YMTHE, Volume *30*

Supplemental Information

The establishment of CDK9/RNA PolII/H3K4me3/DNA methylation feedback promotes HOTAIR expression by RNA elongation enhancement in cancer

Chi Hin Wong, Chi Han Li, Joanna Hung Man Tong, Duo Zheng, Qifang He, Zhiyuan Luo, Ut Kei Lou, Jiatong Wang, Ka-Fai To, and Yangchao Chen

Supplementary Methods and Materials

Microarray analysis

Microarray analysis of lncRNA expression on 4 pairs of PDAC tumor samples was performed on Arraystar Human LncRNA Microarray V4.0 platform. Microarray data are available at ArrayExpress E-MTAB-7305.

siRNA transfection

SiRNAs targeting HOTAIR, CDK2, CDK7, CDK9, PP1, DNMT1, DNMT3A, DNMT3B and MLL1 were purchased from GenePharma (China) and were dissolved in siRNA buffer (Thermo Fisher Scientific). For transient knockdown, cells were transfected with siRNAs using Lipofectamine 3000 transfection reagent (Thermo Fisher Scientific) for 72 h. Cells were then collected for the experiments as described. The efficiency of siRNA knockdown was validated by qRT-PCR and immunoblotting. The siRNA sequences are listed in Table S1.

Luciferase assay

Luciferase reporter plasmid was constructed by cloning 500 bp, 1 kb and 2 kb upstream of HOTAIR transcription start site into pGL3-basic plasmid. HEK293 cells were transfected with the luciferase reporter plasmid, control vector (pGL3-Control) and Renilla luciferase reporter plasmid using Lipofectamine 3000 transfection reagent and P3000 Reagent (Invitrogen). Luciferase activity was analyzed by Dual Luciferase Reporter Assay kit (Promega, Madison, WI, USA).

Quantitative reverse transcription PCR (qRT-PCR)

Total RNA from cell lines was extracted by TRIzol Reagent (Invitrogen). RNA was isolated from formalin-fixed paraffin-embedded (FFPE) tissue samples by miRNeasy FFPE Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. Measurement of gene expression level was performed by qRT-PCR. cDNA was synthesized using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) according to manufacturer's instructions. Quantitative PCR was performed by ABI 7900HT RealTime PCR system (Applied Biosystems, Waltham, MA, USA) using SYBR Green PCR Master Mix (Applied Biosystems). The PCR primer sequences are listed in Table S2.

In Situ Hybridization

In Situ Hybridization of HOTAIR in PDAC FFPE sections was performed as described previously. Briefly, HOTAIR-specific probes (Probe sequence: 5'-TAAGTCTAGGAATCAGCACGAA-3', which was labelled with Digoxigenin (DIG)

(Exiqon, Woburn, MA, USA) at both ends, was applied to label HOTAIR in PDAC FFPE sections after deparaffinization and antigen retrival. The DIG-labelled sections were incubaled with alkaline phosphatase–conjugated anti-DIG antibodies (Roche, Basel, Switzerland) and the signal was developed bynitroblue tetrazolium/5-bromo-4-chloro-3-inodolyl phosphate (NBT/BCIP) (Roche).

MTT cell viability assay

A 3-(4, 5-dimethylthiazol-2-yl)-2, 5-diphenyltetrazolium bromide MTT assay was performed as previously described to measure cell viability.[2] After siRNA transfection, cells were incubated with 0.65 mg/mL MTT at 37°C for 2 h. DMSO was used to dissolve formazan crystals and absorbance was measured at 595 nm.

Colony formation assay

Anchorage- independent colony formation assay was performed as described previously to study the transformation ability of HPDE when HOTAIR was ectopically expressed.[2]

Chromatin immunoprecipitation (ChIP)

ChIP was performed using EZ-Magna ChIP HiSens kit (Millipore, Burlington, MA, USA) according to the manufacturer's protocol. Briefly, cells were cross-linked with 1% formaldehyde for 10 min and the reaction was stopped with 125 mM glycine. Chromatin was then isolated and sonicated into fragments of 200-600 bp. Cross-linked chromatin was incubated at 4°C overnight anti-IgG (negative control), anti-RNA Polymerase II (Millipore, 05-623), anti-RNA Polymerase II Phosphorylated Ser2 (Covance, Princeton, NJ, USA, MMS-129R), anti-MLL1 antibody (Millipore, ABE240), anti-H3K4me3 antibody (Millipore, CS200580), anti-DNMT1 antibody (Cell Signaling Technology, Danvers, MA, USA, 5119), anti-DNMT3A antibody (Cell Signaling Technology, 2160), anti-DNMT3B antibody (Cell Signaling Technology, 2161), anti-CDK9 (phosphor T186) (Cell Signaling Technology, 2459). The precipitated DNA was quantitated by qPCR.

Nuclear Run-on assay

Nuclear run-on assay was performed according to reported protocols with modifications.[63] Cells were washed twice with ice-cold PBS, scraped and pelleted by centrifugation at 800g at 4°C for 10 mins. Nuclei were isolated by NP-40 lysis buffer with proteinase inhibitors cocktail (Roche) and phosphatase inhibitor cocktail (Thermo Fisher Scientific), followed by centrifugation at 800g at 4°C for 5 mins, and was resuspended in 120 μL nuclei storage buffer (40% glycerol, 5mM $MgCl_2$, 0.1mM EDTA

