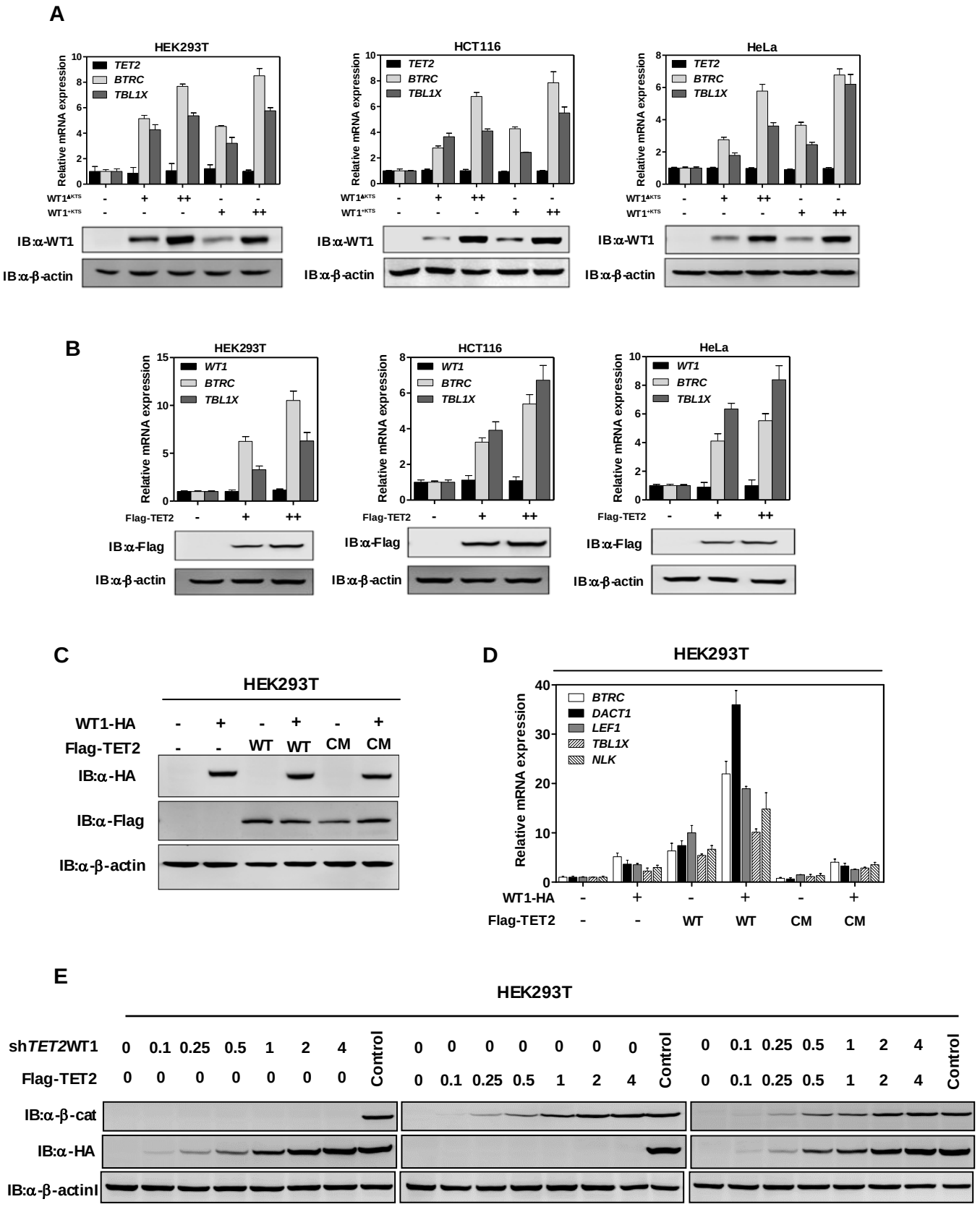
Figure S1. TET2 activates WT1 target genes



