Supplementary Figure legends

Figure S1. Bioinformatics analysis of lncRNA regulated by TGF-β /SMAD3 in lung adenocarcinoma

(A) Flow chart for the selection of lncRNAs overexpressed in LUAD and regulated by TGF-β/SMAD3. **(B)** Venn diagrams showing that 48 lncRNAs were induced by TGF-β and attenuated by SIS3 treatment. **(C)** Hierarchical clustering analysis of the expression of ten lncRNAs in the NC, TGF-β-treated and TGF-β- and SIS3-treated groups. **(D, E)** HCP5 expression in patients with current smoking and not former smokers via analyzing GSE31210 and GSE10072 dataset. *P< 0.05, NS: no statistical significance.

Figure S2. The relationship between *SMAD3* and HCP5 expression and partial sequence of the HCP5 promoter region

(A) *SMAD3* expression in human LUAD tissues and normal lung tissues in the GSE31210 dataset. **(B)** *SMAD3* expression positively correlated with HCP5 expression in LUAD. Linear regression analysis from the GSE37745 dataset using the Pearson correlation coefficient test (R = 0.2631, P < 0.05). **(C)** Partial sequence of the HCP5 promoter. The three putative *SMAD3* binding sites are in red. ***P< 0.001.

Figure S3. HCP5 regulates LUAD cells proliferation and invasion

(A, C) Efficiency of HCP5 knockdown in A549 and Calu3 cells using two small interfering RNAs (siHCP5#001 and siHCP5#002) was detected by qRT-PCR. **(B, D)**

qRT-PCR for HCP5 expression in A549 and Calu3 cells transduced with empty vector (pCDH) or pCDH-HCP5. **(E)** The invasion ability of HCP5-silenced and control Calu3 cells analyzed by Transwell assays. Scale bar: 100 μm. **(F)** Silencing of HCP5 inhibits the colony formation of Calu3 cells. **(G)** Cell Counting Kit-8 (CCK-8) assays were performed in Calu3 cells silenced for HCP5. **(H)** Transwell assay to investigate the invasion ability of HCP5-overexpressing and control Calu3 cells. Scale bar: 100 μm. **(I)** Overexpression of HCP5 inhibits the clonogenic ability of cells. **(J)** Cell Counting Kit-8 (CCK-8) assays were performed in Calu3 cells overexpressing HCP5. All data are shown as the mean ± S.E.M. of three independent experiments (two-tailed Student's t-test). **P*< 0.05, ***P* < 0.01, ****P* < 0.001.

Figure S4. HCP5 knockdown inhibits cell proliferation and metastasis in nude mice

(A) Knockdown efficiency of HCP5 in A549 cells using two designed shRNAs. Data are shown as the mean ± S.E.M. of three independent experiments (two-tailed Student's t-test). **(B)** Mice were subcutaneously injected with cells stably silenced for HCP5 (shHCP5) or control cells (shCtrl). Silencing of HCP5 inhibited A549 xenograft growth in nude mice. Arrows denote tumors in situ. n = 6 mice per group. **(C)** ShHCP5 or shCtrl A549 cells labeled with EGFP-luciferase were intravenously injected into nude mice and IVIS imaging heat maps were obtained in live mice at 8 weeks. **(D)** The luciferase signal of each group show the mean and SEM. Statistical comparison of the means was done by a two-tailed t-test. **P* < 0.05, ****P*< 0.001.

Figure S5. Clinical prognostic value of HCP5 and *SNAI* expression in patients with LUAD

(A, B) Kaplan–Meier curves for overall survival rates of LUAD patients according to the expression level of *SNAI1* and *SNAI2*. **(C, D)** The prognostic value by Kaplan–Meier analysis of the combination of the HCP5 and *SNAI1* or *SNAI2* in the GSE19188 dataset.