and 50mM Tris, pH 8.0). Nuclear Run-on assay was performed by incubating the nuclei at 37°C for 30 mins in transcription buffer (2.5 U RNase inhibitor (Applied Biosystem), 7 % glycerol 5mM $MgCl_2$, 75mM KCl, 50µM EDTA, 25mM Tris and 25mM HEPES, pH 7.5) containing nucleoside triphosphates (NTPs, 0.35 mM ATP, GTP and CTP, 0.4 µM UTP) and 0.25 mM DIG-labelled uridine 5'-triphosphate (UTP) (Roche), followed by RNA isolation by TRIZOL reagent. Nascent transcribed RNA was analyzed by either dot-blot or qPCR. For dot-blot analysis, RNA probe specific to HOTAIR exons and GAPDH were blotted on the positively charged nylon membrane (Roche) for 1 h. The blot was washed with SSC buffer (0.3 M NaCl, 30mM sodium citrate, pH7), followed by incubation at 80°C for 2 h. Then, the blot was incubated with hybridization buffer (50% formamide, 10X Denhardt's solution, 0.2M NaCl, 1mM EDTA, 20mM Tris, pH8.0) at 65°C for 1 h. Before hybridization, DIG-labelled RNA was heated at 95°C for 10 mins to degrade RNA into small fragments. RNA sample was hybridized to RNA probe overnight. After hybridization, the blot was washed three times with SSC buffer and was incubated with anti-DIG antibody (Abcam, Cambridge, UK, ab119345) for 1 h at room temperature. The blot was washed with SSC buffer and the hybridized RNA was visualized by Novex AP Chemiluminescent Substrate (CDP-Star) (Thermo Fisher). For qPCR analysis, DIG-labelled RNA was immunoprecipitated with anti-DIG antibody labelled protein A/G magnetic beads according to previous report. Briefly, the protein A/G magnetic beads were conjugated with anti-DIG antibody for 10 mins at room temperature, followed by blocking with Denhardt's solution. DIG-labelled RNA was immunoprecipitated for 30 mins at room temperature. RNA was extracted by TRIZOL and analyzed by qRT-PCR as described.

Chromatin Accessibility assay

The accessibility of the HOTAIR gene was assayed with MspI digestion, which mimicked the binding of RNA Polymerase II, according to the previous report with minor modifications.[64] Briefly, 1×10^7 cells were washed with cold PBS twice. The cell pellet was resuspended in nuclei isolation buffer (10mM NaCl, 3mM $MgCl_2$, 1mM PMSF, 1% NP-40, and 10mM Tris, pH 7.4) and was incubated on ice for 10 minutes. The cells were homogenized on ice by Dounce homogenizer. The nuclei were pelleted by centrifugation at 1000 g, 10 mins, 4°C and resuspended in MspI buffer. 0U, 25U and 50U MspI (NEB) was added to digest 1×10^6 nuclei at 37°C for 1h. DNA was extracted by phenol: chloroform: isoamylalcohol, followed by ethanol participation in the presence of sodium acetate. The relative accessibility of 3' end to 5' end of the HOTAIR gene was analyzed by qPCR.

Immunoblot analysis

The whole cell extract was prepared by lysing cells in NP-40 lysis buffer with proteinase inhibitors and phosphatase inhibitor (Thermo Fisher Scientific). Protein concentration was determined by BCA assay (Thermo Fisher Scientific). Proteins were resolved by SDS-PAGE at different percentages, transferred to PVDF membrane and immunoblotted overnight at 4°C with antibodies against CDK7 (rabbit; Abcam, ab216437; 1:1000), CDK7 (phosphor T170) (rabbit; abcam Ab155976; 1:1000), CDK9 (rabbit; Cell Signaling Technology, 2316 ; 1:1000), CDK9 (phosphor T186) (rabbit; Cell Signaling Technology, 2459; 1:1000) and GAPDH (rabbit; Cell Signaling Technology, 5174; 1:1000). Then, the blots were washed three times with TBST, followed by incubation with 1:2000 secondary anti-rabbit antibody at room temperature for 1 h. Images of immunoblot were taken with ChemiDoc Imaging System (Bio-Rad, Hercules, CA, USA).

Immunohistochemistry

Immunostaining was performed using 5-µm PDAC FFPE tissue sections. Antigen retrieval was carried out using PT module (Thermo Fisher Scientific). The sectioned tissues were deparaffinized and rehydrated by xylene and a series of graded ethanol. Antigen retrieval was carried out using PT module (Thermo Fisher Scientific). Then, immunostaining was performed using Histostain-Plus IHC Kit, HRP, broad spectrum (Life Technologies, Carlsbad, CA) according to the manufacturer's protocol. Antibodies against CDK7 (rabbit; Abcam, ab216437; 1:100), CDK7 (phosphor T170) (rabbit; abcam, ab59987; 1:50), (rabbit; Cell Signaling Technology, 2316; 1:100); CDK9 (phosphor T186) (rabbit; Cell Signaling Technology, 2459; 1:100) were used for staining at 4°C overnight. Sections were counter-stained with hematoxylin. Images of immunohistochemistry were taken with Spot Digital Camera & Leica Microscope Biological Imaging System (10X magnification). A scoring system, based on the percentage of positive cells and staining intensity under 10X magnification, was used to quantify the staining. 4 categories (0, 1, 2, and 3) were demoted as 0%, 1-10%, 10-50%, and >50%. The staining intensity of the tumor section was compared to adjacent non-tumor tissue. The association between HOTAIR expression and intensity of CDK7, CDK7 (phosphor T170), CDK9 and CDK9 (phosphor T186) protein staining was performed using two-tailed Fisher's exact t-test.

