SUPPLEMENTARY INFORMATION

MicroRNA-26a/b and their host genes cooperate to inhibit the G1/S transition by activating the pRb protein

Zhu, Lu, et al.

Index:

Supplementary Materials and Methods

Vectors

To create luciferase reporter construct, a wild-type 3'UTR fragment of human CDK6 or CCNE1 mRNA that contained putative binding sites for miR-26a/b was PCR-amplified and inserted into the *Eco*RI and *Xba*I sites downstream of the stop codon of the firefly luciferase in pGL3cm vector, which was generated previously (1). The mutant 3'UTR, which carried the mutated sequence in the complementary site for the seed region of miR-26a/b, was generated using fusion PCR based on the construct with wild-type 3'UTR.

The coding sequences of CDK6, CCNE1, c-Myc, CTDSP1, CTDSP2 and CTDSPL were cloned into the *Bam*HI/*Eco*RI or *Hind*III/*Eco*RI sites of pc3-gab to generate expression vectors named pc3-gab-CDK6, pc3-gab-CCNE1, pc3-gab-c-Myc, pc3-gab-CTDSP1, pc3-gab-CTDSP2 and pc3-gab-CTDSPL, respectively. The plasmid pc3-gab was produced based on pcDNA3.0 by replacing the neomycin open reading frame with an expression cassette of Enhanced Green Fluorescent Protein (EGFP) gene (1). All constructs were confirmed by direct sequencing. Primers used for PCR are provided in Supplementary Table 1.

Analysis of gene expression

For semiquantitative RT-PCR and real-time quantitative RT-PCR (qPCR) analyses, total RNA from cells was extracted using the TRIzol reagent (Invitrogen, Carlsbad, CA). For semiquantitative RT-PCR analysis of mRNA, 2 μg of total RNA was

subjected to DNase I digestion (1U/μL, Fermentas, Hanover, MD) at 37℃ for 30 min and then to heat inactivation of DNase I at 65℃ for 10 min, followed by reverse-transcription using Moloney murine leukemia virus reverse transcriptase (Promega, Madison, WI). Semiquantitative RT-PCR was performed as described (1).

For qPCR analysis, 20 ng of total RNA was subjected to DNase I digestion (1U/μL, Fermentas) at 37℃ for 30 min and then to heat inactivation of DNase I at 65℃ for 10 min. qPCR analysis for the expression of CTDSP1/2/L and reference genes (β-actin and GAPDH) was performed using THUNDERBIRD SYBR qPCR Mix (QPS-201, TOYOBO, Japan). The temperature cycle profile for the qPCR reactions was 95℃ for 30s, and then 40 cycles of 95℃ for 5s, 60℃ for 20s, and 72℃ for 10s. To verify the specificity of PCR product, melting curve analysis was performed immediately after amplification, as follows: heating to 95℃ for 20s, cooling to 60℃ for 20s, followed by a temperature increase to 95℃ with a transition rate of 0.11℃/s while continuously collecting the fluorescent signal. Expression of miR-26a/b was quantified using the miRCURY LNA Universal RT microRNA PCR system (Exiqon, Vedbaek, Denmark) when β-actin or GAPDH were used as internal standards, or by a TaqMan MicroRNA Assay kit (Applied Biosystems, Foster City, CA) when U6B was used as reference. All qPCR analysis was performed on a LightCycler 480 (Roche Diagnostics, Germany). The level of target genes was normalized to the expression of internal control genes, which yielded a $2^{-\Delta\Delta Ct}$ value. All reactions were run in triplicate. The cycle threshold (Ct) values should not differ more than 0.5 among triplicates. Sequences for probes and primers are listed in Supplementary Table 1.

The sources of antibodies used for Western blot: mouse monoclonal antibodies (mAb) for CDK6 (cat. 3136, CST, Beverly, MA), pRb (cat. 9309, CST) and c-Myc (cat. ab32, Abcam, Cambridge, MA); Rabbit polyclonal antibody against phospho-Ser780 of pRb (cat. 9307, CST) and cyclin E1 (cat. sc-481, Santa Cruz, CA); mouse mAb for β-actin (cat. BM0627, Boster, Wuhan, China).

For immunohistochemistry, paraffin-embedded, formalin-fixed tissues were cut into 5-μm section, placed on polylysine-coated slide, deparaffinized in xylene, rehydrated through graded ethanol, quenched for endogenous peroxidase activity in 0.3% hydrogen peroxide, and processed for antigen retrieval by high-pressure heating in 1 mM EDTA (pH 8.0). Sections were incubated at 4℃ overnight with mouse mAb for CDK6 (cat. ab54576, Abcam) or cyclin E1 (cat. ZM-0086, Zhongshan Goldenbridge Biotechnology, Beijing, China). Immunostaining was performed using ChemMate DAKO EnVision Detection Kit, Peroxidase/DAB, Rabbit/Mouse (code K 5007, DakoCytomation, Glostrup, Denmark), which resulted in a brown-colored precipitate at the antigen site. Subsequently, the sections were counterstained with hematoxylin (Zymed Laboratories, South San Francisco, CA) and mounted in non-aqueous mounting medium. All runs included a no primary antibody control. CDK6 and cyclin E1 staining was evaluated under a light microscope at a magnification of x400. For each specimen, ten images of representative areas were acquired and a total of 1000 to 2000 tumor cells were counted and the percentage of positive cells was calculated.