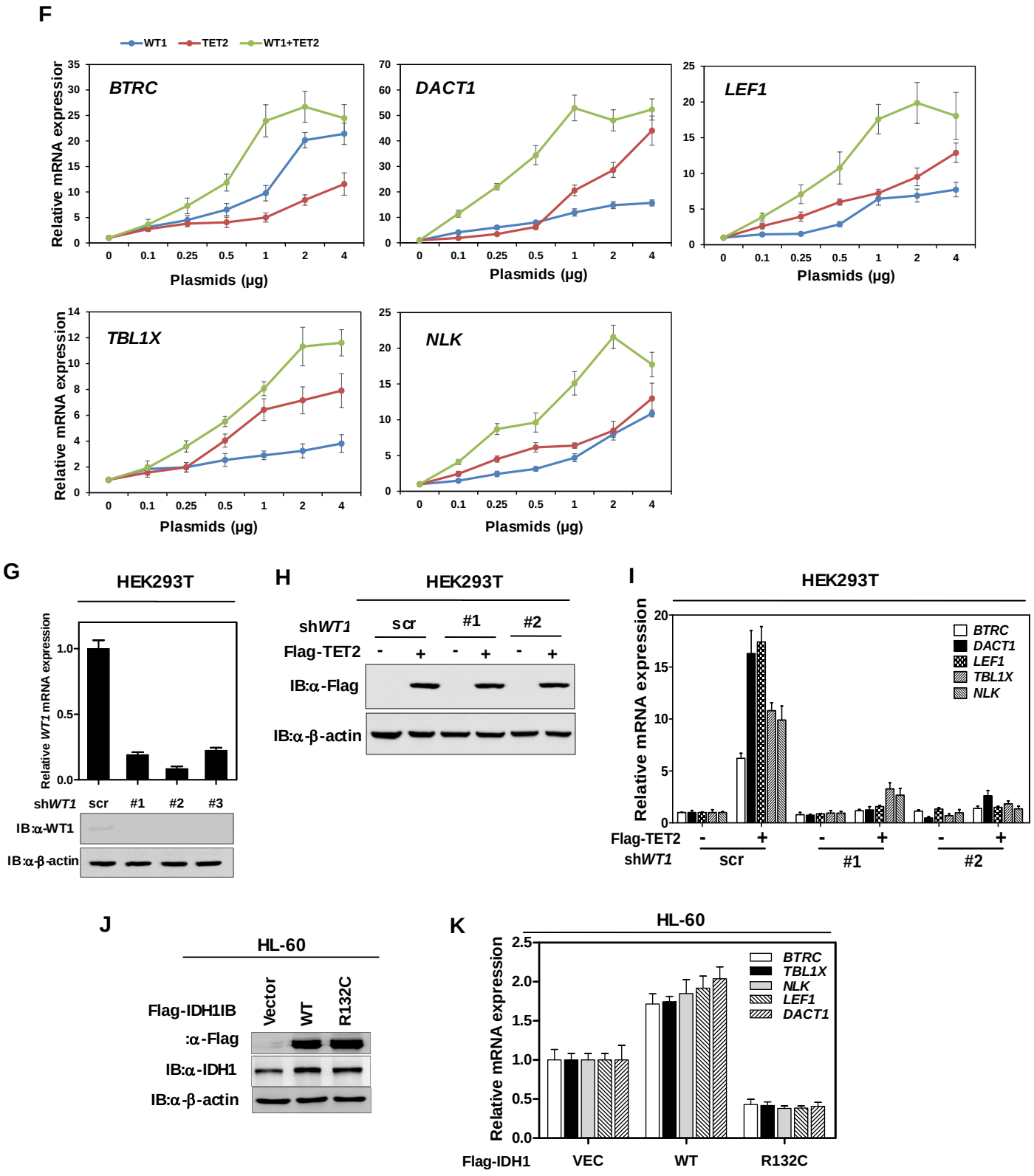

F
WT1
TET2
WT1+TET2
BTRC
DACT1
LEF1
TBL1X
NLK
Relative mRNA expression
Plasmids (μg)
0
0.1
0.25
0.5
1
2
4
G
HEK293T
Relative WT1 mRNA expression
shWT1
scr
#1
#2
#3
IB:α-WT1
IB:α-β-actin
H
HEK293T
shWT1
scr
#1
#2
Flag-TET2
IB:α-Flag
IB:α-β-actin
I
HEK293T
Relative mRNA expression
BTRC
DACT1
LEF1
TBL1X
NLK
Flag-TET2
shWT1
scr
#1
#2
J
HL-60
Flag-IDH1IB
Vector
WT
R132C
:α-Flag
IB:α-IDH1
IB:α-β-actin
K
HL-60
Relative mRNA expression
BTRC
TBL1X
NLK
LEF1
DACT1
Flag-IDH1
VEC
WT
R132C

