Tissue Engineering and Regenerative Medicine, Vol. 9, No. 3, pp 157-169 (2012)
DOI 10.1007/s13770-012-0157-7

| Original Article |

A Bioinformatics Approach for *In Vivo* Imaging of Endogenous MicroRNA Targets During Neurogenesis

Mi-hee Jo[1], Chang Hyun Lee[1], Bahy A Ali[2], Saud A Alarifi[2], Abdulaziz A Al-Khedhairy[2], and Soonhag Kim[1]*

[1]*Laboratory of Molecular Imaging, Department of Biomedical Science, College of Life Science, CHA University, 605-21, Yoeksam 1-dong, Gangnam-gu, Seoul, Korea*

[2]*Department of Zoology, College of Science, King Saud University, Riyadh, Saudi Arabia*

(Received: March 8th, 2012; Revision: April 2nd, 2012; Accepted: April 3rd, 2012)

Abstract : MicroRNAs (miRNAs), a class of small non-coding RNAs, have been reported to be functionally involved with cellular metabolism and a variety of diseases. The importance of miRNA expression and functional targeting has recently become a focus of intense research. However, their endogenous molecular targets have not been clearly identified despite multiple attempts in prior studies using bioinformatics. Our bioinformatics strategy and *in vitro* validation by the PCR, identified 16 out of 337 miR124a-predicted targets and 5 out of 299 miR9*-predicted targets were significantly and directly down-regulated by each of the miRNAs during neurogenesis. *In vitro* and *in vivo* bioluminescent imaging system was used and successfully monitored the miR9*-mediated repression of SOX2 during neuronal differentiation of the P19 cells. The results of this study demonstrate that our bioinformatics approach offers a powerful and precise method for the identification of novel miR124a and miR9* endogenous targets during neuronal differentiation. This bioinformatics approach, using microarray data available from public DBs, provides a practical means for identifying the endogenous targets of other miRNAs.

Key words: *microRNA, bioinformatics, neuronal differentiation, imaging, SOX2*

1. Introduction

MicroRNAs (miRNAs, miRs), a class of small non-coding RNAs, are 22-30 nucleotide (nt) single-strand RNA molecules. miRNAs have been shown to be important negative feedback regulators of many of the biological processes involved in development, differentiation, and proliferation in both animals and plants.[1] MiRNAs completely, or nearly completely, bind to complementary sequences at the 3' untranslated regions (UTR) of target mRNAs and this is followed by mRNA degradation or translational repression.[2,3]

Hundreds of miRNAs have been cloned from a variety of species, and tens of endogenous targets of miRNAs have been identified using molecular experimental approaches. Moreover, bioinformatics analysis has been undertaken, in many studies, to identify novel miRNAs and their cognate targets based on sequence homology and complementarities with mature miRNAs in the genomes of individual species and across species.[4-6] To understand the molecular mechanisms underlying the regulation of cellular metabolism and the pathophysiological effects of various diseases using miRNAs, it is critical to identify miRNA: mRNA duplex formations. Several miRNA target prediction algorithms, such as, PicTar[7] and miRNAMap[8] has been useful in this context, however this approach might not be the most practical for identifying the endogenous targets of miRNAs, because there are hundreds of predicted targets per miRNA and only partial miRNA pairing to such targets occurs.

The combination of sequence-based targeting information on miRNAs from the PicTar database (DB) and microarray expression profiles deposited at public DBs, i.e., Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo), Stanford Microarray DB (http://genome-www5.stanford.edu) and ArrayExpress (http://ebi.ac.uk/arrayexpress) provide powerful tools for endogenous miRNA target discovery.[5,9] In addition, high-throughput screening using DNA chips to identify miRNA targets has been attempted in combination with the transfection of exogenous miRNA precursors of miR1 and miR124a in heterogeneous cells not expressing these miRNAs.[2] However, such screening to identify endogenous miRNA targets should be performed in cells where the

*Tel: +82-2-555-5063; Fax: +82-2-3468-3373
e-mail: kimsoonhag@empal.com (Soonhag Kim)

miRNAs of interest are highly expressed. Furthermore, molecular experiments using DNA chips to detect endogenous target genes of the miRNAs, are challenging for small laboratories with regard to cost and the technology required. Although a variety of approaches for studies using bioinformatics approaches have been carried out to predict miRNA targets, the precise prediction and validation of endogenous miRNA targets remain a major challenge.[10,11]

Here, we propose a method using bioinformatics for the identification of endogenous targets of miR124a and miR9* (5'-end of miR9 precursor) that are highly expressed during neuronal and brain development.[12] This approach is based on the premise that endogenous targets of miR124a and miR9* are biologically down-regulated during the expression of mature miR124a and miR9*, respectively. First, we analyzed the bioinformatics data on miR124a- or miR9*-predicted targets obtained using the PicTar DB; we found 337 orthologues for miR124a and 299 for miR9* in the human and mouse data. Second, to determine the gene expression patterns of the 337 and 299 orthologues during neurogenesis, heterogeneous microarray information sources, related to the neuronal differentiation of cells were acquired from the GEO DB and formatted into our system for bioinformatics analysis in a parallel fashion. Third, using DNA chip information from 52 super normal tissues, from the GEO DB, the gene expression patterns, of miR124a target candidates with significantly low expression during neurogenesis, were evaluated in brain tissue samples. Fourth, to verify the paradigms used to conduct these integrated analyses, RT-PCR analysis was performed on novel candidates of miR124a or miR9* in P19 cells (embryonal carcinoma cell line) treated with exogenously derived miR124a or miR9* precursors. Fifth, both *in vitro* and *in vivo* imaging systems were used to monitor miR9*-mediated repression of SOX2 during neuronal differentiation of P19 cells.

2. Materials and Methods

2.1 Data Preparation

The predicted target genes of miR124a or miR9* as determined by the PicTar webpage were downloaded in HTML (HyperText Makeup Language) format and parsed by PHP to allow import into a local DB. In addition, 5 different gene expression data sets in txt format were downloaded from the GEO DB to a local MySQL DB (Table 1). GEO data sets have three components: a Platform (GPL) file containing information about the probe ID, a Reference sequence ID (Ref ID) and the Gene ontology. In addition, there is a Sample (GSM) file containing information about the probe ID and gene expression values or ratios, and a Series (GSE) file containing information about batches of sample files and sample descriptions. To perform comparisons of this heterogeneous microarray data, in a formatted table, the following steps were taken:

(1) For each data set in a GSE file, a sample matrix and platform files were inserted into the MySQL DB table by: a sample matrix file : insert into table1_name values (probe Id, values1, values2, …); a platform file : insert into table 2_name values (probe Id, Ref ID, annotation2, …);

(2) Then, the sample matrix table and platform table were merged by id:- select * into outfile 'dataset_1.txt' from table1_name, table2_name … where table1_name.id=table2_name.id;

(3) Repeat 5 annotated data sets into DB.

(4) Gene expression values from 5 different GSE files were converted into fold ratios.

(5) The *in silico* predicted targets of miR124a and miR9* from PicTar, with 5 sets of different microarray data, were compared using the procedure described in Fig 1A.

2.2 Scoring Algorithm

The total score (100 points) for the PicTar-predicted targets

Table 1. Microarray datasets from GEO DB.

