3734–3747 Nucleic Acids Research, 2013, Vol. 41, No. 6 Published online 1 February 2013
doi:10.1093/nar/gkt047

Transcribed pseudogene ψ*PPM1K* generates endogenous siRNA to suppress oncogenic cell growth in hepatocellular carcinoma

Wen-Ling Chan[1,2,3], Chung-Yee Yuo[4], Wen-Kuang Yang[5,6], Shih-Ya Hung[3,7], Ya-Sian Chang[3], Chien-Chih Chiu[8], Kun-Tu Yeh[9], Hsien-Da Huang[1,2,*] and Jan-Gowth Chang[3,7,10,11,*]

[1]Institute of Bioinformatics and Systems Biology, National Chiao Tung University, Hsin-Chu 300, Taiwan, [2]Department of Biological Science and Technology, National Chiao Tung University, Hsin-Chu 300, Taiwan, [3]Center of RNA Biology and Clinical Application, China Medical University Hospital, Taichung 404, Taiwan, [4]Department of Biomedical Science and Environmental Biology, Kaohsiung Medical University, Kaohsiung 807, Taiwan, [5]Cell/Gene Therapy Research Laboratory, Department of Medical Research, China Medical University Hospital, Taichung 404, Taiwan, [6]Departments of Biochemistry and Medicine, China Medical University, Taichung 404, Taiwan, [7]Graduate Institute of Integrated Medicine, College of Chinese Medicine, China Medical University, Taichung 404, Taiwan, [8]Department of Biotechnology, Kaohsiung Medical University, Kaohsiung 807, Taiwan, [9]Department of Pathology, Changhua Christian Hospital, Changhua 500, Taiwan, [10]Department of Laboratory Medicine, China Medical University Hospital, Taichung 404, Taiwan and [11]School of Medicine, China Medical Univeristy, Taichung 404, Taiwan

Received June 29, 2012; Revised December 5, 2012; Accepted January 11, 2013

ABSTRACT

Pseudogenes, especially those that are transcribed, may not be mere genomic fossils, but their biological significance remains unclear. Postulating that in the human genome, as in animal models, pseudogenes may function as gene regulators through generation of endo-siRNAs (esiRNAs), antisense RNAs or RNA decoys, we performed bioinformatic and subsequent experimental tests to explore esiRNA-mediated mechanisms of pseudogene involvement in oncogenesis. A genome-wide survey revealed a partial retrotranscript pseudogene ψ*PPM1K* containing inverted repeats capable of folding into hairpin structures that can be processed into two esiRNAs; these esiRNAs potentially target many cellular genes, including *NEK8*. In 41 paired surgical specimens, we found significantly reduced expression of two predicted ψ*PPM1K*-specific esiRNAs, and the cognate gene *PPM1K*, in hepatocellular carcinoma compared with matched non-tumour tissues, whereas the expression of target gene *NEK8* was increased in tumours. Additionally, *NEK8* and *PPM1K* were downregulated in stably transfected ψ*PPM1K*-overexpressing cells, but not in cells transfected with an esiRNA1-deletion mutant of ψ*PPM1K*. Furthermore, expression of *NEK8* in ψ*PPM1K*-transfected cells demonstrated that *NEK8* can counteract the growth inhibitory effects of ψ*PPM1K*. These findings indicate that a transcribed pseudogene can exert tumour-suppressor activity independent of its parental gene by generation of esiRNAs that regulate human cell growth.

INTRODUCTION

Pseudogenes are usually considered to be non-functional copies of protein-coding genes, yet they probably outnumber functional genes in the human genome (1). There are two major classes of pseudogene: one represents processed forms that contain poly-A tails, lack introns and arise through retrotransposition, whereas the other comprises non-processed pseudogenes resulting from gene duplication, which retain exon/intron structure, although occasionally incompletely (2). In apparent contradiction of the assumption that pseudogenes are genomic fossils, genome-wide investigations have recently provided evidence for actively transcribed pseudogenes (TPGs) with potential functional implications (3–10). For instance, the ψNOS transcript acts as a natural antisense

*To whom correspondence should be addressed. Tel: +886 4 22052121, ext. 2008; Fax: +886 4 22033295; Email: d6781@mail.cmuh.org.tw
Correspondence may also be addressed to Hsien-Da Huang. Tel: +886 3 5712121, ext. 56957; Fax: +886 3 5739320; Email: bryan@mail.nctu.edu.tw

regulator of neuronal NOS protein synthesis in snails (11,12), and in mice, reduced expression of *ψmakorin1-p1* because of a transgene insertion caused mRNA instability of its parental gene *Mkrn1*, resulting in polycystic kidneys and bone deformity (10,13), although contradictory results were also reported (14). Additionally, a transcript of *ψPTEN*/*PTENP1*, a highly homologous processed TPG of tumour-suppressor gene *PTEN*, not only interacts with its cognate sequence but also exerts a growth-suppressor role as a decoy by binding to *PTEN*-targeting miRNAs (15). These findings clearly imply that TPGs may play active regulatory roles in cellular functions.

RNA interference (RNAi) is a natural cellular process that defends cells against viruses and transposons, and it also regulates gene expression in a sequence-specific manner (16). Three RNAi pathways can be distinguished on the basis of the biogenesis and functional roles of the classes of small RNA involved, two of which are siRNA, resulting from processing of dsRNA, and miRNA, which derive from shRNA, respectively (17,18). The third category is piRNA: these are ssRNA sequences that interact with piwi protein and seem to be involved in transcriptional gene silencing of retrotransposons and other genetic elements in germ line cells (19). By binding to mRNAs and thereby repressing protein synthesis, miRNAs may regulate cellular development, oncogenesis and apoptosis (20,21). In mice and fruit flies, dsRNAs arising from the antisense/sense transcripts of processed TPGs and their cognates, or hairpin structures resulting from inversion and duplication, are cut by Dicer into 21-nt siRNA that can bind RISC and regulate the expression of the parental gene (17,22–26). In humans, however, the mechanism of generation of such siRNAs remains unclear. Accordingly, the specific aim of this study is the verification of our hypothesis that transcribed pseudogenes in humans can be processed into endogenous siRNAs (endo-siRNAs or esiRNAs) to regulate protein-coding genes.

Hepatocellular carcinoma (HCC) is one of the most common human cancers worldwide, particularly in Asia and Africa (27). In Taiwan, HCC is the leading cause of cancer deaths, with ~8000 new cases diagnosed and 7000 deaths occurring annually (27). Therefore, we are interested in exploring the relationship between TPGs and HCC. The present study incorporated both bioinformatic and experimental approaches to screen >20 000 human pseudogenes, and we found *in silico* 448 TPGs that might regulate protein-coding genes through derivation of esiRNAs. In particular, we focus on *ψPPM1K*, a partial retrotranscript from *PPM1K* (protein phosphatase, Mg^{2+}/Mn^{2+} dependent, 1K). This TPG contains distinct sequences with inverted repeats capable of folding into a hairpin structure for processing into two esiRNAs that may target many cellular genes. To our knowledge, this is the first investigation of an esiRNA-mediated role of human pseudogenes in HCC.

MATERIALS AND METHODS

Data generation

In total, >20 000 human pseudogenes and their cognate genes were obtained from the Ensembl database (Ensembl 63, GRCH37) using BioMart (http://www.ensembl.org/index.html). Functional small RNAs (fsRNAs) with sequence length between 18 and 40 nt were collected from the Functional RNA Database (fRNAdb) (28), which hosts a large collection of known/predicted non-coding RNA sequences from public databases: H-invDB v5.0 (10), FANTOM3 (29), miRBase 17.0 (30), NONCODE v1.0 (31), Rfam v8.1 (32), RNAdb v2.0 (33) and snoRNA-LBME-db rel. 3 (34). Genomic sequences were collected from UCSC hg19 (http://hgdownload.cse.ucsc.edu/downloads.html).

Bioinformatics methods for identifying pseudogene-derived esiRNA–target interactions

Figure 1 depicts the workflow for identifying pseudogene-derived esiRNA–target interactions (eSTIs).

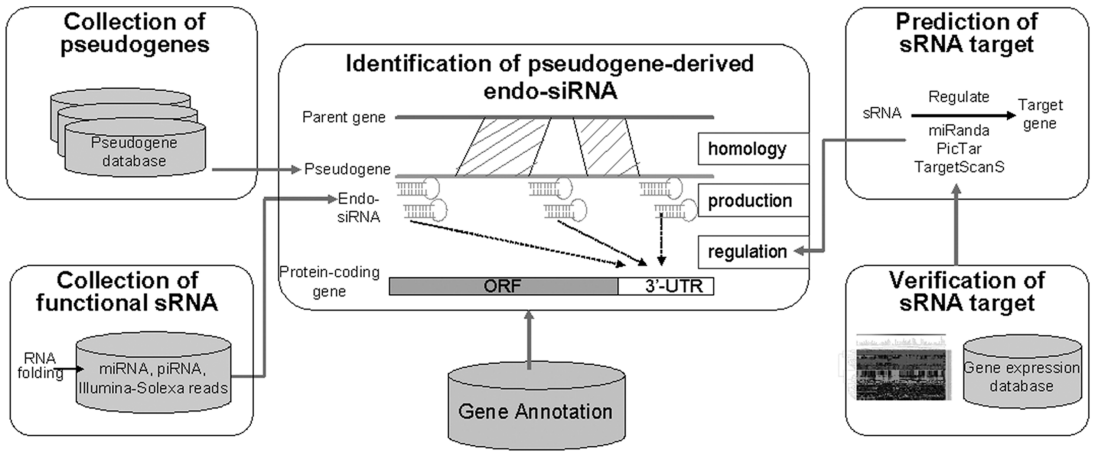

Figure 1. Workflow for identification of pseudogene-derived esiRNA–target interactions. Using a systematic computational procedure of homologous sequence alignment between a collection of transcribed pseudogenes and known functional sRNAs, we identified pseudogene-derived esiRNAs and verified these by reference to available Illumina-Solexa reads, and subsequently by reference to regulated protein-coding target genes (see 'Materials and Methods' section).

After collection of pseudogenes, protein-coding genes and fsRNAs and the pseudogene-specific esiRNAs were examined by aligning the pseudogenes with fsRNAs, excluding alignments with parental genes. Candidate pseudogene-specific esiRNAs were validated by reference to publicly available deep sequencing data from various sRNA libraries. Additionally, eSTIs were analysed by three target prediction tools and verified with gene expression profiles. Detailed procedures are described later in the text.

Identification of pseudogene-derived esiRNAs

To predict candidate pseudogene-derived esiRNAs, we aligned the sequences of pseudogenes and fsRNAs, excluding parental gene alignments. Deep sequencing data of sRNA libraries derived from human embryo stem cells or HCC/liver tissues were used to verify these candidates (35–37). Then, the extended sequences of these candidate esiRNAs were used to predict hairpin structure by Mfold (38). Details of publicly available deep sequencing data are shown in Supplementary Table S1.

