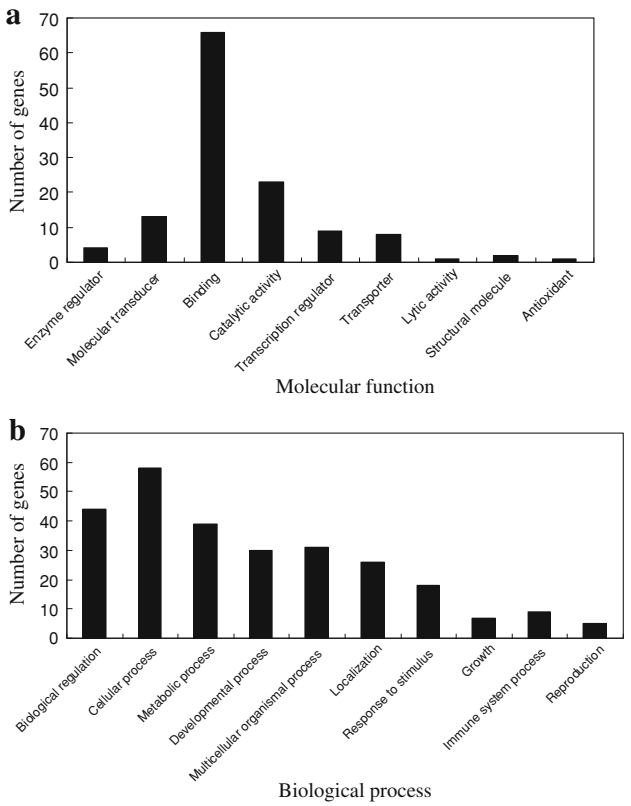

Fig. 4 The gene ontology analysis for predicted target genes was conducted on AmiGO website. Biological classification of predicted target genes by **a** molecular function using gene ontology analysis, **b** biological process using gene ontology analysis

validate the results. From the 24 identified miRNA families, 2 families were selected to construct neighbour-joining trees using MEGA 4 [46] (Fig. 5). The phylogenetic trees clearly showed that *O. aries* miRNAs were located on the same branches with other mammalian miRNAs.

O. aries miRNA targets

miRNAs can regulate gene expression by binding to complementary sites on the target mRNAs, usually located in the 3′UTR. It was found that conserved motifs of 3′UTR were likely under regulation mediated by miRNAs [47]. Here we used 3′UTR sequences conserved in five species searched against *O. aries* mRNA sequences to obtain conserved 3′UTR of *O. aries* genes. Of these 31 identified miRNAs, *oar-miR-374b*, *oar-miR-374b**, *oar-miR-144**, *oar-miR-144 and oar-miR-1826* were failed to predict target sites because of limited information of 3′UTR of *O. aries* genes.

By analysis of these target genes, we found that six of which encoded the regulator in the process of muscle growth. *oar-miR-7-1** targeted *β-actin*, which is an essential regulatory muscle protein, and it also targeted calpastatin (CAST) with oar-miR-1 and oar-miR-206 together, which is involved in cytoskeletal remodeling, myofibrillar turnover, and muscle growth regulation [48, 49]. *oar-miR-425-3p* targeted myogenic factor 6 which is important for myogenic differentiation and the maintenance of terminally differentiated state [50]. Both oar-miR-133 and oar-NEW-2 could target myogenin (myogenic factor 4), which helped to regulate muscle terminal differentiation [51] and also was an important factor for animal growth [52, 53]. oar-miR-133 targeted insulin-like growth factor 2 (IGF2) as well, a regulatory mutation in which caused a major QTL effect on muscle growth in pig [54]. *oar-miR-24* could target insulin-like growth factor-binding protein, which specifically modulates the metabolic and mitogenic actions of insulin-like growth factor (IGF)-I. IGF-I has been reported to stimulate muscle cell differentiation in transgenic mice [55].