GEO series name	Experimental description	Platform	Time series	Reference
GSE2075	HeLa cells with treatment of exogenous miR124 precursor	Rosetta (Merck) Human 25k v2.2.1	12, 24h	(2)
GSE1588	Mouse cranial neural crest cells with the treatment of Retinoic acid	Affymetrix GeneChip Murine Genome U74 Version 2 Set MG-U74A	6, 12, 24, 48 h	(12)
GSE4075	Mouse embryonic stem cells with the treatment of Retinoic acid	NIA Mouse 22K Microarray v2.0 (Development 60-mer Oligo)	1, 2, 3, 4 day	(11)
GSE4076	Mouse embryonic stem cells with the treatment of N2B27	NIA Mouse 22K Microarray v2.0 (Development 60-mer Oligo)	1, 2, 3, 4, 5, 6 day	(11)
GSE740	Gene expression in 52 human tissues	Affymetrix GeneChip Human Genome U133 Plus 2.0 Array	52 different tissues	(13)

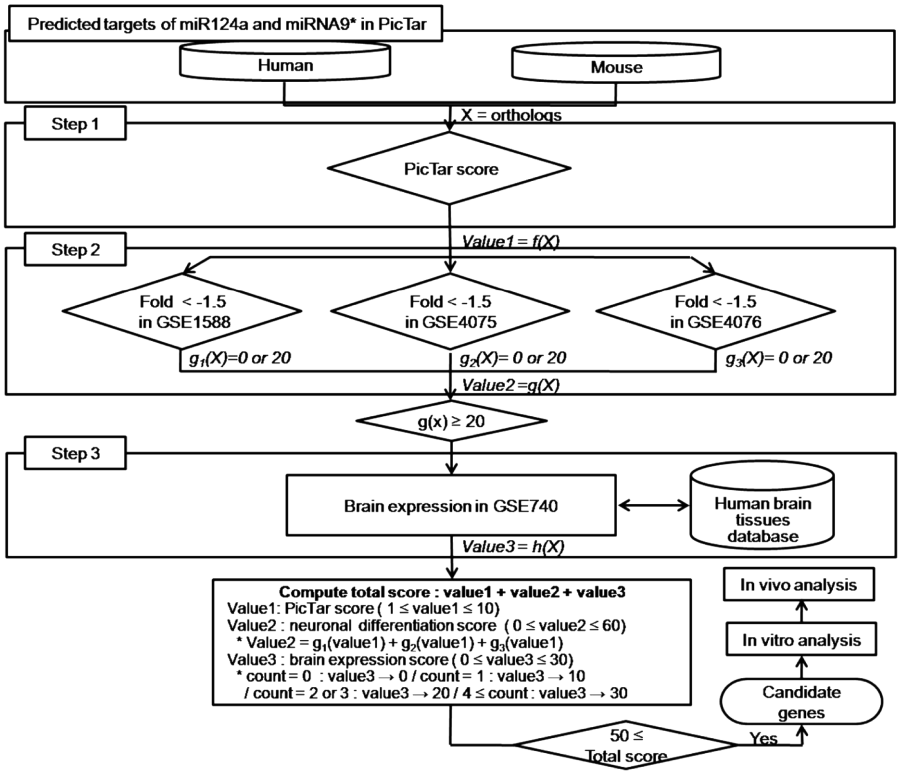

Figure 1. A simple flowchart for miR124a and miR9* endogenous target prediction and validation.

of miR124a or miR9* by the bioinformatics analysis included the PicTar score (10 points) and gene expression values for the heterogeneous microarray data associated with neuronal development (60 points) and normal tissues (30 points) was computed based on the following formula.

Total score of each orthologue = $f(X) + g(X) + h(X)$, where $X = {}^{13 \ldots, xn}$ is a set of orthologues (x) predicted as miR124a or miR9* targets both in the human and mouse targets of the PicTar DB. This approach consisted of the following functions: function f(x) represented the PicTar score for each orthologue in the PicTar DB, function g(x) was applied to GSE1588, GSE4075 and GSE4076 and scored the fold ratio of gene down-regulation for each orthologue during neuronal development, and function h(x) was applied to GSE740 and calculated the rank of gene expression values in brain tissues for only orthologues having g(x) ≥ 20 points compared with the scores of other normal tissues. The total score was 100 and each score quota was allocated as follows. PicTar scores were based on the PicTar score in the PicTar DB and if more than 10, it was scored as 10. The function g(x) had three different datasets ($g_1(x)$ for GSE1588, $g_2(x)$ for GSE4075, $g_3(x)$ for GSE4076) and computed the fold ratio of gene down-regulation for each orthologue for the time series in each dataset by OR operator based-logic (Supplemental Fig. 1B). If there was a fold value of down-regulation ≤ -1.5 for each orthologue, during the time series of neuronal differentiation for each dataset compared with pre-differentiation, then the result was 20 points. If there was a fold value of down-regulation ≥ -1.5 for each orthologue, the score was 0. In total, 60 points for g(x), 20 points for each dataset, were allocated due to the fact that endogenous targets of miR124a or miR9* are directly and specifically related to neurogenesis. The function h(x) was scored as follows: gene expression values of 52 different tissues (including 10 different normal brain tissues) for each orthologue were ranked from 1 (the lowest expression value) to 52 (the highest expression value), the number of brain tissues counted below the range of 5; one was 10, two or three were 20, and more than 4 had scores of 30 points.

2.3 Cell Culture and Transient Transfection Studies

The P19 cells were obtained from the American Type Culture Collection (ATCC). P19 cells have been well-known to be differentiated into neuronal cells by the treatment of all-trans-retinoic acid (RA).[14] Undifferentiated P19 cells were grown at 37°C in α-MEM (Gibco, Grand Island, NY) supplemented with 2.5% fetal bovine serum (Cellgro, Herndon, VA), 7.5% bovine calf serum (Gibco, Grand Island, NY) and 1% antibiotics-antimycotic (Cellgro, Herndon, VA) in a 5% CO_2-humidified

chamber at 37°C. To induce neuronal differentiation using the monoculture differentiation method [1], the P19 cells were plated on gelatin-coated culture plates at a density of 5×10^3 cells/cm^2 in growth medium and cultured for 24 h prior to the removal of growth factor. The cells were then cultured under serum-free conditions in DMEM/12 (1:1) (Gibco, Grand Island, NY) supplemented with 1% insulin, transferrin, and selenium (ITS, Gibco, Grand Island, NY) and 1% antibiotics and then treated with 5×10^{-7} M RA (Sigma,St.Louis, MO., USA). After culturing for 2 days, the RA was removed from the medium and the cells were further cultured under serum-free conditions. The P19 cells were transfected with 40 nM of pre.miR124a or precursor miR9* (Ambion, Austin, TX., USA) using Lipofectamine and Plus reagent (Invitrogen, Grand Island, NY) (4 μl of Plus reagent and 1 μl of Lipofectamine per well). All transfections were carried out in quadruplicate. HeLa cell (an adenocarcinoma cell line) was maintained using RPMI (Jeil biotechservices Inc, Daegu, Korea) media containing 10% fetal bovine serum and 1% antibiotics-antimycotic and transfected with 2.5 nM of precursor miR9* using Lipofectamine and the Plus reagent.

2.4 Identification of miR124a or miR9* Target Genes by RT-PCR

Using Trisol reagent (Invitrogen, Grand Island, NY., USA), total RNA was isolated from the P19 cells: 0d, 1d, 2d, 3d, 4d, 5d, 6d and 7d after RA treatment or 1d after treatment with 40 nM of pre.miR124a or precursor miR9*. First-strand cDNA was reverse transcribed using random-hexamer and SuperScript II reverse transcriptase (Invitrogen, Grand Island, NY., USA), according to the manufacturer's instructions, then used as a template for PCR amplification, which was carried out at 94°C for 5 min, with 30 amplification cycles (94°C for 30 sec, each Tm for 30 sec (Supplemental Table 1 and 2), 72°C for 30 sec), followed by extension at 72°C for 4 min. The sequences of the primers used are listed in Supplemental Table 1 and 2.

2.5 qRT-PCR (Quantitative RT-PCR) to Detect Mature miR124a and miR9* Expression

Small RNA was isolated during the neuronal differentiation of the P19 cells using the *mir*Vana™ miRNA isolation kit (Ambion, Austin, TX., USA). qRT-PCR was performed using the *mir*Vana™ qRT-PCR primer Set (Ambion, Austin, TX., USA) and *mir*Vana™ qRT-PCR miRNA detection kits (Ambion, Austin, TX., USA). The amplification was performed in triplicate using iCycer (Bio-Rad) with SYBR Premix Ex Taq™ (2x) (Takara, Japan) as follows: 95°C for 3 min followed by 40 amplification cycles (95°C for 15 sec, 60°C for 30 sec). To normalize the experimental samples for the RNA content, the U6 snRNA primer set (Ambion, Austin, TX., USA) was used as a control.