Identification of eSTIs

Based on experimentally supported data sets, Sethupathy *et al*. (27) and Baek *et al*. (30) have shown that the intersection of miRNA target prediction tools can yield improved specificity with only a marginal decrease in sensitivity relative to any individual algorithm. We modified our previous approach (39) for identifying pseudogene-derived esiRNA targets. Briefly, three previously developed computational approaches, TargetScan (40–42), miRanda (43) and RNAhybrid (44), were used to identify esiRNA target sites within the conserved regions of the 3′-UTR of genes in 12 metazoan genomes. The minimum free energy (MFE) threshold was −20 kcal/mol with score ≥150 for miRanda; default parameters were used for TargetScan and RNAhybrid. The three criteria for identifying targets were (i) potential target sites must be predicted by at least two tools; (ii) hits with multiple target sites are prioritized; and (iii) target sites must be located in accessible regions. Finally, three gene expression profiles were obtained from NCBI GEO (45) to verify those eSTIs with pseudogene expression higher than their target genes. Gene expression profiles included GDS596 (46), GSE5364 (47) and GSE6222 (48); detailed experimental conditions are described in Supplementary Table S1. The Pearson correlation coefficient was computed for pseudogenes and their target genes.

Prediction of miRNA–target interactions

Potential miRNA–target interactions (MTI) with pseudogenes and parental genes were investigated as described previously (39). Sequences of miRNAs were obtained from miRBase R18 (30).

GO and KEGG enrichment analyses

The function of target genes was examined by performing GO and KEGG pathway enrichment annotation (49) using the DAVID gene annotation scheme (50).

Samples

Resected primary HCC and nearby non-cancerous tissue samples ($n = 41$) were obtained from 41 patients at the Changhua Christian Hospital. The tumour tissues were composed of 90–100% tumour cells and were frozen immediately after surgical resection, then stored in liquid nitrogen until extraction of either RNA or DNA. All studies were approved by the Institutional Review Board of Changhua Christian Hospital.

Cell culture

Human hepatoma Huh-7 and HepG2 cells were grown using standard procedures for all experiments. Cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% FBS, 2 mM glutamine and antibiotics (100 U/ml of penicillin and 100 μg/ml of streptomycin) at 37°C in a humidified atmosphere of 5% CO_2 incubator.

RNA isolation, reverse transcription and real-time quantitative polymerase chain reaction analysis

RNA isolation from specimens or cultured cells and reverse transcription were performed as described (51,52). Reverse-transcriptase quantitative polymerase chain reaction (RT–qPCR) analysis of *PPM1K* and *ψPPM1K* in HepG2 and Huh-7 cells, and of *PPM1K*, *NEK8*, *TBRG1* and *BMPR2* in *ψPPM1K*-expressing Huh-7/HepG2 stable cell lines, was performed using SYBR Green with the ABI 7500 Real-Time PCR System (Applied Biosystems). RT–qPCR of precursor esiRNA1 (24–144 nt), precursor esiRNA2 (170–273 nt), *PPM1K* and *NEK8* in paired HCC tumour and non-tumour tissues was performed using a LightCycle 480 (Roche, Mannheim, Germany) with a primer/probe system. The specific primer/probe sets are shown in Supplementary Table S2. All RNA expression levels were normalized to *GAPDH* (glyceraldehyde-3-phosphate dehydrogenase) RNA with the ΔΔCt method according to Liu *et al*. (53).

RT–PCR of mature esiRNA1 levels in Huh-7 stable cell lines was performed using a TaqMan MicroRNA Assay designed for esiRNA1 according to the manufacturer's instructions (Applied Biosystems) following isolation of small RNA with the mirVana miRNA Isolation Kit. U6 small nuclear RNA was used as an internal control.

Northern blot of pseudogene-derived esiRNAs

Northern blotting was performed according to a previous study (54) with minor modifications. Briefly, 10 μg of total RNA from human hepatoma cell line Huh-7 was dissolved in loading buffer (50 mM ethylenediaminetetraacetic acid, 8 M urea, 20% formamide and xylene cyanol), loaded onto a 2% agarose gel and then run for 1.5 h at 120 V at room temperature. The biotin-labelled esiRNA probes (5′-GTGGCACGCGCCTGTAGTCCCAGC-3′

for esiRNA1 and 5′-GAGGCAGGAGAATGGCGTGA ACC-3′ for esiRNA2, Genomics BioSci & Tech Co., Taipei, Taiwan) were used as the positive control for the avidin–biotin reaction and the size control for esiRNAs. The agarose gel was incubated sequentially in 0.05 M NaOH/NaCl, 0.05 M Tris/NaCl and 2× sodium citrate. Then RNA was transferred to a nitrocellulose membrane (Pall Corporation, East Hills, NY, USA) followed by cross-linking with 254-nm ultraviolet radiation. The membrane was hybridized with the biotinylated esiRNAs overnight, and then membranes were washed sequentially with 2× SSC/0.1% sodium dodecyl sulphate (SDS), 1× SSC/0.1% SDS and 0.5× SSC/0.1% SDS at 42°C. The membrane was incubated with horseradish peroxidase-conjugated avidin (Biolegend, SanDiego, CA, USA) and probe detected by chemiluminescence with the WesternBrightTM-ECL kit (Advansta, Menlo Park, CA, USA).

Fluorescent *in situ* hybridization

Nocodazole and colchicine were added to cell lines before fluorescent *in situ* hybridization FISH was performed. Interphase and metaphase spreads were prepared for FISH using standard methods (55). DNA probes (S1: G TGGCACGCGCCTGTAGTCCCAGC, antisense of esiRNA1; S2: GAGGCAGGAGAATGGCGTGAACC, antisense of esiRNA2; scramble 1: GTGGCTCATGCC TGTAATCCCAGCACTTTG; and scramble 2: TTAAG ACATACAAAGATCTGGCCAGGTGCG) were mixed with hybridization buffer, centrifuged and heated to 73°C for 5 min in a water bath. Slides were immersed in 70% formamide/2× standard saline citrate for 5 min at 73°C, followed by dehydration, dried and hybridized with probe mix in a 42°C incubator for 30 min. Slides were then washed in 0.4× standard saline citrate/0.3% NP-40 for 2 min, air dried in the dark and counterstained with DAPI (4,6-diamidino-2-phenylindole) (1 μg/ml, Abbott, IL, USA). Imaging was performed on a Nikon E600 microscope with cytovision software.

Transduction of the pseudogene transcript in Huh-7 and HepG2 stable cell lines

TPG-expressing (and vector control) Huh-7/HepG2 stable cell lines were established by G418 selection after transfection with the *ψPPM1K*-expression plasmid or blank vector. Total RNA was isolated from cells and subjected to RT–PCR analysis to amplify *ψPPM1K* mRNA (primers shown in Supplementary Table S2). The PCR was performed with a denaturing step at 95°C for 2 min, then 30 cycles of 30 s at 95°C, 1 min at 60°C and 1 min at 72°C, followed by a final 7 min at 72°C.

Cell proliferation assay

To investigate the proliferation of *ψPPM1K*-expressing Huh-7 stable cell lines, 2.5×10^4 cells were plated in each well of a 12-well plate. Cells were trypsinized and counted with a haemocytometer every day until Day 4. Each experiment was repeated twice in triplicate wells, separately. Huh-7 TPG7 cells were transfected with *NEK8*-overexpressing plasmid or empty vector using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA). The *NEK8*-overexpressing plasmid, which contains the human *NEK8* ORF without 3′-UTR in the pCMV6-Entry vector, was obtained from Origene (Rockville, MD, USA). Twenty-four hours after transfection, 2.5×10^4 cells were plated in each well of a 12-well plate. Cells were trypsinized and counted with a haemocytometer every day until Day 4.

Clonogenic activity

For determination of clonogenic activity, we plated 1000 cells of mock2, TPG1, TPG2 and TPG7 in 10 ml of growth medium in 100-mm dishes. Soft agar culture was also performed by inoculating 500 cells/ml in 0.3% agar growth medium over 0.5% agar growth medium in 6-well culture dishes. The dishes were incubated under normoxic 19% O_2 and hypoxic 3% O_2 in 5% CO_2 incubators for 12 days and fixed/stained for counting colony formation. We also took serial photographs of the same colonies at Day 5, Day 7 and Day 9 to visualize the 2D growth of mock2 and three transfected clones.

Transfection of synthetic siRNA1 into Huh-7 cells

siRNA1 was chemically synthesized by Invitrogen (Carlsbad, CA, USA). Oligonucleotides were annealed before use in annealing buffer containing 100 mM potassium acetate, 30 mM HEPES–KOH (pH 7.4) and 2 mM magnesium acetate. Negative Control #1 siRNA was also obtained from Invitrogen. Huh-7 cells in 6-cm culture plates were transfected with 200 pmol siRNA using 10 μl of Lipofectamine 2000 according to manufacturer's instructions.

Construction of the esiRNA1-deleted *ψPPM1K*-expressing plasmid

To delete the esiRNA1 sequence from *ψPPM1K*, the overlap extension method of PCR-based mutagenesis was used. First, two complementary mutagenic primers, forward 5′-CCTCAGCCTCCTGAGTACACCCCTGGC TAATTTT-3′ and reverse 5′-AAAATTAGCCAGGGGT GTACTCAGGAGGCTGAGG-3′, were synthesized. Two PCRs using the mutagenic forward primer/outer *ψPPM1K* reverse primer pair and the mutagenic reverse primer/outer *ψPPM1K* forward primer pair were performed to amplify the right and left *ψPPM1K* fragments, respectively. The two fragments were then mixed and further amplified using the outer *ψPPM1K* primers to generate the esiRNA1-deleted *ψPPM1K* fragment. Finally, this fragment was inserted between the EcoRI and XbaI sites of the pCI-neo vector to generate the esiRNA1-deleted *ψPPM1K*-expressing plasmid.

Mitochondrial activities

For indirect assay of mitochondrial membrane potential and permeability transition pore activity (56), overnight-plated monolayer cells on 100-mm dishes were exposed to 0.5 g/ml or 1.0 g/ml of rhodamine 123 (Rh123) in the growth medium (57). The kinetics of dye uptake were determined by harvesting the cells after 10 min,

30 min, 5 h and 24 h incubation in Rh123-containing media. To determine Rh123 retention activity of cancer cells (56–58), monolayer cells exposed to Rh123-containing medium for 30 min were rinsed 3× with Hanks' balanced salt solution to remove the dye and replenished with fresh medium for further 18 h incubation before harvest. Cells harvested by trypsinization were washed 2× with cold phosphate-buffered saline and collected by centrifugation at 300*g* at 4°C for 5 min. With or without further reaction with fluorescent monoclonal antibody against cell surface markers, e.g. CD133/Prominin (BD Pharmingen), the doubly or triply labelled washed cells were analysed in a fluorocytometer.

miRNA-mediated knockdown of *PPM1K* and *ψPPM1K*

The stable negative control (siCon: 5′FAM-UUC UCC GAA CGU GUC ACG UTT), has-miR-650 (miR-650: 5′-AGG AGG CAG CGC UCU CAG GAC) and has-miR-3174 (miR-3174: 5′-UAG UGA GUU AGA GAU GCA GAG CC-3′) miRNAs were purchased from GeneDireX Inc. (Las Vegas, NV, USA) and transfected into cells by Lipofetamine™ RNAiMax (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. The efficacy of mRNA knockdown after miRNA transfection for 48 h was determined by RT–qPCR.