Figure S6. HCP5 positively regulates EMT via the miR-203/*SNAI* axis

(A) Schematic illustration of the genomic location of HCP5 at 6p21.33; the binding sites of *miR-203* on HCP5 transcript were predicted using the miRDB database. **(B)** Levels of HCP5 in nuclear and cytoplasm of A549 cells. U6 (nuclear) and β-Actin (cytoplasm) were used as controls. Data are shown as the mean ± S.E.M. of three independent experiments (two-tailed Student's t-test). **(C, D)** Western-blot analysis of the EMT markers (E-cadherin, N-cadherin), Snail and Slug expression in A549 cells transfected with *miR-203* mimics or inhibitors. **(E, F)** qRT-PCR for *miR-203* in A549 cells transfected with *miR-203* mimics or inhibitors for 24 h. Data are shown as the mean ± S.E.M. of three independent experiments (two-tailed Student's t-test). **(G, H)** The expression of *miR-203* in A549 cells after silencing or overexpression of HCP5 was detected by qRT-PCR. Data are shown as the mean ± S.E.M. of three independent experiments (two-tailed Student's t-test). **(I)** The expression of *P15*, *P21* and *P57* in A549 cells after silencing or increasing HCP5 was detected by qRT-PCR. Data are

shown as the mean ± S.E.M. of three independent experiments (two-tailed Student's t-test). *P< 0.05, **P < 0.01, ***P < 0.001.

Figure S1

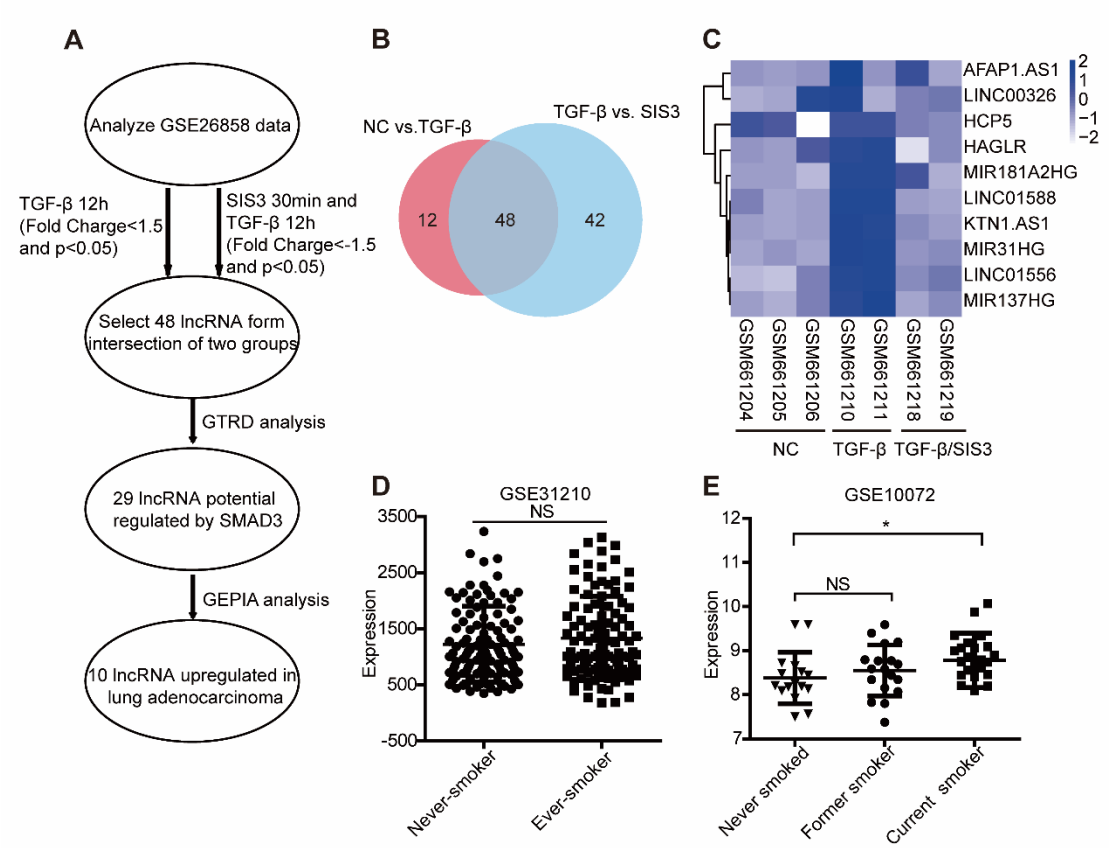
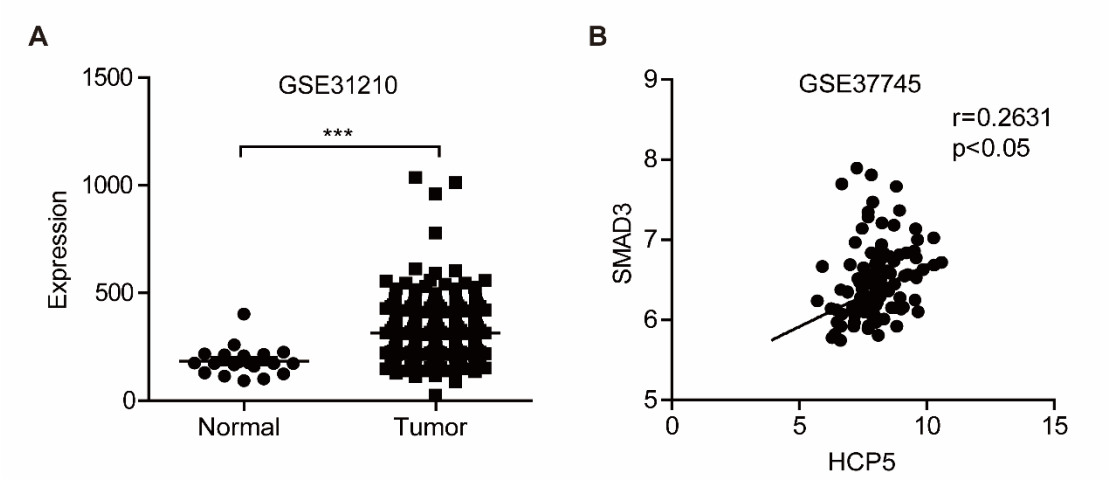