Figure S2. TET2 directly binds to WT1

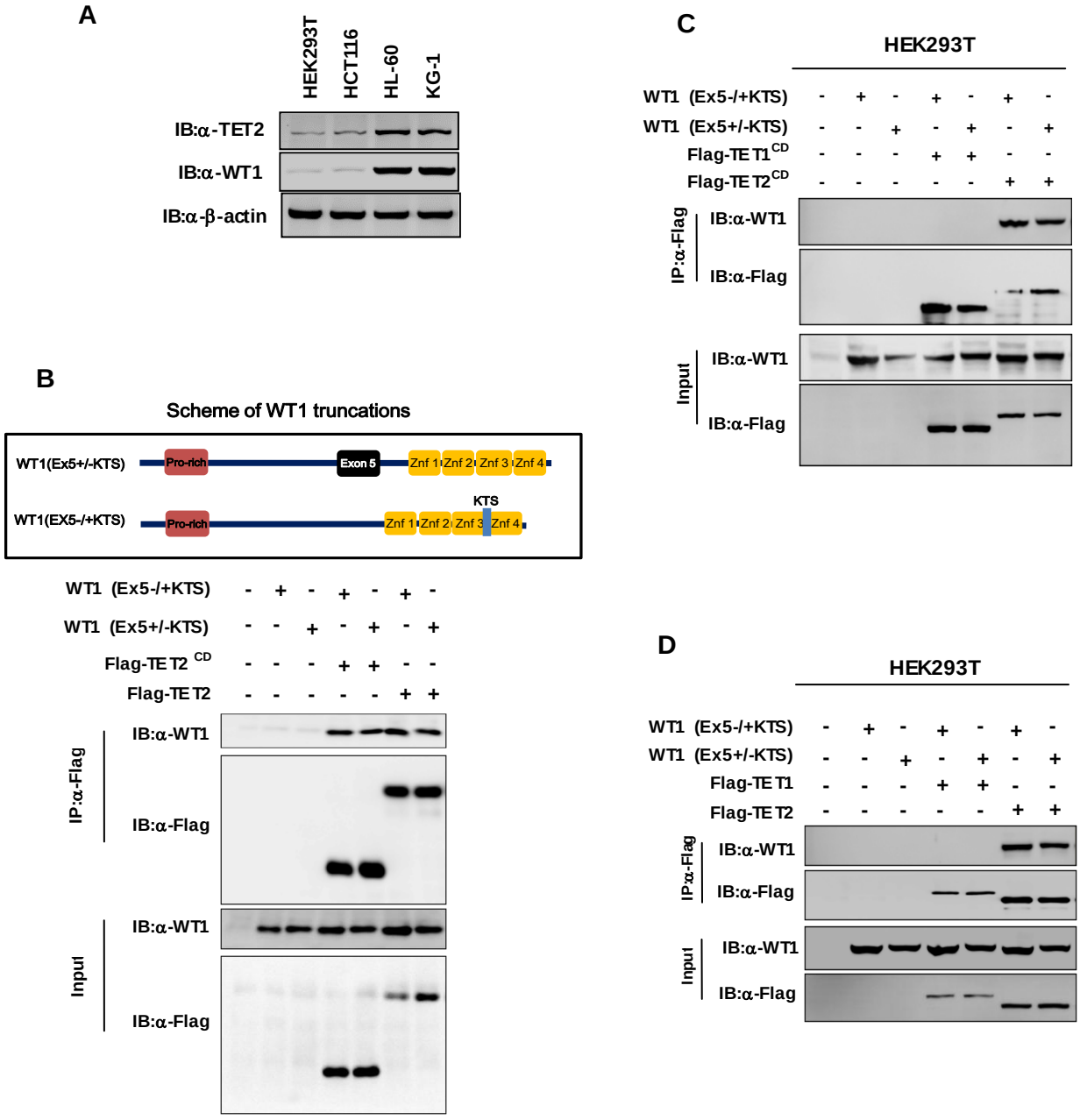
