**High-throughput Screening Identifies Non-coding RNA (ncRNA) *PANC754* as a Novel Pan-Tumor Suppressor ncRNA**

Shicheng Guo1#, Jian Shi2#, Shaoqing Ju2, Steven J. Schrodi1,3,4\*, Weifeng Ding2\*

1.Center for Precision Medicine Research, Marshfield Clinic Research Institute, Marshfield, WI 54449, USA

2.Department of Laboratory Medicine, Affiliated Hospital of Nantong University, Nantong, Jiangsu Province, China

3.Computation and Informatics in Biology and Medicine, University of Wisconsin-Madison, Madison, WI, USA

4.Department of Medical Genetics, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI, USA

# These authors contributed equally to the manuscript.

\* Corresponding Authors

Information for Corresponding Authors:

Steven J. Schrodi, Ph.D.

Department of Medical Genetics

School of Medicine and Public Health

University of Wisconsin-Madison

Madison, WI, 53706

Tel: (608) 262-1069

Email: [schrodi@wisc.edu](mailto:schrodi@wisc.edu)

Weifeng Ding, Ph.D.

Department of Laboratory Medicine

Affiliated Hospital of Nantong University

Nantong, Jiangsu Province,

China, 226001

Tel: +86-152-167-60764

Email: [dingweifeng.ntu@163.com](mailto:dingweifeng.ntu@163.com)

**Highlights**

High-throughput screening identifies ncRNA *PANC754* as being ubiquitously downregulated across human cancers.

Overexpression of PANC754 significantly suppresses tumor growth, migration, and invasion.

Overexpression of PANC754 significantly increases cell apoptosis and downregulation of oncogenes.

High-expression of PANC754 is positively correlated with survival times in cancer patients.

The pan-cancer suppression role of PANC754 is KPNA4 and MTMR8 dependent.

The pan pseudogene *PANC754* can regulate its original gene.

Running title: *PANC754* identified as a novel pan-tumor suppressor lncRNA.

**Abstract**

Tumor suppressor genes are usually down-regulated or deleted in human cancers. However, current tumor suppressor gene scanning strategies preferentially select for protein coding genes. By applying an 11,093 pan-cancer high-throughput mRNA-seq data analysis across 23 human cancers, we identified a novel ncRNA *PANC754* that is ubiquitously down-regulated in human cancers. We validated the low-expression status of *PANC754* in cancer cell lines and cancer tissues collected from colorectal and gastric cancer patients and noted that over-expression of *PANC754* in these cells significantly suppressed the growth, migration, and invasion of multiple human tumor cell lines and significantly increased tumor cell apoptosis. Further signaling analysis revealed that the pan-cancer suppression role of PANC754 was *KPNA4* and *MTMR8* dependent. Furthermore, oncogenes *DNMT1* and *CDK2* were down-regulated when *PANC754* was overexpressed, which may indicate that *PANC754* widely suppresses cancer via the CDK2 signaling pathway. Our study indicates that including non-coding RNA analysis in genome-wide tumor suppressor gene and oncogene screening is likely necessary to better understand cancer pathogenesis and identify novel therapeutic targets.

**Keywords**: Pan-cancer, non-coding RNA, epigenetics, tumor suppressor gene, regulatory

**Background**

Although cancer incidence and mortality rates declined over the past decade,1 cancers are still the second leading cause of death in China.2 There are two primary types of genes involved in tumorigenesis, oncogenes and tumor suppressor genes (TSGs). Oncogenes, such as *DNMT1* (DNA methyltransferase 1)3, *CDK2* (cyclin dependent kinase 2)4, *HRAS* (Human V-Ha-ras Harvey Rat Sarcoma Viral Oncogene Homolog)5, and *HDAC7* (Histone deacetylase 7)6, result in the uncontrolled growth of cells, while tumor suppressor genes like *CDH13* (cadherin 13)7, *ZNF132* (zinc finger protein 132)8, *EPAS1* (endothelial PAS domain protein 1)3, and *RARG* (retinoic acid receptor, gamma)9, typically contribute to the fidelity of the [cell cycle](https://www.sciencedirect.com/topics/medicine-and-dentistry/cell-cycle).10 TSGs commonly act as negative regulators of [oncogenes](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/oncogenes), [cell cycle check points](https://www.sciencedirect.com/topics/medicine-and-dentistry/cell-cycle-checkpoint), or [gene products](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/gene-product) that supply appropriate nutrients or components to complete cell division in the absence of stress.11 Traditionally, tumor suppressor gene scanning was largely limited to protein coding genes. However, more and more evidence indicates that non-coding RNAs also have important regulatory functions that might be novel oncogene or tumor suppressor genes.

Non-coding RNAs (ncRNAs) are RNA molecules that are not translated into proteins and classically include transfer RNAs (tRNAs), ribosomal RNAs (rRNAs), microRNAs, siRNAs, piRNAs, snoRNAs, exRNAs, snRNAs, scaRNAs, circRNA, and lncRNAs12. As potential regulators of gene expression, ncRNA genes are attaining increasing importance in the etiology of various diseases.13-16. Researchers have found that ncRNAs are particularly relevant in playing a role in tumor suppression12,17,18 and may share many regulatory features across cancers, such as aberrant cellular phenotypes and interactions19,20, disrupted pathways21, epigenetic alterations22,23, and genetic damage24,25. Elucidating these common features will provide insight into the underlying molecular mechanisms of carcinogenesis. Therefore, pan-cancer analyses may be an effective tool for understanding the key perturbations that give rise to a wide array of cancers26.

To identify pan-cancer non-coding tumor suppressor RNAs, we collected information from the TCGA project to develop an 11,093 RNA-sequencing dataset across 23 cancer types and applied a meta-analysis based standardized mean difference (SMD) to detect ncRNAs that were significantly downregulated across different cancers. We identified a novel pan-cancer ncRNA, *PANC754 (*ENSG00000213754.2*)*, which was significantly decreased in almost all human cancers analyzed and associated with decreased overall survival in patients. Functional validation via overexpression of *PANC754*in multiple human cancer cell lines indicates that *PANC754* can regulate sequence-similar genes (such as *KPNA4* and *MTMR8*) and is frequently associated with the over-expression of oncogenes including *DNMT1* and *CDK2*.

**Results**

**Ubiquitous down-regulation of non-coding RNA *PANC754* across 23 human cancers.**

To identify novel tumor suppressor ncRNAs, we collected all RNA-seq data from the TCGA database. 11,529 non-coding RNAs were extracted from 11,093 available TCGA RNA-seq data, and a meta-analysis was conducted to identify pan-cancer abnormal-regulated ncRNAs. Using a random effect model, we identified 1,615 ncRNAs that were significantly upregulated or downregulated (**Figure 1,** Bonferroni corrected P<0.05 and **Table S1**). A novel ubiquitously down-regulated ncRNA named PAn-cancer Non-Coding RNA 754 (*PANC754)*, which is an antisense ncRNA and encoded byENSG00000213754, was identified as significantly downregulated across the 23 human cancer tissues analyzed (SMD= -1.2, P<1.0x10-264 in fixed effects model and P=1.78x10-29 in random effects model, **Figure 2 and Table S2**). We validated this result by determining the expression level of *PANC754* in patients’ samples from colorectal cancer (CRC) and gastric cancer (GC) and compared this result to PANC754 expression in adjacent normal tissues (**Figure S1 and S2**). Further correlation analysis of PANC754 expression and patient survival demonstrated that high-expression of PANC754 was a significantly protective outcome factor across all cancer types (HR=0.84, 95%CI: 0.76-0.94, P=0.0024 in fixed effects model versus HR=0.86 95%CI: 0.78-0.94, P=9.0x10-4 in random effects model, **Figure S3 and Table S3**), especially in cervical, kidney, and renal papillary cell carcinoma (HR=0.42, 95%CI=0.21-0.85) as well as liver hepatocellular carcinoma (HR=0.62, 95%CI=0.42-0.93).