Bioinformatic tools

The databases used for bioinformatic analysis included: miRBase (release 16.0) database (ftp://mirbase.org/pub/mirbase/CURRENT/miFam.dat.gz/) for retrieval of miRNA families; Ensembl Genome Browser (release 60) (http://www.ensembl.org/index.html) for retrieval of protein families and genomic sequence; Alibaba2.1 (http://www.gene-regulation.com/pub/programs/alibaba2/index.html) for the prediction of transcription factor binding sites; TargetScan (release 4.1, http://www.targetscan.org/vert_40/) for the prediction of miRNA target genes; Protein Lounge Pathway Database (http://www.proteinlounge.com) for retrieval of defined biological pathways.

Supplementary References:

1. Su, H., Yang, J.R., Xu, T., Huang, J., Xu, L., Yuan, Y. and Zhuang, S.M. (2009) MicroRNA-101, down-regulated in hepatocellular carcinoma, promotes apoptosis and suppresses tumorigenicity. *Cancer Res*, **69**, 1135-1142.

Supplementary Figure Legends

Supplementary Figure 1. miR-26 family genes are colocalized with CTDSP1/2/L gene loci through vertebrate evolution. Schematic representation of miR-26a/b and CTDSP1/2/L gene loci was taken from the UCSC Genome Browser. The annotated transcripts produced from three loci are shown at the top of each panel. The relative positions of exons (blue filled boxes) and introns (solid lines) of CTDSP1/2/L and the miRNAs (red filled triangles) are indicated. The sequence features around the miRNA sequences are presented. Green bars represent the sequence conservation score in the indicated species. The human genome was used as reference. Double-lined regions indicate sequences which are non-homologous to human genome. Black boxes mark the conservation score spanning miR-26a/b loci.

Supplementary Figure 2. miR-26a/b are expressed concomitantly with CTDSP1/2/L under physiological conditions. (A) Analysis of mature miR-26a/b and CTDSP1/2/L expression in human primary fibroblasts by quantitative real-time PCR (qPCR). Cells were serum deprived for 48h, then stimulated to enter S-phase by serum re-addition. Unsync, exponentially growing cells without serum deprivation; 0h denotes serum-deprived cells at the time point when serum was re-added; 2h, 4h, 16h and 24h denote cells in the indicated time points after serum re-addition. (B) Analysis of mature miR-26a/b and CTDSP1/2/L expression in mouse liver by qPCR. 0h denotes two-thirds of the liver that was surgically removed; 24h, 36h, 48h and 72h denote livers that were obtained at the indicated time points after partial-hepatectomy. The levels of miR-26a/b and CTDSP1/2/L were normalized to GAPDH.

Supplementary Figure 3. miR-26a/b are expressed concomitantly with CTDSP1/2/L under pathological condition. (A) Analysis of mature miR-26a/b and CTDSP1/2/L expression in 32 paired HCC and adjacent non-tumor tissues by qPCR. T, HCC tissue and N, adjacent non-cancerous tissue. The *y*-axis indicates the relative level of miR-26a/b and CTDSP1/2/L. The boxes represent the interquartile range (25th to 75th centiles). The horizontal line inside the box indicates the median. The vertical whiskers extend to the maximum and minimum values. (B) The correlation between expression levels of miR-26a/b and their respective host genes in 32 paired HCC and adjacent non-tumor tissues. The expression status of miR-26a/b and CTDSP1/2/L is shown in a log 2 scale. For (A) and (B), statistical analysis was performed by Wilcoxon matched-pairs signed-ranks test and Spearman's correlation coefficient, respectively. The expression levels of miR-26a/b and CTDSP1/2/L were normalized to GAPDH.

Supplementary Figure 4. The correlation between expression levels of miR-26a and miR-26b in 32 paired HCC and adjacent non-tumor tissues. The expression status of miR-26a/b is shown in a log 2 scale. Statistical analysis was performed by Spearman's correlation coefficient. The expression levels of miR-26a and miR-26b were normalized to GAPDH.

Supplementary Figure 5. miR-26a/b overexpression causes accumulation of G1-population. Cells were transfected with 50 nM NC (negative control) or miR-26a/b. Nocodazole (40 ng/ml) was added 32h after transfection and the cells

were cultured for an additional 16h before harvest. The percentage of G1-population is indicated within each histogram.

Supplementary Figure 6. Reduction of endogenous miR-26a/b levels by anti-miR-26. Forty-eight hours after transfection with 200 nM anti-miR-C (negative control) or anti-miR-26, HepG2 cells were analyzed by qPCR. The expression of miR-26 was normalized to U6B. *, $P < 0.05$.

Supplementary Figure 7. Expression levels of miR-26a/b in normal human liver tissue and different cancer cell lines. Expression of miR-26a/b was analyzed by qPCR. The level of miR-26 was normalized to U6B. ***, $P < 0.001$.