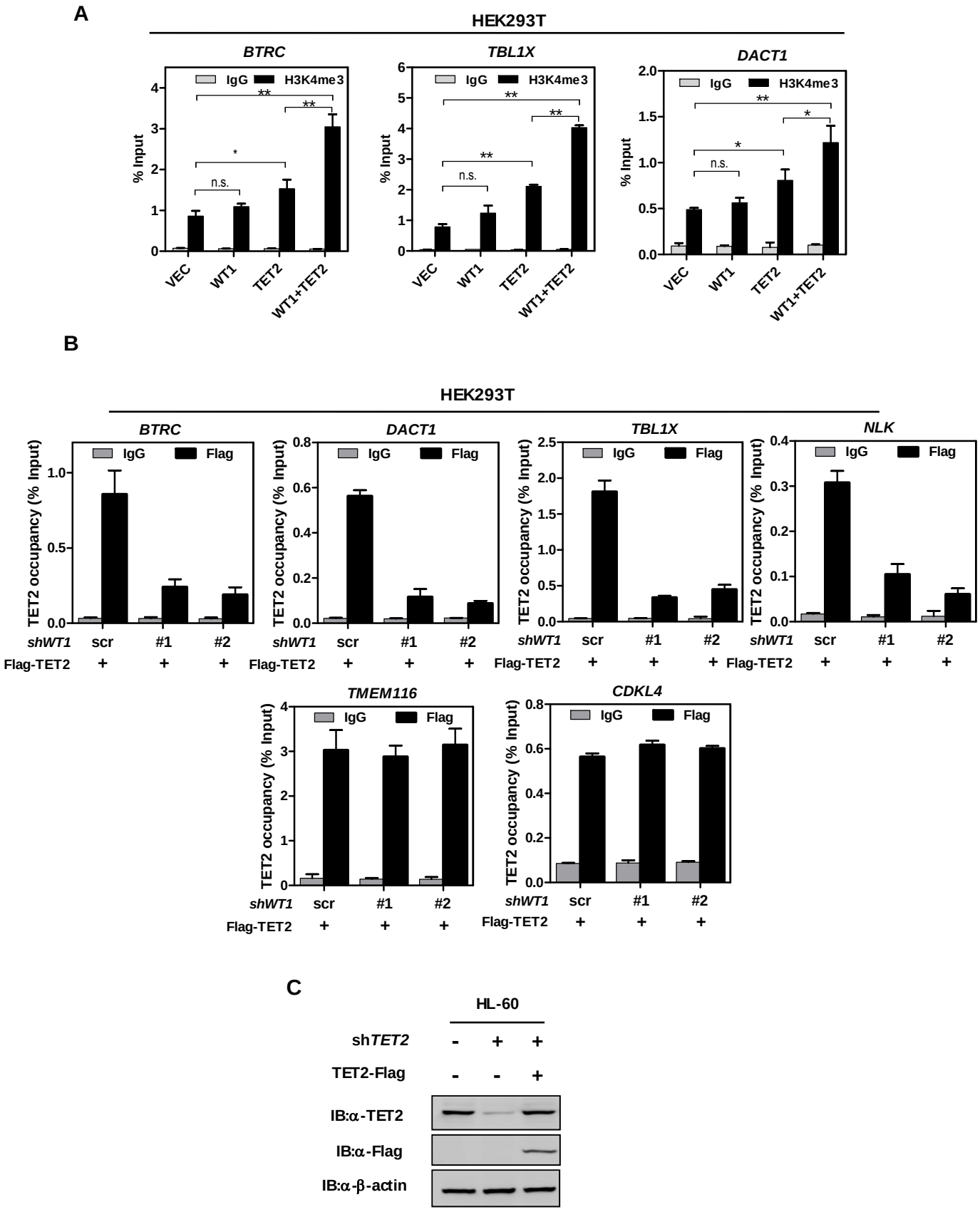
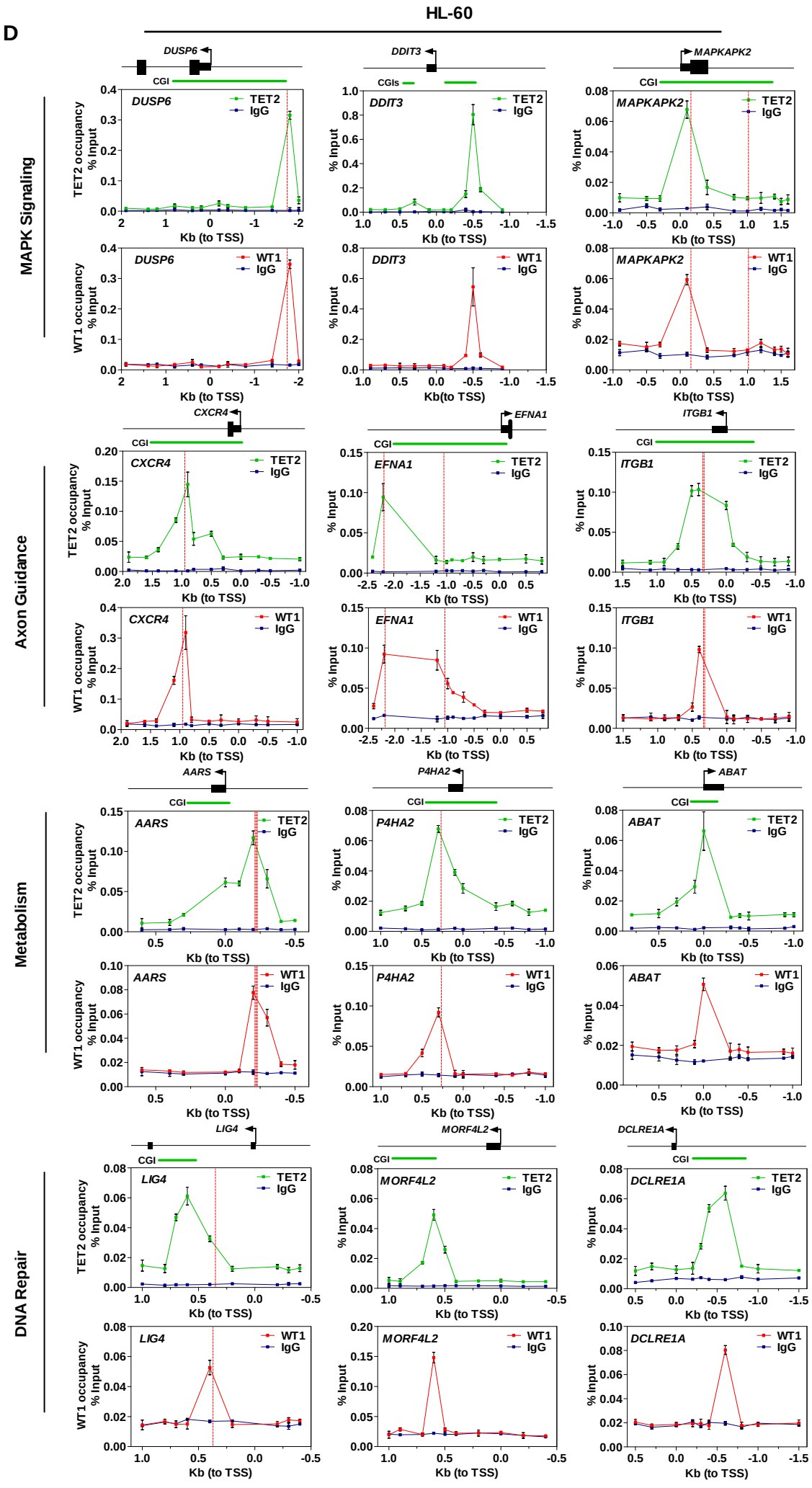