Figure S2



C

The partial sequence of the HCP5 promoter region

GCCAGAAACACCAGCTGTGATTAGAGGATTGGAACTTTCACTGCCATCCCCATCCTCT
GGGGAAGAAAAGGGGGCTGGAAGTTGAGCTCAGTCATCAATGGCCAATGATTTCACC
AATCTTGCCTACACAATGAAACTTCCATAGACACCTCTAGACAGTGAGTTTTGGAGAA
CTTCCCAGTTGGTGAGCACATCCGCATGTCCACGTGCTGGGAGGACGGCACACCTC
ATCTCCATAGAGACAGAGGCTTCTGCGCTTATATCTTTCTGTAAGGCAGACACCCTTG
TTTCTAGGAGGGACCTAGGGTGGACTGTGGATTCTTTCTCTGGGGCAAATAAAAATG
TAGAATCAGAAAATTCAGGCACTTTGCACTCCTCATGGGACACTCCAGCAGCACTCAC
GTGACCATCCTGAGAATGGACAGGACACCTGAGGTGGGGAAGGGAGCACAGAACCC
AGACACCAGCCTGGACACAGGCACCTGGGATAATCTCCTATTCC

Red-Predicted SMAD3 binding sites(site 1, site 2 and site 3)

Figure S3

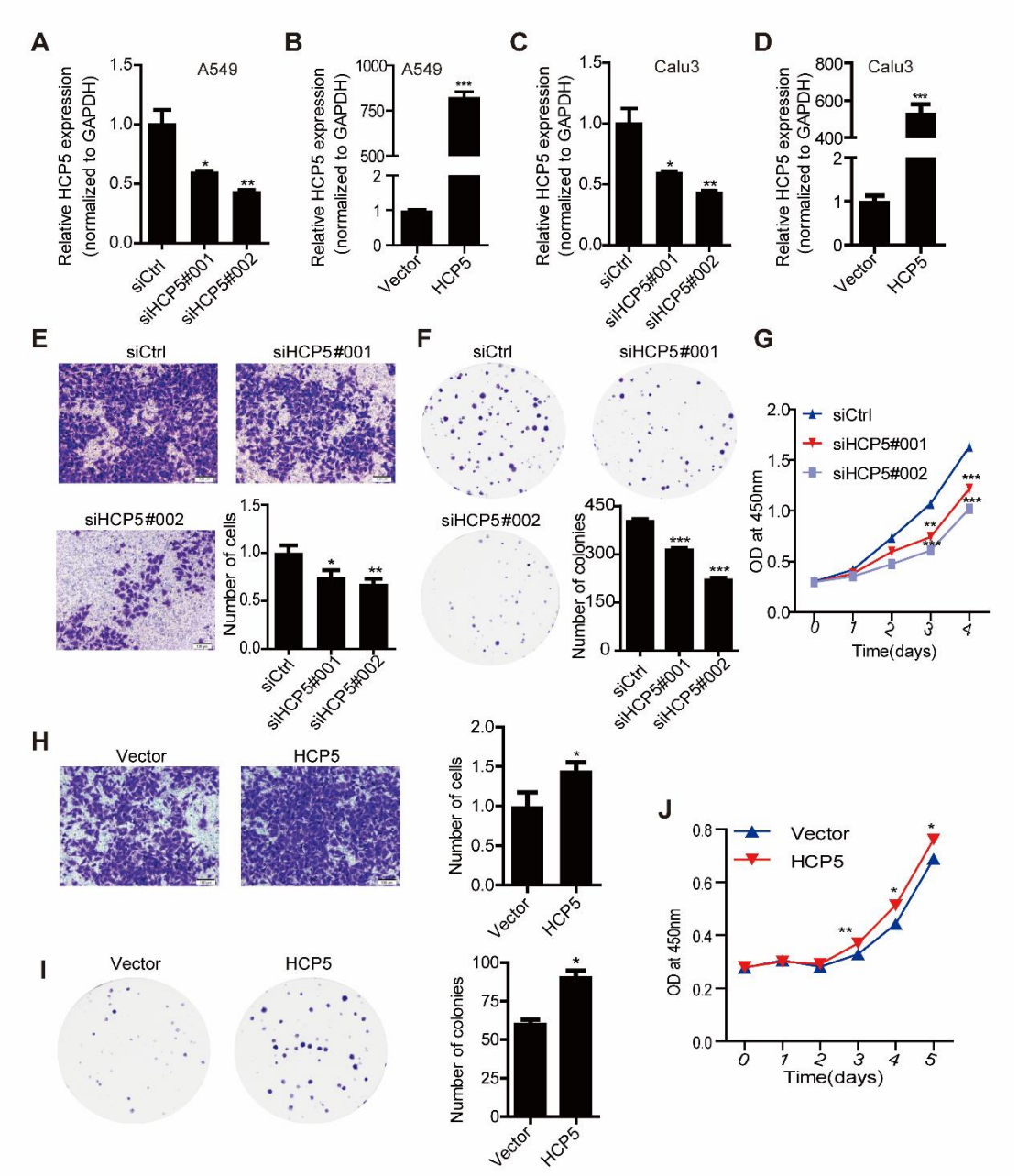

A
A549
Relative HCP5 expression (normalized to GAPDH)
siCtrl
siHCP5#001
siHCP5#002
B
A549
Vector
HCP5
C
Calu3
D
Calu3
E
siCtrl
siHCP5#001
siHCP5#002
Number of cells
F
siCtrl
siHCP5#001
siHCP5#002
Number of colonies
G
siCtrl
siHCP5#001
siHCP5#002
OD at 450nm
Time(days)
H
Vector
HCP5
Number of cells
I
Vector
HCP5
Number of colonies
J
Vector
HCP5
OD at 450nm
Time(days)