**PANC754 overexpression significantly suppressed tumor cell growth and increased cell apoptosis**

To probe the functional role of PANC754 in cancer cell phenotypes, we transfected tumor cell lines with a plasmid containing PANC754 and analyzed downstream CCK-8 expression. Compared to the control (untransfected) or pcDNA3.1 (transfected with empty vector/pcDNA3.1 plasmid) groups, the growth ability of GC cell line MGC803 and esophageal cancer Eca109 cells transfected with pcDNA3.1-PANC754 were gradually inhibited, especially after 48 h and 72 h in culture. (P=0.0004 and P=0.0016, respectively; **Figure 3A and 3B**) Similar growth patterns were also observed in two hepatocellular carcinoma cell lines, HepG2 and Bel-7402, (P=0.0006 and P=0.0045, respectively) and two CRCcell lines,SW480 and DLD1, (P=0.0231 and P=0.0002, respectively) (**Figure S3**).

In order to determine whether over-expression of *PANC754* can induce tumor cell apoptosis, the rate of apoptosis were measured by flow cytometry. As shown in **Figure 3C**, the apoptosis rate of CRC cell line, SW480, was significantly increased following *PANC754* transfection compared to the rates of the control and pcDNA3.1 groups (P=0.0012). Similar patterns of increased apoptosis were observed in the Eca109 (P=0.001), HepG2 (P=0.0045), Bel-7402 (P=0.0018), MGC803 (P<1.0x10-4), and DLD1 (P=5x10-4) cell lines (**Figure S4**).

**Overexpressed PANC754 significantly suppressed the migration and invasion of human tumor cells**

Through a scratch assay experiment, as shown in **Figure 4A and 4B**, the gap distances in the transfected pcDNA3.1-PANC754 group (abbreviated as PANC754) of hepatocellular carcinoma HepG2 cells after culturing for 24 h and 48 h were significantly shorter than those in the control and pcDNA3.1 groups with P<1x10-4 and P<1x10-4, respectively. We observed a similar result in the Eca109, Bel-7402, MGC803, SW480, and DLD1 cells at 48 h culture (**Figure S5-S9**). We also performed a trans-well chamber experiment to determine the effect of PANC754 overexpression on invasion of gastrointestinal tumor cells. After 48 h culture of hepatocellular carcinoma Bel-7402 cells, the PANC754 group was observed to have fewer cells invading the membrane of the transwell chamber compared to the control and pcDNA3.1 groups (P=0.0003, **Figure 4C and 4D**). Furthermore, we also observed a decreased invasion ability of the Eca109 (P=1.64x10-5), HepG2 (P=0.0003), MGC803 (P=0.0001), SW480 (P=0.0012), and DLD1 (P=0.0002) cells (**Figure S10**).

**Pan-cancer suppression role of PANC754** **occurred in a KPNA4- and MTMR8-dependent manner**

In order to investigate the mechanism and regulatory target(s) of PANC754, we conducted an alignment analysis of potential complementary genes to the PANC754 transcript. We identified two genes of high similarity including *KPNA4* (karyopherin alpha 4, also named importin alpha 3) and *MTMR8* (myotubularin related protein 8, also named MTMR9) (**Figure 5**) along with 51 other genes of low similarity to PANC754 (**Table S4**). Based on the sequence complementary theory, these other genes could be targets of *PANC754*27,28. We confirmed these results by conducting an *in silico* co-expression analysis between *PANC754* and *KPNA4* and *MTMR8* as well as the other 51 sequence-similar genes and presented the results as Spearman’s correlation (**Table S4**). As expected, MTMR8 had a significantly positive correlation with PNAC754 in the TCGA pan-cancer dataset (R=0.44, P<2.2x10-16) while *KPNA4* (R= -0.07, P=1.19x10-12), *GPR89A* (G protein-coupled receptor 89A, R= -0.08, P=1.19x10-18), *COPS7B* (COP9 signalosome subunit 7B, R= -0.09, P=1.19x10-18), *PPIE* (peptidylprolyl isomerase E, R= -0.09, P=1.19x10-19), *PMVK* (phosphomevalonate kinase, R= -0.098, P=1.19x10-22), and *SMIM12* (small integral membrane protein 12, R= -0.101, P=1.19x10-24) demonstrated negative correlations with PNAC754 (**Figure S13 and Table S5**).

To validate the effects of PANC754 regulation on downstream protein expression of candidate genes, we analyzed the protein levels of KPNA4 and MTMR8 via Western blot (WB). We found that *PANC754* overexpression significantly decreased protein levels of KPNA429 and MTMR830 compared to those in the control and pcDNA3.1 groups (**Figure 7A-F and Figure S11**). The different magnitude of the down-regulation effect to KPNA4 and MTMR8 proteins were observed in different cells indicating that other cell-specific regulatory factors might be involved (**Figure S11)**. Furthermore, we found that high expression of KPNA4 was a significant factor for increased survival time as indicated by a subset meta-analysis of KPNA4 by overall survival time across all cancer types (HR=1.22 95%CI: 1.12-1.32, P=5.2x10-6 by a fixed effects model and HR=1.24 95%CI: 1.08-1.42, P=1.3x10-3 in a random effects model, **Figure S14 and Table S6**), while MTMR8 expression had no significant association with survival (HR=0.93 95%CI: 0.85-1.01, P=0.099 by fixed effects model and HR=0.93 95%CI: 0.82-1.47, P=0.22 with a random effects model, **Figure S14 and Table S6**). Our results demonstrated that the pan-cancer suppression role of PANC754 might occur in a KPNA4- and MTMR8-dependent manner.

We also investigated the non-direct interaction between PANC754 and several important human cancer onco-genes that are usually over-expressed in cancer. Leveraging the TCGA RNA-sequencing data with oncogene protein expression data, we found that *CDK2* (R = -0.05, P=4.9x10-9), *DNMT1* (R = - 0.08, P=2.5x10-15), *HRAS* (R= -0.1, P=3.9x10-23), and *SRSF2* (Serine/Arginine Rich Splicing Factor 2, R= -0.07, P=1.42x10-11) were negatively correlated with expression of PANC754. We then demonstrated that the mRNA levels of these seven genes were all significantly downregulated after transfecting with *PANC754*-containingplasmids compared to untransfected control and empty vector pcDNA3.1 groups. (P<0.05, <0.03 or <0.01, respectively; **Figure 8 and Figure S12**). Since CDK2 is usually up-regulated in cancer cells, this result indicates that the pan-cancer suppressing role of PANC754 might be involved in the CDK2 pathway. PANC754 is therefore a novel pan-tumor suppressor non-coding RNA that presumably interacts with KPNA4 and MTMR8 via the CDK2 pathway to regulate downstream protein expression and cancer cell phenotype.

**Discussion**

Non-coding RNAs were once considered the “junk” of genomes12. However, a large body of evidence has demonstrated that many ncRNAs are functional elements instead of “junk RNA”31,32. Functional ncRNAs exist in the tumor and play crucial roles during tumorgenesis17,33,34. In this study, we found a novel ubiquitously down-regulated non-coding RNA transcript, *PANC754,* by scanning the transcript files of the TCGA database across 23 human cancer types. ncRNA *PANC754* is encoded by ENSG00000213754.2, which is a proposed RNA of unknown function. The length of the PANC754 transcript is 500 bp, which we mined from GENCODE database [<https://www.gencodegenes.org/>]. Based on the definition of long non-coding RNA (lncRNA), *PANC754* may be a novel lncRNA characterized by pan-cancer suppression and short length. Functional experiments demonstrate that overexpressed *PANC754* could inhibit cell viability, migration, and invasion of gastrointestinal tumors and induces apoptosis in human tumors. *KPNA4* and *MTMR8*, two genes with high sequence similarity, were validated targets of *PANC754* by overexpression experiments. PANC754 also indirectly regulated the expression of vital oncogenes, such as *CDK2*, *DNMT1*, *SRSF2*, and *HRAS*. *PANC745* acts as a novel pan-tumor suppressor ncRNA.

At the genetic level, cancer is caused by oncogene activation, blocking of tumor suppressor gene function, and epigenetic factors. Oncogenes act as the driving force of carcinogenesis; however, tumor-suppressor genes can interact and regulate oncogenes to inhibit cancer occurrence. PANC754 appears to serve as a pan-cancer suppressor based on the results of overexpression experiments. *CDK2* is a cell cycle gene that is over-expressed in multiple human cancers and demonstrated to favor tumorigenesis and tumor development35. Since overexpression of PANC745 can inhibit oncogenic mRNA expression in four types of gastrointestinal tumors, we conclude that that PANC745 is a tumor suppressor lncRNA (**Figure 7A-F**).