Statistical analysis

Student's *t*-test was used for analysis of the cell assays. Significance was accepted at $P < 0.05$.

RESULTS

Overview of TPGs

To verify the hypothesis that human pseudogenes may generate esiRNA to regulate protein-coding genes, we developed a computational pipeline (Figure 1). A total of 16 524 genes including 15 003 pseudogenes and 1521 processed transcripts were collected from BioMart and integrated in the Ensembl Genome Browser, filtering by gene type as 'pseudogene' or 'pseudogene-related gene'. The percentage of pseudogenes on each chromosome is similar to that of protein-coding genes (Supplementary Figure S1a). Of these pseudogenes, 1337 are detectable by the Affymetrix Human Genome U133A/U133Plus2 microarray, and they are thus considered to be TPGs. Among them, 262 TPGs overlapped with its parental gene for detecting the expression profiles are discarded. Consequently, we determined the gene expression profiles (GDS596) by $\log_2$ average intensity of TPGs in 79 healthy human tissues. Hierarchical clustering showed that most TPGs were highly expressed in 19 tissues, especially in blood-related cells where expression levels were ≥4-fold higher than average (Supplementary Figure S1b). Another assessment of gene expression in six different tissues (GSE5364) showed that TPGs were often expressed differently in tumour material compared with paired non-tumour tissue (Supplementary Figure S1c). In particular, most TPGs in HCC samples were underexpressed in comparison with healthy tissue (GSE6222) (Supplementary Figure S1d). Furthermore, serial computational analysis showed that 448 TPGs may generate esiRNAs (Supplementary Table S3). One TPG identified as a strong candidate for functional esiRNA formation, *ψPPM1K*, was the focus of further detailed studies.

ψPPM1K and its cognate gene

PPM1K (NM_152542.3), located on chromosome 4q22.1, produces a 3725-nt mRNA encoding a mitochondrial matrix serine/threonine protein phosphatase shown to regulate the membrane permeability transition pore (MPTP) essential for cell survival and organ development (59). Pseudogene *PPM1K (ψPPM1K)* (Supplementary File S1), 474 bp in length, is partially retrotranscribed from *PPM1K* and is also located on chromosome 4 (Figure 2a). Transcription of *ψPPM1K* is supported by mRNA and EST evidence (Supplementary Figure S2), and bases 155–456 of the *ψPPM1K* transcript show >79% similarity to the antisense-strand of *PPM1K* (Figure 2b). Multiple alignments of the 5′- and 3′-ends of *ψPPM1K* DNA sequences in 44 vertebrate species revealed a high degree of conservation among rhesus, mouse, dog and elephant. The *ψPPM1K* sequence also contains SINE and Alu repeat elements (Supplementary Figure S2).

Determination of *ψPPM1K*-derived esiRNAs and their candidate target genes

To predict *ψPPM1K*-derived esiRNAs, we aligned the sequences of fsRNAs with length 18–40 nt, obtained from fRNAdb, to the *ψPPM1K* sequence and found 76 positive examples that also matched publicly available deep sequencing data from various sRNA libraries. Positive examples can be grouped into three blocks, indicating precursor esiRNA1 as being located in the 24–144-nt region (depicted green), precursor esiRNA2 in the 170–273-nt region (depicted red) and precursor esiRNA3 in the 328–455-nt region (depicted pink), respectively (Figure 3a). The sequences of precursor esiRNAs are shown in Figure 3b, and their respective mature esiRNAs are given in colour. We were mostly interested in the *ψPPM1K*-unique esiRNAs, esiRNA1 and esiRNA2; therefore, we did not study precursor esiRNA3 further, as it is also present in the cognate gene *PPM1K* (Figure 3c). Figure 3d shows that the precursor of esiRNA1 (esiRNA1 depicted in green) folds into a hairpin structure with a MFE value of −40.4 by Mfold (38). Subsequently, we performed a genome-wide search for esiRNA1-target interactions and obtained 57 candidate target genes with scores ≥200 (Supplementary Table S4). These targets could be classified into functional categories involving nucleotide binding, and transferase and serine/threonine protein kinase enzyme activities (Supplementary Figure S3a–c). Analyses of transcription factor binding sites in these target genes revealed that they shared motifs for RSRFC4, GFI1, NF1, GATA1 and ELK1 (Supplementary Figure S3d). As illustrated in Figure 3e, the predicted *ψPPM1K*-specific esiRNAs are expected to regulate cognate gene *PPM1K* and target gene *NEK8* (never in mitosis A-related kinase 8) through

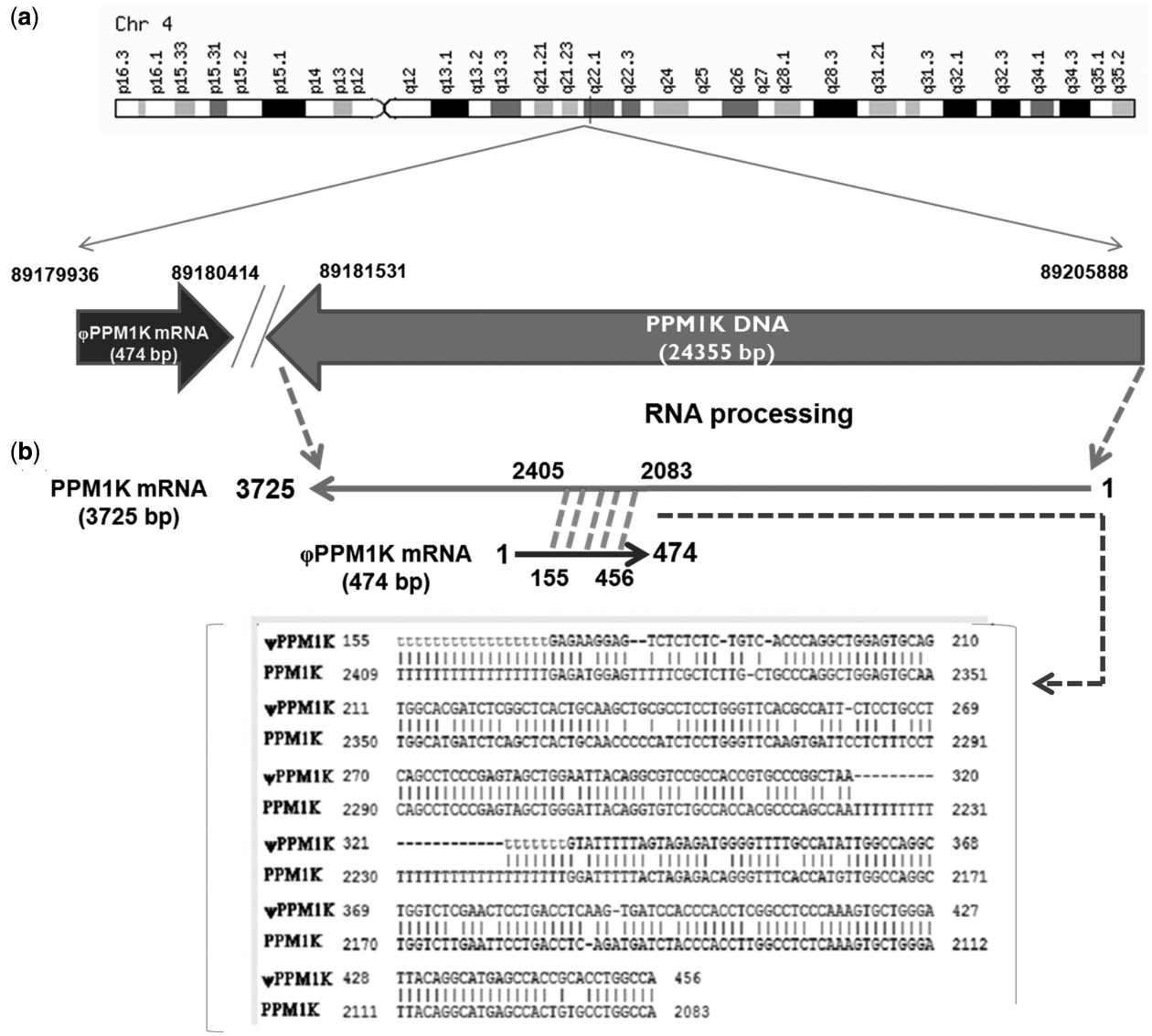

Figure 2. Schematic representation of $\psi PPM1K$ and its parental gene, *PPM1K*. (**a**) $\psi PPM1K$ is located on chromosome 4 proximal to *PPM1K*. (**b**) Alignment of $\psi PPM1K$ and its cognate gene *PPM1K* indicates that $\psi PPM1K$ is a reversed-transcribed partial antisense copy of its parental gene, *PPM1K*, with >79% sequence homology in the 155–456-nt region.

association with multiple target sites. We also observed that $\psi PPM1K$-specific esiRNAs not only complemented the 3′-UTR but also the coding region of *NEK8* (Supplementary Table S5). Additionally, gene expression profiles (GSE6222) also revealed that $\psi PPM1K$ is expressed at lower levels than *NEK8* and *PPM1K* in HCC tissues and cells in comparison with healthy tissues (Figure 3f).

Expression of *PPM1K*, precursors of $\psi PPM1K$-derived esiRNAs and *NEK8* in HCC patient samples

To compare the expression of *PPM1K*, $\psi PPM1K$-derived esiRNAs and target gene, *NEK8*, in clinical samples, we examined 41 HCC surgical specimens together with matched non-tumour tissues by RT–qPCR. Figure 4a shows that the HCC tumour tissues expressed significantly lower levels of both *PPM1K* (0.76-fold; $P = 0.001$) and precursors of $\psi PPM1K$-derived esiRNA1 (24–144 nt; 0.59-fold; $P = 0.007$) and esiRNA2 (170–273 nt; 0.61-fold; $P = 0.007$), than paired non-tumour tissues, whereas expression of the target gene *NEK8* was elevated (1.29-fold) in the tumour samples. These results are consistent with those described earlier in the text (Figure 3f).