2.6 Recombinant Construction and Transfections of Reporter Genes for the Monitoring of miR9*-Mediated Repression of SOX2

To monitor SOX2 regulation, 66-base pair (bp) oligonucleotides near the seed region (SR, 7-mer) of the 3' UTR of SOX2 that is bound by miR9* was explored via the PicTar DB (Fig 4A). Sense and antisense primers of 66-bp of SOX2 containing SR or completely mutated SR were annealed and cloned right after the stop codon of the gaussia luciferase (Gluc) reporter gene regulated by a cytomegalovirus (CMV) promoter (designated as CMV/Gluc_SOX2 or CMV/Gluc_SOX2mt). Annealing was performed with a mixture of 20 pmol of oligonucleotides and 60 ul of annealing buffer containing 1x Tris-ethylenediaminetetraacetic acid buffer and 50 mM sodium chloride, under the following conditions: 95°C for 4 min and 70°C for 10s, followed by slow cooling in a 70°C water bath for 90 min. The annealed oligonucleotide was subcloned into the XhoI and XbaI sites of the CMV/Gluc vector.

The constructs were transfected into HeLa and P19 cells by liposome-mediated transfection using a Lipofectin reagent kit (Invitrogen, Grand Island, NY, USA) and CMV/Fluc activity was monitored during neuronal differentiation of the HeLa and P19 cells treated with retinoic acid. Transfection was performed using Lipofectamine Plus reagent (Invitrogen), 1 mg of DNA, 4 uL of PLUS Reagent (Invitrogen), and 1 uL of Lipofectamine (Invitrogen) per well. The Firefly luciferase (Fluc) reporter gene regulated by the CMV promoter (CMV/Fluc) was co-transfected and used as an internal control and normalized to Gluc activity in each experiment. All transfections were performed in quadruplicate.

2.7 Luciferase Assay

Luciferase assays were carried out on the P19 and HeLa cells by washing with NaCl/P_i buffer and treatment with lysis buffer (200 μl/well). Measurement of the luciferase activity was performed using a luminometer (TR717; Applied Biosystems, Foster City, CA, USA); an exposure time of 1 second after cell lyses was carried out in cells transferred from a six-well plate containing P19 cells that were seeded at 0.6×10^5 (undifferentiated, 1day, 2day) for 24 h, prior to the transfection, to a 96-well white microplate. Gluc was measured with coelenterazine (Biotium, Inc.) as a substrate and a 480 nm emission filter, and Fluc by D-luciferin and a 562 nm emission filter. CMV/Fluc was used as

an internal control to normalize the luciferase activity obtained from each experiment. All transfections were performed in quadruplicate. All data are calculated and presented as means±SD from triplicate wells.

2.8 *In Vivo* Visualization of Primary miR9* Expression in Undifferentiated and Differentiated P19 Cells

All experimental animals were housed under specific pathogen free conditions and handled in accordance with the guidelines issued by the Institutional Animal Care and Use Committee of Seoul National University Hospital.

In vivo imaging of P19 cells transfected with CMV/Gluc_SOX2 or CMV/Gluc_SOX2mt was divided into all-*trans* retinoic acid (RA)-treated in the right thigh and a non-RA treated group in the left thigh. 1×10^6 cells were harvested with 100 ul PBS, and re-suspended with RA for the neuronal differentiation group. The P19 cells were then subcutaneously injected into each the thigh of five-week-old male Balb/c nude mice and monitored for two days. Then, 100 ug of colenterizaine (Biotium, Inc.) in phosphate-buffered saline was injected into the tail-vein of the mice and Gluc images acquired 2 min later. 3 mg of D-luciferin was then administered intraperitoneally. Bioluminescence images of the whole body were acquired for 1 min for Gluc and 5 min for Fluc using an IVIS100 (Xenogen, CA, USA).

3. Results

3.1 Bioinformatics for the Prediction of miR124a Endogenous Targets

The strategy and the coding sources of the bioinformatics method used to identify novel endogenous targets of neuronal miR124a are summarized in Fig 1 and Supplemental Fig 1A and 1B. All procedures for data and text mining were carried out with the MySQL command in the console mode using the PHP web-script and a Windows operating system. For the first step in the identification of miR124a endogenous targets, 787 predicted human targets and 770 mouse targets from the PicTar WEB INTERFACE were downloaded. Based on the gene symbols, both the human and mouse target genes of miR124a were compared. Only 337 unique orthologues were identified due to the abundance of splice variants and genes with unknown function. A PicTar score for 337 genes, in the PicTar DB, was calculated based on a 1 to 10 point scale (Supplemental Table 3). For the second step, heterogeneous microarray data from the GEO ftp site were downloaded including the following (Table 1); genes regulated by miR124a in HeLa cells (GSE2075),[2] genes regulated by neuronal differentiation medium N2B27 (GSE4076)[15] or Retinoic acid (RA) (GSE4075)[15] in mouse embryonic stem cells (ESC), genes regulated by RA in mouse neural crest cells (NCC) (GSE1588),[16] and gene expression in 52 different human tissues (GSE740).[17] All of the different types of normalized expression ratios obtained from GSE2075, GSE4076, GSE4075, and GSE1588 were converted into fold ratios. Since the mRNAs of the miRNA-targeted genes containing miRNA SR, from the PicTar site, are predicted to be destabilized by mature miRNAs, the genes down-regulated during neurogenesis were evaluated as endogenous target candidates of miR124a. To compare the gene expression patterns of the 337 orthologues during neurogenesis, we used the gene expression data from GSE4076, GSE4075, and GSE1588. The genes with mRNA levels significantly down-regulated by at least 1.5 fold in at least one of multiple time-series of each microarray dataset were considered, 20 points were allocated for each dataset. There were 117 out of 337 genes were down-regulated, at least one of three heterogeneous microarray datasets, and 41 genes at two out of three different microarray datasets (Supplemental Table 3). Three genes including QKI, CCL2 and SLC6A15 had 60 points, illustrating significant down-regulation in all three different datasets.

To determine the gene expression of the 117 genes in normal brain tissues where miR124a is highly expressed, GSE740 microarray data was analyzed to identify novel endogenous targets of miR124a with confidence. The gene expression intensities of each candidate in 52 normal human tissues containing 10 different brain tissues were ranked from 1 (the lowest expression value) to 52 (the highest expression value), allocating a score of 10 for 1, 20 for 2 or 3 and 30 points for more than a count of 4 for the brain tissues below the rank of 5. There were 50 out of 117 candidates that had at least 10 points, implying expression at significantly low levels in at least one of the 10 brain tissues, where miR124a was predicted to be highly expressed, in addition to being highly expressed in functionally different tissues, such as, lung, liver, and colon. Meanwhile, gene expression of the remaining 67 genes was not ranked below 5, or their gene expression data was not available, due to no list available from the 10k chip used for the GSE740 (Supplemental Table 3). From 50 genes that had significantly low expression in the brain, 9 genes were scored with 30 points and 20 genes with 20 points.