DNA replication and nuclear translocation of genes significantly influence the timing, duration, and expression of downstream signaling events. KPNA4 functions as an adapter protein for nuclear receptor KPNB1 and can notably influence subsequent gene replication, transcription and translation 36,37 38. From the ensemble database annotation [<http://asia.ensembl.org/>], we found that PANC754 is the pseudogene of KPNA4. Our data show that overexpression of PANC754 can regulate the expression of KPNA4, which may, through blocking of the importin α/β pathway and/or the nuclear localization signal, lead to cell apoptosis and corresponding inhibition of cancer cell migration and growth. This conclusion is supported in the literature where down-regulation of KPNA4 expression was shown to constrain tumor progression of prostate cancer39, neuroblastoma40, breast cancer41, and cutaneous squamous cell carcinoma42. It is worth noting that our research indicates that PANC754 serves as a pseudogene of KPNA4 which may regulate the expression of KPNA4 itself.

Due to its absent dual-specificity phosphatase domain, MTMR8 is an atypical myotubularin-related protein that acts as an adapter for myotubularin-related phosphatases specifically by interacting with phosphatidylinositol 3,5-bisphosphate43,44, negatively regulates DNA damage-induced apoptosis45 and autophagy46, and plays a role in the late stages of macropinocytosis47. It contains a double-helical motif similar to the SET interaction domain, which can control cell proliferation48. High-copy numbers of *MTMR8* in oesophageal adenocarcinoma are associated with a worse prognosis49. Our study demonstrated that overexpression of PANC754 markedly decreased the expression level of MTMR8 in different cancer cells. Although the levels of KPNA4 protein in SW480 cells was not statistically different between the three groups (P = 0.649) and MTMR8 protein in DLD1 cells was marginal (P=0.069; **Figure S11**), *in silico* data support the hypothesis that KPNA4 and MTMR8 are the regulatory targets of ncRNA PANC754. Future work will focus on identifying the downstream CDK2 pathway-regulated genes impacted by PANC754 and to detect any cell-specific differences in PANC754 expression across cancer types.

**Materials and Methods**

# **Cell lines, plasmids and patients**

Four types of human gastrointestinal tumor cell lines, including esophageal cancer Eca109 cell-line, hepatocellular carcinoma HepG2 and Bel7402 cell-lines, gastric cancer MGC803 cell-line, and CRCSW480 and DLD1 cell lines were preserved by our laboratory. Six cell lines were cultured in Dulbecco’s Modified Eagle Medium (DMEM) or RPMI 1640 (KeyGEN Biotech, China) with 10% fetal bovine serum (FBS) (ExCell, USA) and 1% antibiotic solution (penicillin-streptomycin, 10,000 U/mL) at 37°C in a humidified atmosphere containing 5% CO2. ncRNA PANC754 were synthesized by GenScript Biotech Corp. (Nanjing, China) and cloned in pcDNA3.1 plasmid vector, then transfected with Lipofectamine 3000 transfection agent (Invitrogen, USA) into the above-mentioned cells.

26 CRC and 12 GC patients were enrolled for sample collection from the Department of General Surgery in Affiliated Hospital of Nantong University, Jiangsu Province, China. The study was approved by the institutional review boards of the Affiliated Hospital of Nantong University. Written informed consent was obtained from each participant prior to sample collection.

# **Identification of pan-cancer tumor suppressor non-coding RNA**

We downloaded 11,093 gene expression quantification data derived from RNA-seq data from the TCGA database (<https://portal.gdc.cancer.gov/repository>) on February 24, 2019. The RNA-seq data covered 32 cancer types. Nine cancer types were excluded due to low sample size for control samples (N<=1). Log2-transformed fragments per kilobase of transcript per million mapped reads upper quartile (FPKM-UQ) derived from HTSeq50 was applied for differential gene expression analysis. Bayesian generalized linear model (bayesglm) from ARM package (v1.10-1) was applied for differential gene expression analysis. Metafor package (v2.1-0) was applied for meta-analysis across the 23 cancer types. Cox proportional hazards regression model was applied for survival analysis to the TCGA dataset for overall survival times (R survival package v0.9), and results presented as hazard ratios (HR).

**Detection of cell growth viability by CCK-8 assay**

Cells at a concentration of 5 x103 per well were seeded in 96-well plates and incubated for 24 h. At 24, 48, and 72 h after transfection, cell growth and viability was measured with a Cell Counting Kit-8 (Beyotime, China), following manufacturer’s instructions. Absorbance (A) was then recorded at 450 nm using Multiskan Sky Reader (Thermo Electrom, USA).

**Detection of apoptosis by flow cytometry**

Cells were trypsinized and washed twice in ice-cold phosphate buffered saline (PBS). A total of 5 ×105 cells were resuspended in 500 µL of binding buffer (KeyGen). 5 µL of Annexin V-FITC (KeyGen) was then added to the cell suspension which was counterstained with 5 µL of propidium iodide (PI). The mixture was incubated at room temperature for 10 min in the dark followed by apoptosis analysis using the FlowSight Flow Cytometer (Merck, Germany).

**Determination of cell migration by scratch assay**

We seeded cells into 6-well tissue culture plates at a density of ~70-80% monolayer confluence after 24 h of growth. We gently and slowly scratched the monolayer with a 200 µL pipette tip across the center of the well. Then we scratched another straight line perpendicular to the first line to create a cross in each well. After scratching, the well was gently washed twice with medium to remove the detached cells and the gap distance measured at 0 h. To determine the degree of wound healing, the gap distance remaining was measured at 24 h and 48 h post-scratch. Migration distance was calculated as follows: Migration distance = gap distance at 0 h - gap distance at t time (t = 24 h or 48 h).

**Determination of cell invasion by transwell chamber assay**

We performed the cell invasion assay by using an invasion chamber (Chemicon, USA), which is a 24-well tissue culture plate with 12-cell culture inserts. Cell suspensions (0.5×105 cells/mL) transfected by pcDNA3.1-PANC754, untransfected control, and pcDNA3.1 were added to the interior of the inserts with 200 µL serum-free media while 500 µL of media containing 10% FBS was added to the lower chamber as a migration stimulus. The transwell chambers were incubated in a tissue culture incubator for 24 h. The invasive cells on lower surface of the membrane were stained by crystal violet for 10 min and images taken at X number of high magnification fields (HMF) of the microscope.

**RT-PCR and real-time qPCR**

Total RNAs were harvested using TRIzol (Invitrogen, USA) and converted to cDNA using the First-Strand cDNA Synthesis Kit (Vazyme Biotech, China) according to the manufacturer’s instructions. Primer sequences are shown in Supplementary Table S5. Real-time qPCR was performed via SYBR Green Master Mix (Vazyme Biotech) in an ABI 7500 PCR instrument (ABI, USA). As discussed in Jin et al., *GAPDH* was used as the internal reference for ncRNA detection51.

**Sequence Characteristics Analysis**

We conducted sequence alignment to exomes 1 and 2 of PANC754 to 55,270,388 non-redundant sequences collected from GenBank, EMBL, DDBJ, PDB, and RefSeq database with NCBI [blastn suite](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=MegaBlast&PROGRAM=blastn&PAGE_TYPE=BlastSearch&BLAST_SPEC=) tools with default [algorithm parameters](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=MegaBlast&PROGRAM=blastn&PAGE_TYPE=BlastSearch&BLAST_SPEC=#i). Alignment sequences were selected to show only human records (taxid: 9606). T-Coffee was applied for multiple sequence alignment with the default settings52. Slanted cladogram was used to depict the distance i.e., similarity between PNAC754 with significantly similar DNA sequences including KPNA4, MTMR8, and RP11-284B18.3. We also conducted co-expression analysis with Spearman correlation between PANC754 and 51 other genes of low similarity, such as *AFF3*, *DOCK7*, *EMC3*, ITGA6, EXPH5, EIF4G3, ARID4B, and TMCC2 **(Table S4).**

**Western Blotting**

Nuclear and total proteins were extracted by nucleated and total protein extraction kit (Beyotime, China) including PMSF proteinase inhibitor (Beyotime) and quantitated by BCA protein concentration assay kit (Beyotime).