miRNAs and muscle development

Meat production is a vital characteristic of sheep, thus the exploitation of miRNAs involved in regulating the process of muscle development in sheep is very important. To date, a small number of miRNAs are reported to be enriched in muscle tissues and play an important role in the regulation of myoblast proliferation and differentiation [56–58]. For example, miR-133 increases cell proliferation through down-regulation of different target genes [56, 59], miR-181 participates in establishing the muscle phenotype [57], and miR-24 is required for the modulation of TGF-b-inhibited myoblast differentiation [58]. Especially, a G to A transition in the 3′UTR of myostatin gene creates a target site for miR-1 and miR-206, and contributes to the muscular hypertrophy of Texel sheep [28]. To further study the regulatory function of miRNAs in muscle development, we constructed a miRNA cDNA library from sheep skeletal muscle and obtained 12 miRNAs, which almost included all of the muscle development-related miRNAs (miR-1, miR-133, miR-181a and miR-206). Furthermore, based on the result of target genes analysis mentioned above, several miRNAs can target the regulators of the muscle growth process. Thus, all of these miRNAs could be studied as the candidate miRNAs which can regulate the process of muscle development.

About miRNA*

The miRNA-star (miRNA*) is derived from the opposite arm of the hairpin than the more abundantly expressed miRNA, such as miR-56 in *Caenorhabditis elegans* [32], and several vertebrate miRNAs [14, 38]. So far, more than 1,000 miRNA*s have been registered in the database, but the biological function of these is not known yet. Perhaps

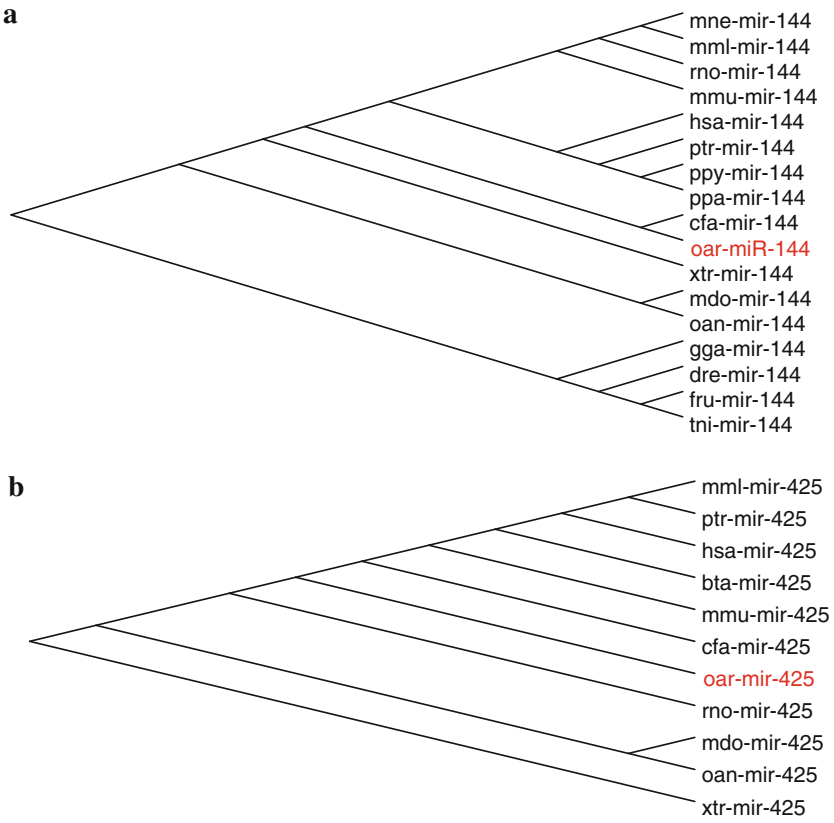

Fig. 5 Two identified miRNA families were selected to construct neighbour-joining trees using MEGA 4. Polygenetic tree of **a** miR-144 family, **b** miR-425 family

the miRNAs are more stable than their miRNA*, or the miRNAs might be stabilized by interactions with their targets, the level of miRNA* present in the cell is lower than the corresponding miRNA [60]. In some cases there is another kind of miRNA present in the cell, named miRNA-3p/miRNA-5p, which is derived from the different arms of one precursor and has equal expression levels. In our study, we compared the expression level of one pair of miRNA/miRNA* and the pair of miRNA-3p/miRNA-5p, and found that, they had a similar relative expression pattern in some tissues. Thus, our results suggest that miRNA* is widely expressed and may play an important role in the genome.

