Gene 512 (2013) 422–428

Contents lists available at SciVerse ScienceDirect

Gene

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

The novel MER transposon-derived miRNAs in human genome

Kung Ahn [a], Jeong-An Gim [a], Hong-Seok Ha [b], Kyudong Han [c], Heui-Soo Kim [a,*]

[a] Department of Biological Sciences, College of Natural Sciences, Pusan National University, Busan 609-735, Republic of Korea
[b] Department of Genetics, Human Genetics Institute of New Jersey, Rutgers, the State University of New Jersey, 145 Bevier Rd, Piscataway, NJ 08854, USA
[c] Department of Nanobiomedical Science & WCU Research Center, Dankook University, Cheonan 330-714, Republic of Korea

ARTICLE INFO

Article history:
Accepted 17 August 2012
Available online 25 August 2012

Keywords:
MER transposon
MicroRNA
MER derived miRNA
Palindromic structure
AGO proteins

ABSTRACT

MicroRNAs (miRNAs) are small RNA molecules (~20–30 nucleotides) that generally act in gene silencing and translational repression through the RNA interference pathway. They generally originate from intergenic genomic regions, but some are found in genomic regions that have been characterized such as introns, exons, and transposable elements (TE). To identify the miRNAs that are derived from palindromic MERs, we analyzed MER paralogs in human genome. The structures of the palindromic MERs were similar to the hairpin structure of miRNA in humans. Three miRNAs derived from MER96 located on chromosome 3, and MER91C paralogs located on chromosome 8 and chromosome 17 were identified in HeLa, HCT116, and HEK293 cell lines. The interactions between these MER-derived miRNAs and AGO1, AGO2, and AGO3 proteins were validated by immunoprecipitation assays. The data suggest that miRNAs derived from transposable elements could widely affect various target genes in the human genome.

1. Introduction

MicroRNAs (miRNAs) are single-stranded RNAs that are ~20–30 nucleotides long in their mature forms. They are generally responsible for gene silencing by binding to target messenger RNA (mRNA) transcripts. miRNAs are produced in several steps: they are transcribed as long primary transcripts, which are referred to as primary miRNA (pri-miRNA) transcripts. Next, they are processed by 2 well-known proteins, namely, Drosha and Dicer. Drosha cleaves the pri-miRNA down to ~70 nucleotides in length with a short double-stranded hairpin structure (precursor miRNA) in the nucleus. The precursor miRNA is exported to the cytoplasm and cleaved by Dicer to generate the mature miRNA form (~20–30 nucleotides). The mature miRNA comprises an active RNA-induced silencing complex (RISC) along with argonaute protein (AGO), which binds to the mature miRNA and mediates the interaction between the miRNA and its target mRNA (Denli et al., 2004; Lee et al., 2004; Okamura et al., 2004, 2009; Yang and Lai, 2010). miRNAs have been highlighted as key regulators in numerous types of cells and organismal processes (Krichevsky et al., 2003; Piriyapongsa and Jordan, 2008; Xu et al., 2010). In addition, it was reported that alterations in miRNA expression are associated with the initiation, progression, and metastasis of human tumors (Shalgi et al., 2010). Although most miRNAs originate from intergenic genomic sequences, some of them originate from sequences that have been characterized such as genes and transposable elements (Piriyapongsa and Jordan, 2007, 2008; Smalheiser and Torvik, 2005; Zhang et al., 2007). Furthermore, retroviral transposable elements (TEs) such as human immunodeficiency virus and bovine leukemia virus encode viral miRNAs (Althaus et al., 2012; Kincaid et al., 2012; Klase et al., 2007). In particular, because of the wide distribution of the TEs, the miRNAs derived from them could harmfully affect host cells and their genomic stability. If a miRNA that is derived from TEs with a high copy number is involved in gene silencing, it would have a huge impact on an individual's survival. However, it is still unclear how miRNA is produced from a TE. In the case of Made 1, it has been reported that its palindromic structure (the origin of hsa-mir-548) is produced by either orientation (+/−) and that it can form itself into a hairpin structure. Then, it is inserted into transcriptionally active genomic regions, and it takes the form of pri-miRNA structures that can be processed by the RNA interference enzymatic machinery (Piriyapongsa et al., 2007). MEdium Reiteration frequency (MER) refers to interspersed repeats in the genomes of primates, rodentia, and lagomorpha. Age estimations place the origin of most MER repeats at the time of decline in Mammalian-wide Interspersed Repeats (MIR) retroposition and before the origin of the Alu family.