1 µg nuclear protein or total proteins from each group were separated by 10% sodium dodecyl sulfate/polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to 0.45 μm polyvinylidene fluoride (PVDF) membrane (Millipore, USA). Following blocking with 5% non-fat milk (Beyotime), membranes were incubated with anti-KPNA4 (R&D, USA), anti-MTMR8 (Novus bio., USA), anti-Lamin B (ABclonal, USA) and anti-GAPDH (Proteintech, USA) overnight at 4 °C and then incubated with HRP-conjugated secondary antibodies (Abcam, USA) for 1 h at room temperature. Finally, protein bands were detected using enhanced chemiluminescence (ECL) detection kit (Beyotime, China) and developed on x-ray film. The band grey-intensities were analyzed via Gel-Pro 32 software with the default settings. Lamin B and GAPDH were used as the loading controls for KPNA4 and MTMR8 normalization, respectively. Normalized band densities are presented as ratios of the relative band density of the target protein: relative band density of the loading control.

**Statistical Analysis**

High-throughput differential gene expression and meta-analysis were conducted using R software package (version 3.5.1). Data are presented as the mean ± standard deviation (SD). A two-tailed p-value of <0.05 was used as the threshold for statistical significance, and statistical analyses were performed usingSPSS 20.0.

**Acknowledgements**

This work was supported by the National Natural Science Foundation of China (81974313), Postdoctoral Science Foundation of China (2019M651930), Jiangsu Province's Key Young Medicine Talents Program (QNRC2016688), and Nantong People's Livelihood Science and Technology Plan (MS12018032). The authors gratefully acknowledge Emily A. Andreae, PhD, for reviewing and editing their manuscript.

**Author contribution**

SG designed the study, performed analyses, interpreted the results, and drafted the manuscript. JS conducted molecular and cell biology experiments. SJ aided in drafting and editing the manuscript and advising on the study. SJS aided in drafting and editing the manuscript and supervised the analyses and results interpretation. SJS also designed and monitored the study, contributed to the drafting of this manuscript, and interpreted the results. WD designed the study, conducted molecular and cell biology experiments, performed analyses, interpreted results, designed the functional experiments, and contributed to the drafting of this manuscript.

**Disclosure of Conflicts of Interest**

The authors declare no conflict of interest.

**Data and Code Available**

All related software and script used in the manuscript are available through GitHub at: [https://github.com/Shicheng-Guo/HowtoBook/tree/master/PANC.](https://github.com/Shicheng-Guo/HowtoBook/tree/master/PANC.%20)

**Website for data and R package:**

TCGA: <https://portal.gdc.cancer.gov/repository>

Metafor: <https://cran.r-project.org/web/packages/metafor/index.html>

**Abbreviations**

PANC754: PAn-cancer Non-Coding RNA encoded byENSG00000213754.2

CESC: cervical squamous cell carcinoma and endocervical adenocarcinoma

CHOL: cholangiocarcinoma

HNSC: head-neck squamous cell carcinoma

TCGA: The Cancer Genome Atlas

KPNA4: Karyopherin Alpha 4, also named Importin Alpha 3

MTMR8: Myotubularin Related Protein 8, also named MTMR9

HR: Hazard ratio

**References**

1. Siegel, R.L., Miller, K.D. & Jemal, A. Cancer statistics, 2019. *CA Cancer J Clin* **69**, 7-34 (2019).

2. Chen, W. *et al.* Cancer statistics in China, 2015. *CA Cancer J Clin* **66**, 115-32 (2016).

3. Xu, X.H. *et al.* Hypoxic-stabilized EPAS1 proteins transactivate DNMT1 and cause promoter hypermethylation and transcription inhibition of EPAS1 in non-small cell lung cancer. *FASEB J*, fj201700715 (2018).

4. Campaner, S. *et al.* Cdk2 suppresses cellular senescence induced by the c-myc oncogene. *Nat Cell Biol* **12**, 54-9; sup pp 1-14 (2010).

5. Menke, J. *et al.* Uniparental Trisomy of a Mutated HRAS Proto-Oncogene in Embryonal Rhabdomyosarcoma of a Patient With Costello Syndrome. *J Clin Oncol* **33**, e62-5 (2015).

6. Lei, Y. *et al.* Hdac7 promotes lung tumorigenesis by inhibiting Stat3 activation. *Mol Cancer* **16**, 170 (2017).

7. Pu, W. *et al.* Aberrant methylation of CDH13 can be a diagnostic biomarker for lung adenocarcinoma. *J Cancer* **7**, 2280-2289 (2016).

8. Jiang, D. *et al.* Epigenetic silencing of ZNF132 mediated by methylation-sensitive Sp1 binding promotes cancer progression in esophageal squamous cell carcinoma. *Cell Death Dis* **10**, 1 (2018).

9. Priestley, P. *et al.* Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature* **575**, 210-216 (2019).

10. Lee, S.H. *et al.* Widespread intronic polyadenylation inactivates tumour suppressor genes in leukaemia. *Nature* **561**, 127-131 (2018).

11. Tsherniak, A. *et al.* Defining a Cancer Dependency Map. *Cell* **170**, 564-576 e16 (2017).

12. Palazzo, A.F. & Lee, E.S. Non-coding RNA: what is functional and what is junk? *Front Genet* **6**, 2 (2015).

13. Zhou, J. *et al.* Whole-genome deep-learning analysis identifies contribution of noncoding mutations to autism risk. *Nat Genet* **51**, 973-980 (2019).

14. Slack, F.J. & Chinnaiyan, A.M. The Role of Non-coding RNAs in Oncology. *Cell* **179**, 1033-1055 (2019).

15. Sarropoulos, I., Marin, R., Cardoso-Moreira, M. & Kaessmann, H. Developmental dynamics of lncRNAs across mammalian organs and species. *Nature* **571**, 510-514 (2019).

16. Matsui, M. & Corey, D.R. Non-coding RNAs as drug targets. *Nat Rev Drug Discov* **16**, 167-179 (2017).

17. Kim, J. *et al.* Long noncoding RNA MALAT1 suppresses breast cancer metastasis. *Nat Genet* **50**, 1705-1715 (2018).

18. Anastasiadou, E., Jacob, L.S. & Slack, F.J. Non-coding RNA networks in cancer. *Nat Rev Cancer* **18**, 5-18 (2018).

19. Varn, F.S., Wang, Y., Mullins, D.W., Fiering, S. & Cheng, C. Systematic Pan-Cancer Analysis Reveals Immune Cell Interactions in the Tumor Microenvironment. *Cancer Res* **77**, 1271-1282 (2017).

20. Kumar, S. & Weaver, V.M. Mechanics, malignancy, and metastasis: the force journey of a tumor cell. *Cancer Metastasis Rev* **28**, 113-27 (2009).

21. Sanchez-Vega, F. *et al.* Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell* **173**, 321-337 e10 (2018).

22. Saghafinia, S., Mina, M., Riggi, N., Hanahan, D. & Ciriello, G. Pan-Cancer Landscape of Aberrant DNA Methylation across Human Tumors. *Cell Rep* **25**, 1066-1080 e8 (2018).

23. Bennett, R.L. & Licht, J.D. Targeting Epigenetics in Cancer. *Annu Rev Pharmacol Toxicol* **58**, 187-207 (2018).

24. Bailey, M.H. *et al.* Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* **173**, 371-385.e18 (2018).

25. Zack, T.I. *et al.* Pan-cancer patterns of somatic copy number alteration. *Nat Genet* **45**, 1134-40 (2013).

26. Lawrence, M.S. *et al.* Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* **505**, 495-501 (2014).

27. Esposito, R. *et al.* Oncogenic Properties of the Antisense lncRNA COMET in BRAF- and RET-Driven Papillary Thyroid Carcinomas. *Cancer Res* **79**, 2124-2135 (2019).

28. Canzio, D. *et al.* Antisense lncRNA Transcription Mediates DNA Demethylation to Drive Stochastic Protocadherin alpha Promoter Choice. *Cell* **177**, 639-653.e15 (2019).