Analysis of identification results

From the small RNA cDNA library, we did not obtain the miRNAs identified by computational method likely because of their low expression level, or the expression time–space specificity. As for the computational approach, some similar attempts were made earlier [58, 61–63]. These studies largely were to identify the plant miRNAs and the number of allowed mismatches between the known miRNAs and the ESTs were set to three or four. In our study, we set the number of mismatches to two because the conservation of miRNAs in animals was higher than that in plant. Furthermore, earlier reports [61–63] just used the mRNA sequences of the species that were studied to predict the target genes. Here, we used the conservative 3′UTR sequences of sheep extracted by using UTR Sequences from 5-way alignments downloaded from targetscan database searched against *O. aries* mRNA sequences to make our results more convictive. The number of miRNAs obtained by EST analysis was related with the number of ESTs. It's about 10,000 ESTs contain one miRNA. In our study, we identified 14 miRNAs from 209808 ESTs, nearly 1:14000, maybe because of our stricter criteria. However, more miRNAs will be found by this method with the increase of genomic information.

Acknowledgements This work was supported by national High Technology Research and Development Program (863 Program) (No. 2006AA10Z199) and New Special Sheep Varieties with Top Quality Selection Program (No. 2006BAD01A11). We would like to thank Yanchun Yan, Hunter Linda, Zhengkui Zhou, Kejia Wu, Tinghua Huang, Shengsong Xie for language help and technical assistance.

References

1. Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116:281–297
2. Pillai RS, Bhattacharyya SN, Filipowicz W (2007) Repression of protein synthesis by miRNAs: how many mechanisms? Trends Cell Biol 17:118–126