For the 100 point total, used for our bioinformatics analysis to identify endogenous miR124a targets during neurogenesis, we considered 50 points as a cut-off threshold for miR124a targets. Since genetic expression related to neuronal differentiation might be a critical factor determining whether

PicTar-predicted miR124a targets are endogenous targets of neuronal specific miR124a during neurogenesis, 40 points from the neuronal microarray data, GSE4076, GSE4075 and GSE1588 plus 10 points from the gene expression of brain tissue of GSE740 were counted in our bioinformatics analysis. On the basis of 50 points, 26 genes from 337 orthologues including LRCH4, LMNB1,ANXA11, QKI, SNTB2, TRIB3, IQGAP1, PLP2, KLF4, CCL2, CYB5A, SLC6A15, DDX3X, LAMC1, ETS1, F11R, PTBP1, HBP1, EGR1, G3BP1, PECR, MBOAT5, MAT2A, RFX1, KPNB1 and KLF6, were highly compatible with our bioinformatics analysis (Supplemental Table 3); the gene expression was significantly down-regulated in the brain samples with the exception of QKI where the data was not available in GSE740 (Supplemental Fig 1C).

To analyze the accuracy of our bioinformatics analysis for the identification of endogenous miRNA targets, microarray data from miR124a-treated HeLa cells where mature miR124a is not expressed was analyzed and scored. From GSE2075, 174 annotated genes in the HeLa cells showed more than a 1.5 fold down-regulation by miR124a, and 35 miR124a-down-regulated genes were found among a list of miR124a-predicted targets obtained from the PicTar DB. Only 8 out of 35 genes were significantly down-regulated in at least one of the three different neuronal microarray datasets, GSE4076, GSE4075 and GSE1588 (Supplemental Table 4). Fifteen out of 35 genes had significantly lower expression in the brain tissues from GSE740. Only 4 genes, LMNB1, IQGAP1, LAMC1, and PTBP1 that were predicted by our bioinformatics analysis were scored to be more than 50 points.

3.2 Multiple Endogenous Targets of miR124a During the Neuronal Differentiation of P19 Cells

To monitor the gene expression patterns of the 26 genes predicted by the bioinformatics analysis as being endogenous targets of miR124a during neurogenesis, real-time RT-PCR was performed. A pair of specific primers for mature miR124a was used first with cDNAs synthesized from the corresponding small RNA of neuronal differentiated P19 cells 0d, 1d, 2d, 3d, 4d, 5d, 6d and 7d post-RA treatment. U6 small RNA was used to normalize the quantities of mature miR124a amplified each day post-RA. Real-time RT-PCR demonstrated that endogenous mature miR124a was highly and gradually expressed during the neuronal differentiation of P19 cells induced by RA (Fig 2A). To determine and quantify the endogenous mRNA levels of the 26 novel candidates as targets of mature miR124a, a RT-PCR using the pairs of primers listed in Supplemental Table 1 was performed. Seven out of 26 novel candidates did not produce their amplicons by PCR due to weak expression in P19 cells in the presence or absence of RA. Similar to the gradual increase in gene expression of endogenous mature miR124a observed during the neuronal differentiation of P19 cells, the endogenous gene expression patterns of the 4 candidates, LMNB1, IQGAP1, KLF4 and F11R were clearly visualized with a gradual decrease over 7 days of neuronal differentiation, and the gene expression of 5 genes, LRCH4, ANXA11, TRIB3, LAMC1 and ETS1 were dramatically decreased right after neuronal differentiation (Fig 2B). In addition, when gene expression was normalized to β-actin (internal control), due to a relatively weak expression in non-RA treated cells versus cells treat with RA for 7 days, the region of interest (ROI) analysis for each gene surprisingly showed that all novel miR124a candidates were gradually down-regulated during RA-induced neurogenesis (Supplemental Fig 1D). This result indicates that the down-regulation of the 19 endogenous targets of miR124a was in part caused by the high expression of endogenous mature miR124a during neurogenesis.

To determine whether the 19 novel candidates were down-regulated by mature miR124a, 40 nM of exogenous miR124a precursor was transfected into the P19 cells in the absence of RA. RT-PCR was then performed to determine the gene expression patterns of the 19 candidates. All candidates except for SNTB2, DDX3X, ETS1, F11R and RFX1 appeared to be significantly and dramatically down-regulated at the mRNA level by mature miR124a, suggesting that at least 14 genes out of the 19 candidates were endogenous targets of miR124a during neurogenesis (Fig 2D). Interestingly, ETS1 and F11R which were dramatically down-regulated during neuronal differentiation of P19 cells showed no significant decrease in expression by the transfection of exogenous miR124a precursor in P19 cells. This result may be due to either of the amount of transfected exogenous miR124a precursor in P19 cells or the expression of unknown RNA-binding proteins that bind to the 3' UTR of both genes and inhibit the function of miR124a.[18]

For further verification of our bioinformatics analysis, 5 genes that scored below 10 points in total were randomly selected and are shown in Supplemental Table 4. Both RT-PCR using the pairs of primers and the ROI analysis demonstrated that none of the 5 genes show dramatic or significant down-regulation during neurogenesis except on the seventh day of neuronal differentiation (Supplemental Fig 2A and 2B). In addition, P19 cells with 40 nM of exogenous miR124a precursor demonstrated that none of the genes appeared to be