Figure S4

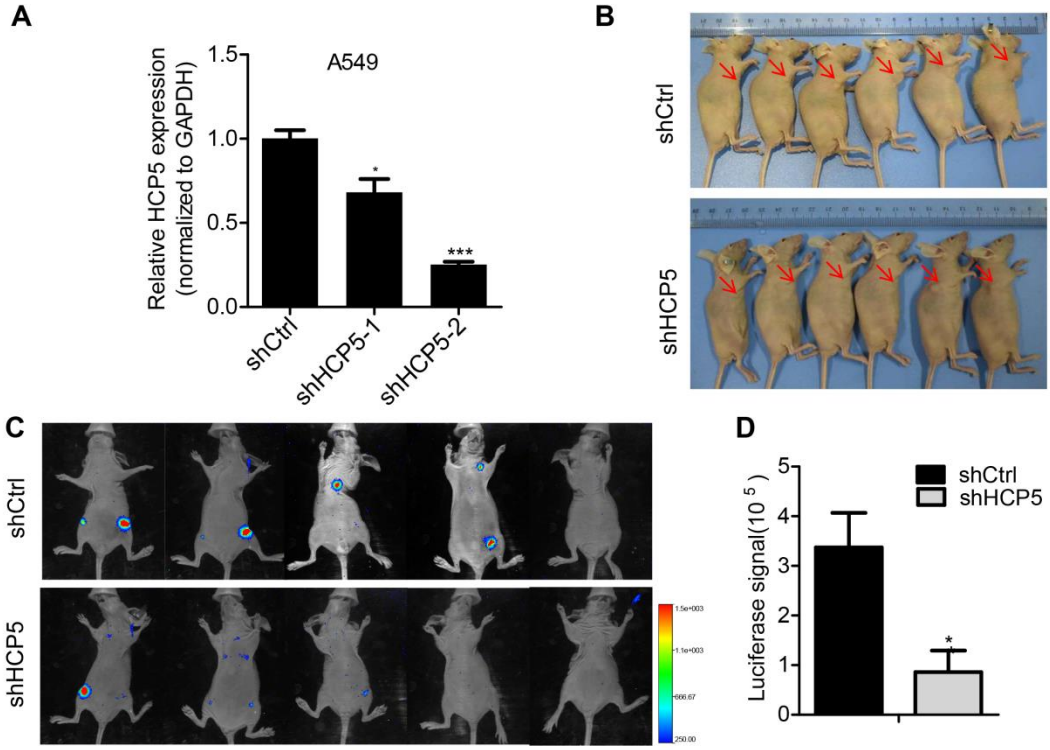

A
A549
Relative HCP5 expression (normalized to GAPDH)
1.5
1.0
0.5
0.0
*
***
shCtrl
shHCP5-1
shHCP5-2
B
shCtrl
shHCP5
C
shCtrl
shHCP5
D
shCtrl
shHCP5
Luciferase signal(10 5)
5
4
3
2
1
0
*

Figure S5

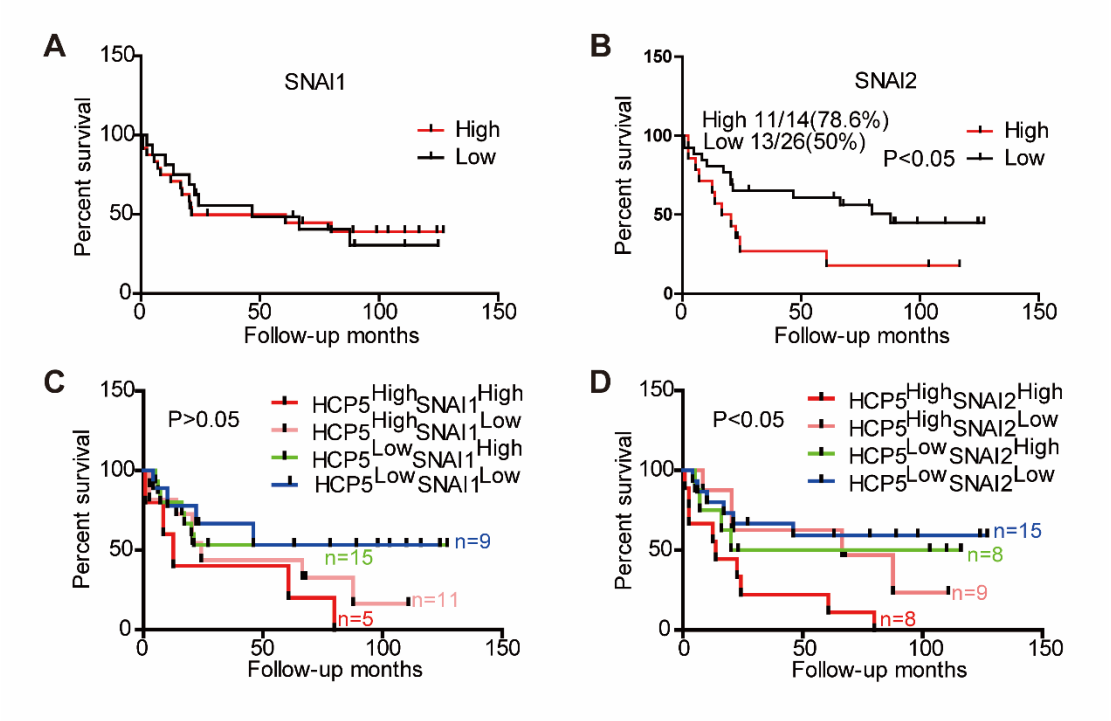

A
SNAI1
High
Low
Percent survival
Follow-up months
B
SNAI2
High 11/14(78.6%)
Low 13/26(50%)
P<0.05
High
Low
Percent survival
Follow-up months
C
P>0.05
HCP5HighSNAI1High
HCP5HighSNAI1Low
HCP5LowSNAI1High
HCP5LowSNAI1Low
n=9
n=15
n=11
n=5
Percent survival
Follow-up months
D
P<0.05
HCP5HighSNAI2High
HCP5HighSNAI2Low
HCP5LowSNAI2High
HCP5LowSNAI2Low
n=15
n=8
n=9
n=8
Percent survival
Follow-up months