Figure S3. TET2 is recruited by WT1 to its target genes




D
HL-60
MAPK Signaling
Axon Guidance
Metabolism
DNA Repair
DUSP6
DDIT3
MAPKAPK2
CXCR4
EFNA1
ITGB1
AARS
P4HA2
ABAT
LIG4
MORF4L2
DCLRE1A
CGI
CGIs
TET2
WT1
IgG
TET2 occupancy % Input
WT1 occupancy % Input
% Input
Kb (to TSS)
Kb(to TSS)

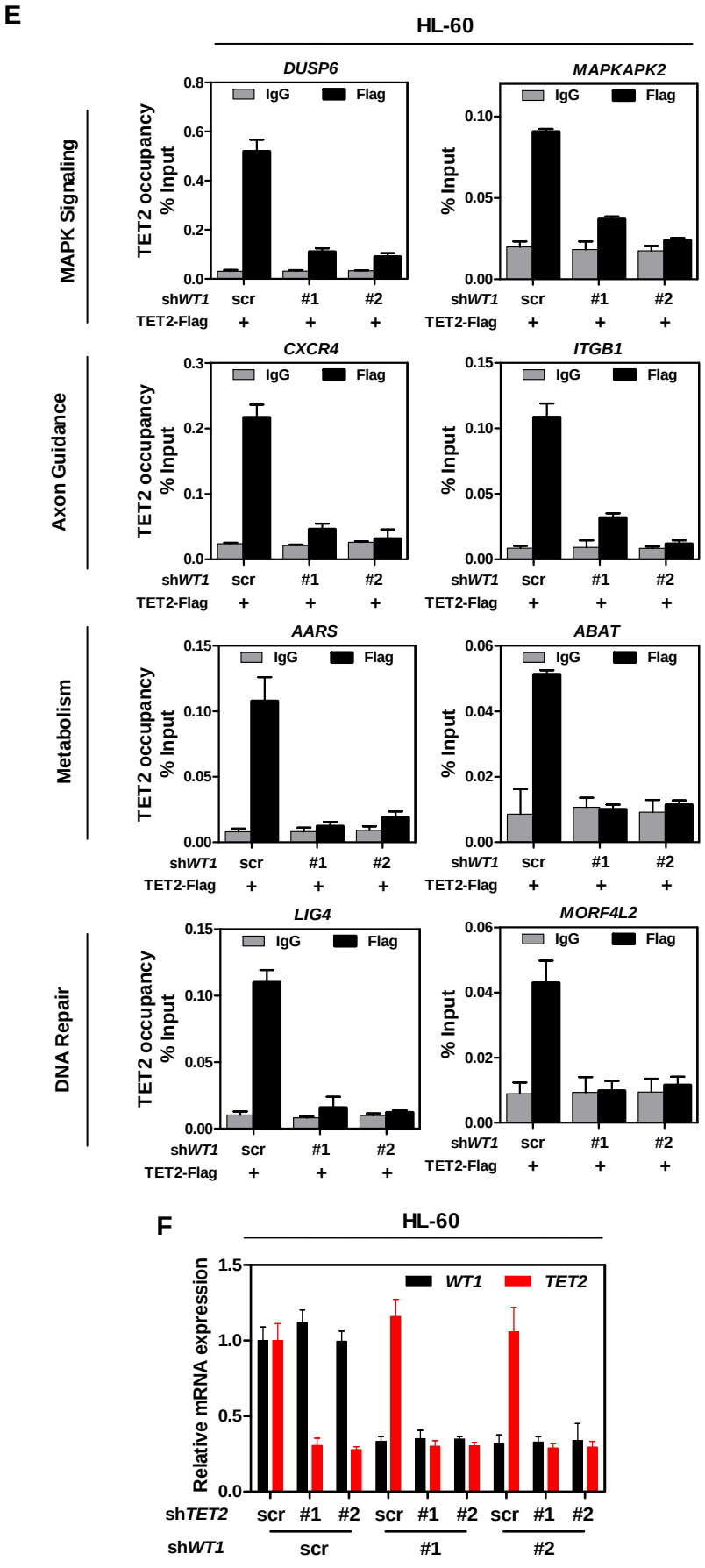

E
HL-60
MAPK Signaling
Axon Guidance
Metabolism
DNA Repair
DUSP6
MAPKAPK2
CXCR4
ITGB1
AARS
ABAT
LIG4
MORF4L2
IgG
Flag
TET2 occupancy % Input
% Input
shWT1
scr
#1
#2
TET2-Flag
+
F
HL-60
WT1
TET2
Relative mRNA expression
shTET2
shWT1