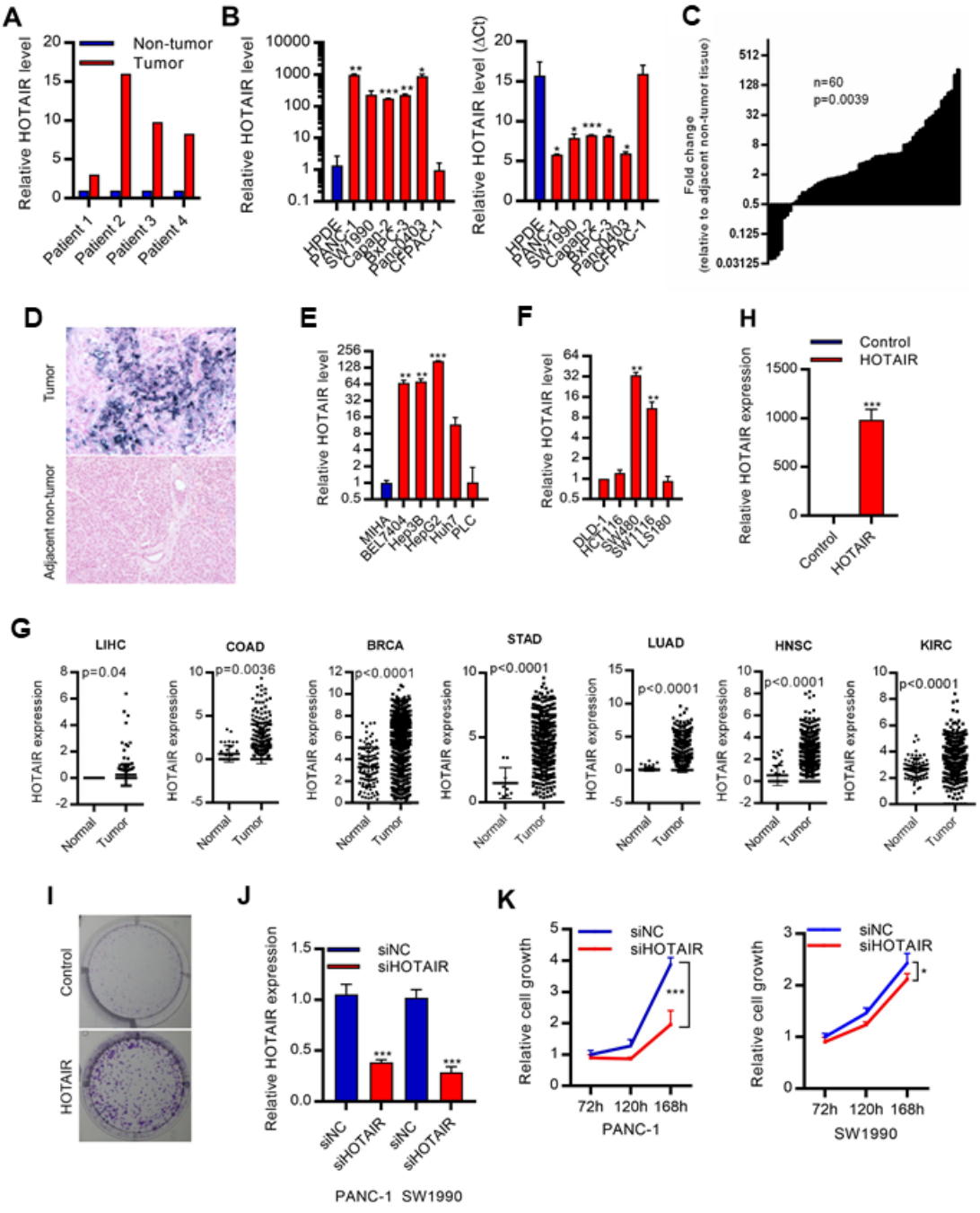

Figure S1. HOTAIR is upregulated in cancers and promotes PDAC progression. **(A)** Microarray analysis identified the upregulated HOTAIR in pancreatic ductal adenocarcinoma (PDAC) primary tumors, compared to adjacent non-tumor tissues. **(B-C)** HOTAIR expression was upregulated in PDAC **(B)** cells and **(C)** primary tumors. Expressions of HOTAIR in PDAC cells and tumors were compared to non-tumorigenic human pancreatic ductal epithelial (HPDE) cells or adjacent non-tumor tissues respectively. **(D)** Representative *In situ* hybridization (ISH) images showing the expression of HOTAIR in PDAC primary tumor but no HOTAIR expression in adjacent non-tumor tissue. **(E-F)** HOTAIR expression was upregulated in **(E)** hepatocellular

carcinoma (HCC) and **(F)** colorectal carcinoma (CRC) cells. Expression of HOTAIR in HCC cells was compared to non-tumorigenic MIHA cells. Expression of HOTAIR in CRC cells was compared to DLD-1 cells. **(G)** HOTAIR expression was frequently upregulated in liver hepatocellular carcinoma (LIHC), colorectal adenocarcinoma (COAD), breast carcinoma (BRCA), stomach adenocarcinoma (STAD), esophageal cancer (ESCA), lung adenocarcinoma (LUAD), head and neck squamous cell carcinoma (HNSC), and kidney renal clear cell carcinoma (KIRC). HOTAIR expression was analyzed using TCGA datasets, n = 3,234 samples. **(H)** qPCR analysis of the overexpression efficiency of HOTAIR in HPDE cells. **(I)** Overexpression of HOTAIR promoted colony formation in HPDE cells. **(J)** qPCR analysis of the knockdown efficiency of HOTAIR in cells. **(K)** Knockdown of HOTAIR inhibited cell growth in PANC-1 and SW1990 cells. Data were from at least three independent experiments and plotted as means ± SD. * $P < 0.05$, * * $P < 0.01$, * * * $P < 0.001$.