3. Place RF, Li LC, Pookot D, Noonan EJ, Dahiya R (2008) MicroRNA-373 induces expression of genes with complementary promoter sequences. Proc Natl Acad Sci USA 105:1608–1613
4. Vasudevan S, Tong Y, Steitz JA (2007) Switching from repression to activation: microRNAs can up-regulate translation. Science 318:1931–1934
5. Johnston RJ, Hobert O (2003) A microRNA controlling left/right neuronal asymmetry in *Caenorhabditis elegans*. Nature 426: 845–849
6. Chang S, Johnston RJ Jr, Frokjaer-Jensen C, Lockery S, Hobert O (2004) MicroRNAs act sequentially and asymmetrically to control chemosensory laterality in the nematode. Nature 430: 785–789
7. Hatfield SD, Shcherbata HR, Fischer KA, Nakahara K, Carthew RW, Ruohola-Baker H (2005) Stem cell division is regulated by the microRNA pathway. Nature 435:974–978
8. Poy MN, Eliasson L, Krutzfeldt J, Kuwajima S, Ma X, Macdonald PE, Pfeffer S, Tuschl T, Rajewsky N, Rorsman P, Stoffel M (2004) A pancreatic islet-specific microRNA regulates insulin secretion. Nature 432:226–230
9. Xu P, Vernooy SY, Guo M, Hay BA (2003) The Drosophila microRNA Mir-14 suppresses cell death and is required for normal fat metabolism. Curr Biol 13:790–795
10. Lee YS, Dutta A (2007) The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. Genes Dev 21:1025–1030
11. Tavazoie SF, Alarcon C, Oskarsson T, Padua D, Wang Q, Bos PD, Gerald WL, Massague J (2008) Endogenous human microRNAs that suppress breast cancer metastasis. Nature 451: 147–152
12. Calin GA, Ferracin M, Cimmino A, Di Leva G, Shimizu M, Wojcik SE, Iorio MV, Visone R, Sever NI, Fabbri M, Iuliano R, Palumbo T, Pichiorri F, Roldo C, Garzon R, Sevignani C, Rassenti L, Alder H, Volinia S, Liu CG, Kipps TJ, Negrini M, Croce CM (2005) A microRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. N Engl J Med 353:1793–1801
13. Liu T, Tang H, Lang Y, Liu M, Li X (2009) MicroRNA-27a functions as an oncogene in gastric adenocarcinoma by targeting prohibition. Cancer Lett 273:233–242
14. Lagos-Quintana M, Rauhut R, Meyer J, Borkhardt A, Tuschl T (2003) New microRNAs from mouse and human. RNA 9:175–179
15. Fu H, Tie Y, Xu C, Zhang Z, Zhu J, Shi Y, Jiang H, Sun Z, Zheng X (2005) Identification of human fetal liver miRNAs by a novel method. FEBS Lett 579:3849–3854
16. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB (2003) Prediction of mammalian microRNA targets. Cell 115:787–798
17. Chang YM, Juan HF, Lee TY, Chang YY, Yeh YM, Li WH, Shih AC (2008) Prediction of human miRNAs using tissue-selective motifs in 3′ UTRs. Proc Natl Acad Sci USA 105:17061–17066
18. Lim LP, Glasner ME, Yekta S, Burge CB, Bartel DP (2003) Vertebrate microRNA genes. Science 299:1540
19. Huang TH, Zhu MJ, Li XY, Zhao SH (2008) Discovery of porcine microRNAs and profiling from skeletal muscle tissues during development. PLoS ONE 3:e3225
20. Alexander Stark PK, Parts L, Brennecke J, Hodges E, Hannon GJ, Kellis M (2008) Systematic discovery and characterization of fly microRNAs using 12 Drosophila genomes. Genome Res 17:1865–1879
21. Singh J, Nagaraju J (2008) In silico prediction and characterization of microRNAs from red flour beetle (*Tribolium castaneum*). Insect Mol Biol 17:427–436
22. Adai A, Johnson C, Mlotshwa S, Archer-Evans S, Manocha V, Vance V, Sundaresan V (2005) Computational prediction of miRNAs in *Arabidopsis thaliana*. Genome Res 15:78–91
23. Pan XP, Zhang BH, Wang QL, Cobb GP, Anderson TA (2005) Identification and characterization of new plant microRNAs using EST analysis. Cell Res 15:336–360
24. Weibo Jin NL, Zhang B, Wu F, Li W, Guo A, Deng Z (2008) Identification and verification of microRNA in wheat (*Triticum aestivum*). J Plant Res 121:351–355
25. Griffiths-Jones S (2004) The microRNA registry. Nucleic Acids Res 32:D109–D111
26. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ (2006) miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res 34:D140–D144
27. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008) miRBase: tools for microRNA genomics. Nucleic Acids Res 36:D154–D158
28. Clop A, Marcq F, Takeda H, Pirottin D, Tordoir X, Bibe B, Bouix J, Caiment F, Elsen JM, Eychenne F, Larzul C, Laville E, Meish F, Milenkovic D, Tobin J, Charlier C, Georges M (2006) A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. Nat Genet 38:813–818
29. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98
30. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31:3406–3415
31. Mathews DH, Sabina J, Zuker M, Turner DH (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J Mol Biol 288:911–940
32. Lau NC, Lim LP, Weinstein EG, Bartel DP (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. Science 294:858–862
33. Chen C, Ridzon DA, Broomer AJ, Zhou Z, Lee DH, Nguyen JT, Barbisin M, Xu NL, Mahuvakar VR, Andersen MR, Lao KQ, Livak KJ, Guegler KJ (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. Nucleic Acids Res 33:e179
34. Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120:15–20
35. Rehmsmeier M, Steffen P, Hochsmann M, Giegerich R (2004) Fast and effective prediction of microRNA/target duplexes. RNA 10:1507–1517
36. Ambros V, Bartel B, Bartel DP, Burge CB, Carrington JC, Chen X, Dreyfuss G, Eddy SR, Griffiths-Jones S, Marshall M, Matzke M, Ruvkun G, Tuschl T (2003) A uniform system for microRNA annotation. RNA 9:277–279
37. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T (2001) Identification of novel genes coding for small expressed RNAs. Science 294:853–858
38. Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W, Tuschl T (2002) Identification of tissue-specific microRNAs from mouse. Curr Biol 12:735–739
39. Jin W, Grant JR, Stothard P, Moore SS, Guan LL (2009) Characterization of bovine miRNAs by sequencing and bioinformatics analysis. BMC Mol Biol 10:90
40. Kim J, Cho IS, Hong JS, Choi YK, Kim H, Lee YS (2008) Identification and characterization of new microRNAs from pig. Mamm Genome 19:570–580
41. Ruby JG, Jan C, Player C, Axtell MJ, Lee W, Nusbaum C, Ge H, Bartel DP (2006) Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. Cell 127:1193–1207
42. Ruby JG, Stark A, Johnston WK, Kellis M, Bartel DP, Lai EC (2007) Evolution, biogenesis, expression, and target predictions of a substantially expanded set of Drosophila microRNAs. Genome Res 17:1850–1864