Figure S4. TET2 inhibits leukemia cell proliferation in a WT1-dependent manner

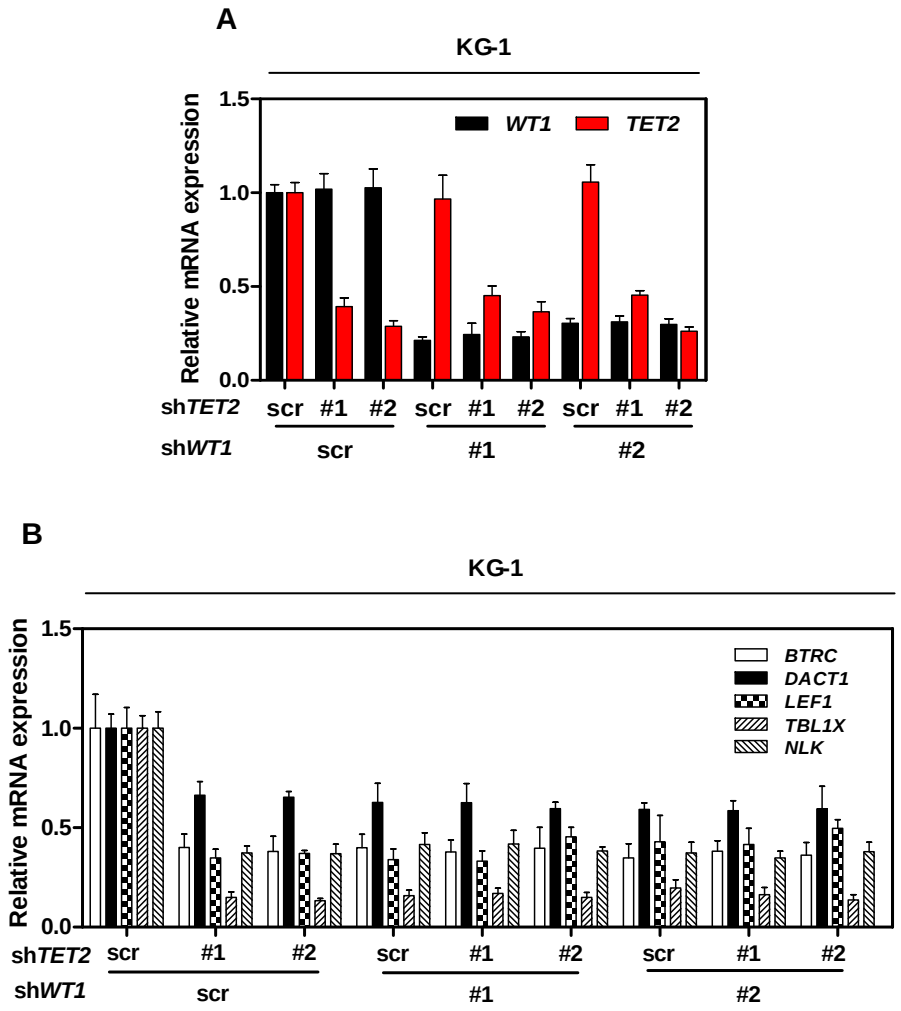
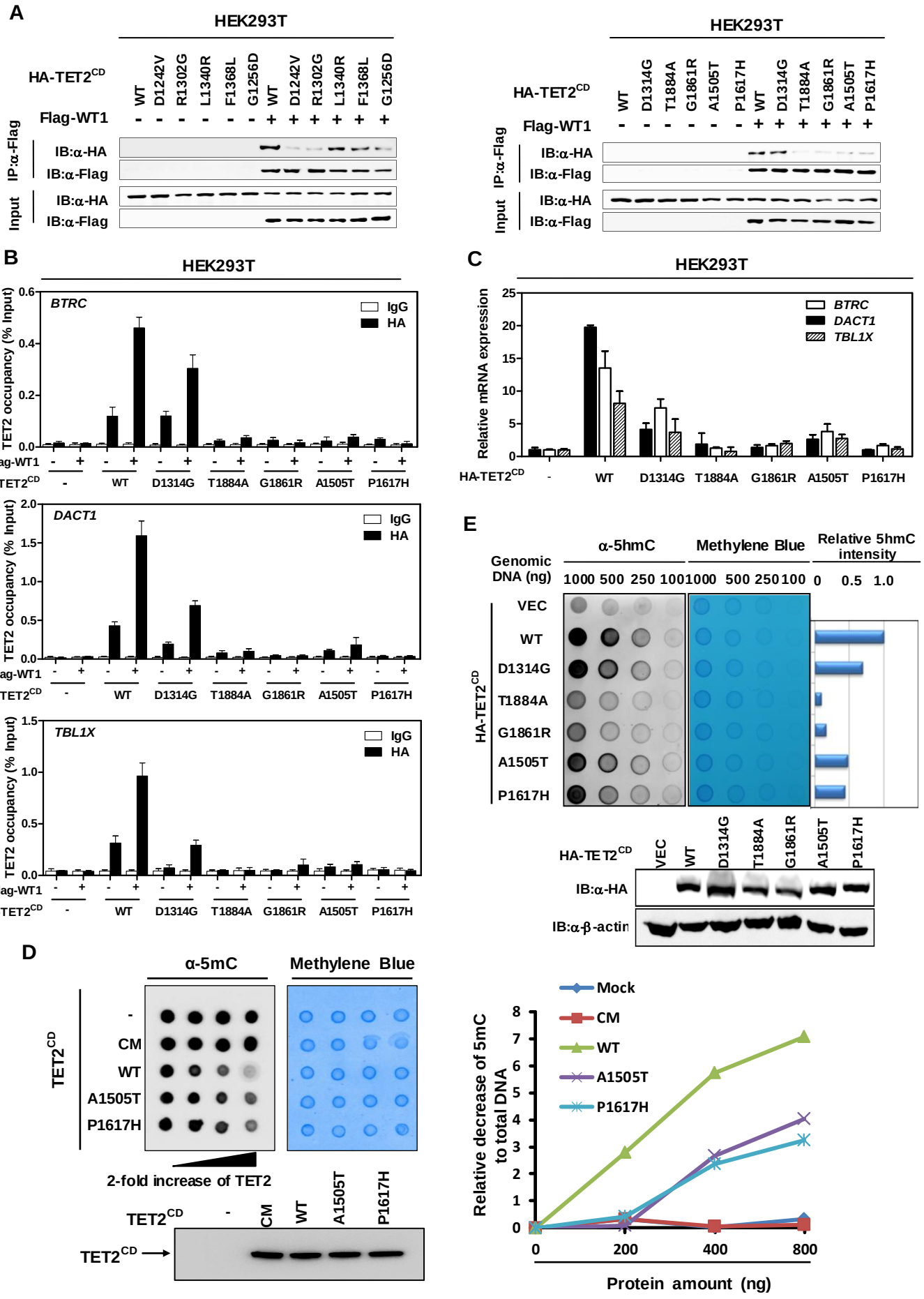