Supplementary Figure 8. miR-26a/b and their putative binding sites in the 3'UTR of target genes. The mutant miR-26 binding site was generated in the complementary site (*underlined*) that binds to the seed region of miR-26.

Supplementary Table 1. Sequences of RNA and DNA oligonucleotides

Name	Sense strand/Sense primer (5'-3')	Antisense strand/Antisense primer (5'-3')
miRNA and siRNA duplexes		
siCDK6	CUGGAAAGGUGCAAAGAAAdTdT	UUUCUUUGCACCUUUCCAGdGdT
siCCNE1	UGGCCAAAAUCGACAGGACdTdT	GUCCUGUCGAUUUUGGCCAdTdT
siMyc	GGUCAGAGUCUGGAUCACCdTdT	GGUGAUCCAGACUCUGACCdTdT
siCTDSP1	UAGCUGACCUGCUGGACAAdTdT	UUGUCCAGCAGGUCAGCUAdCdT
siCTDSP2	CGUAUAAGGAGGAAGCAAAdTdT	UUUGCUUCCUCCUUAUACGdCdA
siCTDSPL	GGAGGAGAAUGGUGGGCUUdTdT	AAGCCCACCAUUCUCCUCCdAdC
miRNA inhibitors		
anti-miR-C	GUGGAUAUUGUUGCCAUCA	
anti-miR-26	ACCUAUCCUGAAUUACUUGAA	
Primers for gene or 3'UTR cloning (restriction enzyme sites are underlined)		
CDK6	AGTGGATCCATGGAGAAGGACGGCCTGT	AGTGAATTCTCAGCTTAAGGCGGCTGCT
cyclin E1	AGTAAGCTTATGCCGAGGGAGCGCAGG	AGTGAATTCTCACGCCATTTCCGGCCC
Myc	AGTAAGCTTCTGGATTTTTTTCGGGTAGTG	AGTGAATTCTTACGCACAAGAGTTCCGTAG
CTDSP1	AGTAAGCTTATGGACAGCTCGGCCGTCATT	AGTGAATTCCTAGCTCCCTGGCCGTGGCTG
CTDSP2	AGTAAGCTTATGGAACACGGCTCCATCATC	AGTGAATTCCTAAGGGGCCCGCAGCTG
CTDSPL	AGTAAGCTTATGGACGGCCCGGCCATCAT	AGTGAATTCCTACCTATTGCAGAGTCTGTGCA
CDK6-1 3'UTR	AGTGAATTCAGCAGGGGATTTTCATGTTG	AGTTCTAGACTGGGGGTAGCTGATGCTAT
CDK6-2 3'UTR	AGTGAATTCAGAGATGGTTCCTGATGCTG	AGTTCTAGAGGCCACTGTGGTAACTCTCAA
cyclin E1 3'UTR	GATGAATTCCATCAAACAGGGCAAAGTGTT	GATTCTAGACCAGCTGTCTCAAAAACAGTATTA
Primers for RT-PCR		
CDK6	TGCACAGTGTCACGAACAGA	ACCTCGGAGAAGCTGAAACA
cyclin E1	CAGATTGCAGAGCTGTTGGA	TCCCCGTCTCCCTTATAACC
hPRT	GCCCTGGCGTCGTGATTAG	AAGCTTGCGACCTTGACCATC
Primers for qPCR		
CTDSP1 (human)	GCGAGCTCTTTGAATGTGTG	GGCTCAGGTCCTTCACGTAG
CTDSP2 (human)	ATGTTGGCCAGTCAAGTTCC	CTGTCACCTCTGGGAGCAG
CTDSPL (human)	TGCTGAGGGAGGGGAGTGAG	GCAGCATGCCACAGGTTGTC
β-actin (human)	ACTGGAACGGTGAAGGTGAC	AGAGAAGTGGGGTGGCTTTT
GAPDH (human)	GAGTCAACGGATTTGGTCGT	GACAAGCTTCCCGTTCTCAG
CTDSP1 (mouse)	TTTGAATGTGTGCTGTTCACTG	CAGCCTGCTCAGGTCCTTTA
CTDSP2 (mouse)	TGAGTTCCTGAGACGGATGG	CCTGGTGGAACACACAAGC
CTDSPL (mouse)	AGAGGATGGGGCAGCTCT	GGTGGAACACGCAGGATTC
β-actin (mouse)	CCCTGAAGTACCCCATTGAA	CTTTTCACGGTTGGCCTTAG
GAPDH (mouse)	AACTTTGGCATTGTGGAAGG	CACATTGGGGGTAGGAACAC