Figure S6

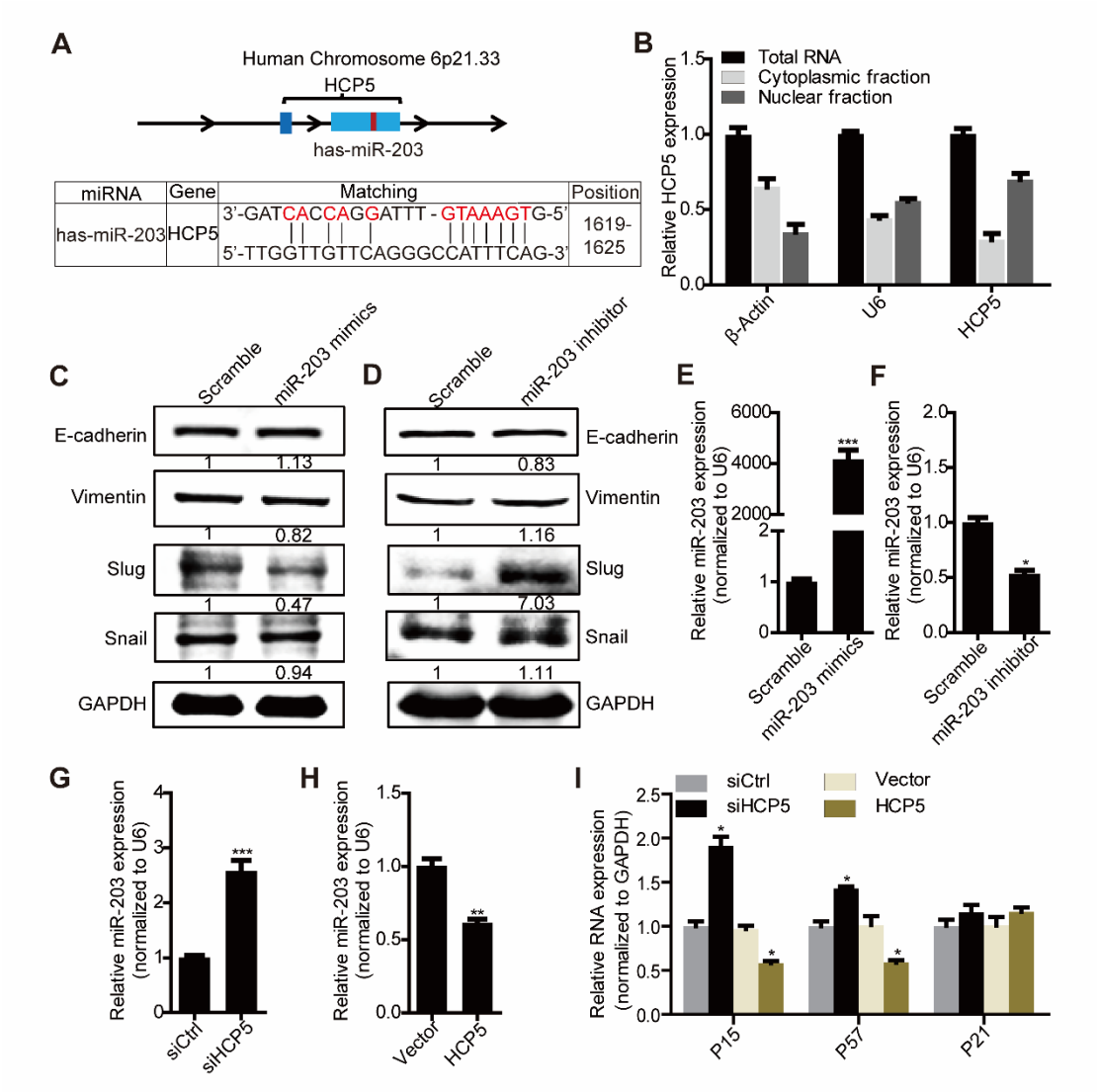

A
Human Chromosome 6p21.33
HCP5
has-miR-203
miRNA
Gene
Matching
Position
has-miR-203
HCP5
3'-GATCACCAGGATTT - GTAAAGTG-5'
5'-TTGGTTGTTCAGGGCCATTTCAG-3'
1619-1625
B
Total RNA
Cytoplasmic fraction
Nuclear fraction
Relative HCP5 expression
β-Actin
U6
HCP5
C
Scramble
miR-203 mimics
E-cadherin
Vimentin
Slug
Snail
GAPDH
1
1.13
0.82
0.47
0.94
D
Scramble
miR-203 inhibitor
E-cadherin
Vimentin
Slug
Snail
GAPDH
1
0.83
1.16
7.03
1.11
E
Relative miR-203 expression (normalized to U6)
Scramble
miR-203 mimics
F
Relative miR-203 expression (normalized to U6)
Scramble
miR-203 inhibitor
G
Relative miR-203 expression (normalized to U6)
siCtrl
siHCP5
H
Relative miR-203 expression (normalized to U6)
Vector
HCP5
I
siCtrl
siHCP5
Vector
HCP5
Relative RNA expression (normalized to GAPDH)
P15
P57
P21

Table S1. Data of 30 clinicopathologic LUAD patients for analyzing the expression of HCP5.

No.	Age	Sex	Grade	TX	NX	MX
1	49	female	moderately-poorly differentiated	2	0	0
2	58	female	well-moderately differentiated	1	0	0
3	55	female	moderately differentiated	3	2	0
4	60	female	well-moderately differentiated	1	0	0
5	40	male	moderately-poorly differentiated	4	2	0
6	43	male	moderately-poorly differentiated	2	1	0
7	58	male	moderately-poorly differentiated	4	0	0
8	73	male	moderately differentiated	2	0	0
9	61	male	moderately-poorly differentiated	4	0	0
10	60	male	moderately-poorly differentiated	4	0	0
11	58	male	moderately-poorly differentiated	4	0	0
12	47	female	moderately differentiated	3	0	0