Figure S5. AML-derived mutations in TET2 disrupt WT1



Supplemental Figure Legends

Figure S1. TET2 activates WT1 target genes, related to Figure 1.

(A) Over-expression of WT1 does not affect *TET2* mRNA expression. The (-KTS) and (+KTS) isoforms of WT1 were transiently overexpressed in HEK293T (left), HCT116 (middle), and HeLa (right) cells, and the mRNA expression of *TET2*, *BTRC* and *TBL1X* was determined by qRT-PCR.

(B) Over-expression of TET2 does not affect *WT1* mRNA expression. Flag tagged full-length TET2 was transiently overexpressed in HEK293T (left), HCT116 (middle), and HeLa (right) cells, and the mRNA expression of *WT1*, *BTRC* and *TBL1X* was determined by qRT-PCR.

(C) Flag-tagged wild-type (WT) or catalytic inactive mutant (CM) of TET2 was expressed either singularly or with HA-tagged WT1 in HEK293T cells. Whole cell lysate was subjected to western blot to detect ectopic proteins of TET2 and WT1.

(D) HEK293T cells were transiently transfected with plasmids expressing indicated genes as described in (D), and the mRNA expression of indicated WT1-target genes was determined by qRT-PCR.

(G) Flag-tagged TET2 and/or HA-tagged WT1 were overexpressed in HEK293T cells as indicated. Whole cell lysate was subjected to western blot to detect the protein levels of ectopically expressed TET2 and WT1. Control indicates the same loading control sample in different panels.

(F) HEK293T cells were transiently transfected with plasmids expressing

WT1-HA or Flag-TET2 as described in (F), and the mRNA expression of indicated WT1-target genes was determined by qRT-PCR.

(G) HEK293T cells were transduced with retrovirus expressing different shRNAs against *WT1*, and the *WT1* knockdown efficiency was verified by both qRT-PCR and western blot.

(H, I) Stable HEK293T cells with or without *WT1* knockdown were generated as described above in (H), and these cells were overexpressed with Flag-tagged full-length TET2. The presence of ectopically expressed TET2 protein was determined by western blot (I), and the mRNA expression of indicated WT1-target genes was determined by qRT-PCR (J).

(J, K) HL-60 cells were transduced with retrovitrus expressing Flag-tagged wild-type or R132C mutant IDH1. The presence of ectopically expressed IDH1 protein was determined by western blot (J), and the mRNA expression of indicated WT1-target gene was determined by qRT-PCR (K).

Shown are average values of triplicated results with standard deviation (S.D.).

Figure S2. TET2 directly binds to WT1, related to Figure 2.

(A) Endogenous protein level of TET2 and WT1 in HEK293T, HCT116, HL-60 and KG-1 cells was determined by western blot. β-actin was detected as a loading control.

(B) Both the (Ex5-/+KTS) and (Ex5+/-KTS) isoforms of WT1 bind to TET2. The indicated two WT1 isoforms were overexpressed in HEK293T cells together

with full-length or the catalytic domain of TET2, and the protein association of WT1 with TET2 was determined by western blot.

(C) WT1 binds to the catalytic domain of TET2, but not TET1. The catalytic domain of TET1 or TET2 was co-expressed with the (-KTS) and (+KTS) isoforms of WT1 in HEK293T cells, and the protein association of WT1 with TET^{CD} proteins was determined by western blot.