Supplementary Table 2. Mapping of miRNA and protein families

miRNA family [a]	Protein family ID[b]	Colocalization of miRNA and protein family members	miRNA members not Involved[c]	Protein members not Involved[d]
miR-10	ENSFM00500000269958	hsa-miR-10a -- HOXB3 hsa-miR-10b -- HOXD3		HOXA3
miR-26	ENSFM00500000269811	hsa-miR-26a-1 -- CTDSPL hsa-miR-26b -- CTDSP1 hsa-miR-26a-2 -- CTDSP2		
mR-33	ENSFM00250000001438	hsa-miR-33a -- SREBF2 hsa-miR-33b -- SREBF1		
miR-103	ENSFM00260000050538	hsa-miR-103-2 -- PANK2 hsa-miR-103-1 -- PANK3 hsa-miR-107 -- PANK1		
miR-128	ENSFM00600000921181	hsa-miR-128-1 -- R3HDM1 hsa-miR-128-2 -- ARPP21		R3HDM2
miR-148	ENSFM00500000270533	hsa-miR-152 -- COPZ2 hsa-miR-148b -- COPZ1	hsa-miR-148a	
miR-153	ENSFM00260000050522	hsa-miR-153-1 -- PTPRN hsa-miR-153-2 -- PTPRN2		
miR-204	ENSFM00250000000305	hsa-miR-204 -- TRPM3 hsa-miR-211 -- TRPM1		TRPM6, TRPM7
miR-208	ENSFM00250000000024	hsa-miR-208a -- MYH6 hsa-miR-208b -- MYH7		MYH8, MYH7B, MYH15, MYH4, MYH3, MYH2, MYH1, MYH13
miR-218	ENSFM00250000000421	hsa-miR-218-1 -- SLIT2 hsa-miR-218-2 -- SLIT3		SLIT1

(continued on next page)

Supplementary Table 2. Mapping of miRNA and protein families *(continued)*

miRNA family [a]	Protein family ID[b]	Colocalization of miRNA and protein family members	miRNA members not Involved[c]	Protein members not Involved[d]
miR-506	ENSFM00570000851105	hsa-miR-511-1 -- MRC1L1 hsa-miR-511-2 -- MRC1	hsa-miR-514-1, 512-1, 514-2, 506, 513a-2, 513b, 513c, 509-2, 514-3, 509-1, 513a-1, 507, 509-3, 512-2, 514b, 508, 510	PLA2R1, MRC2
miR-548	ENSFM00250000001060	hsa-miR-548h-3 -- HS3ST3A1 hsa-miR-548w -- HS3ST4	hsa-miR-548t, 548f-2, 548i-4, 548p, 548f-4, 548aa-1, 548i-2, 603, 548aa-2, 548x, 570, 548q, 548c, 548h-4, 548f-1, 548f-5, 548u, 548z, 579, 548l, 548m, 548e, 548j, 548i-1, 548i-3, 548a-1, 548g, 548s, 548a-3, 548v, 548h-2, 548a-2, 548f-3, 548y, 548b, 548d-2, 548k, 548d-1, 548n, 548h-1	HS3ST6, HS3ST2, HS3ST5, HS3ST1, HS3ST1
miR-1233	ENSFM00600000921309	hsa-miR-1233-1 -- GOLGA8A hsa-miR-1233-2 -- GOLGA8B		GOLGA8H, GOLGA6C, GOLGA6D, GOLGA6B, GOLGA8DP, GOLGA8J, GOLGA6A, GOLGA8C, GOLGA8F, GOLGA8G, GOLGA8E, ENSG00000225033, ENSG00000196102, ENSG00000214394, ENSG00000249931, ENSG00000186399, ENSG00000187812, ENSG00000206127

[a] miRNA family was derived from miRBase (release 16.0);

[b] Protein family ID was retrieved from Ensembl Genome Browser (release 60);

[c] Members of a miRNA family that were not colocalized with the gene loci of the protein family;

[d] Members of a protein family that were not colocalized with the gene loci of the miRNA family.

Supplementary Table 3. Enrichment of G1/S-phase transition-related genes among the predicted targets of miR-26a/b

	All genes[a]	G1/S related genes[c]	Fold of enrichment	*P* [e]
All genes[a]	19365	33	3.9	0.0184
Predicted targets[b]	601	4[d]		

[a] All HUGO Gene Nomenclature Committee (HGNC) genes;