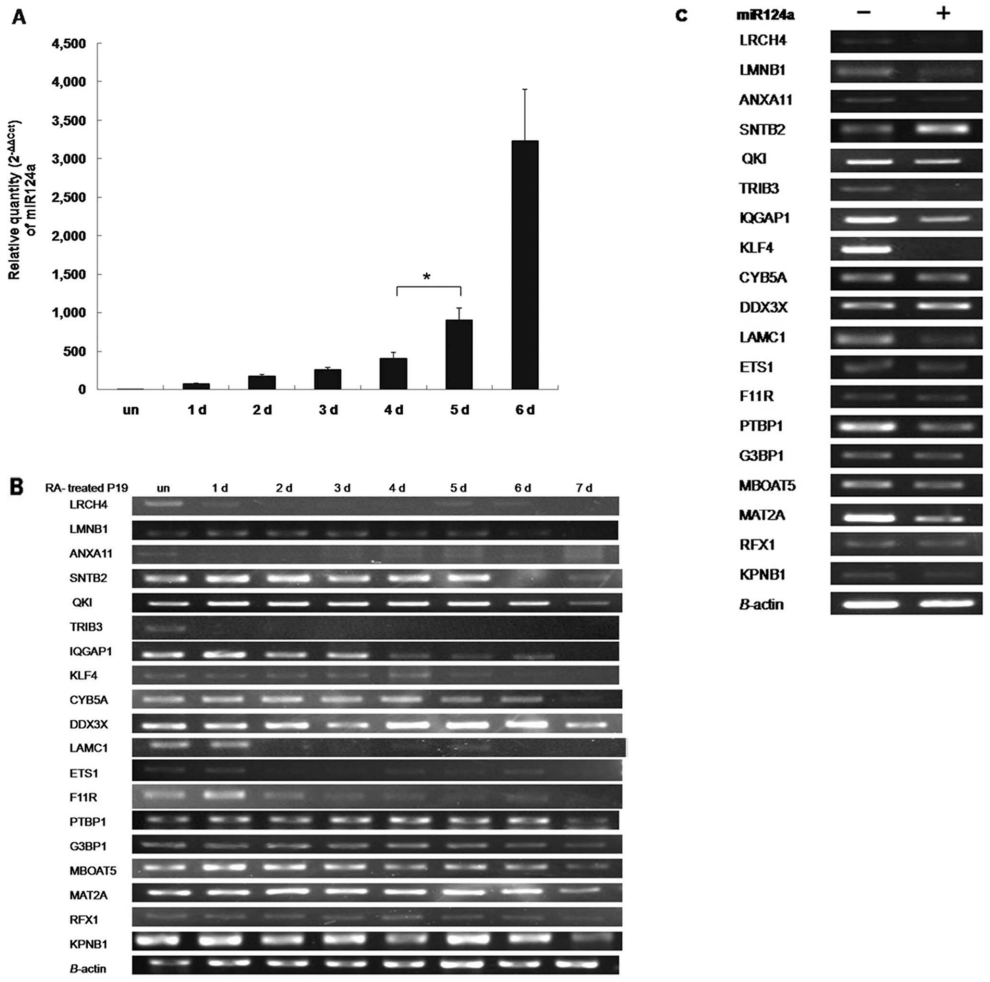

Figure 2. Gene expression patterns of 19 novel candidates for miR124a. (A) qRT-PCR analysis of the expression of mature miR124a during neuronal differentiation of P19 cells induced by RA. Gene expression of mature miR124a followed daily for 6 days post-RA treatment. Endogenous mature miR124a levels were found to gradually increase during neuronal differentiation. Data are normalized versus U6 snRNA ($C_T = C_{T\text{-before}} - C_{T\text{-day}}$, $C_T = C_{T\text{-miRNA}} - C_{T\text{-U6RNA}}$) and are presented as means of triplicate determined ± SDs (*$p < 0.05$). (B) RT-PCR analysis of the endogenous levels of the 19 predicted target genes during neurogenesis. The endogenous expression levels of the 19 candidates in the P19 cells were monitored over 7 days of neuronal differentiation. β-actin was used as an internal control. (C) RT-PCR analysis of the 19 predicted target genes in the P19 cells treated with exogenous pre-miR124a. The MiR124a precursor (40 nM) was transfected into the P19 cells. Most of the endogenous target candidates were found to be significantly down-regulated by miR124a. β-actin was used as an internal control.

directly repressed by mature miR124a (Supplemental Fig 2C).

3.3 Extension of the Bioinformatics Analysis for the Discovery of Endogenous Target Genes of miR9* During Neurogenesis

The bioinformatics strategy used to identify endogenous targets of miR124a was applied for the research of endogenous targets of miR9*. The biogenesis of miR9* has recently been associated with the neuronal differentiation of P19 cells treated with retinoic acid.[19] In the PicTar DB, there are 535 and 305 predicted targets of miR9* for humans and mice, respectively. Comparison of miR9* targets between both species identified 299 orthologues. The function, f(x), for the PicTar score was initially calculated based on 1 to 10 point scoring for each orthologue (Supplemental Table 5). For the second step, the function g(x) of 299 orthologues was investigated using heterogeneous microarray data, GSE7076, GSE4075 and GSE1588. Among 299 genes, 22 PicTar-predicted targets of miR9* were significantly down-regulated during neuronal differentiation in one of three different microarray data sets by more than 1.5 fold, scoring 20 points in each of the 22 targets (Supplemental Table 5). The gene expression of 14 targets was

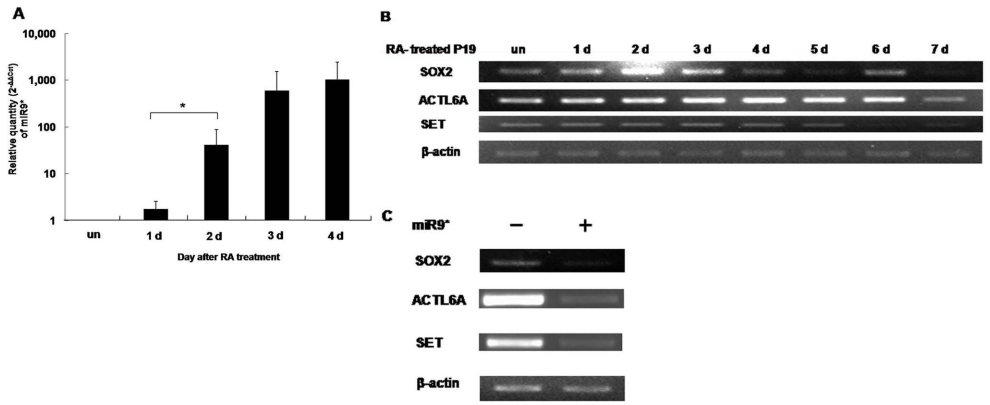

Figure 3. Gene expression pattern of 3 novel endogenous targets of miR9*. (A) qRT-PCR analysis of the expression of mature miR9* during neuronal differentiation of P19 cells induced by RA. Gene expression of mature miR9* was determined for 4 days post-RA treatment. Endogenous mature miR9* levels were found to gradually increase during neuronal differentiation. Data are normalized versus U6 snRNA ($C_T = C_{T\text{-before}} - C_{T\text{-day}}$, $C_T = C_{T\text{-miRNA}} - C_{T\text{-U6RNA}}$) and are presented as the means of triplicate determined ± SDs ($*p < 0.05$). (B) RT-PCR analysis of the endogenous levels of the 3 predicted target genes during neurogenesis. The endogenous expression levels of the 3 candidates in the P19 cells were monitored over 7 days of neuronal differentiation. β-Actin was used as an internal control. (C) RT-PCR analysis of the 3 predicted target genes in the P19 cells were treated with exogenous pre-miR9*. The miR9* precursor (40 nM) was transfected into the P19 cells. Most of the endogenous target candidates were found to be significantly down-regulated by miR124a. β-actin was used as an internal control.

highly repressed in two of three different microarray datasets, with allocating 40 points allocated for each gene. In addition, two genes including QKI and SOX2, had 60 points and were significantly down-regulated in all three different microarray datasets. For the third step, GSE740 microarray data was used to determine gene expression in normal brain tissues. The function h(x) was further analyzed with 38 targets that had at least 20 points from the function g(x). Seven out of the 38 possible endogenous targets of miR9* had relatively lower expression in one of 10 different brain tissues compared to the other tissues allocating 10 points for each target, which two genes in two or three brain tissues scored 20 points and no gene in four brain tissues scoring 30 points (Supplemental Table 5). However, the remaining 29 targets were not scored because their gene expression values in brain tissues were not ranked below 5 in the 52 normal tissues or were not available on the 10k chip used for the GSE740.

To identify the endogenous targets of miR9*, during neuronal differentiation, 50 out of a total 100 points was used as a cut-off threshold (Supplemental Table 5). Five genes, ACTL6A, QKI, SOX2, KLF4 and SET, were strong candidates for endogenous targets of miR9* during neurogenesis when miR9* was highly expressed.

3.4 *In Vitro* Validation of the Endogenous Targets of miR9* Identified by Bioinformatics Analysis

To determine whether the five candidates for endogenous targets of miR9*, that is; ACTL6A, QKI, SOX2, KLF4 and SET, as suggested by our bioinformatics analysis, were directly repressed by miR9* during neuronal differentiation of the P19 cells, the expression of the miR9* was first investigated by real-time PCR. Similar to our findings reported previously,[19] mature miR9* gradually increased after neuronal differentiation of the P19 cells induced by RA treatment (Fig 3A). Among the five targets, both our bioinformatics and *in vitro* PCR analysis demonstrated that QKI and KLF4 were down-regulated by miR124a during neuronal differentiation of the P19 cells (Fig 2B and 2C). Therefore, to monitor the endogenous transcript levels of the miR9* targets during neuronal differentiation of the P19 cells, only ACTL6A, SOX2 and SET genes were amplified by PCR using the pairs of primers listed in Supplemental Table 2. When miR9* was highly expressed during neuronal differentiation of the P19 cells, all of three targets were decreased gradually or suddenly due to the fact that the expression time-point and cellular distribution of miR9* and its targets in cells may result in different patterns of the functional interaction between miR9* and its targets (Fig 3B). To determine if the three targets were directly repressed by mature miR9*, 40 nM of exogenous miR9* precursor was induced in the P19 cells, in the absence of RA, and the gene expression patterns of the three candidates were determined by RT-PCR (Fig 3C). The transcript levels of all of three candidates appeared to be dramatically repressed by exogenous mature miR9*. This finding implies that ACTL6A, SOX2 and