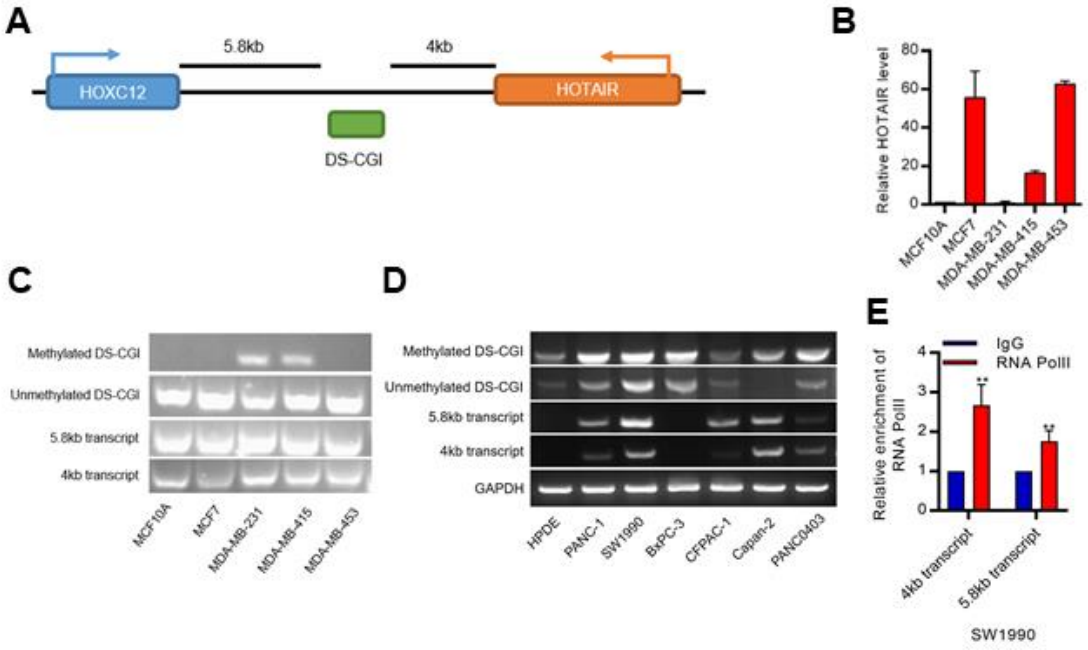

Figure S2. Methylation of downstream CpG (DS-CGI) is not associated with HOTAIR expression. (A) Location of DS-CGI island between HOTAIR and HOXC12. DS-CGI is located 5.8 kb downstream of HOXC12 gene and 4 kb downstream of HOTAIR gene. **(B)** qRT-PCR analysis of HOTAIR expression in breast cancer cells. Expressions of HOTAIR in breast cancer cells were compared to non-tumorigenic human breast epithelial cells MCF10A. **(C, D)** PCR analysis of the methylation status of DS-CGI and expression of both 5.8kb and 4kb transcripts in (**C**) breast cancer cells and MCF10A cells, and (**D**) PDAC cells and HPDE cells. **(E)** ChIP analysis of RNA PolII occupancy at both 5.8kb and 4kb regions in SW1990 cells. Data were from at least three independent experiments and plotted as means ± SD. * * $P < 0.01$

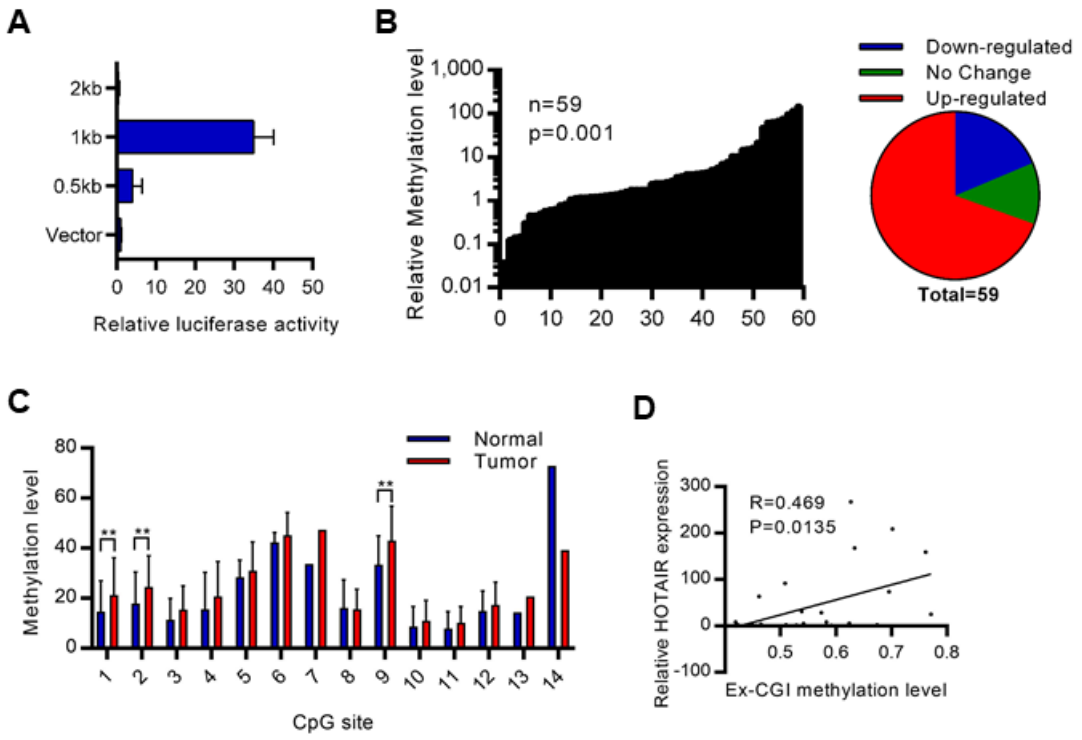

Figure S3. Ex-CGI is hypermethylated in PDAC primary tumors and is associated with HOTAIR expression. (A) Luciferase assay was performed to locate the promoter of HOTAIR gene. **(B)** qMSP revealed that Ex-CGI was hypermethylated in PDAC primary tumors. **(C)** Pyrosequencing analysis revealed that individual CpG sites in Ex-CGI were hypermethylated in PDAC primary tumors. N=27, ** $P < 0.01$. **(D)** Ex-CGI methylation level was positively correlated to HOTAIR expression in PDAC primary tumors. Data were from at least three independent experiments and plotted as means ± SD.