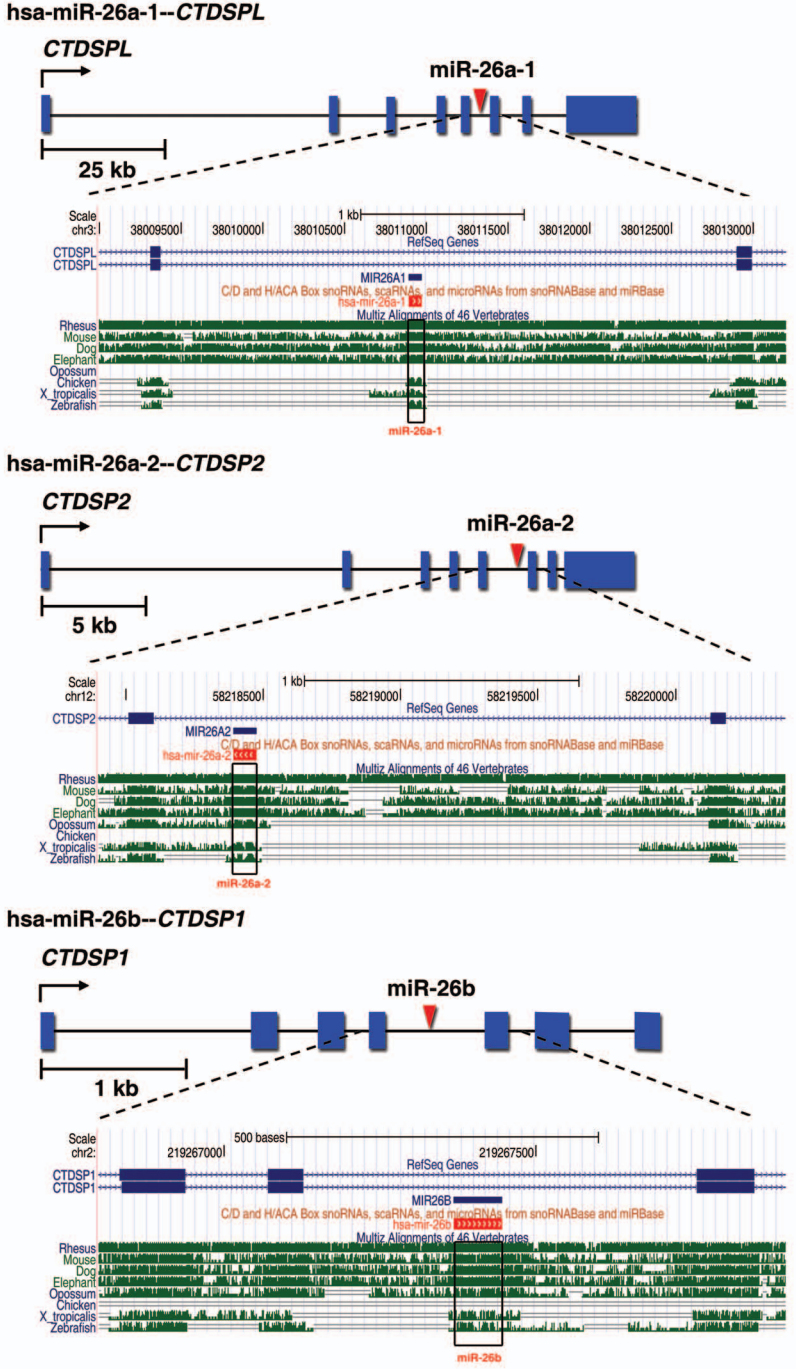
[b] HGNC genes predicted to be targets of miR-26 by TargetScan program;

[c] HGNC genes in the "G1-S Phase Transition" pathway according to the Protein Lounge Pathway database;

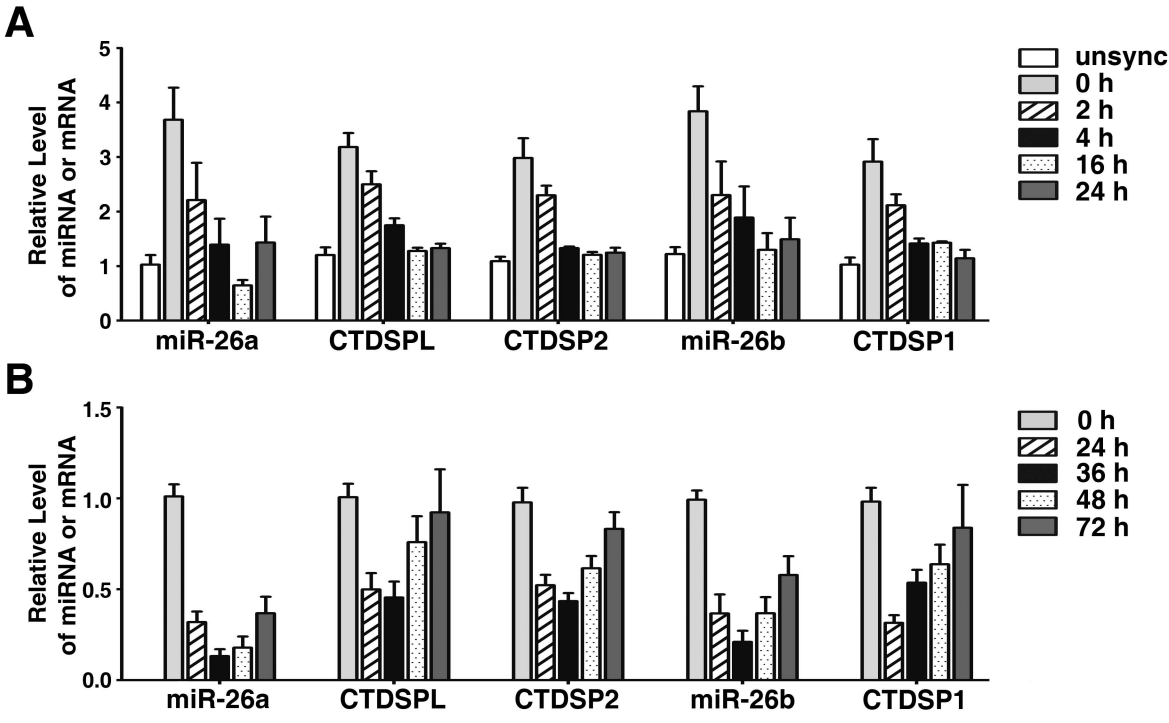
[d] Four predicted targets of miR-26a/b in the "G1-S Phase Transition" pathway are CDK6, cyclin E1, cyclin E2 and cyclin D2;

[e] *P*-value was calculated by hypergeometric test.