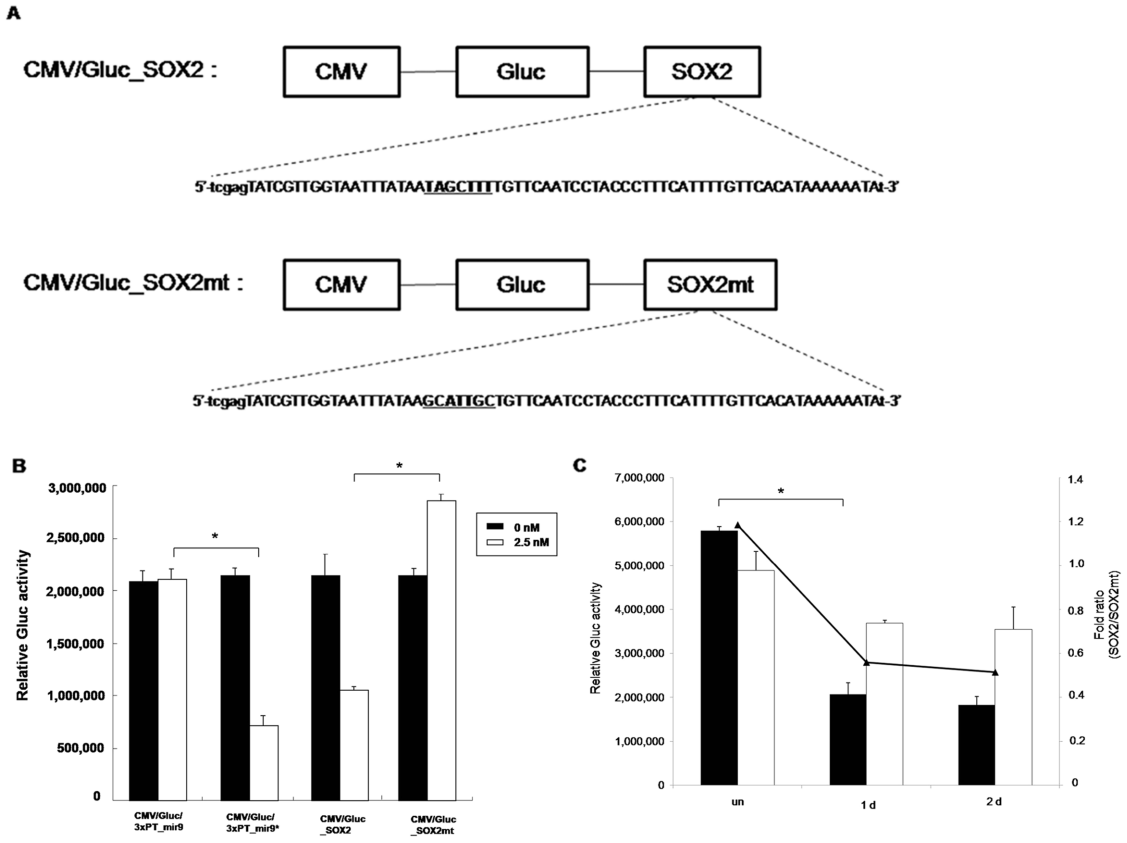

Figure 4. *In vitro* luciferase assay of SOX2 regulated by miR9* during neuronal differentiation. (A) A schematic diagram of CMV/Gluc_SOX2 and CMV/Gluc_SOX2mt for monitoring SOX2 regulated by miR9*. The underlined sequences indicated a seed region of miR9* at the 3' UTR of SOX2. The small letters at the 5' end indicate restriction enzyme recognition sites for XhoI and XbaI at the 3' end. (B) *In vitro* Gluc activity of CMV/Gluc_SOX2 regulated by miR9* in HeLa cells. The Gluc activities of 4 different recombinant constructs, CMV/Gluc/3xPT_mir9, CMV/Gluc/3xPT_miR9*, CMV/Gluc_SOX2 and CMV/Gluc_SOX2mt were measured in response to 0 and 2.5 nM of exogenous miR9* precursor in the HeLa cells. The CMV/Fluc vector was co-transfected to normalize Gluc activity. Experiments were performed in triplicate and results are shown as the means ± SD (*$p < 0.05$). (C) *In vitro* luciferase assay of endogenous SOX2 during the neuronal differentiation of the P19 cells. CMV/Gluc_SOX2 or CMV/Gluc_SOX2mt was co-transfected with the CMV/Fluc vector. The CMV/Fluc-normalized Gluc activity of each construct was measured over 2 days. The right y-axis represents the Gluc fold ratio of CMV/Gluc_SOX2 divided by CMV/Gluc_SOX2mt before (un), 1 day (1d) and 2 days (2d) after neuronal differentiation of the P19 cells induced by RA. The line with a triangular head indicates the ratio for each day. The black bar represents the construct of CMV/Gluc_SOX2 and the gray bar indicates CMV/Gluc_SOX2mt. Gluc activities were performed in triplicate and results are expressed as the mean ± SD (*$p < 0.05$).

SET genes are directly regulated endogenous targets of miR9* during neuronal differentiation of P19 cells.

3.5 *In Vitro* and *In Vivo* Imaging of miR9*-Mediated Repression of SOX2 During Neuronal Differentiation of P19 Cells

Among the three endogenous targets directly regulated by miR9* during neurogenesis, the sex determining region Y-box 2 (SOX2) was randomly selected to further study visualize miR9*-mediated repression of a target gene both *in vitro* and *in vivo*. SOX2 is well known to function as a transcription factor that is essential for the maintenance of self-renewal of undifferentiated embryonic stem cells during the regulation of embryonic development and the determination of the cell's fate.[20]

To investigate whether the 3' UTR of SOX2 containing the SR was possibly bound and destabilized at the transcript level by miR9*, we searched the PicTar DB. The 66-bp region from the 3' UTR of SOX2 containing the SR was cloned into a reporter vector containing the open reading frame of Gluc regulated by the CMV promoter (designated as CMV/Gluc_SOX2) (Fig 4A). To compare the bioluminescent signals of CMV/Gluc_SOX2, 7-bp of the 3' UTR of SOX2 in the SR

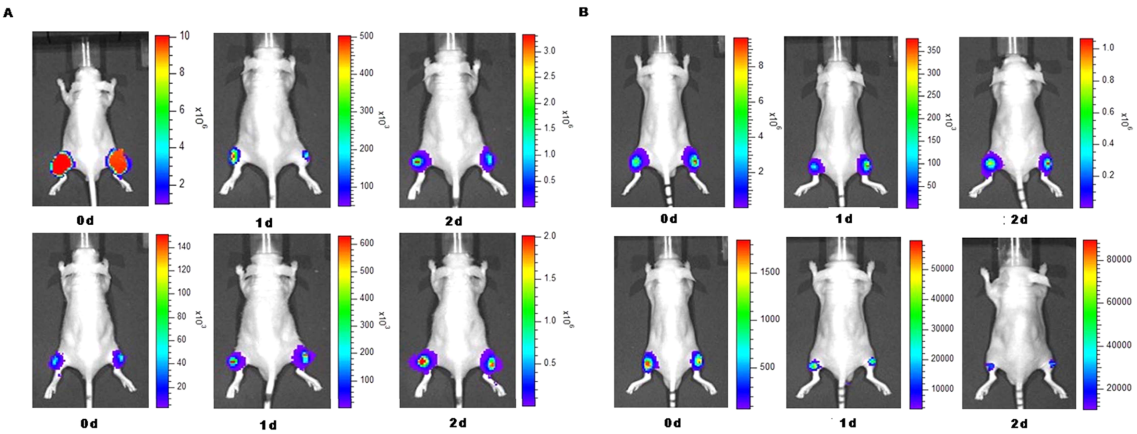

Figure 5. *In vivo* luciferase imaging of miR9*-mediated repression of SOX2 during neuronal differentiation in nude mice. Images were acquired for 2 days after implanting the P19 cells into the left thigh without RA treatment and the right thigh with RA treatment of nude mice. The CMV/Fluc was used as an internal control and was co-transfected to normalize the Gluc activity of CMV/Gluc_SOX2 (A) or CMV/Gluc_SOX2mt (B). The lower panels of A and B show *in vivo* Fluc imaging of CMV/Fluc and the upper panels show the Gluc activity of CMV/Gluc_SOX2 or CMV/Gluc_SOX2mt.