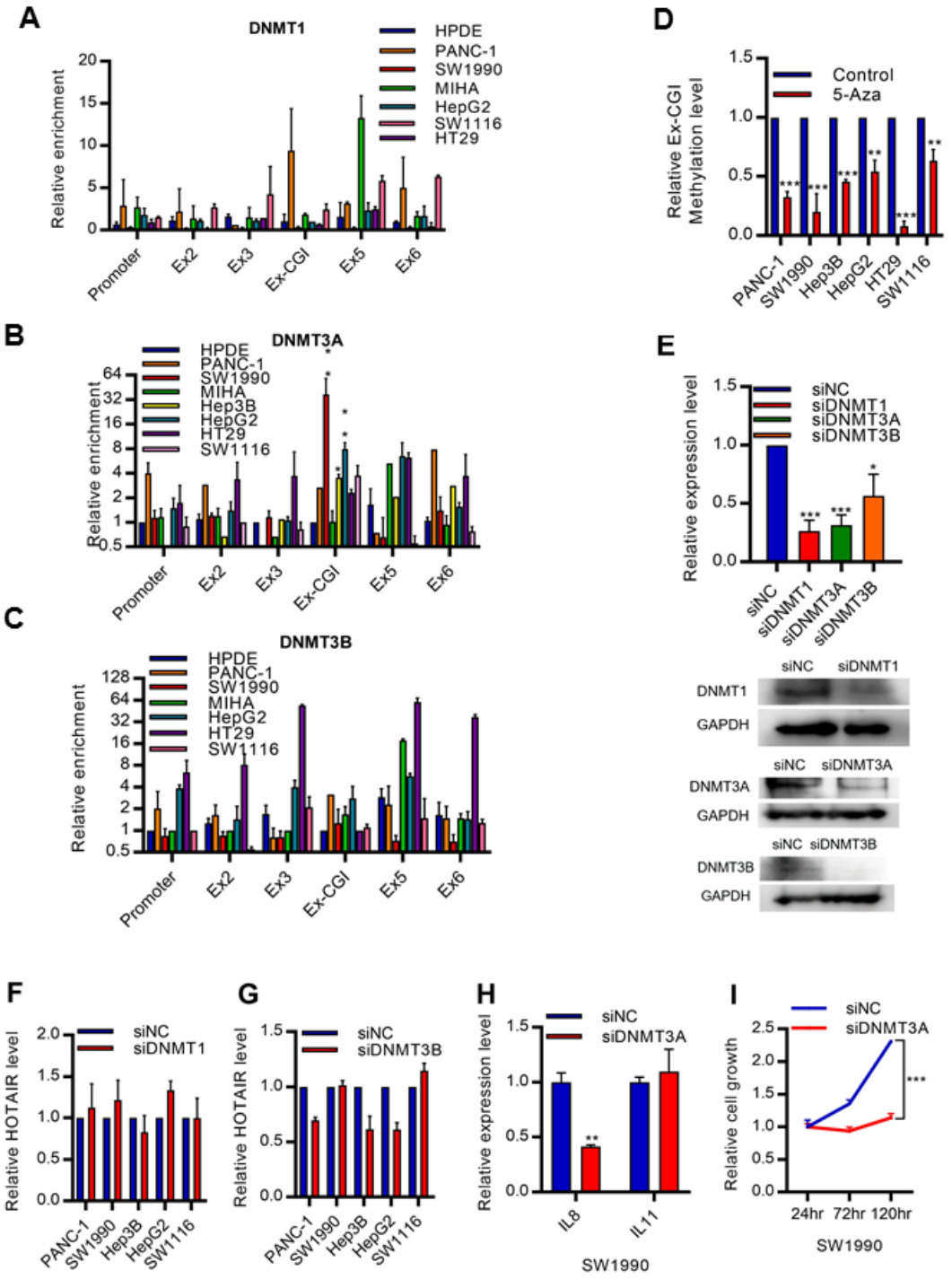

Figure S4. Inhibition of Ex-CGI methylation decreases HOTAIR expression. (A-C) ChIP analysis of **(A)** DNMT1, **(B)** DNMT3A and **(C)** DNMT3B on HOTAIR gene in PDAC, HCC, and CRC cells. The occupancy of DNMT3A binding at the Ex-CGI was increased in PDAC, HCC, and CRC cells. **(D)** Ex-CGI methylation level was decreased after treatment with 5-Aza in PDAC, HCC, and CRC cells. **(E)** Knockdown efficiency of siRNAs targeting DNMT1, DNMT3A and DNMT3B respectively. **(F-G)** Knockdown of **(F)** DNMT1 or **(G)** DNMT3B did not affect HOTAIR expression in PDAC, HCC and CRC cells. (**H**) Controls of DNMT3A in SW1990 cells. It has been

demonstrated that DNMT3A promoted the expression of IL8, without affecting the expression of IL11. Consistently, knockdown of DNMT3A inhibited Interleukin 8 (IL8) expression, but not Interleukin 11 (IL11) expression. (**I**) Knockdown of DNMT3A inhibited cell growth in SW1990 cells. Data were from at least three independent experiments and plotted as means ± SD. * $P < 0.05$, * * $P < 0.01$, * * * $P < 0.001$.