13	53	female	moderately differentiated	3	1	0
14	47	male	moderately differentiated	2	0	0
15	60	male	moderately differentiated	3	0	0
16	59	female	moderately differentiated	2	0	0
17	75	female	moderately-poorly differentiated	2	0	0
18	64	male	moderately differentiated	3	2	0
19	57	female	moderately-poorly differentiated	2	1	0
20	34	male	poorly differentiated	3	2	0
21	40	male	Mucinous adenocarcinoma	3	2	0
22	50	female	moderately-poorly differentiated	3	2	0
23	65	male	moderately differentiated	2	0	0
24	46	male	moderately-poorly differentiated	3	1	0
25	67	female	moderately-poorly differentiated	2	0	0
26	84	male	poorly differentiated	2	1	0
27	50	female	moderately-poorly differentiated	3	2	0
28	57	female	moderately-poorly differentiated	2	1	0
29	55	male	moderately-poorly differentiated	4	1	0
30	51	male	poorly differentiated	2	2	0

Table S2. Information on antibodies used in this study.

Antibody	Company	Catalog #	Species	Dilution	Note
				WB	
GAPDH	Proteintech	10494-1-AP	Rabbit	1:5000	
IgG	Santa Cruz Blotechnology	SC-2025	Mouse	1:2000	CHIP(1:500)
Snail	Cell Signaling Technology	3879	Rabbit	1:1000	
Slug	Cell Signaling Technology	9585	Rabbit	1:1000	
Vimentin	Cell Signaling Technology	5741	Rabbit	1:1000	IF(1:100)
N-Cadherin	Cell Signaling Technology	13116	Rabbit	1:1000	
E-Cadherin	Cell Signaling Technology	3195	Rabbit	1:1000	IF(1:100)
SMAD3	Thermo Fisher	MA5-14939	Rabbit	1:1000	CHIP(1:50)
Phospho-Smad3	Thermo Fisher	MA5-14936	Rabbit	1:1000	
Snail	OriGene	TA506430	Mouse	1:800	
Slug	Proteintech	12129.1.AP	Rabbit	1:1000	

Table S3. Primers used in this study.