was mutated (designated as CMV/Gluc_SOX2mt). To determine the specificity of CMV/Gluc_SOX2, it was transfected into HeLa cells where miR9* is not expressed. In addition, CMV/Gluc/3xPT_mir9 (containing 3 copies of perfectly complementary sequences of mature miR9 in CMV/Gluc vector) and CMV/Gluc/3xPT_mir9* (containing 3 copies of perfectly complementary sequences of mature miR9* in CMV/Gluc vector), were used to monitor miR9 and miR9*, as a negative and a positive control respectively.[19] The reporter gene, CMV/Fluc, hardly affected by miR9*, was co-transfected into the HeLa cells to normalize the Gluc activities of CMV/Gluc_SOX2, CMV/Gluc_SOX2mt, CMV/Gluc/3xPT_mir9 or CMV/Gluc/3xPT_mir9* in response to exogenous miR9*. The Gluc activities of both CMV/Gluc_SOX2 and CMV/Gluc/3xPT_mir9* were significantly decreased by 2.5 nM of exogenous miR9* compared to 0 nM of miR9* (Fig 4B). However, CMV/Gluc_SOX2mt and CMV/Gluc/3xPT_mir9 not show any significant repression of Gluc expression in the presence of exogenous miR9* in the HeLa cells. The response of Gluc activity in CMV/Gluc/3xPT_mir9 and CMV/Gluc/3xPT_mir9* to exogenous miR9* was exactly the same as previous reported.[19]

To perform *in vitro* imaging of endogenously repressed SOX2 by miR9*, CMV/Gluc_SOX2 or CMV/Gluc_SOX2mt was co-transfected into P19 cells with CMV/Fluc for normalization; then neuronal differentiation was induced by RA for 2 days. When gene expression of miR9* gene was gradually increased, during neuronal differentiation of the P19 cells (Fig 3A), the normalized Gluc activity of CMV/Gluc_SOX2 was significantly decreased compared to that of the CMV/Gluc_SOX2mt (Fig 4C). When the P19 cells were differentiated by RA, the Gluc fold ratio of CMV/Gluc_SOX2 divided by CMV/Gluc_SOX2mt was decreased about three fold compared to the undifferentiated P19 cells.

The successfully developed luciferase reporter systems, CMV/Gluc_SOX2 and CMV/Gluc_SOX2mt were used to monitor *in vivo* the endogenously mediated repression of SOX2 by miR9* during neuronal differentiation. 3×10^6 of P19 cells were first co-transfected with CMV/Fluc and subcutaneously implanted into the right thigh of nude mice with RA treatment and the left thigh without RA treatment as a control. Fluc activity in the CMV/Fluc construct was measured from the left thighs not treated with RA and the treated right thighs (treatment for 2 days) by the IVIS 100 and showed no significant changes before and after neuronal differentiation (Fig 5A and 5B, lower panel). Gluc activity of CMV/Gluc_SOX2 from the right thigh was gradually and clearly decreased and more rapidly repressed by RA treatment compared to CMV/Gluc_SOX2 without RA treatment in the left thighs. However, the CMV/Gluc_SOX2mt was not significantly reduced in the right thighs after treatment with RA compared to the left thighs without treatment of RA (Fig 5A and 5B, upper panel). The region of interest (ROI) analysis showed that Gluc activity of CMV/Gluc_SOX2 was reduced almost 3-fold on day 1 and 2 versus that of CMV/Gluc_SOX2mt (Supplemental Fig 3). The *in vivo* Gluc signals and ROI analysis demonstrated that SOX2 was significantly and directly down-regulated by miR9* during neuronal differentiation of

the P19 cells.

4. Discussion

The expression and regulation of miRNAs have been reported to be involved in various cellular processes, such as, the regulation of cell differentiation, proliferation, apoptosis and development, as well as in the pathogeneses of clinically important diseases, such as, cancer, cardiovascular diseases, neurological diseases, viral diseases and metabolic disorders.[21,22] In general, tens of coding mRNAs in a cell are directly regulated by a single miRNA, which results in the down-regulations of hundreds of other metabolic genes.[2] However, the identification of the endogenous targets of miRNAs, using bioinformatics, has not provided useful information due to the absence of experimental confirmation of endogenous miRNA targets.

In this study where the endogenous targets of miRNAs were investigated, the scoring system used for the bioinformatics analysis resulted in 60 out of 100 points for the gene expression data related to neuronal differentiation. This was based on the assumption that the bioinformatics analysis depended on gene expression in cells where miRNAs are expressed. If this system is accurate and reliable, the microarray data deposited in public DB can be easily used for determining endogenous targets of other miRNAs. The gene expression in tissues, where miRNAs are expressed, is of central importance in determining the genetic expression of miRNA targets. Therefore, 30 points were allocated for the gene expression data in 52 normal tissues; this is because the tissues used for microarray experiments are frequently contaminated by blood or other adjacent tissues. Our rapid and successful bioinformatics strategy, based on data mining and *in vitro* experimentation using multiple heterogeneous information sources of microarray data, in public DBs, provides a promising method for identifying endogenous targets of miR124a during neurogenesis. The results showed that 84.2% (16 of 19 targets) of the endogenous targets, predicted by our bioinformatics analysis, were found to be significantly down-regulated by miR124a during neurogenesis. About 95% of the 337 predicted miR124a targets, in the human and mouse PicTar DB, were excluded from the bioinformatics analysis because they did not meet the inclusion criteria adopted, i.e., only those genes that were down-regulated in the heterogeneous microarray data for neuronal differentiation, and showed significantly low expression in human brain tissue samples. Furthermore, only 2% of miR124a targets, experimentally predicted in the HeLa cells,[2] where mature miR124a is not expressed, satisfied the high-stringent criteria of our bioinformatics analysis. Our findings showed that 93.1% of the endogenous miR124a targets, scoring more than 20 points, by bioinformatics computation, were found not to be significantly down-regulated by miR124a in the HeLa cells, due to weak or barely perceptible expression of 109 out of 117 endogenous miR124a candidates; this finding implies that endogenous miRNA targets should be screened for and monitored where miRNAs are expressed. Among the 16 novel endogenous miR124a targets, PTPB1 (highly expressed in non-neuronal cells) and LAMC1 (a heterodimer component) have previously been shown to be directly regulated by miR124a.[23,24]

The extension of this bioinformatics approach for finding endogenous miR9* targets demonstrated that 38 out of 299 genes predicted to be miR9* targets, in both the human and mouse PicTar DB, might be significantly regulated by miR9* during neurogenesis. Among them, we selected genes with more than 50 points as true endogenous targets of miR9*; the *in vitro* validation by RT-PCR confirmed that the transcript levels of ACTL6A, QKI, SOX2, KLF4 and SET genes were significantly repressed during neuronal differentiation in the P19 cells when expression of the miR9* gradually increased.

If the DNA chips for gene expression data of normal tissues including brain had more annotated genes, the number of endogenous targets analyzed by our bioinformatics approach could be expanded. Since GSE740 (52 normal tissues) contained 10,000 annotated genes on the DNA chip, many of the targets with g(X) scores were missed from the gene list of GSE740, resulting in scores of less than 50 points; such scores were not considered as targets for endogenous miRNAs by our bioinformatics analysis. The association between scores analyzed by our bioinformatics approach and the PicTar scores demonstrated that the groups of genes scoring more than 50, 49-40, 39-30, and 29-20 points had in average of 5.7, 3.8, 4.0 and 3.6 PicTar scores for miRNA124a targets, and 6.2, 4.9, 3.7 and 4.6 for miRNA9* targets, respectively. Our bioinformatics analysis show a tendency for individual genes with a higher number of points ($\geq$ 50 points) by our bioinformatics analysis to have PicTar scores $\geq$ 5.0.