43. Weber MJ (2005) New human and mouse microRNA genes found by homology search. FEBS J 272:59–73
44. Altuvia Y, Landgraf P, Lithwick G, Elefant N, Pfeffer S, Aravin A, Brownstein MJ, Tuschl T, Margalit H (2005) Clustering and conservation patterns of human microRNAs. Nucleic Acids Res 33:2697–2706
45. He PA, Nie Z, Chen J, Chen J, Lv Z, Sheng Q, Zhou S, Gao X, Kong L, Wu X, Jin Y, Zhang Y (2008) Identification and characteristics of microRNAs from *Bombyx mori*. BMC Genomics 9:248
46. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599
47. Xie X, Lu J, Kulbokas EJ, Golub TR, Mootha V, Lindblad-Toh K, Lander ES, Kellis M (2005) Systematic discovery of regulatory motifs in human promoters and 3′ UTRs by comparison of several mammals. Nature 434:338–345
48. Goll DE, Thompson VF, Li H, Wei W, Cong J (2003) The calpain system. Physiol Rev 83:731–801
49. Wendt A, Thompson VF, Goll DE (2004) Interaction of calpastatin with calpain: a review. Biol Chem 385:465–472
50. Perry RL, Rudnick MA (2000) Molecular mechanisms regulating myogenic determination and differentiation. Front Biosci 5: D750–D767
51. Zhang H, Stavnezer E (2009) Ski regulates muscle terminal differentiation by transcriptional activation of Myog in a complex with Six1 and Eya3. J Biol Chem 284:2867–2879
52. Wang X, Zhang J, Li F, Gu J, He T, Zhang X, Li Y (2005) MicroRNA identification based on sequence and structure alignment. Bioinformatics 21:3610–3614
53. Anton I, Zsolnai A, Komlosi I, Kiraly A, Fesus L (2006) Effect of MYOG genotypes on growth rate and production traits in Hungarian large white pigs. Acta Vet Hung 54:393–397
54. Van Laere AS, Nguyen M, Braunschweig M, Nezer C, Collette C, Moreau L, Archibald AL, Haley CS, Buys N, Tally M, Andersson G, Georges M, Andersson L (2003) A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. Nature 425:832–836
55. Coleman ME, DeMayo F, Yin KC, Lee HM, Geske R, Montgomery C, Schwartz RJ (1995) Myogenic vector expression of insulin-like growth factor I stimulates muscle cell differentiation and myofiber hypertrophy in transgenic mice. J Biol Chem 270:12109–12116
56. Chen JF, Mandel EM, Thomson JM, Wu Q, Callis TE, Hammond SM, Conlon FL, Wang DZ (2006) The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. Nat Genet 38:228–233
57. Naguibneva I, Ameyar-Zazoua M, Polesskaya A, Ait-Si-Ali S, Groisman R, Souidi M, Cuvellier S, Harel-Bellan A (2006) The microRNA miR-181 targets the homeobox protein Hox-A11 during mammalian myoblast differentiation. Nat Cell Biol 8: 278–284
58. Sunkar R, Jagadeeswaran G (2008) In silico identification of conserved microRNAs in large number of diverse plant species. BMC Plant Biol 8:37
59. Kim HK, Lee YS, Sivaprasad U, Malhotra A, Dutta A (2006) Muscle-specific microRNA miR-206 promotes muscle differentiation. J Cell Biol 174:677–687
60. Lim LP, Lua NC, Weinstein EG et al (2003) The microRNAs of *Caenorhabditis elegans*. Genes Dev 17(8):991–1008
61. Qiu CX, Xie FL, Zhu YY, Guo K, Huang SQ, Nie L, Yang ZM (2007) Computational identification of microRNAs and their targets in *Gossypium hirsutum* expressed sequence tags. Gene 395:49–61
62. Yin Z, Li C, Han X, Shen F (2008) Identification of conserved microRNAs and their target genes in tomato (*Lycopersicon esculentum*). Gene 414:60–66
63. Zhang B, Wang Q, Wang K, Pan X, Liu F, Guo T, Cobb GP, Anderson TA (2007) Identification of cotton microRNAs and their targets. Gene 397:26–37