29. Sun, X. *et al.* MicroRNA-181b regulates NF-kappaB-mediated vascular inflammation. *J Clin Invest* **122**, 1973-90 (2012).

30. Mochizuki, Y. & Majerus, P.W. Characterization of myotubularin-related protein 7 and its binding partner, myotubularin-related protein 9. *Proc Natl Acad Sci U S A* **100**, 9768-73 (2003).

31. Hu, Y. *et al.* Long noncoding RNA GAPLINC regulates CD44-dependent cell invasiveness and associates with poor prognosis of gastric cancer. *Cancer Res* **74**, 6890-902 (2014).

32. Gupta, R.A. *et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**, 1071-6 (2010).

33. Peng, L. *et al.* Super-Enhancer-Associated Long Noncoding RNA HCCL5 Is Activated by ZEB1 and Promotes the Malignancy of Hepatocellular Carcinoma. *Cancer Res* **79**, 572-584 (2019).

34. Hosono, Y. *et al.* Oncogenic Role of THOR, a Conserved Cancer/Testis Long Non-coding RNA. *Cell* **171**, 1559-1572 e20 (2017).

35. Lapenna, S. & Giordano, A. Cell cycle kinases as therapeutic targets for cancer. *Nat Rev Drug Discov* **8**, 547-66 (2009).

36. Ota, T. *et al.* Complete sequencing and characterization of 21,243 full-length human cDNAs. *Nat Genet* **36**, 40-5 (2004).

37. Kohler, M. *et al.* Cloning of two novel human importin-alpha subunits and analysis of the expression pattern of the importin-alpha protein family. *FEBS Lett* **417**, 104-8 (1997).

38. Aratani, S. *et al.* The nuclear import of RNA helicase A is mediated by importin-alpha3. *Biochem Biophys Res Commun* **340**, 125-33 (2006).

39. Yang, J. *et al.* Inhibition of KPNA4 attenuates prostate cancer metastasis. *Oncogene* **36**, 2868-2878 (2017).

40. Menard, M. *et al.* Hey1- and p53-dependent TrkC proapoptotic activity controls neuroblastoma growth. *PLoS Biol* **16**, e2002912 (2018).

41. Bertoli, G. *et al.* MicroRNA-567 dysregulation contributes to carcinogenesis of breast cancer, targeting tumor cell proliferation, and migration. *Breast Cancer Res Treat* **161**, 605-616 (2017).

42. Zhang, M., Luo, H. & Hui, L. MiR-3619-5p hampers proliferation and cisplatin resistance in cutaneous squamous-cell carcinoma via KPNA4. *Biochem Biophys Res Commun* **513**, 419-425 (2019).

43. Huttlin, E.L. *et al.* Architecture of the human interactome defines protein communities and disease networks. *Nature* **545**, 505-509 (2017).

44. St-Denis, N. *et al.* Phenotypic and Interaction Profiling of the Human Phosphatases Identifies Diverse Mitotic Regulators. *Cell Rep* **17**, 2488-2501 (2016).

45. Zou, J., Chang, S.C., Marjanovic, J. & Majerus, P.W. MTMR9 increases MTMR6 enzyme activity, stability, and role in apoptosis. *J Biol Chem* **284**, 2064-71 (2009).

46. Zou, J. *et al.* Myotubularin-related protein (MTMR) 9 determines the enzymatic activity, substrate specificity, and role in autophagy of MTMR8. *Proc Natl Acad Sci U S A* **109**, 9539-44 (2012).

47. Maekawa, M. *et al.* Sequential breakdown of 3-phosphorylated phosphoinositides is essential for the completion of macropinocytosis. *Proc Natl Acad Sci U S A* **111**, E978-87 (2014).

48. Appel, S. *et al.* Identification and localization of a new human myotubularin-related protein gene, mtmr8, on 8p22-p23. *Genomics* **75**, 6-8 (2001).

49. Goh, X.Y. *et al.* Integrative analysis of array-comparative genomic hybridisation and matched gene expression profiling data reveals novel genes with prognostic significance in oesophageal adenocarcinoma. *Gut* **60**, 1317-26 (2011).

50. Anders, S., Pyl, P.T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-9 (2015).

51. Jin, C. *et al.* Long non-coding RNA HULC as a novel serum biomarker for diagnosis and prognosis prediction of gastric cancer. *Oncotarget* **7**, 51763-51772 (2016).

52. Notredame, C., Higgins, D.G. & Heringa, J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**, 205-17 (2000).

**Figure Legends**

**Figure 1. Pan-cancer down-regulation of PANC754 cross 23 human cancers.** Cancer samples were collected from the TCGA project (N=10,490). Gene expression level was log2 transformed before the meta-analysis. Both fixed effect and random effect models were applied for the aggregation. 95% confidence interval (CI) was applied to show the risk and protective effect to overall survival time. In order to show more details for different studies, any standardized mean difference (SMD) higher than 3 and lower than -3 was denoted with an arrow. Blue-filled parallelograms represent the SMD for the fixed effect and random effect models. The expression levels of *PANC754* were validated in tumor tissues and paratumorous tissues of 26 paired CRC **(B)** and 12 paired GC patients **(C)** by quantitative PCR.

**Figure 2. Forest plot to show the Cox regression based survival analysis.** Cancer samples were collected from the TCGA project (N=8,074). Both fixed effect and random effect models were applied for the aggregation. 95% confidence interval was applied to show the risk and protective effect to overall survival time. In order to show more details for different studies, any hazard ratio (HR) higher than 3 and lower than 0.1 was denoted with an arrow. Blue-filled parallelograms represent the HR for fixed effect and random effect models.

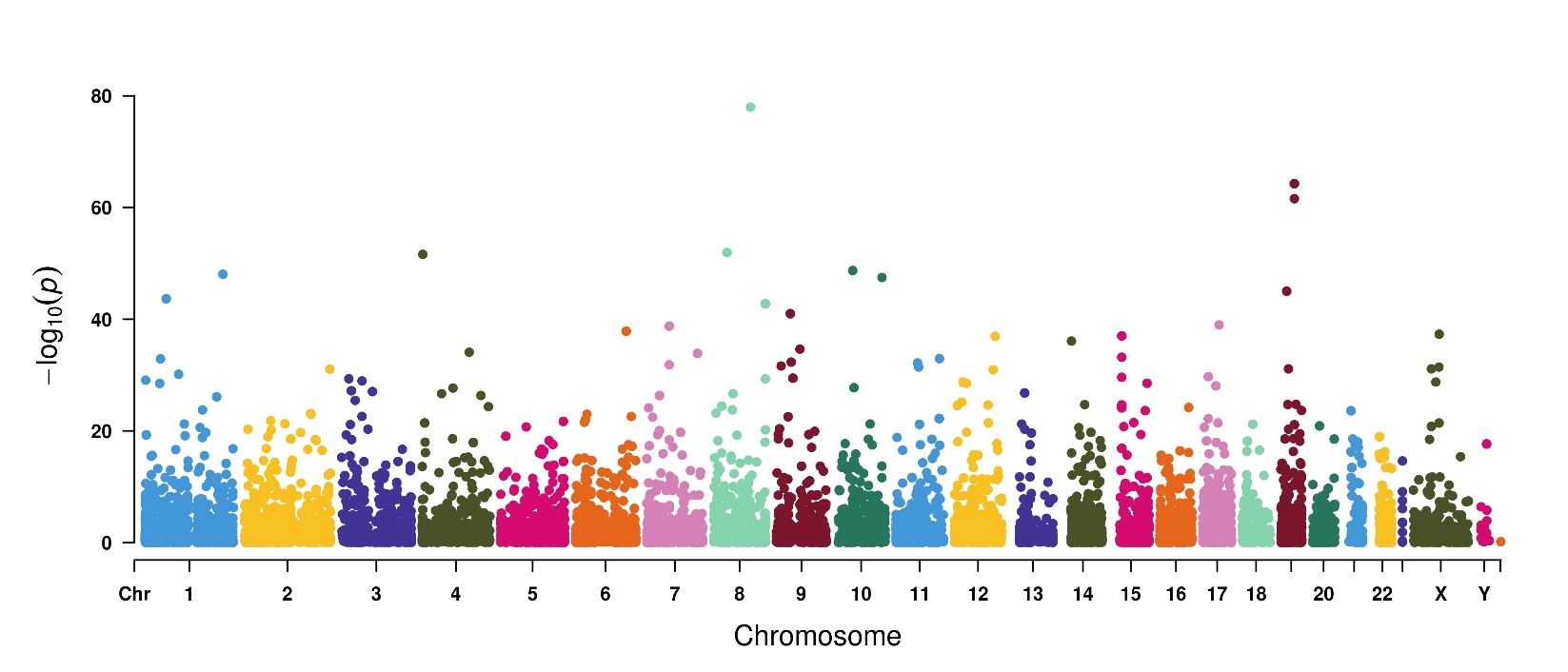
**Figure 3. PANC754 significantly suppressed tumor cell growth and induced apoptosis.**PANC754 significantly suppressed the cell growth of esophageal cancer Eca109 cells **(A)** and gastric cancer MGC803 cells **(B)** by CCK-8 assay. **(C)** PANC754 induced cell apoptosis of CRC cell lineSW480 by flow cytometry analysis. Control, untransfected group; pcDNA3.1, transfected with empty pcDNA3.1 plasmid group; PANC754, transfected with pcDNA3.1-PANC754 overexpression plasmid group; \*\*\*, P<0.01 (PANC754 versus control and pcDNA3.1).