There are four critical transcription factors that are important for maintaining the pluripotency of embryonic stem cells (ESC): OCT4, SOX2, NANOG, and TCF3.[20] SOX1, SOX2 and SOX3 are expressed in a redundant fashion to maintain the expression of the progenitor cell characteristics and to prevent cells from up-regulating post-mitotic neuronal markers, and counteract neurogenesis during central nervous system development.[25] Reduction of SOX2 expression promotes neural progenitor cells to initiate differentiation prematurely.[25]

In this study, both *in vitro* and *in vivo* bioluminescent reporter systems containing miR9* SR at the 3' UTR of SOX2, successfully and specifically demonstrated that Gluc expression of CMV/Gluc_SOX2 was significantly repressed by the direct binding of miR9* during neuronal differentiation of P19 cells, implying that SOX2 is an endogenous target of miR9* during neurogenesis.

Therefore, the gene expression data deposited into public DBs and bioinformatically predicted targets of miRNAs from the PicTar and miRNAMap DBs were found to be useful for identifying novel endogenous miRNA targets. The availability of microarray data from cells or tissues where miRNAs of interest are expressed is critical, and the gene expression data of *in silico* predicted targets is likely to be the most important method for filtering out false positive candidates predicted by the PicTar and miRNAMap DBs.

One limitation of the bioinformatics analysis for identifying endogenous targets of miRNA, was the reduction of mRNA levels, which resulted in exclusion of targets regulated at the translational level by miR124a and miR9*. These findings imply that more endogenous targets of miR124a and miR9*, even those with scores less than 50 points, might be present during neurogenesis but our bioinformatics system is not sensitive enough to identify such targets. Nevertheless, our bioinformatics approach was found to provide a useful and powerful means for identifying endogenous targets of miR124a and miR9* during neurogenesis. This form of data mining, involving the screening of miRNA targeted endogenous genes using microarray data from public DBs, could be generalized to globally identify targets and functions of other miRNAs related to cellular metabolism and disease.

5. Conclusion

The results of this study demonstrate a simple bioinformatics approach for identifying endogenous miRNA targets in the laboratory by investigating miRNA regulation as follows: 1) download the predicted miRNA targets from public miRNA DBs, 2) if possible, refer to microarray experiments using cells treated with the miRNAs of interest, 3) compare the gene expression ratios of *in silico* targets using heterogeneous microarray data resources available from public DBs to determine whether these genes are significantly down-regulated, and 4) monitor the gene expression patterns of down-regulated targets, in healthy normal or abnormal tissues where the miRNAs of interest might be highly expressed, to determine whether they are tissue- or cell-specifically repressed by the cell- or tissue-specific miRNAs of interest by *in vitro* and *in vivo* visualization.

In conclusion, analysis of important information from public raw data using the bioinformat approach presented in this paper can easily and rapidly assists biologists and biomedical researchers in the diagnosis of or treatment of disease by identifying and monitoring genes of interest.

Supplementary information: Supplementary information is available at the Korea Tissue Engineering and Regenerative Medicine Society website (www.kterms.or.kr).

Acknowledgments: This work was supported by a grant of the Korea Healthcare technology R&D Project, Ministry for Health, Welfare & Family Affairs, Republic of Korea (A085136) and the Bio & Medical Technology Development Program of the National Research Foundation (NRF) funded by the Korean government (MEST) (No. 2011-0019270).

References

1. Y Lee, K Jeon, JT Lee, *et al.*, MicroRNA maturation: stepwise processing and subcellular localization, *Embo J*, **21**, 4663 (2002).
2. LP Lim, NC Lau, P Garrett-Engele, *et al.*, Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs, *Nature*, **433**, 769 (2005).
3. HJ Kim, JK Chung, W Hwang do, *et al.*, In vivo imaging of miR-221 biogenesis in papillary thyroid carcinoma, *Mol Imaging Biol*, **11**, 71 (2009).
4. X Yan, T Chao, K Tu, *et al.*, Improving the prediction of human microRNA target genes by using ensemble algorithm, *FEBS Lett*, **581**, 1587 (2007).
5. JG Joung, KB Hwang, JW Nam, *et al.*, Discovery of microRNA-mRNA modules via population-based probabilistic learning, *Bioinformatics*, **23**, 1141 (2007).
6. DP Bartel, MicroRNAs: genomics, biogenesis, mechanism, and function, *Cell*, **116**, 281 (2004).
7. A Krek, D Grun, MN Poy, *et al.*, Combinatorial microRNA target predictions, *Nat Genet*, **37**, 495 (2005).
8. PW Hsu, HD Huang, SD Hsu, *et al.*, miRNAMap: genomic maps of microRNA genes and their target genes in mammalian genomes, *Nucleic Acids Res*, **34**, D135 (2006).
9. X Wang, X Wang, Systematic identification of microRNA functions by combining target prediction and expression profiling, *Nucleic Acids Res*, **34**, 1646 (2006).
10. HJ Kim, YH Kim, DS Lee, *et al.*, In vivo imaging of functional targeting of miR-221 in papillary thyroid carcinoma, *J Nucl Med*, **49**, 1686 (2008).
11. HY Ko, DS Lee, S Kim, Noninvasive imaging of micro RNA124a-mediated repression of the chromosome 14 ORF 24 gene during neurogenesis, *FEBS J*, **276**, 4854 (2009).
12. M Deo, JY Yu, KH Chung, *et al.*, Detection of mammalian microRNA expression by in situ hybridization with RNA oligonucleotides, *Dev Dyn*, **235**, 2538 (2006).

13. TY Leon, ES Ngan, HC Poon, *et al.*, Transcriptional regulation of RET by Nkx2-1, Phox2b, Sox10, and Pax3, *J Pediatr Surg*, **44**, 1904 (2009).
14. J Pachernik, V Bryja, M Esner, *et al.*, Neural differentiation of pluripotent mouse embryonal carcinoma cells by retinoic acid: inhibitory effect of serum, *Physiol Res*, **54**, 115 (2005).
15. K Aiba, AA Sharov, MG Carter, *et al.*, Defining a developmental path to neural fate by global expression profiling of mouse embryonic stem cells and adult neural stem/progenitor cells, *Stem Cells*, **24**, 889 (2006).
16. SS Williams, JP Mear, HC Liang, *et al.*, Large-scale reprogramming of cranial neural crest gene expression by retinoic acid exposure, *Physiol Genomics*, **19**, 184 (2004).
17. JM Johnson, J Castle, P Garrett-Engele, *et al.*, Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays, *Science*, **302**, 2141 (2003).
18. M Kedde, R Agami, Interplay between microRNAs and RNA-binding proteins determines developmental processes, *Cell Cycle*, **7**, 899 (2008).
19. MH Ko, S Kim, W Hwang do, *et al.*, Bioimaging of the unbalanced expression of microRNA9 and microRNA9* during the neuronal differentiation of P19 cells, *Febs J*, **275**, 2605 (2008).
20. A Marson, SS Levine, MF Cole, *et al.*, Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells, *Cell*, **134**, 521 (2008).
21. WP Kloosterman, RH Plasterk, The diverse functions of microRNAs in animal development and disease, *Dev Cell*, **11**, 441 (2006).
22. J Weiler, J Hunziker, J Hall, Anti-miRNA oligonucleotides (AMOs): ammunition to target miRNAs implicated in human disease?, *Gene Ther*, **13**, 496 (2006).
23. EV Makeyev, J Zhang, MA Carrasco, *et al.*, The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing, *Mol Cell*, **27**, 435 (2007).
24. X Cao, SL Pfaff, FH Gage, A functional study of miR-124 in the developing neural tube, *Genes Dev*, **21**, 531 (2007).
25. M Bylund, E Andersson, BG Novitch, *et al.*, Vertebrate neurogenesis is counteracted by Sox1-3 activity, *Nat Neurosci*, **6**, 1162 (2003).