Biochemical and Biophysical Research Communications 420 (2012) 391–396

Contents lists available at SciVerse ScienceDirect

Biochemical and Biophysical Research Communications

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

# Identification of methylation-dependent regulatory elements for intergenic miRNAs in human H4 cells

Kwang Hee Lee [a,c], Hyunyoung Kim [b], Byeong Jae Lee [c,d,*], Kiejung Park [a,*]

[a] Division of Bioinformatics and Bio-Medical Informatics, Center for Genome Science, National Institute of Health, Republic of Korea
[b] Division of Brain Disease, Center for Bio-Medical Science, National Institute of Health, Republic of Korea
[c] Laboratory of Molecular Genetics and Genomics, Interdisciplinary Program in Bioinformatics, Institute of Molecular Biology and Genetics, Seoul National University, Seoul 151-742, Republic of Korea
[d] School of Biological Sciences, Institute of Molecular Biology and Genetics, Seoul National University, Seoul 151-742, Republic of Korea

ARTICLE INFO

*Article history:*
Received 27 February 2012
Available online 8 March 2012

*Keywords:*
Intergenic microRNA
DNA methylation
Regulatory element

ABSTRACT

MicroRNAs (miRNAs) are important post-transcriptional regulators of various biological processes. Although our knowledge of miRNA expression and regulation has been increased considerably in recent years, the regulatory elements for miRNA gene expression (especially for intergenic miRNAs) are not fully understood. In this study, we identified differentially methylated regions (DMRs) within 1000 bp upstream from the start site of intergenic miRNAs in human neuroglioma cells using microarrays. Then we identified a unique sequence pattern, $C[N]_6CT$, within the DMRs using motif analysis. Interestingly, treatment of cells with a methyl transferase inhibitor (5-aza-2-deoxycytidine, DAC) significantly increased expression of miRNA genes with a high frequency of the $C[N]_6CT$ motif in DMRs. Statistical analysis showed that the frequency of the $C[N]_6CT$ motif in DMRs is highly correlated with intergenic miRNA gene expression, suggesting that $C[N]_6CT$ motifs associated with DNA methylation regions play a role as regulatory elements for intergenic miRNA gene expression.



## 1. Introduction

MicroRNAs (miRNAs) are small, non-coding RNA molecules that act as post-transcriptional regulators of gene expression by inhibiting translation or degrading mRNA genes through partial or complete base pairing with complementary sequences of target genes [1]. In addition, some miRNAs participate in the remodeling of chromatin structures [2]. miRNAs are initially transcribed as large precursor RNAs, or primary miRNAs (pri-miRNA), and sequentially processed by Drosha and Dicer to produce ~22-nucleotide-long active mature miRNAs [3–5]. miRNAs are highly conserved in multiple organisms and play crucial roles in development, cell differentiation, determination of cell fate, and cancer [6,7].

miRNA genes can be classified into two categories according to their genomic contexts: intronic and intergenic miRNAs. Intronic miRNAs are embedded within other genes. Therefore, they are thought to be transcribed by sharing promoters with host genes [8]. On the other hand, intergenic miRNAs are believed to have independent transcription units because they are positioned within flanking regions or in antisense orientation to annotated genes [9]. Intronic miRNAs are generally believed to be transcribed by RNA polymerase II (pol II); however, it remains unclear what type of RNA polymerase is responsible for intergenic miRNA transcription, although pol II and RNA polymerase III (pol III) are obvious candidates. For example, pri-miR-23a~27a~24-2 and pri-miR-21 are transcribed by pol II and have a 5′-7-methylguanosine cap structure and a 3′-polyadenylated [poly(A)] tail similar to the structure of mRNAs [10,11], while miR-517a and miR-517c, which are interspersed among Alu repeats in the human chromosome 19, are transcribed by pol III [12].

The transcriptional start site of intergenic miRNA genes usually occurs within 2 kb upstream from the start site of miRNAs [13]. Using computational methods, several conserved sequence patterns for intergenic miRNA genes, including putative promoters, have been proposed from various species [14,15]. Among these, CT repeats are most well known. They are highly conserved among four species, such as *Caenorhabditis elegans*, *Homo sapiens*, *Arabidopsis thaliana* and *Oryza sativa*, and are abundant within 1000 bp upstream sequences from miRNA hairpins [14]. Another sequence pattern, GANNNNGA, was identified within 1000 bp upstream of

*Abbreviations:* DMPs, differentially methylated probes; DMRs, differentially methylated regions; DAC, 5-aza-2-deoxycitidine; pol II, RNA polymerase II; pol III, RNA polymerase III; pri-miRNA, primary miRNA.

* Corresponding authors. Address: Division of Bioinformatics and Bio-Medical Informatics, Center for Genome Science, National Institute of Health, Building 2, Room 521, Osong 363-951, Republic of Korea. Fax: +82 43 719 8869 (K. Park).
*E-mail address:* kjpark63@gmail.com (K. Park).

0006-291X/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bbrc.2012.03.005

worm miRNAs [15]. However, there is no direct evidence that these conserved patterns play a role as promoter or regulatory elements. Currently, the transcriptional mechanisms of most intergenic miRNAs are largely unknown.