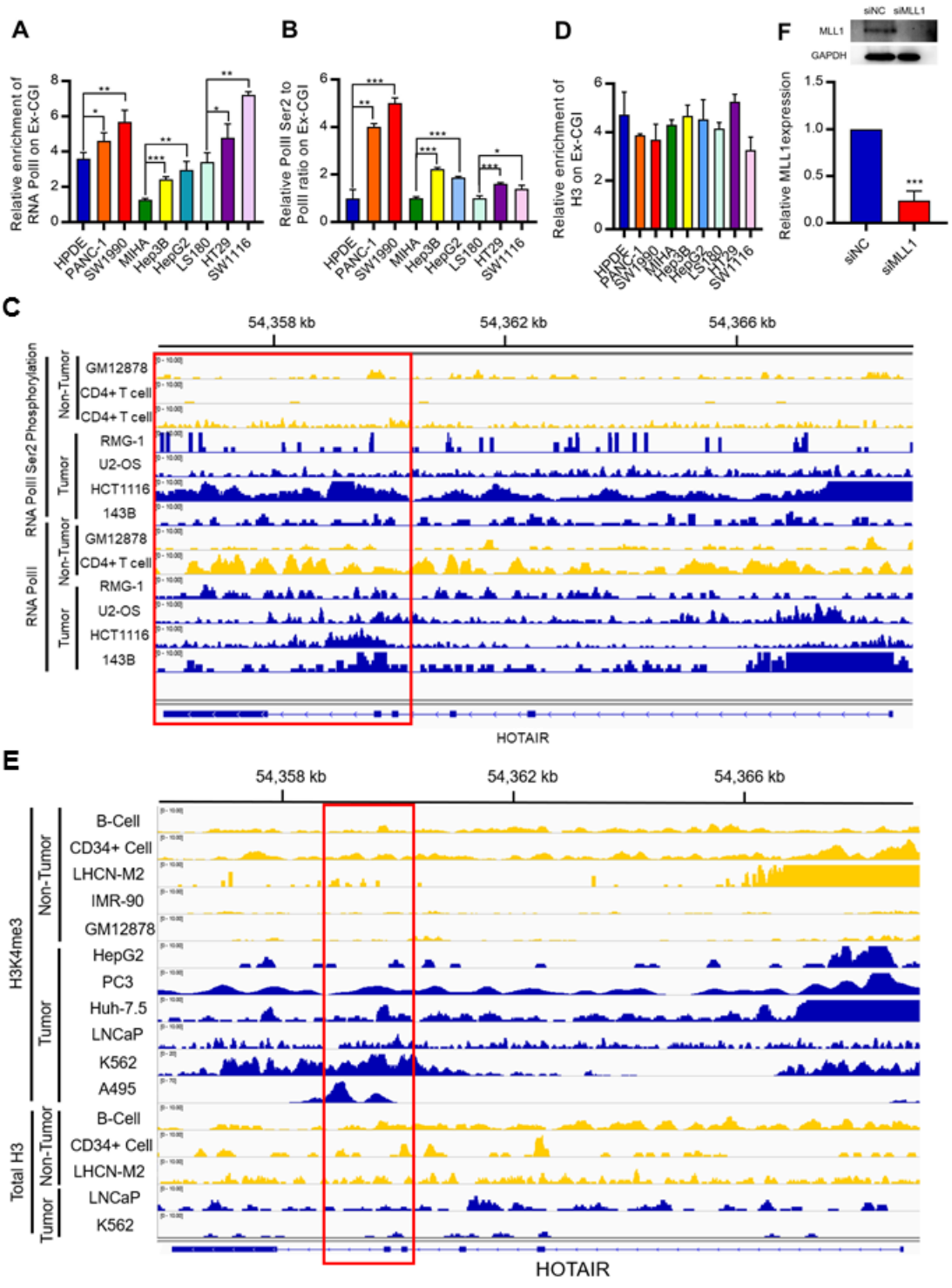

Figure S5. RNA PolII Ser2 phosphorylation and H3K4me3 were upregulated at Ex-CGI in cancer. (**A**) The occupancy of total RNA PolII at the Ex-CGI was increased in cancer cells, as compared to non-tumor cells. (**B**) Ratio of RNA PolII Ser2 phosphorylation to total RNA PolII at Ex-CGI was increased in cancer cells, as compared to non-tumor cells. (**C**) ChIP-seq analysis revealed that RNA PolII Ser2 phosphorylation along HOTAIR gene was increased in cancer cells HepG2 (HCC), PC3

(prostate cancer), Huh-7.5 (HCC), LNCaP (prostate cancer), K592 (leukemia) and A495 (lung cancer), as compared to non-tumor cells GM12878 and CD4+ T cells. (Results were from ENCODE, GEO GSE33281, GSE20040, GSE100040, GSE97589, GSE132233 and GSE104545) (17-23). (**D**) Total H3 was not significantly enriched at the Ex-CGI in cancer cells, as compared to non-tumor cells. (**E**) ChIP-seq analysis revealed that H3K4me3, but not the total H3, at Ex-CGI was increased in cancer cells RMG-1 (ovarian cancer), U2-OS (osteosarcoma), 143B (osteosarcoma), and HCT1116 (HCC), as compared to non-tumor cells B-cells, CD34+ cells, LHCN-M2, GM12878 and IMR-90 cells. (Results were from ENCODE, GEO GSE142579, GSE113336, GSE78158, GSE76344, GSE91401, GSE103728, GSE117306, and GSE103734) (23-31). (**F**) Knockdown efficiency of siRNAs targeting MLL1. Data were from at least three independent experiments and plotted as means ± SD. * $P < 0.05$, * * $P < 0.01$, * * * $P < 0.001$.