Intriguingly, short palindromic sequences contain a hairpin structure, which is very similar to the structure of miRNA. The miR-1302 gene family was derived from the MER53 transposon that retains its palindromic structure (Yuan et al., 2010). In this study, we computationally identified miRNAs that originated from MER elements and experimentally confirmed them. To identify the miRNA, we compared palindromic MERs with miRNA precursors and mature miRNAs from miRBase (http://www.mirbase.org/). Then, we carried out miRNA

Abbreviations: TEs, transposable elements; MER, MEdium Reiteration frequency; AGO proteins, argonaute proteins; miRNA, microRNA.

* Corresponding author. Tel.: +82 51 510 2259; fax: +82 51 581 2962.
E-mail address: khs307@pusan.ac.kr (H.-S. Kim).

0378-1119/$ – see front matter.
http://dx.doi.org/10.1016/j.gene.2012.08.028

prediction to identify potential miRNA derived from the paralogs of palindromic MER consensus sequences. Finally, we experimentally verified the existence of MERs that originated from miRNAs.

2. Materials and methods

2.1. Computational analysis

MER consensus sequences containing palindromic sequences were obtained from the Repbase browser of the Genetic Information Research Institute database. The genomic locations and sequences of repetitive elements, including MERs, were extracted from the human genome by using UCSC Genome Browser hg18 (http://genome.cse.ucsc.edu/) and they were analyzed using RepeatMasker 3.27 (http://www.repeatmasker.org/). The sequences and coordinates of human pre-miRNAs and mature miRNAs were downloaded from miRBase v13.0 (http://www.mirbase.org/) and mapped to the human genome (hg18). Using the UCSC Table Browser, the genomic locations of the miRNAs and the repetitive elements from the human genome were compared. To identify the potential MER-derived miRNA genes, we developed a three-step operational scheme. We blasted the human genome with the palindromic consensus sequences of MER families using the BLAT program. Then, we carried out predictions of the MiRNA that originated from MER on MER paralogs by using miPred [http://www.bioinf.seu.edu.cn/miRNA/] and miRNAFinder [http://bioinfo3.noble.org/mirna/]. miPred is a search tool for which all pre-miRNAs have a characteristic stem-loop hairpin structure. Therefore, the hairpin structures provide key clues to the ab initio prediction of pre-miRNAs. In addition, miRNAFinder can predict potential intronic miRNA in intron regions of the expressed genes (ESTs/cDNAs), find possible miRNA in genomic sequences, or predict if the input small RNA is mature miRNA.

A mature miRNA is processed from the left or right arm of a potential precursor miRNA sequence. We examined the secondary structures of single stranded RNA sequences by using the RNAfold web server, which has limits of 7500 nt for partition function calculations and 10,000 nt for minimum free energy. The TargetScan 6.2 [http://www.targetscan.org/] was used to identify the potential target sites of miRNAs. To determine the functional categories to which the target genes belonged, we used GOmir [http://www.bioacademy.gr/bioinformatics/projects/GOmir/] and displayed the GO categories of the target genes.

2.2. Cell cultures and lysates preparation

HEK293, HCT116, and HeLa cells were seeded in Dulbecco's modified Eagle's medium that was supplemented with 10% (v/v) heat-inactivated fetal bovine serum and 1% (v/v) antibiotics–antimycotic solution, and then incubated at 37 °C in a 5% (v/v) CO_2 incubator. Cells (5×10^6) from each cell line were washed with DPBS and collected by trypsinization. The cells were incubated in 1 ml of lysis buffer (20 mM Tris, pH 7.4, 2.5 mM $MgCl_2$, 200 mM NaCl, 0.05% NP40) for 10 min on ice, and the cell lysates were cleared by centrifugation at $15{,}000\times g$ for 20 min at 4 °C.