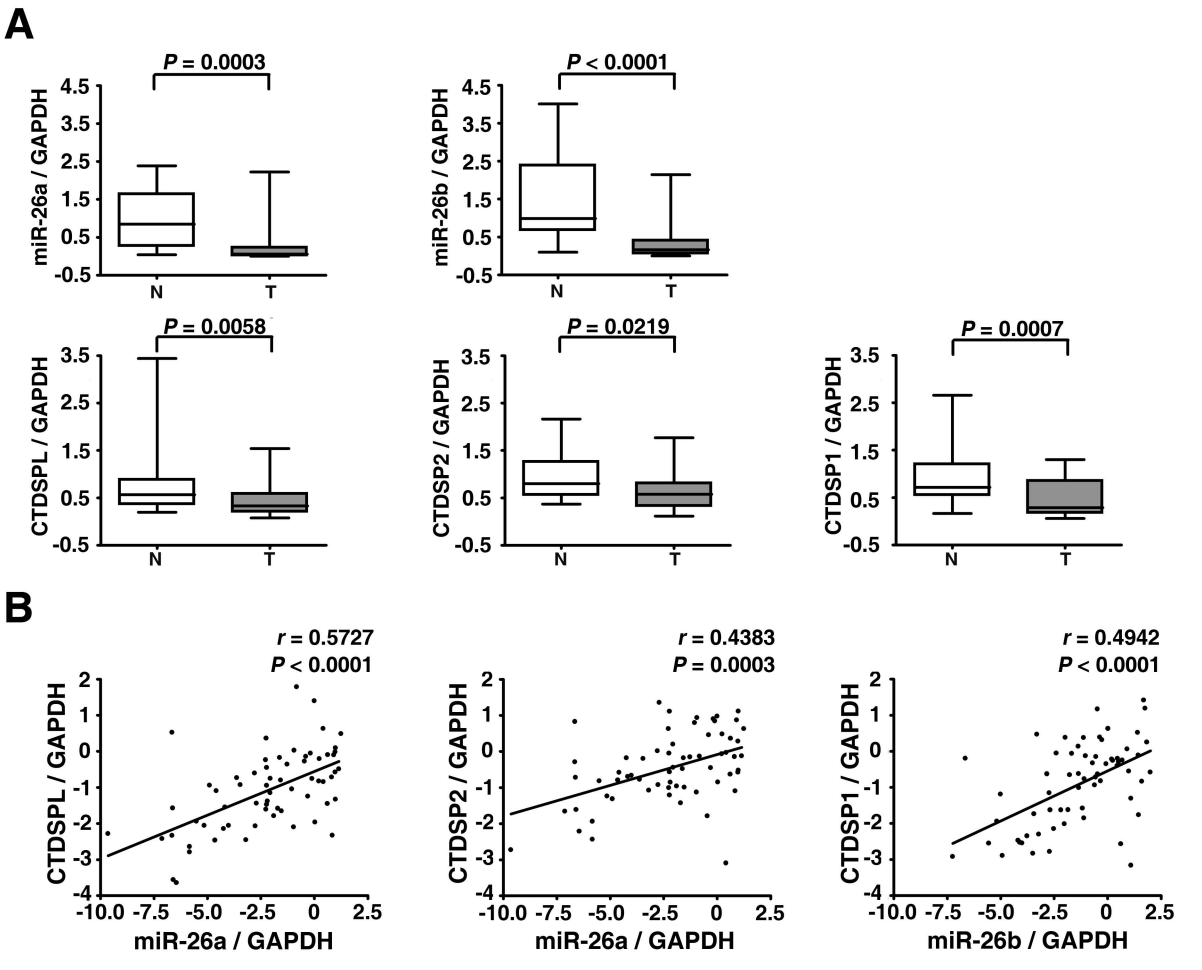
Supplementary Fig. 1



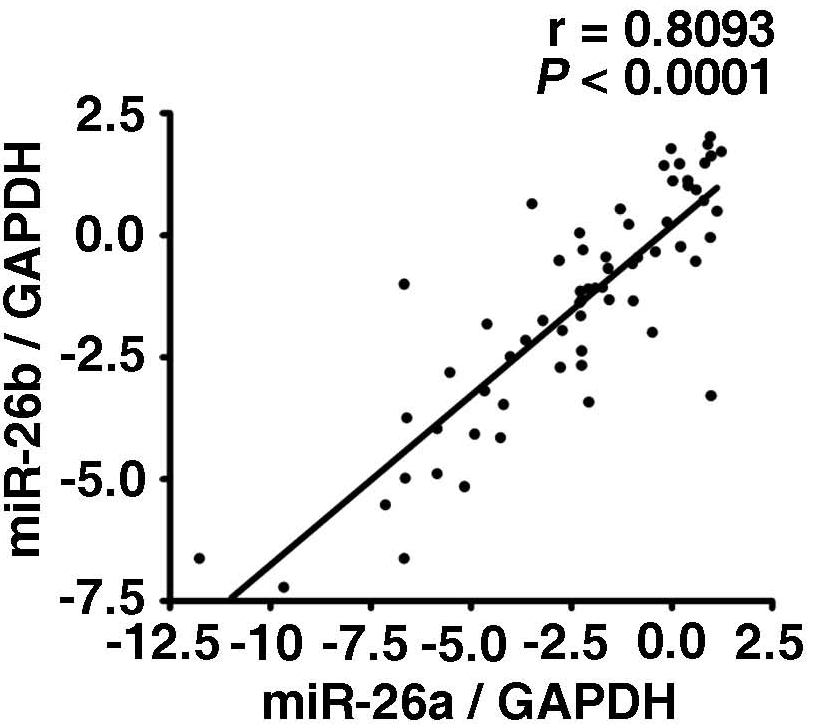
Supplementary Fig. 2



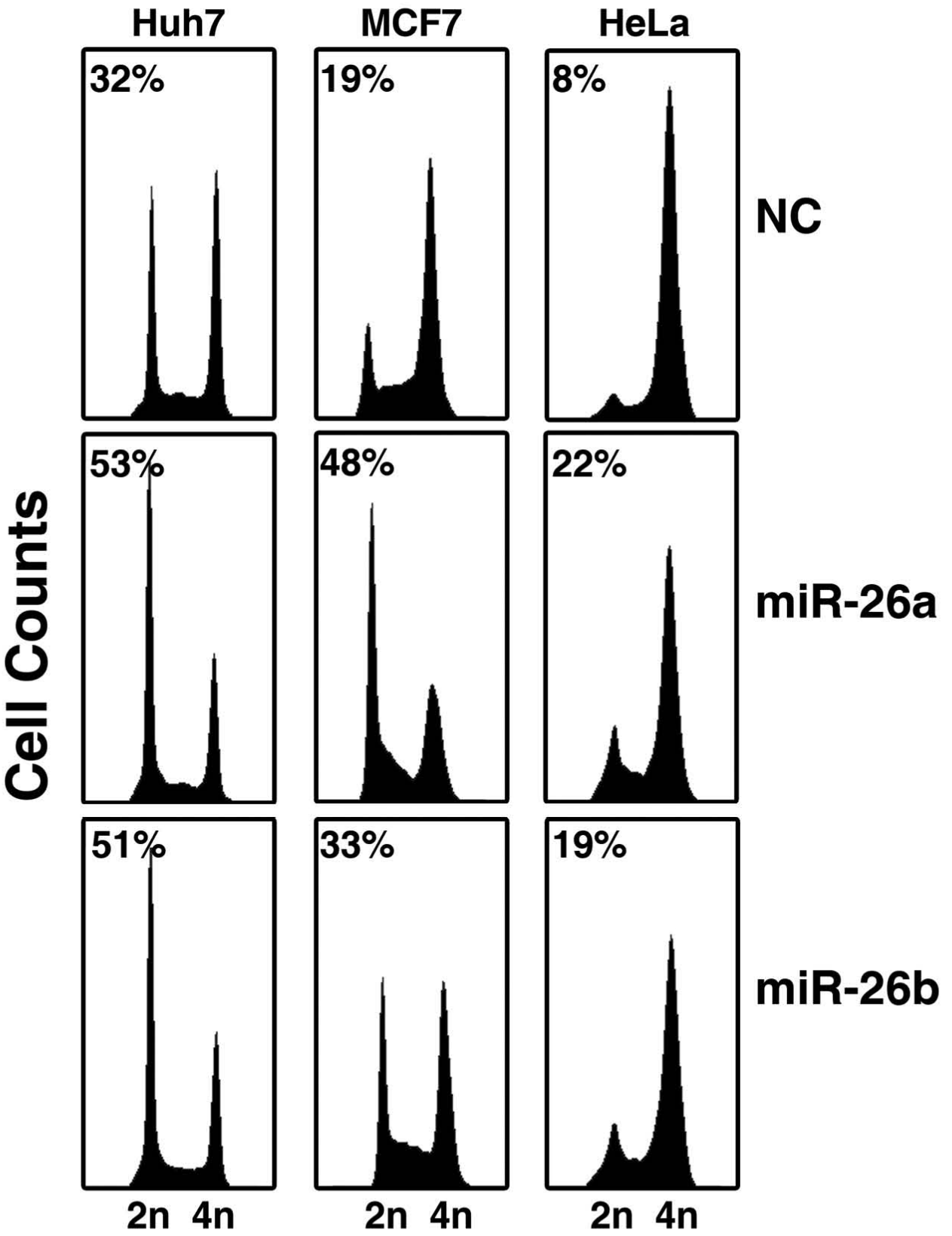
Supplementary Fig. 3



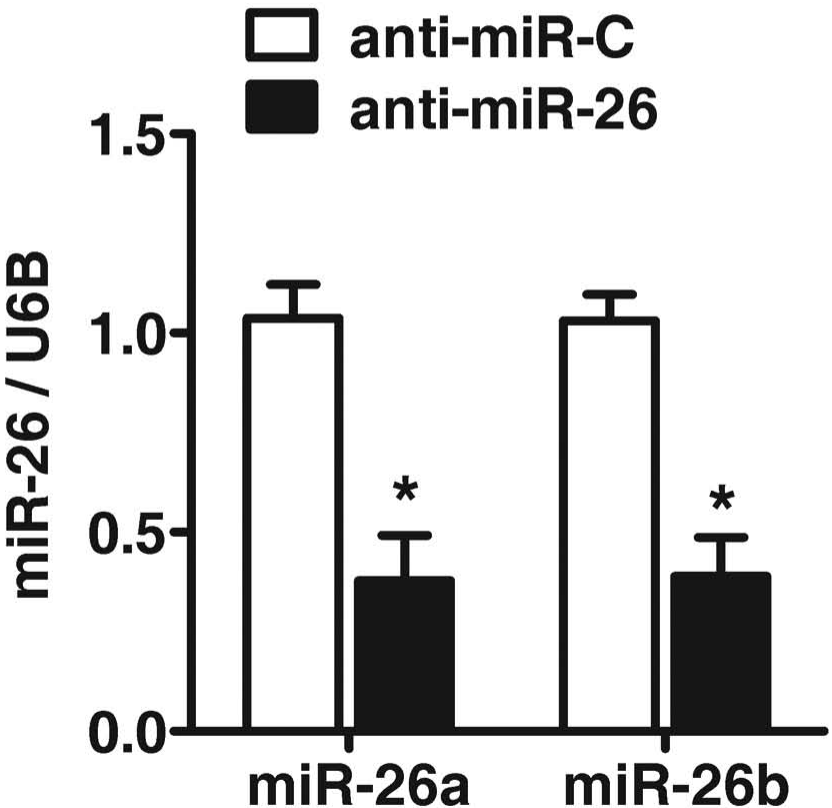