Decreased cell growth and clonogenic activity in $\psi PPM1K$-transduced Huh-7 cells

RT–qPCR analyses (Figure 4b) showed that *PPMIK* mRNA levels are higher in Huh-7 than in HepG2 cells, and that the expression of $\psi PPM1K$ mRNA is higher than that of *PPMIK* in both cell lines ($n = 7$). To examine the effect of overexpression of $\psi PPM1K$, Huh-7 and HepG2 cells were transfected with $\psi PPM1K$-harbouring recombinant plasmid vector, and four stable

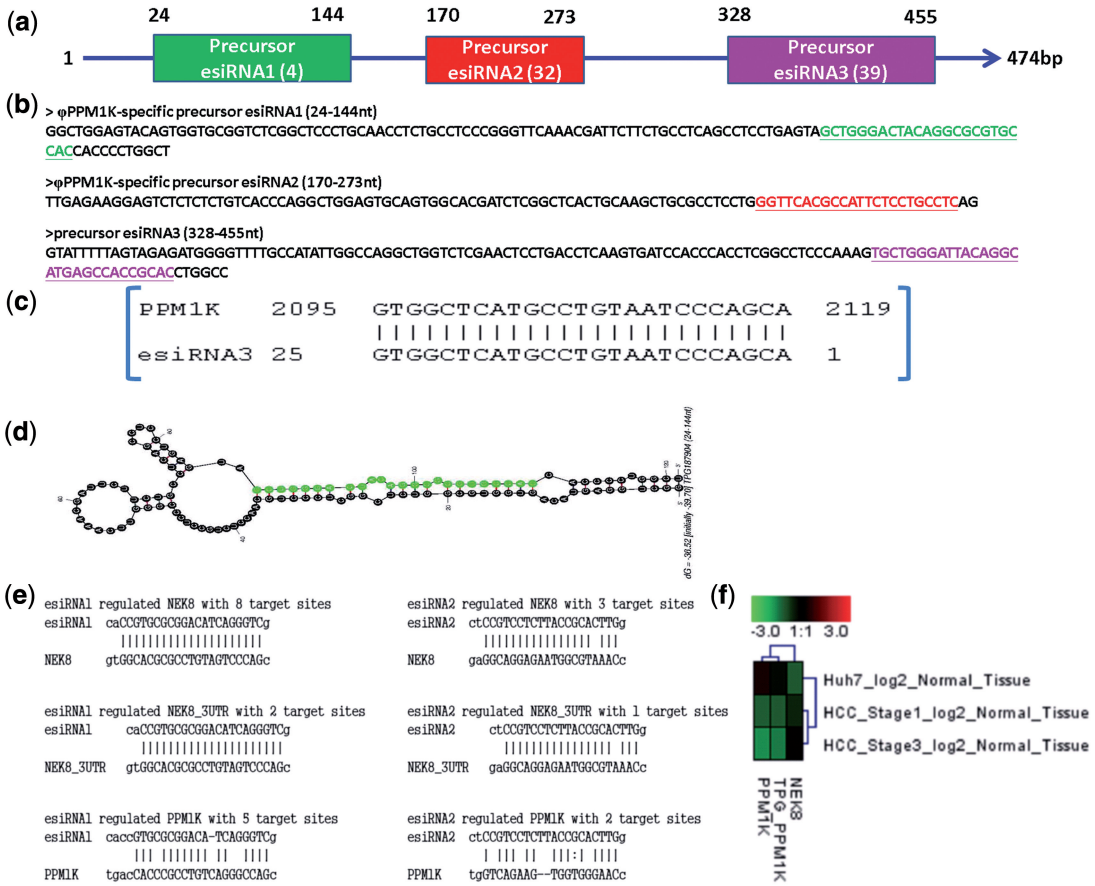

Figure 3. Candidate *ψPPM1K*-derived esiRNAs and their targets. (**a**) Location and read counts of transcribed *ψPPM1K* RNA from sRNA deep sequencing data. The candidate precursor esiRNAs derived from the *ψPPM1K* transcript are shown by green (24–144 nt) and red rectangles (170–273 nt) with read counts 4 and 32, respectively. The other putative esiRNA in the 328–455-nt region (pink rectangle) with 39 read counts was excluded from further analysis, as it is located within the protein-coding gene region that shows a high degree of similarity between *ψPPM1K* and *PPM1K*. (**b**) Sequences of precursor esiRNAs, with esiRNAs shown in colour. (**c**) The alignment of esiRNA3 mapping to *PPM1K* gene. (**d**) Mfold was used to predict the hairpin structure of precursor esiRNA1 with MFE −40.4. The predicted mature esiRNA1 sequence is depicted in green. (**e**) Matches of esiRNA1 and esiRNA2 sequences with target gene *NEK8* and parental gene *PPM1K*. Canonical pairings, solid lines; non-canonical pairings (G:U), dotted lines. (**f**) Expression profiles of *ψPPM1K*, *PPM1K* and *NEK8* in HCC tissues/cells (GSE6222).

transfectant clones, TPG1, TPG2, TPG7 (for Huh-7) and TPG (for HepG2) were isolated; a control clone transfected with empty vector was also produced (Figure 5a). Because of the effects of *NEK8*/*NEK9* on cell cycle regulation (60,61) and their relatively high expression in human breast cancer (61), we examined cell growth behaviour of *ψPPM1K*-transfected versus control Huh-7 cells. A cell proliferation assay showed that *ψPPM1K*-overexpressing cell lines have a slower proliferation rate than the vector control cells (TPG7 cells: $P = 0.036$ for Day 2, $P = 0.018$ for Day 4; TPG1 cells: $P = 0.045$ for Day 4) (Figure 5b). Rates of colony expansion after plating also indicated slower proliferation of the three transfected Huh-7 TPG cell clones in comparison with the control clone (Figure 5c). Furthermore, assessment of colony formation in soft agar suspension cultures under normoxic and hypoxic conditions (Figure 5d) showed that the three stably transfected *ψPPM1K*-expressing cell clones all consistently exhibited less clonogenic activity than the vector control (TPG1cells: $P = 1.54E\text{-}06$, TPG2 cells: $P = 0.0001$ and TPG7 cells: $P = 0.013$ for normoxic conditions; TPG1cells: $P = 2.22E\text{-}05$, TPG2 cells: $P = 5.71E\text{-}05$ and TPG7 cells: $P = 0.0011$ for hypoxic conditions). These effects on cell proliferation suggest that *ψPPM1K* might act as a tumour suppressor in Huh-7 cells.

Expression of *NEK8* in *ψPPM1K*-transfected HCC cells

The *in silico* results indicated that *ψPPM1K*-derived esiRNAs may regulate protein-coding genes, such as *NEK8*, *TBRG1* and *BMPR2* (Supplementary Table S4). Respectively, *NEK8* encodes a member of the serine/threonine protein kinase family related to NIMA of *Aspergillus nidulans* and is overexpressed in human breast tumour (61); *TBRG1* acts as a growth inhibitor that activates *TP53* to cause G_1 arrest and collaborates with *CDKN2A* to restrict proliferation, but it does not require either protein to inhibit DNA synthesis (62,63); whereas *BMPR2* encodes a member of the bone morphogenic protein receptor family of transmembrane serine/threonine kinases (64). To examine the effect of overexpression of *ψPPM1K* in Huh-7 and HepG2 stable transfectant clones, the mRNA levels of *NEK8*, *TBRG1*

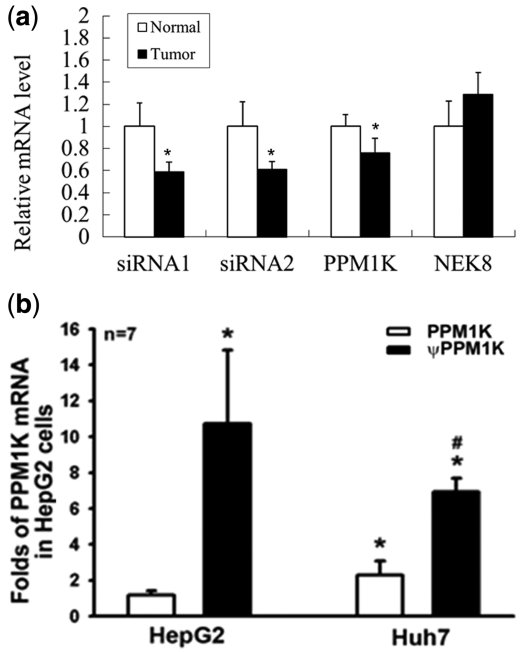

Figure 4. Expression patterns of HCC tissues and cell lines. (**a**) RT–qPCR results showing relative expression levels of two esiRNA precursors, *PPM1K* and *NEK8*, in HCC tumour and paired non-tumour tissues from 41 HCC patients ($^{*}P \leq 0.01$). (**b**) RNA levels of *PPM1K* and *ψPPM1K* in HepG2 and Huh-7 cells. RT–qPCR results show that *PPMIK* expression was higher in Huh7 than in HepG2 cells. The expression of *ψPPM1K* was higher than *PPMIK* in both cell lines ($n = 7$). $^{*}P < 0.05$ compared with *PPM1K* in HepG2; $^{\#}P < 0.05$ compared with *PPM1K* in Huh-7.

and *BMPR2* were determined by RT–qPCR. The results showed that *NEK8* expression was reduced in all *ψPPM1K*-expressing HCC cell lines relative to the vector control cells ($P = 0.033$ in TPG7 Huh-7 cells; $P = 0.047$ in TPG HepG2 cells) (Figure 6a and b). Thus, *ψPPM1K* downregulated *NEK8* in both Huh-7 and HepG2 cell lines.

Expression of *ψPPM1K*-derived esiRNAs

Before verifying downregulation of *NEK8* by *ψPPM1K*-derived esiRNAs, we first demonstrated esiRNA expression in the Huh-7 cell line and TPG-expressing cell clones derived from it. The northern blots showed that both of esiRNA1 and esiRNAs2 were expressed in Huh-7 cells, and that the expression of esiRNA1 is higher than that of esiRNA2 (Figure 7a). Next, a TaqMan MicroRNA Assay (Applied Biosystems) was used to show that the esiRNA1 level in TPG7 cells was significantly higher than that in the vector control cells (2-fold; $P < 0.05$) (Figure 7b). These results suggest that esiRNA1 is expressed in Huh-7 cells, but is produced at higher levels when *ψPPM1K* is overexpressed from a recombinant plasmid.

FISH localization of *ψPPM1K*-derived esiRNAs

FISH results showed that *ψPPM1K*-derived esiRNA1 was located in the cytoplasm of interphase cells (Supplementary Figure S4, upper panels), and that esiRNA2 was mainly present the in nuclear area of both interphase