Epigenetic signatures such as DNA methylation and histone modification, and the regulation of expression of miRNA genes are tightly linked similarly to other genes [16–19]. Recently it was reported that hypermethylation of the human miR-124 loci, which is the most abundant miRNA in the adult brain and plays a key role in neurogenesis, inhibits miR124a expression and results in brain tumors [20–22]. Interestingly, some miRNAs control the expression of epigenetic regulators, including DNA methyltransferases and histone deacethylases [23,24]. The fact that miRNA gene expression can be regulated by DNA methylation indicates the feasibility of using methylated sequences to predict miRNA gene promoters or regulatory elements.

In this study, we found a novel sequence motif, $C[N]_6CT$, for intergenic miRNA gene expression by predicting sequence patterns in the differentially methylated regions (DMRs), and by examining the relationship between the occurrence of this motif and methylation dependence of gene expression.

## 2. Materials and methods

### 2.1. Cell lines and culture

H4 cells, a human neuroglioma cell line, were purchased from the American Type Culture Collection and cultured in Dulbecco's modified Eagle's medium (DMEM; Invitrogen, Carlsbad, CA, USA) supplemented with 10% fetal bovine serum (FBS; Invitrogen, Carlsbad, CA, USA) and 1% antibiotics-antimycotics (Invitrogen, Carlsbad, CA, USA) at 37 °C in a humidified incubator containing 5% $CO_2$.

### 2.2. Identification of miRNAs from sequence and annotation data

The genomic coordinates of 1049 human miRNAs were obtained from the miRBase (ver. 16.0) [25], and all sequences and the annotated data were from the UCSC genome browser (http://genome.ucsc.edu). A total of 1049 miRNAs were classified into 621 intronic and 428 intergenic miRNAs according their genome contexts.

### 2.3. Probe design

Sequences up to 1000 bp upstream from the start site of 428 intergenic miRNAs were retrieved and cleaved into 60-bp-long sequences overlapped by 40 bp of adjacent sequence (Fig. 1A). Chopped sequences were filtered based on sequence redundancy, low GC ratios (GC ratio < 0.6), and low melting temperatures (Tm < 85 °C). A total of 7646 sequences were selected as probes for printing on an Agilent 15K array platform to build a customized array chip (Chip No. 253347810001).

### 2.4. Microarray experiment

Genomic DNA was isolated from H4 cells cultured in the presence or absence of 5 μM DAC, an inhibitor of DNA methyltransferase. Briefly, after sonicating the genomic DNA (0.5 μg), the fragments were incubated with 2 μg recombinant methylation-specific binding protein (MBD2bt) at 4 °C for 4 h on a rocking platform. The enriched methylated DNA was amplified using a Whole Genome Amplification Kit (GenomePlex®, Sigma–Aldrich, St. Louis, MO, USA) as recommended by the manufacturer's instructions. The amplified DNA from DAC-treated and untreated cells were labeled with cyanine 5 (Cy5) and cyanine 3 (Cy3), respectively. The labeled DNA samples were purified using a PCR Purification Kit (QIAquick, Qiagen, Valencia, CA, USA) and co-hybridized to the customized microarrays according to the manufacturer's protocol. The microarrays contained a total of 7646 oligonucleotide probes, including control probes and those covering the sequences upstream of the miRNA genes.

### 2.5. Microarray data analysis

The hybridized images were analyzed using an Agilent DNA Microarray Scanner and data quantification was performed using Feature Extraction software version 10.7.3.1 (Agilent Technologies, Palo Alto, CA, USA). Preprocessing of raw data and normalization steps were performed using R software (http://www.r-project.org). Background-corrected intensity data were normalized using the intensity-dependent LOWESS method to remove the dye bias within each array. The *p*-values for each probe were calculated using linear fit models implemented in the Limma package (http://bioconductor.org/), and the probes within the threshold ($p$-value < 0.05) were selected as differentially methylated probes.

### 2.6. Reverse transcription PCR reaction

Total RNA was isolated using the RecoverAll, Total Nucleic Acid Isolation Kit (Ambion, Austin, TX, USA), according to the manufacturer's protocol. RNA quantity and purity was determined using the NanoDrop 1000 spectrophotometer (Thermo Scientific, Rockford, IL, USA). Reverse transcription was performed using the TaqMan MicroRNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA, USA). All reactions were performed as per the manufacturer's protocol. *Rnu6b* was used as a negative control.

### 2.7. Quantitative real-time PCR

Quantitative real-time PCR was performed to amplify miRNAs with specific primer sets against target miRNAs (Applied Biosystems, Foster City, CA, USA) using the ABI-7500 Real Time PCR system according to the manufacturer's protocol. Template (10 ng) was amplified in 20 μl reaction volumes. PCR conditions were as follows: 50 °C for 2 min, 95 °C for 10 min, 50 cycles of 95 °C for 15 s, and 60 °C for 1 min. Experiments were performed in triplicate.

### 2.8. Motif analysis

MEME software [26] was used to search for top-ranking degenerate motifs within the probe sequences in each cluster, and its optional parameters were set as follows: optimum motif width was set to 8–12 bp, occurrence of motif in the input sequences was set to any number of repetitions in the input sequence, and other parameters were left as default. Alignment of DMR sequences was performed using ClustalX (ver. 2.0.12) software [27] with the default parameters.

## 3. Results

### 3.1. Identification of DMRs

To identify DMRs in the upstream of miRNA genes, microarray analysis was performed using our custom chips. A total of 7646 probes against the 5′-flanking region of 428 intergenic miRNA genes were designed (Fig. 1A) and implemented on the Agilent 15K array chip platform. Genomic DNA was isolated from H4 cells cultured in the absence or presence of 5-aza-2-deoxycytidine (DAC). Methylated sequences were enriched using MBD2bt