Supplementary Fig. 4



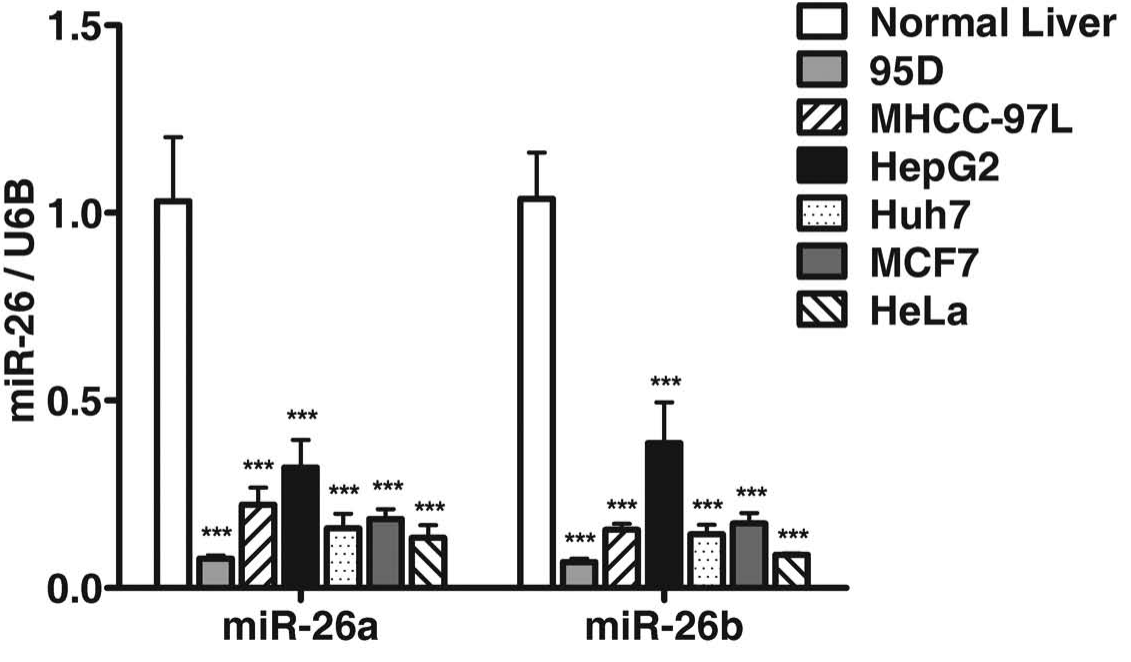
Supplementary Fig. 5



Supplementary Fig. 6



Suplementary Fig. 7



Supplementary Fig. 8

hsa-miR-26a	**5 ' - UUCAAGUAAUCCAGGAUAGGCU - 3 '**
hsa-miR-26b	**5 ' - UUCAAGUAAUUCAGGAUAGGU - 3 '**
CDK6	
binding site 1	**3 ' - AAGUUCAUUCCCAUGUUGCCCAU - 5 '**
binding site 2	**3 ' - AAGUUCAUACUUCUUUAUGUCAU - 5 '**
CCNE1	
binding site	**3 ' - CAGUUCAUCGUGGAAGGUAUCGU - 5 '**
seed mutant	**UUCAAGUA**