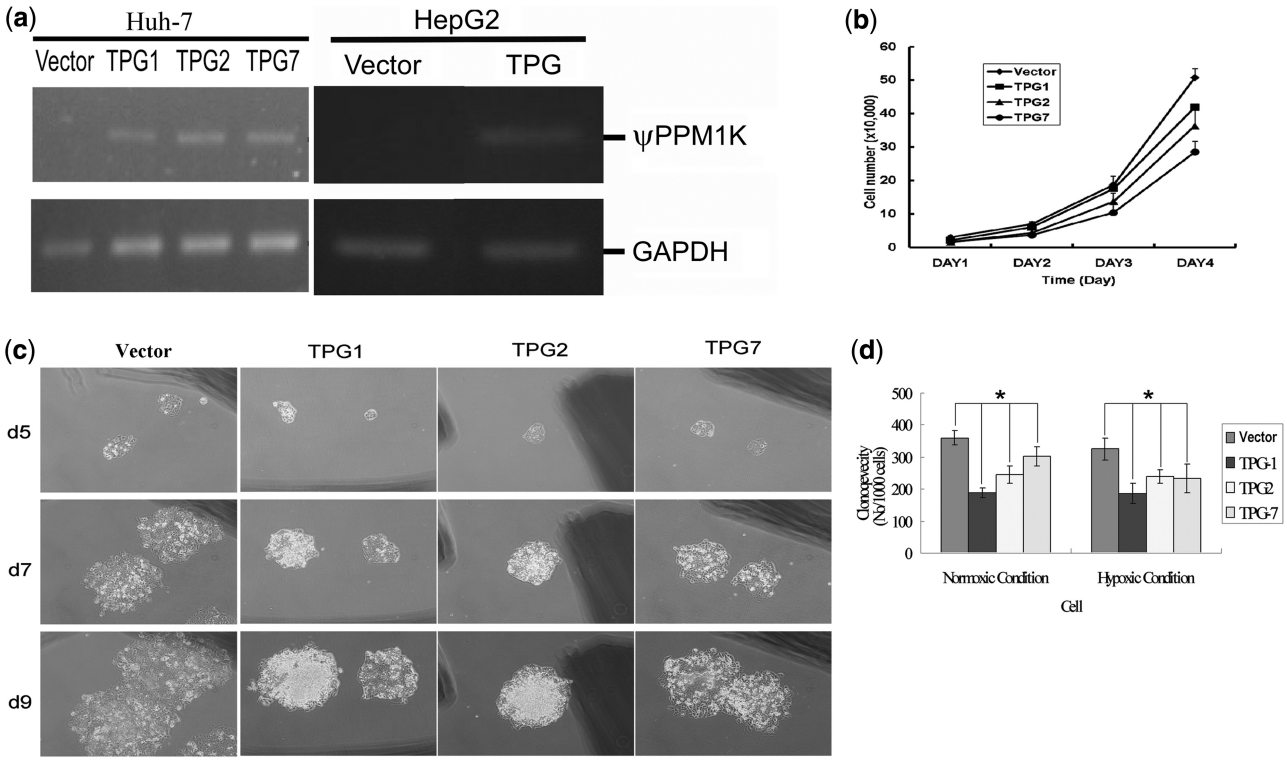

Figure 5. Effect of overexpressed *ψPPM1K* on cell growth and clonogenic activity in transfected Huh-7 cell clones. (**a**) HCC line Huh-7 and HepG2 cells were transfected with *ψPPM1K*-expressing recombinant plasmid to isolate stably transfected cell clones (TPG1, TPG2 and TPG7 for Huh-7; TPG for HepG2) by G418 selection, or with empty vector plasmid for G418-selected control mock2 cells (Vector). Equivalent amounts of total RNA from each cell clone were used for RT–PCR/gel electrophoresis analysis of *ψPPM1K* and *GAPDH* mRNA. (**b**) All three *ψPPM1K*-expressing cell

(continued)

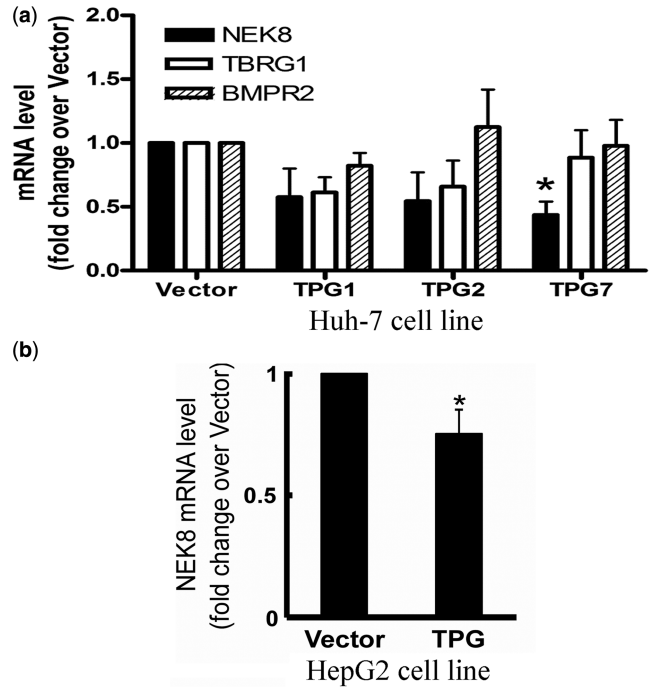

Figure 6. Expression of predicted target genes in Huh-7 and HepG2 HCC cell clones transfected with *ψPPM1K*. (**a**) The mRNA levels of three potential esiRNA1-targeted genes (*NEK8*, *TBRG1* and *BMPR2*) were analysed by RT–qPCR. *NEK8* was downregulated in all three *ψPPM1K*-expressing Huh-7 cell lines relative to the vector control cells ($*P = 0.033$ in TPG7 cells). (**b**) *NEK8* mRNA levels were reduced in the *ψPPM1K*-expressing HepG2-TPG cell line relative to the vector control cells ($*P = 0.047$).

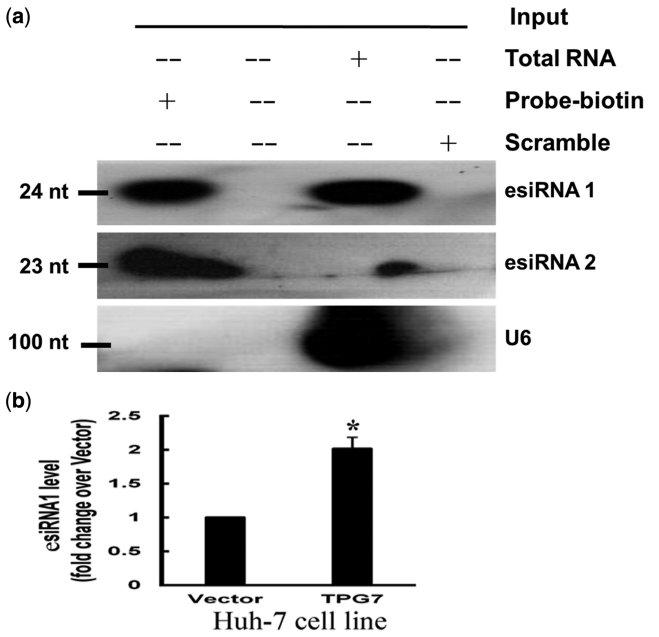

Figure 7. Expression of *ψPPM1K*-dervied esiRNAs. (**a**) *ψPPM1K*-dervied esiRNAs detection by northern blots. Total RNA from human hepatoma cell Huh-7 was resolved by agarose electrophoresis and then transferred to an NC membrane. The NC membrane was hybridized with a biotinylated esiRNA probe overnight, washed and incubated with horseradish peroxidase-conjugated avidin. Finally, esiRNAs signals were detected by chemiluminescence. (**b**) Expression of esiRNA1 in the *ψPPM1K*-expressing Huh-7 stable cell line (TPG7) and the control cell line (Vector). The esiRNA1 levels were determined using a TaqMan MicroRNA Assay (Applied Biosystems) and were significantly higher in TPG7 cells than in vector control cells ($*P < 0.05$).

and prometaphase cells (Supplementary Figure S4, lower panels), suggesting that esiRNA1, like miRNAs, regulates its targets in the cytoplasm, whereas esiRNA2 might not.

Expression of *NEK8* in Huh-7 cells transfected with synthetic siRNA1

To verify the effect of *ψPPM1K*-dervied esiRNA1 on *NEK8* expression, we transfected Huh-7 cells with synthetic siRNA1 identical in sequence to esiRNA1 (Figure 8a) and determined *NEK8* mRNA levels at 48 h. *NEK8* mRNA levels were significantly lower in Huh-7 cells transfected with the synthetic siRNA1 than those transfected with negative control siRNAs ($P = 0.001$), implying that synthetic siRNA homologous to the esiRNA1 can directly downregulate *NEK8* gene expression in Huh-7 cell lines (Figure 8b).

NEK8 expression and cell growth activation in esiRNA1-deleted *ψPPM1K*-overexpressing cells

To establish that *NEK8* is the major target of *ψPPM1K* through esiRNA1, we constructed an esiRNA1-deletion mutant *ψPPM1K*-expression plasmid to transfect HepG2 cells (named mTPG cells) and determined *NEK8* expression levels. The results showed that *NEK8* was downregulated in TPG HepG2 cells, but not in mTPG cells (Figure 8c). Additionally, we also co-expressed *NEK8* in *ψPPM1K*-transfected cells and carried out the cell proliferation assay used previously. This demonstrated that *NEK8* can counteract the inhibitory effects of *ψPPM1K* on HCC cell proliferation (Figure 8d).

ψPPM1K-derived esiRNA1 downregulation of PPM1K

ψPPM1K-derived esiRNA1 is also predicted to target *PPM1K* (Figure 2d); therefore, we examined the

Figure 5. Continued
lines have a slower proliferation rate than the vector control cell line (TPG7 cells: $P = 0.036$ for Day 2, $P = 0.018$ for Day 4; TPG1 cells: $P = 0.045$ for Day 4). (**c**) Serial photographs of the same colonies at Day 5, Day 7 and Day 9 showing the 2D growth of mock2, TPG1, TPG2 and TPG7 transfected Huh-7 clones on plastic culture dishes. (**d**) Clonogenic activity of mock2, TPG1, TPG2 and TPG7 was determined in soft-agar cultures incubated under normoxic 19% O2 and hypoxic 3% O2 conditions in 5% CO2 incubators for 12 days. Numbers of colonies formed per 1000 cells were counted for mock2, TPG1, TPG2 and TPG7 cell clones. Decreased clonogenicity compared with the mock2 control was calculated for TPG1 ($P = 1.54E-06$), for TPG2 ($P = 0.0001$) and for TPG7 ($P = 0.013$) in normoxic conditions; for TPG1 ($P = 2.22E-05$), for TPG2 ($P = 5.71E-05$) and for TPG7 ($P = 0.0011$) in hypoxic conditions. ($*P < 0.05$).