2.3. Immunoprecipitation, poly (A) tailing, and reverse-transcription

To extract small RNAs from the human cells, AGO proteins AGO1, AGO2, and AGO3 were immunoprecipitated. AGO1, 2, 3-Antibody Beads were purchased from Wako Pure Chemical Industry, Ltd. Fifty microliters of Anti-AGO1, AGO2, and AGO3 Antibody Bead solutions and human cell lysates were incubated at 4 °C for 2 h. After the antigen–antibody reactions, the incubated mixtures were washed twice, and to each was added 50 μl of the elution solution. The eluted microRNAs were purified using phenol:chloroform:isoamyl alcohol mixture (25:24:1) and visualized on a polyacrylamide gel (data not shown). The isolated miRNAs were subjected to poly (A) tailing; 4 μg of each was polyadenylated by 5 U of poly (A) polymerase (Ambion). The reactions were incubated at 37 °C for 30 min. After incubation, poly (A) tailed miRNAs were recovered by phenol/chloroform extraction and ethanol precipitation.

Reverse transcription of the poly (A) tailed miRNAs was performed using M-MLV reverse transcriptase (RT, Invitrogen) following the manufacturer's instructions.

A solution (9.5 μl) including 4 μg of the poly (A) tailed miRNAs and 1 μl of 10 mM RT linker sequence (CTGTGAATGCTGCGACTACGA-18 dTs) was incubated at 65 °C for 5 min to remove any RNA secondary structures. After incubation, 5 μl of 10 mM dNTP mix, 4 μl of 5× RT buffer (including 250 mM Tris–HCl, 375 mM KCl, 15 mM $MgCl_2$, and 50 mM DTT), 0.5 μl of RNase inhibitor and 1 μl of M-MLV reverse transcriptase were added. The reaction was then incubated at 42 °C for 90 min. After incubation, the reaction was allowed to proceed at 4 °C for 60 min.

2.4. RT-PCR

MER91C located on chromosome 8 was amplified using the sense primers 5′-TGAAGGGGTTACAATTGGCAT-3′ (Chr.8, MER91C derived mature miRNA 5P) and 5′-TGCGGATGGCACCTCCTGAG-3′ (Chr.8, MER91C derived mature miRNA 3P). MER91C located on chromosome 17 was amplified using the sense primers 5′-TACAACTGGAAGGATGTTCAT-3′ (Chr.17, MER91C derived mature miRNA 5P) and 5′-GTGACATCCCTTGAGTTGTGC-3′ (Chr.17, MER91C derived mature miRNA 3P). MER96 located on chromosome 3 was amplified using the sense primers 5′-TGCAGCATTTAAGGAAGCACC-3′ (Chr.3, MER96 derived mature miRNA 5P) and 5′-GAGTGCCTCCTTAAATGTTTT-3′ (Chr.3, MER96 derived mature miRNA 3P). All the amplifications were carried out using the antisense primer (Universal primer) 5′-CTGTGAATGCTGCGACTACGAT-3′. RT-PCRs were carried out for 30 cycles at 94 °C for 40 s, 56 °C for 40 s, and 72 °C for 7 min.

3. Results

3.1. Comparative analysis of palindromic MERs and miRNAs from miRBase

Eight palindromic MER consensus sequences were found (Fig. 1). Among them, the consensus sequence of MER91C matched the precursor of hsa-miRNA-652 in the miRBase and the E-value of the match was 4.00E−08. Through a UCSC blat search, 80 paralogs aligned with the consensus sequence of MER91C were identified in the human genome (hg18). In addition, the consensus sequence of MER96 was matched with a hsa-miR-3680 and the E-value was 5.00E−18. Ninety-seven paralogs of the MER96 consensus sequence were found in the human genome (Table 1). The palindromic consensus MER sequences were matched with the precursor miRNA sequence of the miRBase. As shown in Fig. 2, (a) MER53 was matched with precursor miR-1302-1 and (b) MER91C was matched with precursor miR-652, while (c) MER96 was matched with precursor miR-3680/3680*.

We aligned the palindromic consensus sequences of the MER elements with the precursor miRNA sequences in the miRBase (Fig. 2). The consensus sequence of MER53 had 83.6% similarity with the precursor of has-mir-1302-1. In addition, the consensus sequence of MER91C had 75.8% similarity with the precursor of hsa-mir-652, and the consensus sequence of MER96 had 85.1% similarity with the precursor of has-mir-3680. The target genes of hsa-mir-1302 and hsa-mir-652 seemed to be intricately involved in cellular processes, intracellular membrane, catalytic activity, and binding by gene ontology (GO) cluster analysis (Fig. S1). However, the consensus sequences of other MERs, including MER123, MER124, MER126, MER133B, and MER134 did not exactly match any of the precursor miRNAs in the miRBase. Although there were no matches with the precursor miRNAs, we could not exclude the possibility that