**Figure 4. PANC754 significantly inhibited hepatocellular carcinoma HepG2 cell migration and Bel7402 cell invasion.** **(A)** Scratch assay of PANC754 overexpressed cells and corresponding **(B)** statistical histogram of (A). **(C)** Representative images of the invasion assay of PANC754 overexpressed cells and corresponding **(D)** statistical histogram of (C). Control, untransfected group; pcDNA3.1, transfected with empty pcDNA3.1 plasmid group; PANC754, transfected with pcDNA3.1-PANC754 overexpression plasmid group; \*\*\*, P<0.01 (PANC754 VS Control and pcDNA3.1).

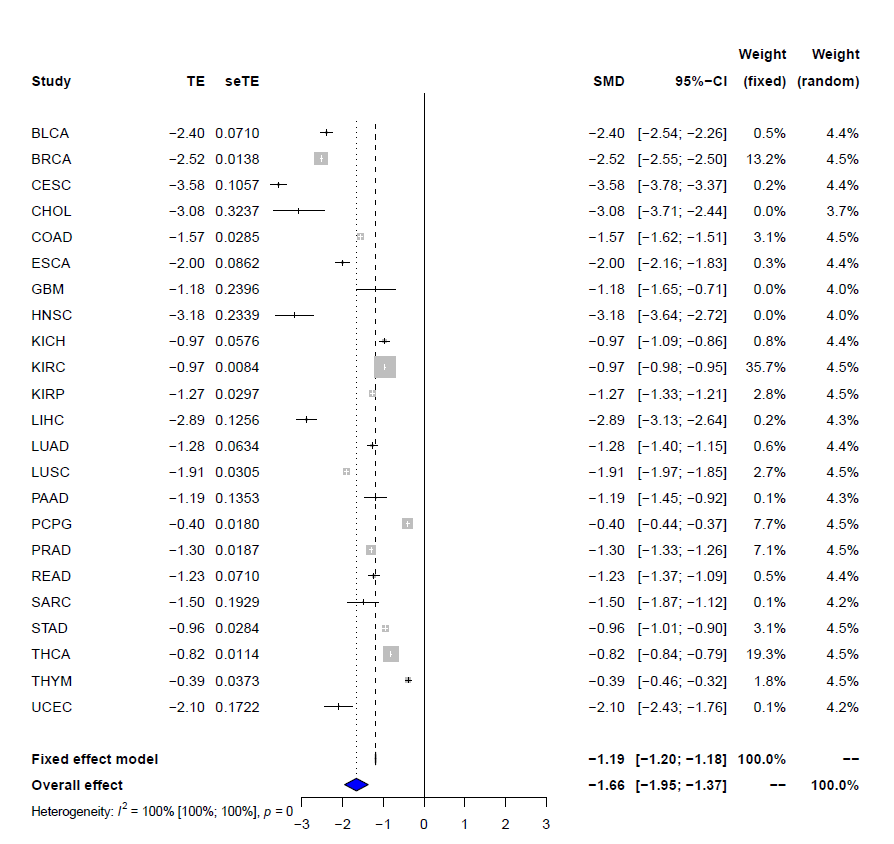
**Figure 5. The regulation targets of PANC754 were investigated according to complementary sequence.(A)** the sequence alignment analysis of PANC754 to 55,270,388 non-redundant sequences collected from GenBank, EMBL, DDBJ, PDB and RefSeq with NCBI [blastn suite](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=MegaBlast&PROGRAM=blastn&PAGE_TYPE=BlastSearch&BLAST_SPEC=) tools with default [algorithm parameters](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=MegaBlast&PROGRAM=blastn&PAGE_TYPE=BlastSearch&BLAST_SPEC=#i). **(B)** Slanted cladogram was conducted to display the distance or similarity between PANC754 and KPNA4, MTMR8 or others. **(C)** The heatmap of the co-expression analysis was performed to discover the regulated targets of antisense ncRNA *PANC754* based on the sequence complementary theory.

**Figure 6. Western blotting experiments confirmed the potential targets of PANC754.** Nuclear protein of KPNA4 and total protein of MTMR8 were detected by Western blotting in esophageal cancer cell line Eca109 **(C)**, hepatocellular carcinoma HepG2 **(D)** and Bel-7402 **(A)** cell lines, gastric cancer MGC803 cell line **(E),** and CRC cell lines SW480 **(F)** and DLD1 **(B)**, respectively. Lamin B served as the loading control for the nuclear protein KPNA4 while GAPDH served as the loading control forthe cytoplastic protein MTMR8; the number under each band represents the normalized ratio of relative band intensity of the target protein/ the relative band intensity of control group; Control, untransfected group; pcDNA3.1, transfected with empty pcDNA3.1 plasmid group; PANC754, transfected with pcDNA3.1-PANC754 overexpression plasmid group.

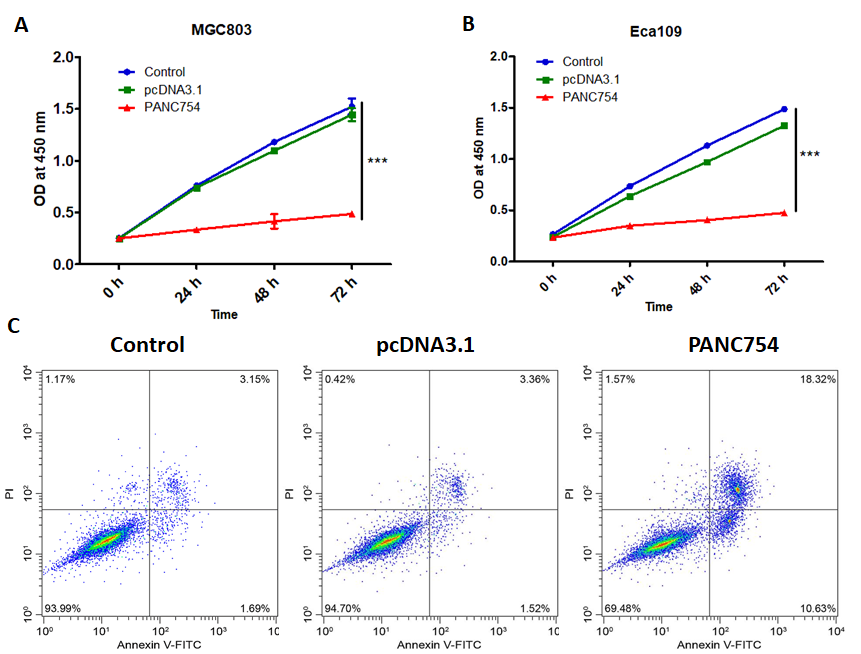
**Figure 7. PANC754 significantly down-regulated the mRNA expression of oncogenes by QPCR detection.** The mRNA levels of CDK2 **(A)**, CSMD3 **(B)**, KRAS **(C)**, MUC16 **(D)**, TTN **(E)**, and PKI3CA **(F)** oncogenes were detected by real-time quantitative PCR (QPCR). Control, untransfected group; pcDNA3.1, transfected with empty pcDNA3.1 plasmid group; PANC754, transfected with pcDNA3.1-PANC754 overexpression plasmid group; \*, P<0.05; \*\*, P<0.03; \*\*\*, P<0.01 (PANC754 VS Control and pcDNA3.1).