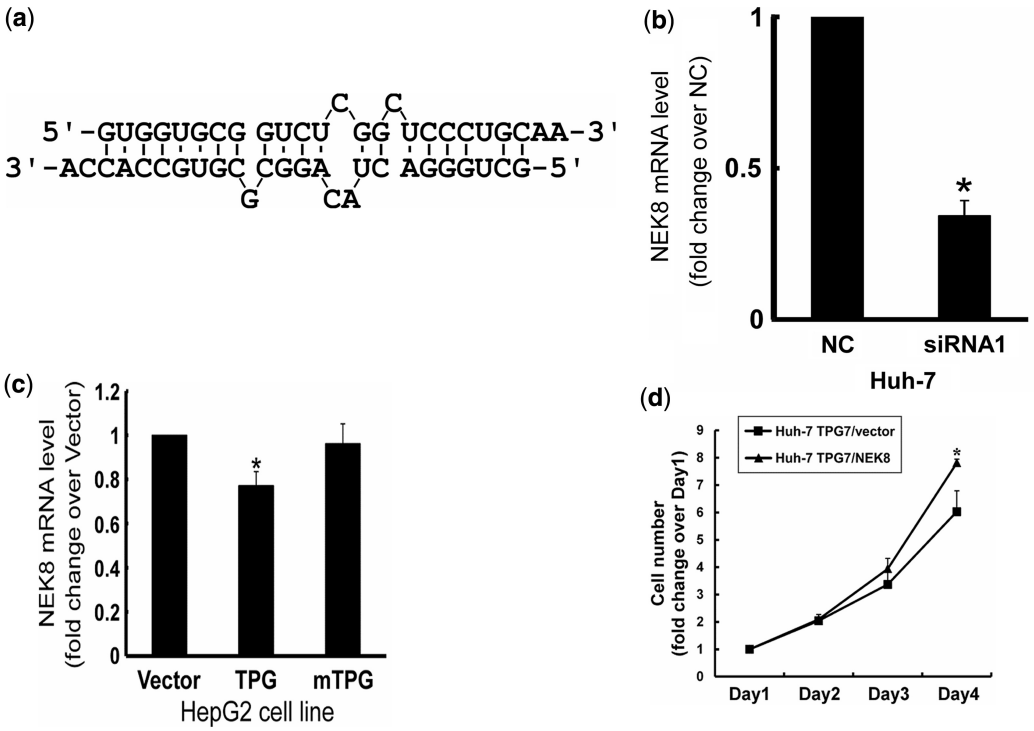

Figure 8. Expression of *NEK8* in Huh-7 cells transfected with synthetic siRNA1. (**a**) The sequence of synthetic siRNA1. (**b**) Huh-7 cells were transfected with synthetic siRNA1 or with negative control siRNA (NC) for 48 h. Total RNA was isolated, and the *NEK8* mRNA level was analysed by RT–qPCR. *NEK8* was downregulated in the Huh-7 cells transfected with the synthetic siRNA1 relative to those transfected with negative control siRNA ($*P = 0.001$), implying that synthetic siRNA homologous to esiRNA1 can directly downregulate *NEK8* gene expression in Huh-7 cells. (**c**) Expression of *NEK8* in an esiRNA1-deletion mutant cell line. *NEK8* mRNA levels in the *ψPPM1K*-expressing HepG2 cell line (TPG), the esiRNA1-deleted *ψPPM1K*-expressing cell line (mTPG) and the vector control cell line (Vector) were analysed by RT–qPCR. The results showed that *NEK8* was downregulated in TPG cells ($*P = 0.025$ relative to vector control cells), but not in mTPG cells ($P = 0.534$ relative to vector control cells). (**d**) Growth of Huh-7 TPG7 cells transfected with either *NEK8*-overexpressing plasmid or empty vector analysed by cell proliferation assay. The proliferation rate of *NEK8*-overexpressing TPG7 cells was significantly higher than vector-transfected control cells ($*P = 0.041$ at Day 4).

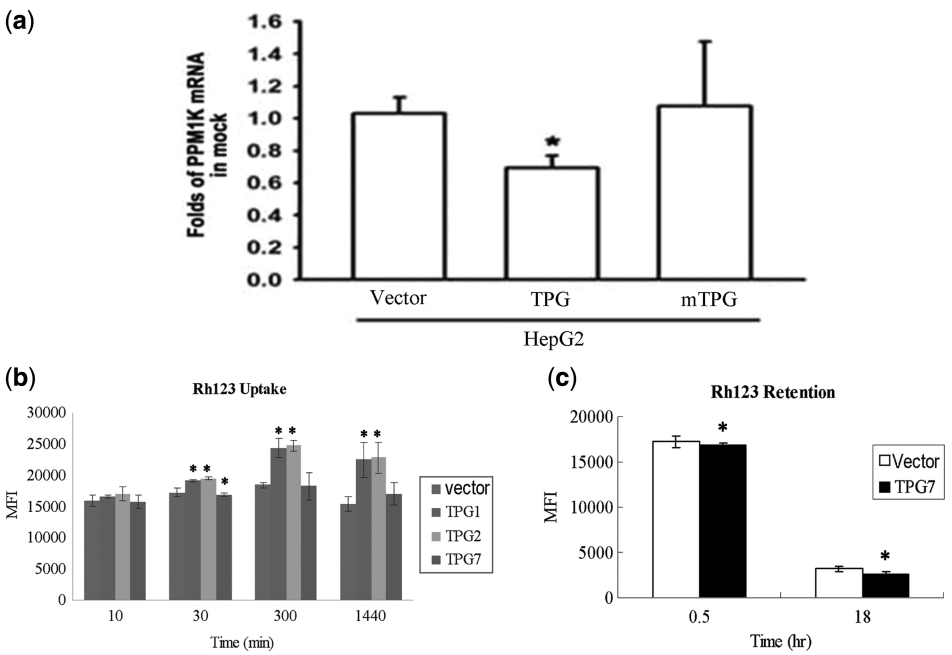

Figure 9. *ψPPM1K* alters *PPM1K* expression and mitochondrial function. (**a**) *ψPPM1K*-derived esiRNA1 downregulates *PPM1K* expression. *PPM1K* mRNA levels in the *ψPPM1K*-expressing HepG2 cell line (TPG), the esiRNA1-deleted *ψPPM1K*-expressing cell line (mTPG) and the vector control cell line (Vector) were analysed by RT–qPCR. The results showed that *PPM1K* was downregulated in TPG cells ($*P = 0.018$ relative to vector control cells), but not in mTPG cells. (**b**) Fluorescence-activated cell sorting analysis shows no significant differences of mitochondrial Rh123 uptake by the four transfected Huh-7 clones. ($*P < 0.01$). (**c**) Release of Rh123 from mitochondria (0.5–18 h) was significantly faster in *ψPPM1K*-expressing Huh-7 cell line (TPG7) cells (84.59%) than Vector control cells (81.32%) ($P = 0.031$ for 0.5 h; $P = 1.3E\text{-}06$ for 18 h).

expression of *PPM1K* in mTPG HepG2 cells. *PPM1K* mRNA levels were significantly lower in TPG HepG2 cells than in control cells ($P = 0.001$), but the levels were unaffected in mTPG HepG2 cells, implying that *ψPPM1K*-derived esiRNA1 can directly downregulate *PPM1K* expression (Figure 9a).

ψPPM1K alters mitochondrial functions in ψPPM1K-transduced Huh-7 cells

It has been shown that, following cellular uptake of Rh123 and specific accumulation of this fluorescent dye in mitochondria, human cancer cells tend to retain significantly more Rh123 than healthy cells (57,58). It is thought that this difference in the uptake and retention of Rh123 may reflect MPTP activity (56). As the phosphatase encoded by *PPM1K*/*PP2Cm* is located in the mitochondrial matrix and regulates MPTP activity, the effect of *ψPPM1K* overexpression is of interest. We found that rates of Rh123 uptake were higher for TPG1 and TPG2, whereas remained approximately the same as the control for TPG7 (Figure 9b and Supplementary Figure S5). The mitochondrial Rh123 retention assay showed significantly reduced Rh123 retention in TPG7 cells than in control vector cells (Figure 9c), suggesting faster Rh123 release or relatively normalized MPTP stability in TPG7 cells. However, no such change in Rh123 retention was observed in TPG1 or TPG2 cells in comparison with control cells.

miRNA regulation of *ψPPM1K* and *PPM1K*

Considering the recently reported novel findings that pseudogenes *PTENP1* and *KRAS1P* can de-repress their cognate genes by an miRNA decoy mechanism (15), we took a similar approach to investigate MTI in *ψPPM1K* and *PPM1K*. The *in silico* results revealed many miRNAs that potentially interact with multiple possible target sites in both *ψPPM1K* and *PPM1K* (Supplementary Table S6). Furthermore, seed matches for *PPM1K*-targeting miRNAs were perfectly conserved in *ψPPM1K* (Figure 10a). To verify this prediction, we assessed the expression levels of *PPM1K* and *ψPPM1K* in miRNA-transfected TPG7 Huh-7 cells and vector control cells. The results showed that has-miR-3174 (miR-3174) downregulated *PPM1K* in both cell lines, whereas *ψPPM1K* was also downregulated in vector control cells compared with a negative control sRNA (siCon) (Figure 10b).

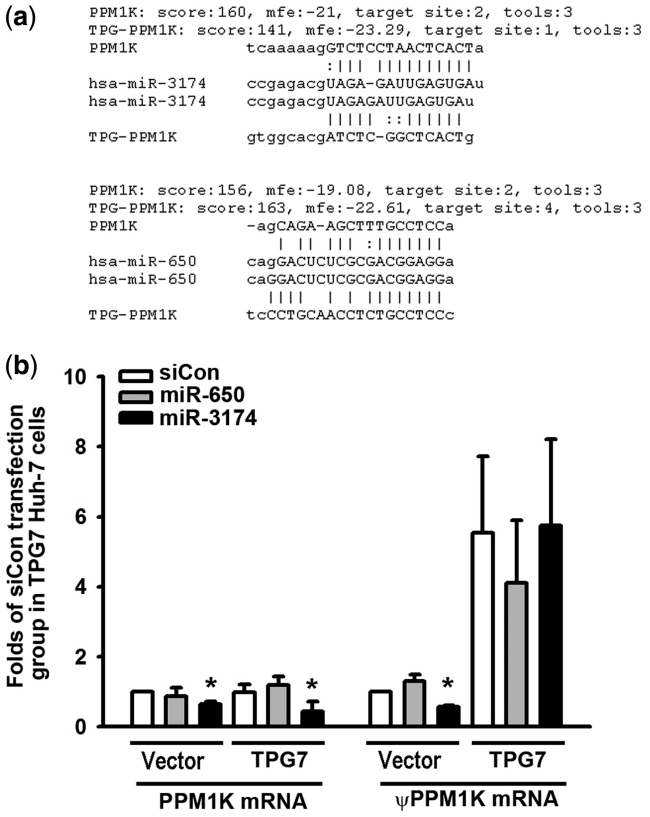

Figure 10. miRNA regulation of *PPM1K* and *ψPPM1K*. (**a**) Alignments of miRNA sequences with *PPM1K* and *ψPPM1K*. (**b**) Expression of *PPM1K* and *ψPPM1K* in miRNA-transfected TPG7 and mock cells. *PPM1K* is significantly downregulated in TPG7 and mock cells, and *ψPPM1K* is downregulated in mock cells, relative to a stable negative control (siCon) after miR-3174 transfection for 48 h ($^*P \leq 0.01$).