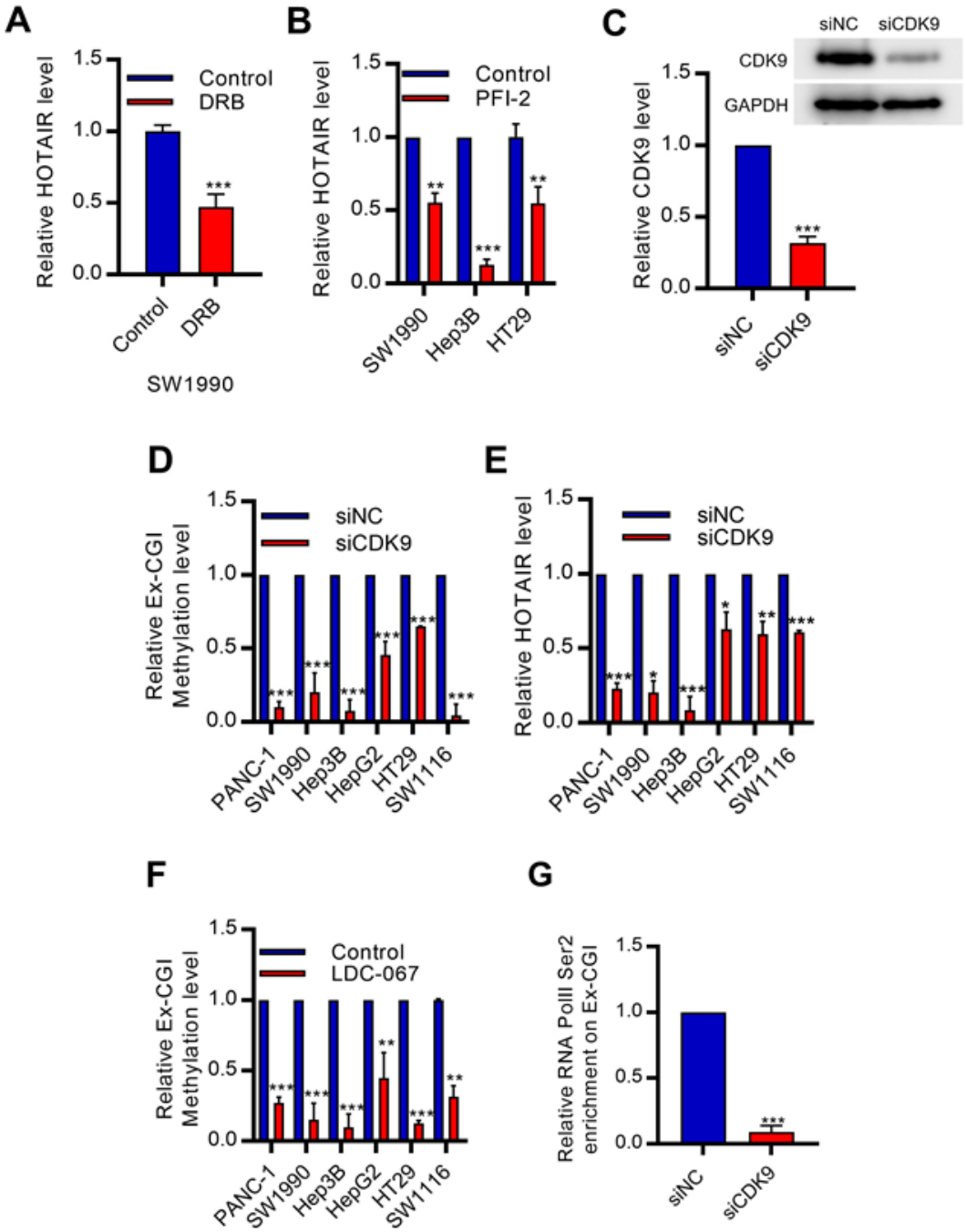

Figure S6. CDK9 promotes the expression of HOTAIR in cancers. (A-B) HOTAIR expression was inhibited after inhibition of pTEFb by **(A)** DRB or **(B)** PFI-2 in PDAC, HCC, and CRC cells. **C**, Knockdown efficiency of siRNAs targeting CDK9. **(D-E)** knockdown of CDK9 by siRNA reduced **(D)** Ex-CGI methylation and (**E**) HOTAIR expression in PDAC, HCC, and CRC cells. **(F)** Inhibition of CDK9 by LDC-067 reduced Ex-CGI methylation level in PDAC, HCC, and CRC cells. **(G)** RNA PolII Ser2 phosphorylation level on Ex-CGI were reduced after CDK9 knockdown in PANC-1 cells. Data were from at least three independent experiments and plotted as means ± SD. * $P < 0.05$, * * $P < 0.01$, * * * $P < 0.001$

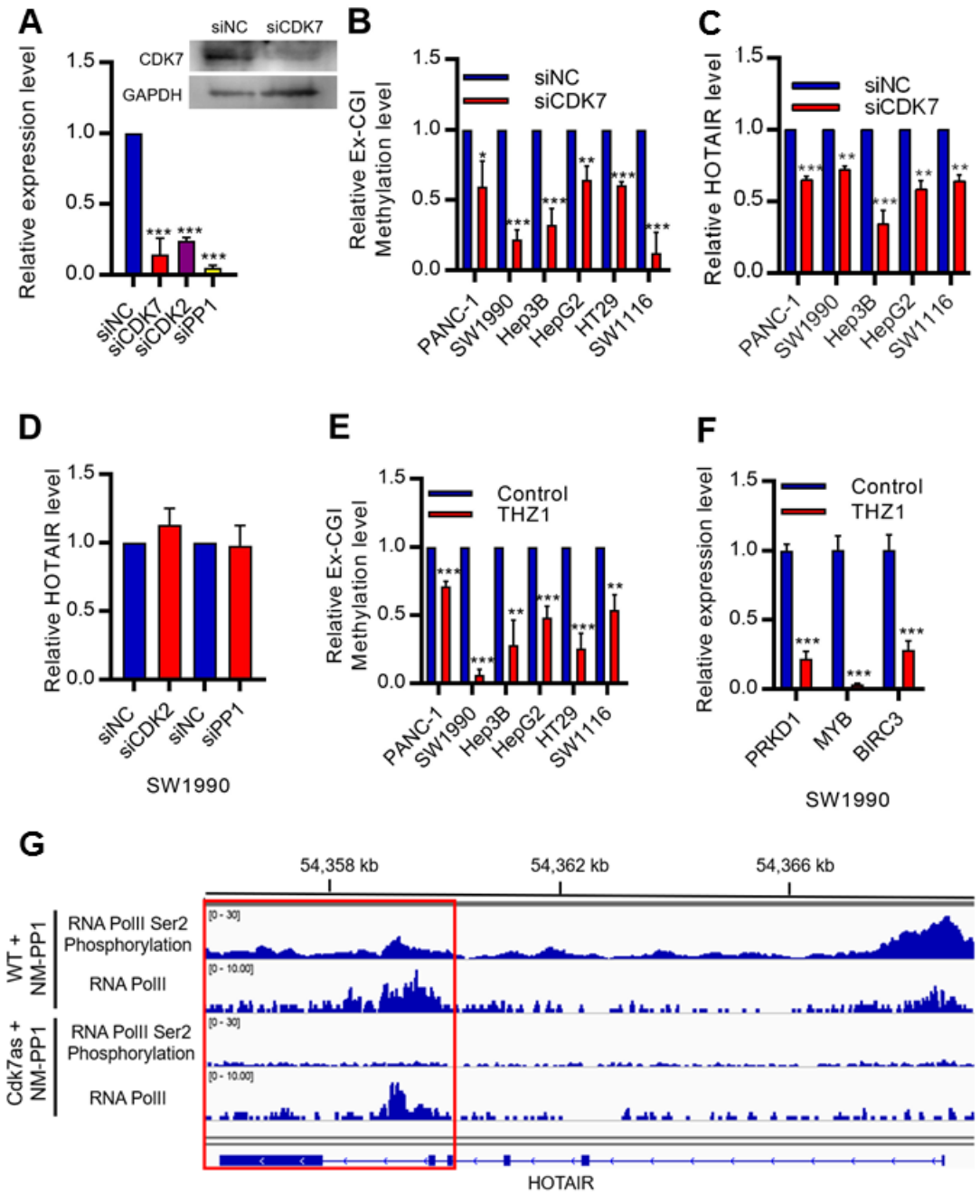

Figure S7. CDK7 promotes the expression of HOTAIR in cancers. (A) Knockdown efficiency of siRNA targeting CDK7, CDK2 and PP1 respectively. **(B-C)** Knockdown of CDK7 decreased (**B**) Ex-CGI methylation and (**C**) HOTAIR expression in PDAC, HCC and CRC cells. (**D**) Knockdown of CDK2 and PP1 did not affect HOTAIR expression in SW1990 cells. **(E)** Inhibition CDK7 by THZ1 reduced methylation level of Ex-CGI in PDAC, HCC, and CRC cells. (**F**) Controls of CDK7 inhibition in SW1990 cells. It has been demonstrated that CDK7 promoted the expressions of PRKD1, MYB and BIRC3 in cancers. Consistently, inhibiting CDK7 by THZ1 inhibited Protein Kinase D1 (PRKD1), MYB and Baculoviral IAP Repeat Containing 3 (BIRC3)