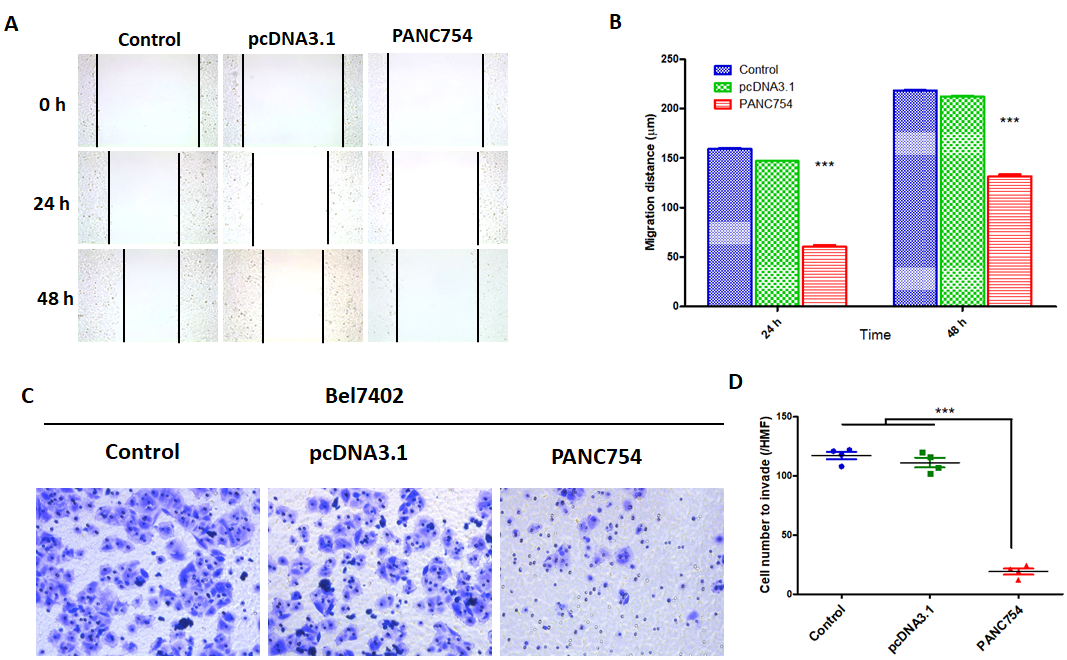
**Figure 1. Manhattan plot of a genome-wide meta-analysis to identify abnormally regulated ncRNAs in a TCGA pan-cancer RNA-seq dataset.** P-values from a random-effect model were applied to identify ncRNAs that were associated with statistically significant changes in gene expression.



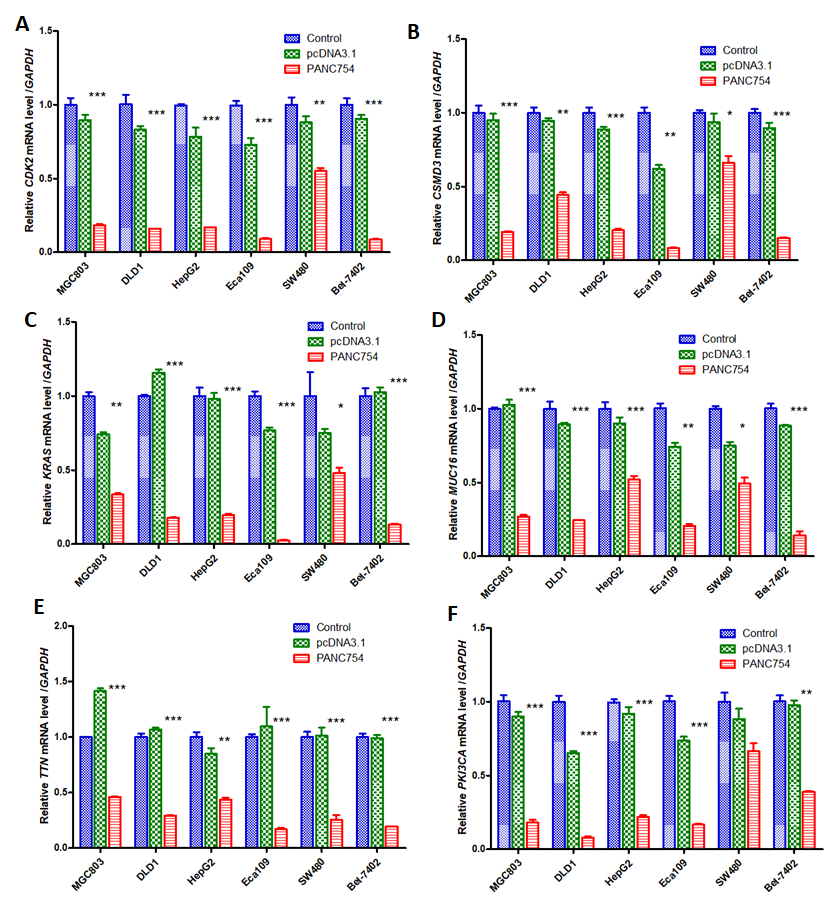
**Figure 2**. **Pan-cancer down-regulation of PANC754 across 23 human cancers.** Gene expression data from tumor samples were collected from the TCGA project (N=10,490). Gene expression level was log2 transformed before the meta-analysis. Both fixed effect random effect models were applied for analysis of the aggregated data, and 95% CI was applied to show the risk and protective effect of gene expression to overall survival time. In order to show more details for different studies, any standardized mean difference (SMD) higher than 3 and lower than -3 was denoted with an arrow, and the blue filled parallelograms represent the SMD for fixed effect model and random effect model. Abbreviations: TE=tissue expression, seTE=standardized tissue expression.