DISCUSSION

In this study, we developed and performed a systematic bioinformatics pipeline for identifying TPG-derived esiRNAs and their interacting gene targets in the human genome. We selected *ψPPM1K* to verify the *in silico* analysis by experiment and obtained five major results. First, we found that *ψPPM1K*-derived esiRNAs were expressed at higher levels in human liver tissue than in paired HCC tumour samples (Figure 4a), and this is reflected in publicly available gene expression profile data (Figure 3f). Second, growth inhibitory effects were observed in *ψPPM1K*-transfected cells (Figure 5), implicating *ψPPM1K*-derived esiRNAs in the control of cell growth. This also demonstrates that esiRNAs can be derived from pseudogene transcripts in human somatic cells, thereby extending similar findings in mouse germ cells (oocytes) (22,23). Third, at least two esiRNAs that are expressed in liver tissue and in transfected HCC cells are derived from distinct *ψPPM1K* sequences not present in the cognate *PPM1K* gene (Figures 3–5 and 7). Fourth, we provided direct evidence that *ψPPM1K*-derived esiRNA1 downregulates target gene *NEK8* as well as the parental *PPM1K* gene, and it inhibits cell growth through downregulation of *NEK8* (Figures 8 and 9). Moreover, the proliferation rate of *ψPPM1K*-transfected cells in which *NEK8* was co-expressed was higher than control cells, supporting the hypothesis that *NEK8* can counteract the inhibitory effects of *ψPPM1K* on HCC cell proliferation (Figure 8d). This is consistent with reports that *NEK* family members regulate cell cycle progression (60), and that *NEK8*/*NEK9* are overexpressed in human breast cancer (60,61). Thus, in a broad sense, *ψPPM1K* could be considered a tumour-suppressor gene. Fifth, an miRNA transfection assay showed that miR-3174 significantly downregulated *PPM1K* in vector control cells and

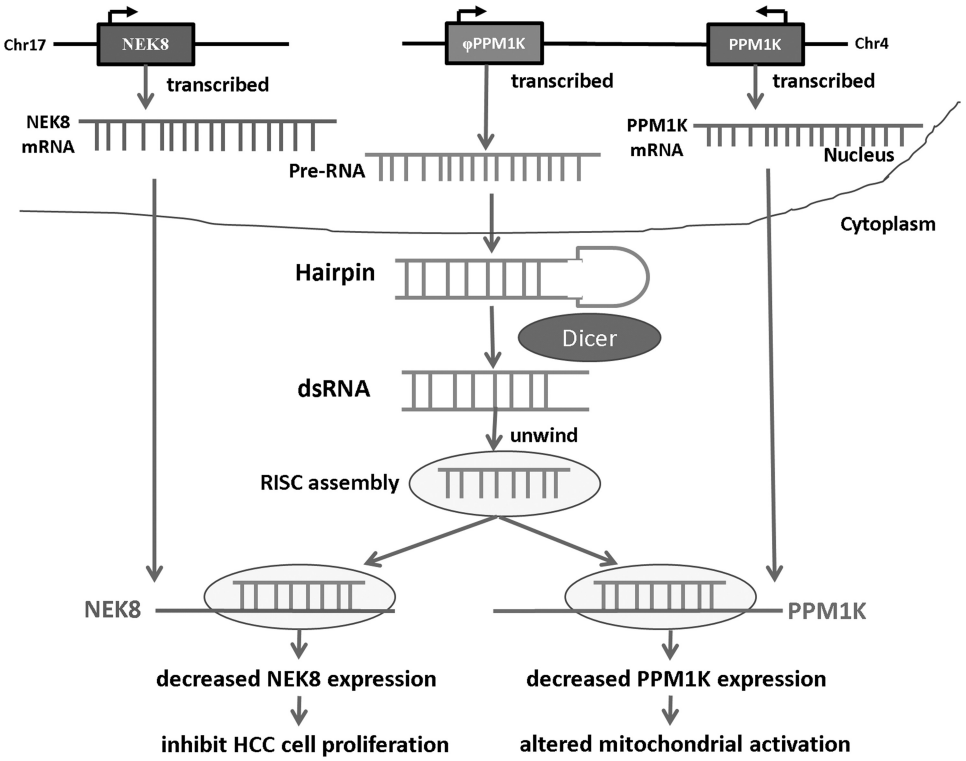

Figure 11. Possible genetic regulatory mechanisms involving *ψPPM1K*. Transcripts of *ψPPM1K* are exported to the cytoplasm where dsRNAs are cut from hairpin structures by Dicer into esiRNAs, which then interact with protein-coding target genes. Our results indicate that *ψPPM1K*-derived esiRNA1 may inhibit HCC cell proliferation through downregulation of *NEK8*, as well as by decreasing expression of *PPM1K* and alteration of mitochondrial activation.

TPG7 cells; it also downregulated *ψPPM1K* in vector control cells compared with a stable negative control (siCon) (Figure 10b). The miR-3174 downregulation of *PPM1K* and *ψPPM1K*, however, is distinct from the recently elucidated decoy mechanism by which the tumour-suppressor pseudogene, *PTENP1*, helps to regulate its cognate tumour suppressor gene, *PTEN* (15).

This study indicates two probable pathways for the generation of *ψPPM1K*-derived esiRNAs: one from dsRNAs formed by the antisense transcript of *ψPPM1K* and its cognate gene sequence in the 420–449-nt region, which gives rise to esiRNA3 (Figure 3c), and the other from a hairpin structure resulting from inverted repeats in *ψPPM1K*, giving esiRNA1 (Figures 3b and 11).

PPM1K, which encodes a mitochondrial matrix serine/threonine protein phosphatase containing an N-terminal mitochondrial localization signal and a central PP2C catalytic domain, regulates the mitochondrial MPTP and is essential for cellular survival and development (59). It is, therefore, possible that, through the predicted dsRNA-generated esiRNA, *ψPPM1K* may affect the pivotal regulatory function of its cognate *PPM1K* gene on MPTP activity. Another report showed that human cancer cells have an abnormally high uptake and retention of the mitochondria-specific dye, Rh123 (57,58), and this is thought to reflect MPTP activity (56). It is uncertain whether *PPM1K* is associated with carcinogenesis and/or oncogenesis, or how it may interact with its tumour-suppressor pseudogene *ψPPM1K*. Our findings indicating that *ψPPM1K* downregulates *PPM1K* by producing esiRNA1 (Figure 9a) and decreases uptake and retention of Rh123, suggesting altered MPTP activity (Figure 9b), may be significant in this regard.

In conclusion, we have shown that pseudogene-derived esiRNAs can, independently of the cognate gene, regulate cell growth-related target genes in HCC. Specifically, our bioinformatics pipeline, and subsequent biological experiments demonstrated that *ψPPM1K*-derived esiRNA1 modulates HCC cell proliferation via direct downregulation of *NEK8* and may also regulate mitochondrial activation through decreased *PPM1K* cognate gene function (Figure 11). This suggests a tumour-suppressor role for the esiRNAs derived from *ψPPM1K*. To explore whether other pseudogene-derived esiRNAs identified in somatic human cells (65) also possess regulatory functions, it will be of interest to perform similar experimental tests on the 447 transcribed pseudogenes and their esiRNA derivatives predicted from our *in silico* analyses.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–6, Supplementary Figures 1–5 and Supplementary File 1.

FUNDING

National Science Council of the Republic of China [NSC 98-2311-B-009-004-MY3, NSC 99-2627-B-009-003, NSC100-2627-B-009-002, NSC101-2311-B-009-003-MY3,

to H.-D.H., NSC99-2320-B-039-038-MY3, to J.-G.C., NSC98-2314-B-039-010-MY3, to W.-K.Y.]; UST-UCSD International Center of Excellence in Advanced Bio-engineering sponsored by the Taiwan National Science Council I-RiCE Program [NSC101-2911-I-009-101 in part]; Veterans General Hospitals and University System of Taiwan (VGHUST) Joint Research Program [VGHUST101-G5-1-1], MOE ATU (in part). Funding for open access charge: National Science Council of the Republic of China.

Conflict of interest statement. None declared.

REFERENCES

1. Torrents,D., Suyama,M., Zdobnov,E. and Bork,P. (2003) A genome-wide survey of human pseudogenes. *Genome Res.*, **13**, 2559–2567.
2. Mighell,A.J., Smith,N.R., Robinson,P.A. and Markham,A.F. (2000) Vertebrate pseudogenes. *FEBS Lett.*, **468**, 109–114.
3. Balasubramanian,S., Zheng,D., Liu,Y.J., Fang,G., Frankish,A., Carriero,N., Robilotto,R., Cayting,P. and Gerstein,M. (2009) Comparative analysis of processed ribosomal protein pseudogenes in four mammalian genomes. *Genome Biol.*, **10**, R2.
4. Harrison,P. and Yu,Z. (2007) Frame disruptions in human mRNA transcripts, and their relationship with splicing and protein structures. *BMC Genomics*, **8**, 371.
5. Harrison,P.M., Zheng,D., Zhang,Z., Carriero,N. and Gerstein,M. (2005) Transcribed processed pseudogenes in the human genome: an intermediate form of expressed retrosequence lacking protein-coding ability. *Nucleic Acids Res.*, **33**, 2374–2383.
6. Vinckenbosch,N., Dupanloup,I. and Kaessmann,H. (2006) Evolutionary fate of retroposed gene copies in the human genome. *Proc. Natl Acad. Sci. USA*, **103**, 3220–3225.
7. Zheng,D., Frankish,A., Baertsch,R., Kapranov,P., Reymond,A., Choo,S.W., Lu,Y., Denoeud,F., Antonarakis,S.E., Snyder,M. *et al.* (2007) Pseudogenes in the ENCODE regions: consensus annotation, analysis of transcription, and evolution. *Genome Res.*, **17**, 839–851.
8. Zheng,D., Zhang,Z., Harrison,P.M., Karro,J., Carriero,N. and Gerstein,M. (2005) Integrated pseudogene annotation for human chromosome 22: evidence for transcription. *J. Mol. Biol.*, **349**, 27–45.
9. Khachane,A.N. and Harrison,P.M. (2009) Assessing the genomic evidence for conserved transcribed pseudogenes under selection. *BMC Genomics*, **10**, 435.
10. Imanishi,T., Itoh,T., Suzuki,Y., O'Donovan,C., Fukuchi,S., Koyanagi,K.O., Barrero,R.A., Tamura,T., Yamaguchi-Kabata,Y., Tanino,M. *et al.* (2004) Integrative annotation of 21,037 human genes validated by full-length cDNA clones. *PLoS Biol.*, **2**, e162.
11. Korneev,S.A., Park,J.H. and O'Shea,M. (1999) Neuronal expression of neural nitric oxide synthase (nNOS) protein is suppressed by an antisense RNA transcribed from an NOS pseudogene. *J. Neurosci.*, **19**, 7711–7720.
12. Korneev,S. and O'Shea,M. (2002) Evolution of nitric oxide synthase regulatory genes by DNA inversion. *Mol. Biol. Evol.*, **19**, 1228–1233.
13. Hirotsune,S., Yoshida,N., Chen,A., Garrett,L., Sugiyama,F., Takahashi,S., Yagami,K., Wynshaw-Boris,A. and Yoshiki,A. (2003) An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene. *Nature*, **423**, 91–96.
14. Gray,T.A., Wilson,A., Fortin,P.J. and Nicholls,R.D. (2006) The putatively functional Mkrn1-p1 pseudogene is neither expressed nor imprinted, nor does it regulate its source gene in trans. *Proc. Natl Acad. Sci. USA*, **103**, 12039–12044.
15. Poliseno,L., Salmena,L., Zhang,J., Carver,B., Haveman,W.J. and Pandolfi,P.P. (2010) A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature*, **465**, 1033–1038.
16. Kim,D.H. and Rossi,J.J. (2007) Strategies for silencing human disease using RNA interference. *Nat. Rev. Genet.*, **8**, 173–184.
17. Ghildiyal,M., Seitz,H., Horwich,M.D., Li,C., Du,T., Lee,S., Xu,J., Kittler,E.L., Zapp,M.L., Weng,Z. *et al.* (2008) Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. *Science*, **320**, 1077–1081.
18. Yi,R., Qin,Y., Macara,I.G. and Cullen,B.R. (2003) Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.*, **17**, 3011–3016.
19. Brennecke,J., Aravin,A.A., Stark,A., Dus,M., Kellis,M., Sachidanandam,R. and Hannon,G.J. (2007) Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell*, **128**, 1089–1103.
20. Bartel,D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281–297.
21. Jovanovic,M. and Hengartner,M.O. (2006) miRNAs and apoptosis: RNAs to die for. *Oncogene*, **25**, 6176–6187.
22. Tam,O.H., Aravin,A.A., Stein,P., Girard,A., Murchison,E.P., Cheloufi,S., Hodges,E., Anger,M., Sachidanandam,R., Schultz,R.M. *et al.* (2008) Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature*, **453**, 534–538.
23. Watanabe,T., Totoki,Y., Toyoda,A., Kaneda,M., Kuramochi-Miyagawa,S., Obata,Y., Chiba,H., Kohara,Y., Kono,T., Nakano,T. *et al.* (2008) Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature*, **453**, 539–543.
24. Czech,B., Malone,C.D., Zhou,R., Stark,A., Schlingeheyde,C., Dus,M., Perrimon,N., Kellis,M., Wohlschlegel,J.A., Sachidanandam,R. *et al.* (2008) An endogenous small interfering RNA pathway in Drosophila. *Nature*, **453**, 798–802.
25. Kawamura,Y., Saito,K., Kin,T., Ono,Y., Asai,K., Sunohara,T., Okada,T.N., Siomi,M.C. and Siomi,H. (2008) Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. *Nature*, **453**, 793–797.
26. Okamura,K., Chung,W.J., Ruby,J.G., Guo,H., Bartel,D.P. and Lai,E.C. (2008) The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs. *Nature*, **453**, 803–806.
27. Sethupathy,P., Megraw,M. and Hatzigeorgiou,A.G. (2006) A guide through present computational approaches for the identification of mammalian microRNA targets. *Nat. Methods*, **3**, 881–886.
28. Mituyama,T., Yamada,K., Hattori,E., Okida,H., Ono,Y., Terai,G., Yoshizawa,A., Komori,T. and Asai,K. (2009) The Functional RNA Database 3.0: databases to support mining and annotation of functional RNAs. *Nucleic Acids Res.*, **37**, D89–D92.
29. Carninci,P., Kasukawa,T., Katayama,S., Gough,J., Frith,M.C., Maeda,N., Oyama,R., Ravasi,T., Lenhard,B., Wells,C. *et al.* (2005) The transcriptional landscape of the mammalian genome. *Science*, **309**, 1559–1563.
30. Baek,D., Villen,J., Shin,C., Camargo,F.D., Gygi,S.P. and Bartel,D.P. (2008) The impact of microRNAs on protein output. *Nature*, **455**, 64–71.
31. He,S., Liu,C., Skogerbo,G., Zhao,H., Wang,J., Liu,T., Bai,B., Zhao,Y. and Chen,R. (2008) NONCODE v2.0: decoding the non-coding. *Nucleic Acids Res.*, **36**, D170–D172.
32. Griffiths-Jones,S., Moxon,S., Marshall,M., Khanna,A., Eddy,S.R. and Bateman,A. (2005) Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res.*, **33**, D121–D124.
33. Pang,K.C., Stephen,S., Dinger,M.E., Engstrom,P.G., Lenhard,B. and Mattick,J.S. (2007) RNAdb 2.0–an expanded database of mammalian non-coding RNAs. *Nucleic Acids Res.*, **35**, D178–D182.
34. Lestrade,L. and Weber,M.J. (2006) snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. *Nucleic Acids Res.*, **34**, D158–D162.
35. Morin,R.D., O'Connor,M.D., Griffith,M., Kuchenbauer,F., Delaney,A., Prabhu,A.L., Zhao,Y., McDonald,H., Zeng,T., Hirst,M. *et al.* (2008) Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res.*, **18**, 610–621.
36. Seila,A.C., Calabrese,J.M., Levine,S.S., Yeo,G.W., Rahl,P.B., Flynn,R.A., Young,R.A. and Sharp,P.A. (2008) Divergent transcription from active promoters. *Science*, **322**, 1849–1851.
37. Yeo,G.W., Xu,X., Liang,T.Y., Muotri,A.R., Carson,C.T., Coufal,N.G. and Gage,F.H. (2007) Alternative splicing events