expressions. **(G)** CRC HCT1116 cells were mutated (Cdk7as) to sensitize to CDK7 inhibitor MM-PP1. ChIP-seq analysis revealed that RNA PolII Ser2 phosphorylation along HOTAIR gene was decreased after CDK7 inhibition in Cdk7as cells, as compared to MM-PP1-resistant WT cells (Results were from GEO GSE100040) (19). Data were from at least three independent experiments and plotted as means ± SD. * $P < 0.05$, * * $P < 0.01$, * * * $P < 0.001$.

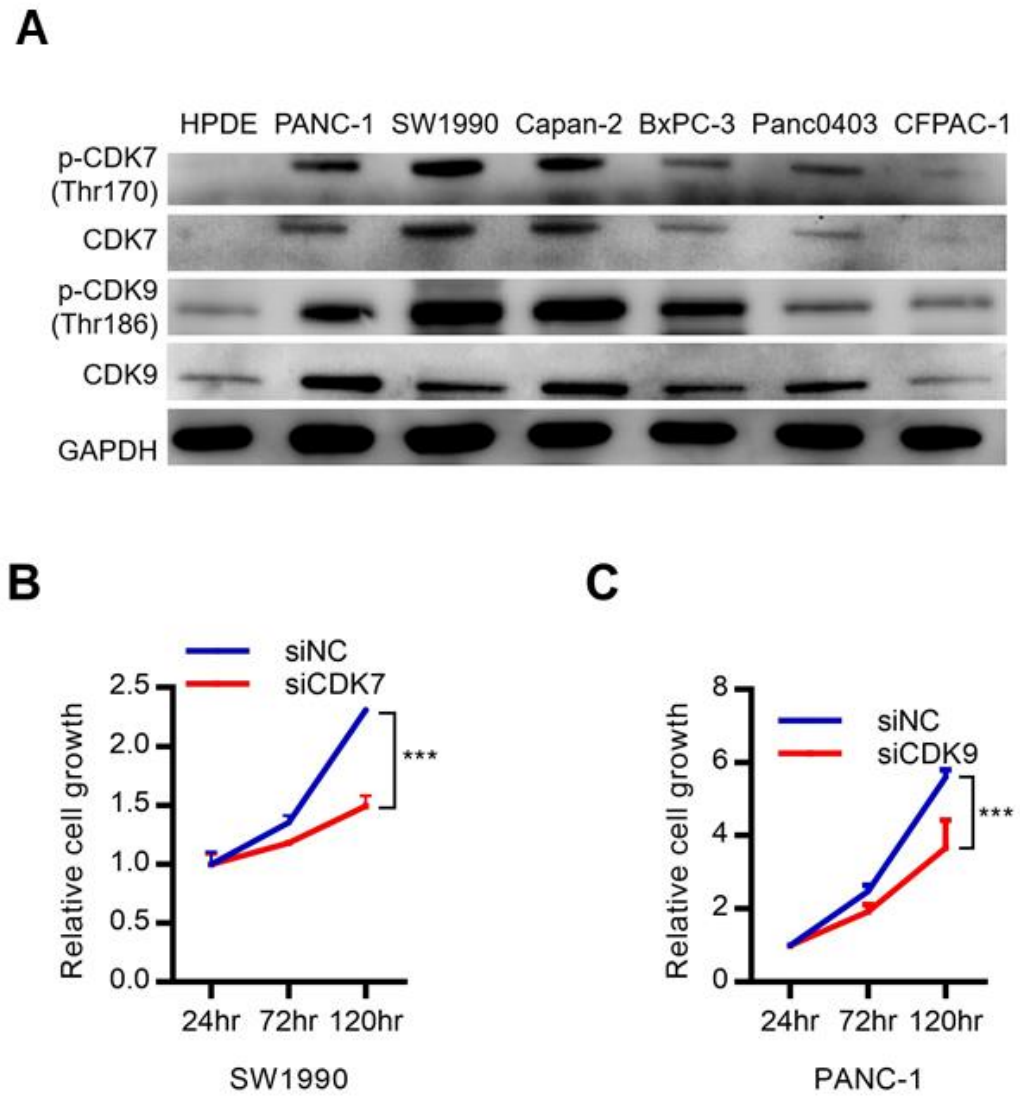

Figure S8. Upregulated CDK7-CDK9 promotes PDAC progression. (**A**) CDK7, p-CDK7, CDK9 and p-CDK9 were increased in PDAC cells with high HOTAIR expression. (**B-C**) Knockdown of (**B**) CDK7 and (**C**) CDK9 inhibited cell growth in PDAC cells. Data were from at least three independent experiments and plotted as means ± SD. * * * $P < 0.001$.

Table S1. siRNAs and sgRNAs used in this study

siHOTAIR	GAACGGGAGUACAGAGAGATT
siDNMT1	GGAAGAAGAGUUACUAUAATT
siDNMT3A	GCACUGAAAUGGAAAGGGUUU
siDNMT3B	GAAAGUACGUCGCUUCUGAUU
siMLL1	CGAUCAAAUGCCCGCCUAATT
siCDK9	GGGACAUGAAGGCUGCUAATT
siCDK7	GACUCUUCAAGGAUUAGAATT
siCDK2	GCUGAAGAGGGUUGGUAUAUU
siPP1	CUGGCAAGAAUGUACAGCUTT
sgEx-CGI	GGCCGGCTCACCCCCGGTAAAGG

Table S2. Primers used in this study