Available online at www.sciencedirect.com
ScienceDirect
Computational Biology and Chemistry 32 (2008) 61–66
Computational Biology and Chemistry
www.elsevier.com/locate/compbiolchem

Brief Communication

Genome-wide computational analyses of microRNAs and their targets from *Canis familiaris*

Donggen Zhou [1], Siguang Li [1], Jian Wen, Xi Gong, Ling Xu, Yuping Luo *

College of Life Sciences, Nanchang University, Nanchang 330031, PR China

Received 30 March 2007; accepted 10 August 2007

Abstract

By computational analyses, we identified 357 miRNA candidates from *Canis familiaris* genome, among which 300 are homology of characterized human miRNAs, the remains are not reported in any other animal. Of the 357 miRNA genes, 142 are organized into 53 clusters, and two clusters locate in the paternally imprinted region. These dog miRNAs may regulate more than 800 possible targets, which are involved in a wide range of cellular processes. Remarkably, miR-186 resides in the eighth intron of its target gene in the same orientation, suggesting a feedback regulation of miRNA on its host gene.

Keywords: miRNA; *Canis familiaris*; Computational analyses; Paralogus cluster

1. Introduction

MicroRNAs are small non-coding RNAs with 21–25 nucleotides in length and believed to be involved in a wide range of developmental and cellular processes by base-pairing with complementary sites of target transcript either to repress translation or to trigger mRNA degradation, depending on the extent of complementarity between miRNA and mRNA (Ambros et al., 2003). Since the first identification of *Caenorhabditis elegans* lin-4, hundreds of microRNAs have been characterized experimentally (Kim et al., 2004; Sunkar et al., 2005). More recently, computational algorithms have been developed to identify new miRNAs precursors from sequenced genome (Lindow and Krogh, 2005).

It is popular thought that many miRNAs are evolutionarily conserved among a variety of organism, therefore, it is possible to predict orthologues of known microRNAs in different species by computational RNomics (Shahi et al., 2006; Bartel, 2004). So far, a large number of miRNAs (about 4039 entries) have been characterized in miRBase database (Release 8.2) from different animals (Watanabe et al., 2006).

Domestic dog is a premier species for the study of morphology, behavior, and disease (Ostrander and Wayne, 2005). In addition, dogs and humans have similar levels of overall nucleotide diversity, and the *Canis familiaris* system was proposed to hold the power to map and clone disease genes that had proven intractable through studies of human families (Parker et al., 2004). The use of *C. familiaris* for comparative studies should provide a broad perspective in miRNA molecular biology. However, so far only six miRNAs were characterized in dog genome. In the present study, we report the identification of *C. familiaris* miRNAs by genome-wide computational analyses. The structure and targets of these miRNAs and their genomic organization were further characterized.

2. Material and Method

2.1. Prediction of Dog microRNA Genes

The sequences of human mature and precursor miRNAs were obtained from the miRNA Registry (Release 8.2) and the entire set of sequences was subjected to a BLAT search in the dog genomes on the UCSC Genome Browser server (http://genome.ucsc.edu/cgi-bin/hgBlat). Sequences cor-

Abbreviations: UTR, un-translated region; LINE, long interspersed nuclear element; SINE, short interspersed nuclear element; nt, nucleotide.

* Corresponding author. Fax: +86 791 3969103.
E-mail address: luoyuping@163.com (Y. Luo).

[1] These authors contributed equally to this work.

1476-9271/$ – see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compbiolchem.2007.08.007

responding to potential novel miRNA precursors were obtained from the UCSC Genome Browser, and their hairpin structures were assessed with the mfold program (http://www.bioinfo.rpi.edu/applications/mfold/old/rna/).

At the same time, we downloaded Human/dog genome pairwise alignments file from UCSC Genome Browser, then import these alignment file with conservation more than 90% in clustalw format into RNAz program to predict functional RNA in linux-like environment (Washietl et al., 2005). To avoid designating other RNAs as miRNA candidates, these potential miRNA precursors should fit the criteria as previously described (Ambros et al., 2003; Lagos-Quintana et al., 2003; Lai et al., 2003).

2.2. Analysis of Genomic Organization of Dog miRNA

The genomic locations of all dog miRNA genes identified in this study were derived from the UCSC Genome Browser (http://genome.ucsc.edu). We evaluated and extracted the general clustering of miRNA genes according to the method described by Altuvia et al. (2005), and analyzed the transcription start sites and promoter sequences using the following three analysis softwares, McPromoter MM:II (http://genes.mit.edu/McPromoter.html), Promoter Prediction 2.0 (http://www.cbs.dtu.dk/services/Promoter/) and NNPP (http://www.fruitfly.org/seq_tools/promoter.html).