identified in human embryonic stem cells and neural progenitors. *PLoS Comput. Biol.*, **3**, 1951–1967.
38. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
39. Hsu,S.D., Chu,C.H., Tsou,A.P., Chen,S.J., Chen,H.C., Hsu,P.W., Wong,Y.H., Chen,Y.H., Chen,G.H. and Huang,H.D. (2008) miRNAMap 2.0: genomic maps of microRNAs in metazoan genomes. *Nucleic Acids Res.*, **36**, D165–D169.
40. Friedman,R.C., Farh,K.K., Burge,C.B. and Bartel,D.P. (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.*, **19**, 92–105.
41. Grimson,A., Farh,K.K., Johnston,W.K., Garrett-Engele,P., Lim,L.P. and Bartel,D.P. (2007) MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol. Cell*, **27**, 91–105.
42. Lewis,B.P., Burge,C.B. and Bartel,D.P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, **120**, 15–20.
43. John,B., Enright,A.J., Aravin,A., Tuschl,T., Sander,C. and Marks,D.S. (2004) Human MicroRNA targets. *PLoS Biol.*, **2**, e363.
44. Kruger,J. and Rehmsmeier,M. (2006) RNAhybrid: microRNA target prediction easy, fast and flexible. *Nucleic Acids Res.*, **34**, W451–W454.
45. Barrett,T. and Edgar,R. (2006) Gene expression omnibus: microarray data storage, submission, retrieval, and analysis. *Methods Enzymol.*, **411**, 352–369.
46. Su,A.I., Wiltshire,T., Batalov,S., Lapp,H., Ching,K.A., Block,D., Zhang,J., Soden,R., Hayakawa,M., Kreiman,G. *et al.* (2004) A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl Acad. Sci. USA*, **101**, 6062–6067.
47. Yu,K., Ganesan,K., Tan,L.K., Laban,M., Wu,J., Zhao,X.D., Li,H., Leung,C.H., Zhu,Y., Wei,C.L. *et al.* (2008) A precisely regulated gene expression cassette potently modulates metastasis and survival in multiple solid cancers. *PLoS Genet.*, **4**, e1000129.
48. Liao,Y.L., Sun,Y.M., Chau,G.Y., Chau,Y.P., Lai,T.C., Wang,J.L., Horng,J.T., Hsiao,M. and Tsou,A.P. (2008) Identification of SOX4 target genes using phylogenetic footprinting-based prediction from expression microarrays suggests that overexpression of SOX4 potentiates metastasis in hepatocellular carcinoma. *Oncogene*, **27**, 5578–5589.
49. Okuda,S., Yamada,T., Hamajima,M., Itoh,M., Katayama,T., Bork,P., Goto,S. and Kanehisa,M. (2008) KEGG Atlas mapping for global analysis of metabolic pathways. *Nucleic Acids Res.*, **36**, W423–W426.
50. Huang da,W., Sherman,B.T. and Lempicki,R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.*, **4**, 44–57.
51. Chan,C.H., Ko,C.C., Chang,J.G., Chen,S.F., Wu,M.S., Lin,J.T. and Chow,L.P. (2006) Subcellular and functional proteomic analysis of the cellular responses induced by Helicobacter pylori. *Mol. Cell. Proteomics*, **5**, 702–713.
52. Yuo,C.Y., Lin,H.H., Chang,Y.S., Yang,W.K. and Chang,J.G. (2008) 5-(N-ethyl-N-isopropyl)-amiloride enhances SMN2 exon 7 inclusion and protein expression in spinal muscular atrophy cells. *Ann. Neurol.*, **63**, 26–34.
53. Liu,T.C., Lin,S.F., Chang,J.G., Yang,M.Y., Hung,S.Y. and Chang,C.S. (2003) Epigenetic alteration of the SOCS1 gene in chronic myeloid leukaemia. *Br. J. Haematol.*, **123**, 654–661.
54. Torres,A.G., Fabani,M.M., Vigorito,E. and Gait,M.J. (2011) MicroRNA fate upon targeting with anti-miRNA oligonucleotides as revealed by an improved Northern-blot-based method for miRNA detection. *RNA*, **17**, 933–943.
55. Fleming,E.J., Langdon,A.E., Martinez-Garcia,M., Stepanauskas,R., Poulton,N.J., Masland,E.D. and Emerson,D. (2011) What's new is old: resolving the identity of Leptothrix ochracea using single cell genomics, pyrosequencing and FISH. *PLoS One*, **6**, e17769.
56. Rottenberg,H. and Wu,S. (1998) Quantitative assay by flow cytometry of the mitochondrial membrane potential in intact cells. *Biochim. Biophys. Acta*, **1404**, 393–404.
57. Nadakavukaren,K.K., Nadakavukaren,J.J. and Chen,L.B. (1985) Increased rhodamine 123 uptake by carcinoma cells. *Cancer Res.*, **45**, 6093–6099.
58. Darzynkiewicz,Z., Traganos,F., Staiano-Coico,L., Kapuscinski,J. and Melamed,M.R. (1982) Interaction of rhodamine 123 with living cells studied by flow cytometry. *Cancer Res.*, **42**, 799–806.
59. Lu,G., Ren,S., Korge,P., Choi,J., Dong,Y., Weiss,J., Koehler,C., Chen,J.N. and Wang,Y. (2007) A novel mitochondrial matrix serine/threonine protein phosphatase regulates the mitochondria permeability transition pore and is essential for cellular survival and development. *Genes Dev.*, **21**, 784–796.
60. O'Regan,L., Blot,J. and Fry,A.M. (2007) Mitotic regulation by NIMA-related kinases. *Cell Div.*, **2**, 25.
61. Bowers,A.J. and Boylan,J.F. (2004) Nek8, a NIMA family kinase member, is overexpressed in primary human breast tumors. *Gene*, **328**, 135–142.
62. Tompkins,V., Hagen,J., Zediak,V.P. and Quelle,D.E. (2006) Identification of novel ARF binding proteins by two-hybrid screening. *Cell Cycle*, **5**, 641–646.
63. Tompkins,V.S., Hagen,J., Frazier,A.A., Lushnikova,T., Fitzgerald,M.P., di Tommaso,A., Ladeveze,V., Domann,F.E., Eischen,C.M. and Quelle,D.E. (2007) A novel nuclear interactor of ARF and MDM2 (NIAM) that maintains chromosomal stability. *J. Biol. Chem.*, **282**, 1322–1333.
64. Nohno,T., Ishikawa,T., Saito,T., Hosokawa,K., Noji,S., Wolsing,D.H. and Rosenbaum,J.S. (1995) Identification of a human type II receptor for bone morphogenetic protein-4 that forms differential heteromeric complexes with bone morphogenetic protein type I receptors. *J. Biol. Chem.*, **270**, 22522–22526.
65. Okamura,K., Chung,W.J. and Lai,E.C. (2008) The long and short of inverted repeat genes in animals: microRNAs, mirtrons and hairpin RNAs. *Cell Cycle*, **7**, 2840–2845.