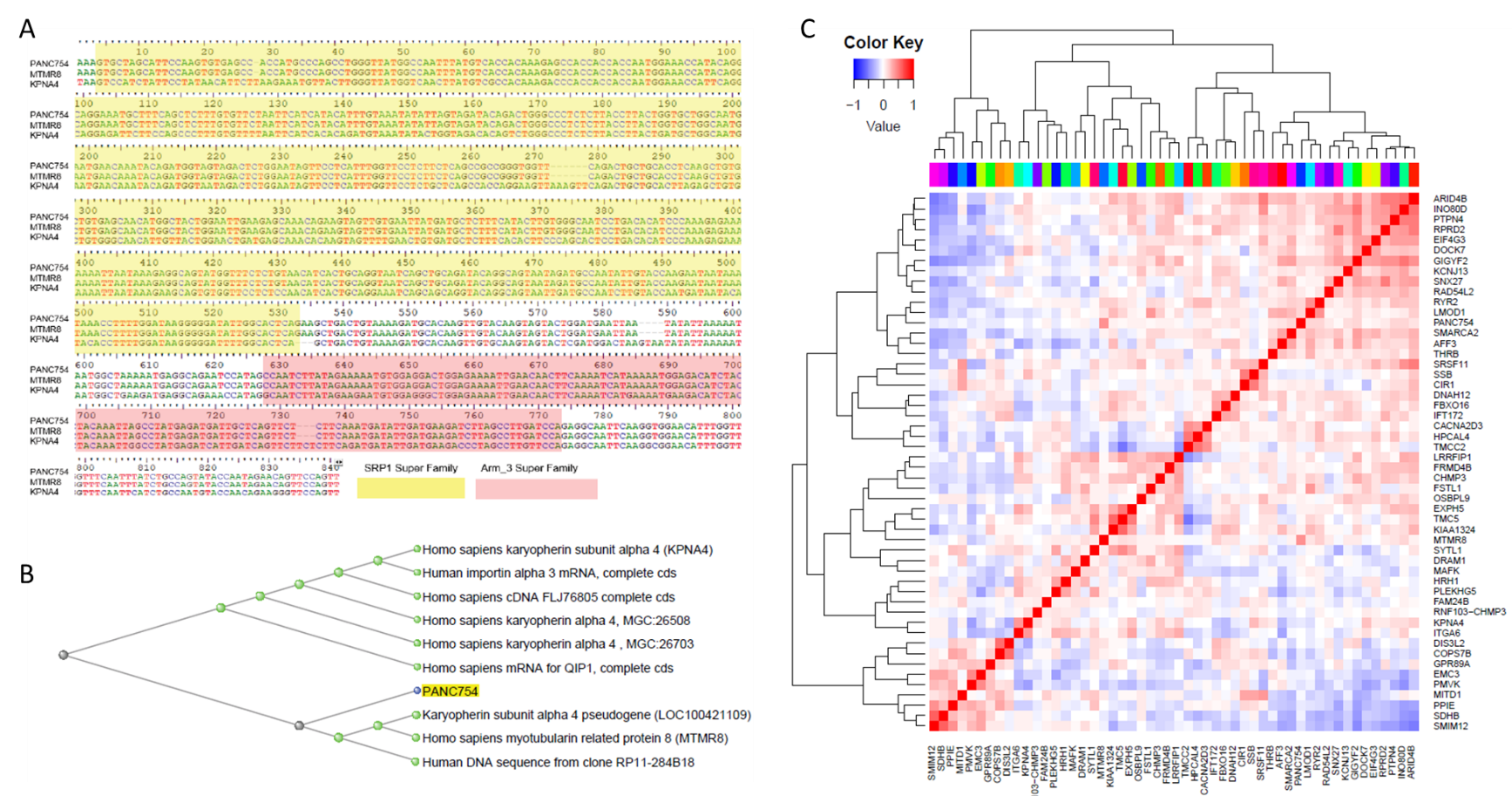
**Figure 3. PANC754 significantly suppressed cell growth and induced apoptosis of tumor cells** PANC754 significantly suppressed cell growth of esophageal cancer Eca109 cells (A) and gastric cancer MGC803 cells (B) by CCK-8 assay. (C) PANC754 induced cell apoptosis of colorectal cancer cell line SW480 by detection of flow cytometry. Control, untransfected group; pcDNA3.1, transfected with empty pcDNA3.1 plasmid group; PANC754, transfected with pcDNA3.1-PANC754 overexpression plasmid group; \*\*\*, P<0.01.

****

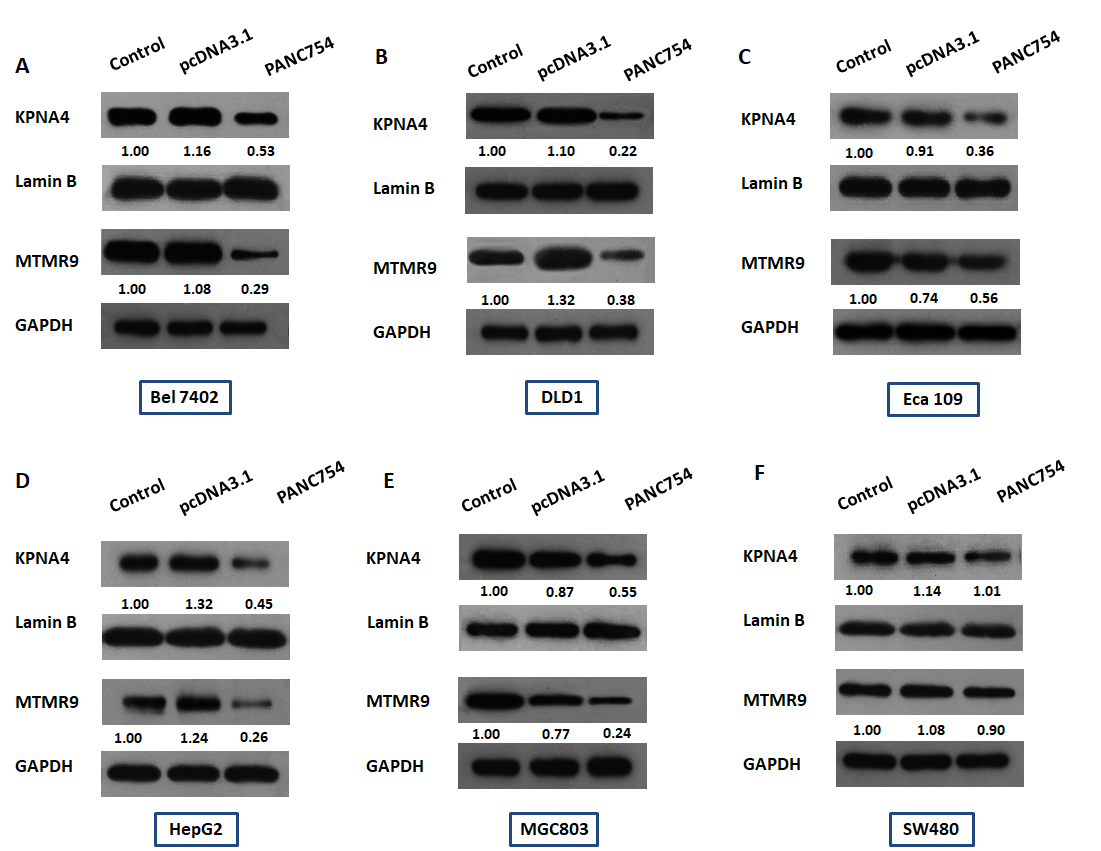
**Figure 4. PANC754 significantly inhibited hepatocellular carcinoma HepG2 cell migration and Bel7402 cell invasion.** (A) Scratch assay for untransfected, empty vector, and PANC754 overexpressed cells; (B) Statistical histogram of (A); (C) a representative invasion assay of untransfected, empty vector, and PANC754 overexpressed cells. (D) Statistical histogram of (C). Control, untransfected group; pcDNA3.1, transfected with empty pcDNA3.1 plasmid group; PANC754, transfected with pcDNA3.1-PANC754 overexpression plasmid group; \*\*\*, P<0.01.



**Figure 6. PANC754 significantly down-regulated the mRNA expression of oncogenes by QPCR detection** (A) CDK2; (B) CSMD3; (C) KRAS; (D) MUC16; (E) ; and (F) PKI3CA mRNA levels were compared between untransfected, empty vector, and PANC754 overexpressed cells. Control, untransfected group; pcDNA3.1, transfected with empty pcDNA3.1 plasmid group; PANC754, transfected with pcDNA3.1-PANC754 overexpression plasmid group; \*, P<0.05; \*\*, P<0.03; \*\*\*, P<0.01.



**Figure 5. The regulation targets of PANC754 were investigated based on the sequence complementary theory.** (A) Sequence alignment analysis of *PANC754* to 55,270,388 non-redundant sequences collected from GenBank, EMBL, DDBJ, PDB, and RefSeq databases with NCBI [blastn suite](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=MegaBlast&PROGRAM=blastn&PAGE_TYPE=BlastSearch&BLAST_SPEC=) tools with default [algorithm parameters](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=MegaBlast&PROGRAM=blastn&PAGE_TYPE=BlastSearch&BLAST_SPEC=#i). (B) A slanted cladogram displays the distance i.e., degree of similarity between PANC754 and KPNA4, MTMR8, and other genes with complementary sequences. (C) A heatmap of the co-expression analysis depicts the degree of association between PANC754 overexpression and potential gene targets.



**Figure 7. Western blot analysis of nuclear extracts and lysates from transfected tumor cell lines confirmed potential targets of PANC754.** Nuclear protein of KPNA4 and total protein of MTMR8 were detected by Western blot in esophageal cancer Eca109 cell-line **(C)**, hepatocellular carcinoma HepG2 **(D)** and Bel-7402 **(A)** cell-lines, gastric cancer MGC803 cell-line **(E),** and colorectal cancer cell lines SW480 **(F)** and DLD1 **(B)**, respectively. Lamin B was the loading control for nuclear protein KPNA4 while GAPDH served as the loading control of the cytoplasmic protein MTMR8. Number under the band represent the normalized ratio of band intensity of the target protein: band intensity of the loading control. Control, untransfected group; pcDNA3.1, transfected with empty pcDNA3.1 plasmid group; PANC754, transfected with pcDNA3.1-PANC754 overexpression plasmid group.