2.3. Prediction of miRNA Target

We downloaded 3′ UTR sequences of all protein-encoding genes from dog genome in UTResource (http://www.ba.itb.cnr.it/UTR/UTRHome.html). The miRanda algorithm (version 1.0b) was used to scan all available miRNA sequences for dog genome against 3′ UTR sequences (Enright et al., 2003). The parameter is following: score threshold is 120, energy threshold is −17.0 kcal/mol, scaling factor is 2.0. After finding optimal local matches above these thresholds between a particular miRNA and the set of 3′ UTRs in each genome, we analyzed whether target site position and sequence for this miRNA were conserved in the 3′ UTRs of orthologous genes among dog, human and mouse. To analyse function of target genes and functional enrichment, we search GO term for all target genes (http://www.geneontology.org/).

2.4. RNA Isolation and RT-PCR Assay

Total cellular RNA was isolated and purified from liver tissue of *C. familiaris* according to the method described by Chomczynski et al. (Chomczynski and Sacchi, 1987). Reverse transcription was carried out in 20 μL reaction mixture containing 15 μg of total cellular RNA treated with DNase, reverse primer and 500 mmol L^{-1} dNTPs. After being denatured at 65 °C for 5 min and then cooled to 42 °C, 200 units of M-MLV reverse transcriptase (Promega) were added and the extension was carried out at 42 °C for 45 min. After reverse transcription, PCR was carried out with the reverse primer and corresponding forward primer.

3. Results

3.1. Identification of 357 Novel miRNA from Dog Genome

By computational algorithms, we identified 357 microRNA candidates from dog genome, among which 300 are homologs of known human miRNAs, the remains are not reported in any other animal (Supplementary Table A). The length of miRNA precursors varied from 65 to 120 nucleotides with an average of 80 nt. Even so, the 357 potential microRNA precursors can be folded to the typical secondary structure of the miRNA family. The secondary structures of some novel miRNAs whose orthologues are not reported before are shown in Supplementary Figure A. The 300 dog miRNAs were named after their counterparts in human, and the remaining 57 novel miRNAs that were not reported before were named in succession with a serial number from miR-800 to miR-856.

The 300 dog miRNAs which are homologs of known human miRNAs belong to 149 miRNA families. miR-154, miR-17 and let-7 each has 15, 13 and 10 members, respectively, while, for most of other miRNA families, only one member was identified. For members of each family, the mature miRNAs are either identical (miR-15, miR-218), closely related (miR-199a and b, miR-26a and b), or display significant homology (miR-148b and miR-152, miR-107, miR-103-1 and miR-103-2).

3.2. Genomic Organization of Dog miRNAs

Of the 357 microRNA candidates, 125 located in intronic regions of protein-coding genes, 75 in the sense orientation and 50 in the antisense orientation of their host genes (Supplementary Table A), the others reside in UTR or exon of mRNA or intergenic spacers. Among all these dog miRNA, 142 are organized into 53 clusters. Sequences alignment showed that some miRNA clusters are formed by paralogous miRNAs, and some miRNA clusters have multi-copies. For example, 28 dog miRNA clusters contain homologous miRNAs (data not shown) and 18 clusters form 8 distinct paralogous cluster groups (Fig. 1).

Interestingly, in dog chromosome 8 there is a large region, whose homologous region in human and mouse are annotated as paternally imprinted regions (Cavaille et al., 2001). In this region there are two miRNA gene clusters, one snoRNA gene cluster and some protein-encoding genes (Fig. 2). The miR-127 cluster resides in a retrotransposon-like gene, Rtl1, in antisense orientation. It is not clear whether the miRNAs in miR-127 cluster have functional role as antisense regulators for the expression of the Rtl1 (Lin et al., 2003).

Remarkably, 13 dog miRNAs appear derived from transposable elements and other genome repeats. Seven of which derived from LINE/L2 repeats, and the others derived from SINE/Alu and SINE/MIR (MIRb). Most of repeat-associated microRNA hairpin foldbacks are formed by the junction of two adjacent LINE-2 segments apposed in opposite orientation and this cases are conservative in mammalian (Fig. 3A). However, dog miR-325 made an exception that its precursor was folded by joining one LINE-2 element to one SINE element apposed in opposite way (Fig. 3B). Interestingly, miR-378 and miR-422b lie in a