Primers used in qRT-PCR analysis	
Primer name	**Primer sequence(5'to3')**
β-Actin	GTCACCGGAGTCCATCACGAT
	TCACCAACTGGGACGACATG
GAPDH	CAAGGTCATCCATGACAACTTTG
	GTCCACCACCCTGTTGCTGTAG
U6	CTCGCTTCGGCAGCACA
	AACGCTTCACGAATTTGCGT
HCP5	GACTCTCCTACTGGTGCTTGGT
	CACTGCCTGGTGAGCCTGTT
SMAD3	CACGCAGAACGTGAACACC
	GGCAGTAGATAACGTGAGGGA
Snai1	TGCGTCTGCGGAACCTG
	GGACTCTTGGTGCTTGTGGA
Slug	TGTGACAAGGAATATGTGAGCC
	TGAGCCCTCAGATTTGACCTG
Vimentin	CCTGAACCTGAGGGAAACTAA
	GCAGAAAGGCACTTGAAAGC
E-Cadherin	GCCCCATCAGGCCTCCGTTT
	ACCTTGCCTTCTTTGTCTTTGTTGGA
TWIST	AAGCTGCAGCTATGTGGCTCACG
	AATCACTGTCCACGGGCCTGTCT
ZEB1	ACTCTGATTCTACACCGC
	TGTCACATTGATAGGGCTT
P15	CGGGGTCGGGTAGAGGA
	GCGCTGCCC ATCATCAT
P57	ACATCCACGATGGAGCGTC
	GGAAGTCGTAATCCC AGCGG
P21	CCTCATCCCGTGTTCTCCTTT
	GTACCACCCAGCG GACAAGT
miR-203a-3p F	CGCGGTGAAATGTTTAGGACCACTAG

Table S4. HCP5 siRNA, shRNA and primers used in plasmid construction used in this study.

siRNAs and shRNAs designed for silencing HCP5(5'to3')	
siRNA#001	sense-CCAACAUAUUUCUCUGCUU
	Antisense-AAGCAGAGAAAUAUGUUGG
siRNA#002	sense-GCUGAUGAGUAGGACAUUU
	Antisense-AAAUGUCCUACUCAUCAGC
sh-HCP5-1	CGCGGCTGATGAGTAGGACATTTCTCGAGAAATGTCCTACTCATCAGCTTTTTTG
sh-HCP5-2	CGCGGATCTATTACCTGTGCCTGGACTCGAGTCCAGGCACAGGTAATAGATCTTTTTTG

Primers used in plasmid construction(5'to3')	
pCDH-HCP5	CCG GAATTCGACTCA GATTCTCCCCAG AC
	TTTTCCTTTTGCGGCCGCTTCATGTGGGATCCACAAC
H1	CTAGCTAGCGGTTGAAGCCGTATGTTGCTGAGACC
	CCGGCGCGCCAAGCttTAATTGTAATCTGTAATTAAATATATGTGC
H2	CTAGCTAGCGGTTGAAGCCGTATGTTGCTGAGACC
	CCCAAGCTTTATCTAGGAGCCCCTCACCCCATAGT
H3	CTAGCTAGCGTTTCAGGATGGAGGCTGCT
	CCCAAGCTTTGCGGATGTGCTCACCAACT
H4	CGGCTAGCTAGCGGGCAAATAAAAATGTAG
	CCCAAGCTTCAAGGAATAGGAGATTATCCC
H5	CTAGCTAGCGAAAGTTCCAGTATCTGAGGGA
	CGCGCCCCCAAGCTTTAATTGTAATCTGTAATTA
H6	CTAGCTAGCGTTTCAGGATGGAGGCTGCT
	CCCAAGCTTTGCGGATGTGCTCACCAACT
psiCHECK2-HCP5 Wild-Type	TAATTCTAGGCGATCGCTCGAGGACTCAGATTCTCCCCAGACGC
	TTTTATTGCGGCCAGCGGCCGCTTCATGTGGGATCCACAACACT
psiCHECK2-HCP5 Mutant	CTTGGTTGTTCAGGGCGTAAAGTGGTTTGGGTGTTTTCTGGGGATG
	CATCCCCAGAAAACACCCAAACCACTTTACGCCCTGAACAACCAAG
pcDNA3.1-HCP5	CTAGCTAGCGACTCAGATTCTCCCCAGACG
	CCCAAGCTTTTCATGTGGGATCCACAACAC
pcDNA3.1-HCP5-antisense	CTAGCTAGCAAGTACACCCTAGGTGTTGTG
	CCCAAGCTTCTGAGTCTAAGAGGGGTCTGC

Table S5. HCP5 RNA-FISH probes used in this study.

Name	Probe sequence(5'to3')	Labeled fluorescein
HCP5-1	AGAACAGCAGGAGGAGGGTT	5'CY3
HCP5-2	TAAT+TGTAATC+TGCCCAGGT	
HCP5-3	GAGA+TCATTT+CTGCC+TTGAT	
has-miR-203a-3p	CTAGTGG+TCCTAAACATT+TC AC	5'FAM