(D) WT1 binds to full-length TET2, but not TET1. The full-length of TET1 or TET2 was co-expressed with (-KTS) and (+KTS) isoforms of WT1 in HEK293T cells, and the protein interaction of WT1 with full-length TET proteins was determined by western blot.

Figure S3. TET2 is recruited by WT1 to its target genes, related to Figure 3.

(A) Flag-TET2 was overexpressed either singularly or with WT1-HA in HEK293T cells, and the H3K4me3 enrichment at the promoter regions of indicated WT1-target genes was determined by ChIP-qPCR. Rabbit IgG was included as a negative control.

(B) Flag-TET2 was transiently overexpressed in stable *WT1* knockdown HEK293T cells. The occupancy of TET2 on the promoter regions of indicated WT1-target genes and two non-WT1 target genes (i.e. *TMEM116* and *CDKL4*) (Deplus et al., 2013) was determined by ChIP-qPCR. Mouse IgG was included as a negative control.

(C) HL-60 cells were transduced with retrovirus expressing shRNA against *TET2* either singularly or together with retrovirus re-expressing Flag-tagged TET2 as indicated. Protein levels of endogenous and ectopically expressed TET2 were determined by western blot.

(D) In stable HL-60 cells in (C), the occupancy of TET2-Flag and endogenous WT1 at the promoter regions of indicated WT1-target genes was determined by ChIP-qPCR. Mouse IgG and rabbit IgG were included as negative controls. Arrow denotes promoter orientation and CGI (green line) indicates CpG islands, red dash line indicate WT1 binding motif.

(E) HL-60 cells were transduced with retrovirus expressing different shRNAs against *WT1* and retrovirus expressing TET2-Flag. The occupancy of TET2-Flag at the promoter regions of indicated WT1-target genes was determined by ChIP-qPCR. Mouse IgG was included as a negative control.

(F) HL-60 cells were transduced with retrovirus expressing different shRNAs against *WT1* and/or *TET2*. The mRNA expression of *WT1* and *TET2* genes was determined by qRT-PCR.

Shown are average values of triplicated results with standard deviation (S.D.)

Figure S4. TET2 inhibits leukemia cell proliferation in a WT1-dependent manner, related to Figure 4.

(A) KG-1 cells were transduced with retrovirus expressing different shRNAs against *WT1* and/or *TET2*. The knockdown efficiency of *WT1* and *TET2* genes

was determined by qRT-PCR.

(B) In stable KG-1 cells in (A), the mRNA expression of indicated WT1-target genes was determined by qRT-PCR.

Shown are average values of triplicated results with standard deviation (S.D.). n.s. = not significant.

Figure S5. **AML-derived mutations in TET2 disrupt WT1 binding, related to Figure 5.**

(A) Wild-type and a panel of AML-derived TET2 mutants were generated in the catalytic domain of TET2 ($TET2^{CD}$), and were transiently co-expressed with Flag-WT1 in HEK293T cells. The interaction of WT1 with TET2 mutants was determined by IP-western analysis.

(B) Wild-type and several AML-derived, WT1-binding defective TET2 mutants were overexpressed in HEK293T cells singularly or together with Flag-WT1. The TET2 occupancy on the promoter regions of indicated WT1-target genes was determined by ChIP-qPCR. Mouse IgG was included as a negative control. Shown are average values of triplicated results with standard deviation (S.D.)

(C) Wild-type and several AML-derived, WT1-binding defective TET2 mutants were overexpressed in HEK293T cells. The mRNA expression of indicated WT1-target genes was determined by qRT-PCR. Shown are average values of triplicated results with standard deviation (S.D.)

(D) *In vitro* TET2 catalytic activity assay. HA-tagged wild-type, catalytic-inactive mutant (CM) or AML-derived mutants of TET2 were overexpressed in HEK293T. Genomic DNA from MDM cells were incubated with HA-beads purified TET2 proteins at 37°C for 1 hr as described in "Experimental Procedures". TET2-catalyzed oxidation was measured by 5mC decrease by dot-blot assay. Methylene blue staining indicates equal genomic DNA loading.

(E) *In vivo* TET2 catalytic activity assay. Wild-type and several AML-derived mutants of TET2 were overexpressed in HEK293T cells. Genomic DNA was extracted and 5hmC level was determined by dot-blot assay with series of dilutions of the genomic DNA. Protein levels of ectopically expressed TET2 were determined by western blot.

Supplemental Table Legends

Table S1. List of oligonucleotides used in this study

Primers used for retroviral-mediated shRNA against *WT1*, qRT-PCR, ChIP-qPCR, MeDIP-qPCR, hMeDIP-qPCR, and GluMS-qPCR analyses are listed.

Table S2. Mutual exclusive mutations of TET2 and WT1 in various cancers, related to Figure 6.

Mutation analysis of *WT1* and *TET2* in various human cancers in addition to AML. Data of TET2 and WT1 mutations was collected using cBioPortal (http://www.cbioportal.org) (Cerami et al., 2012; Gao et al., 2013) from previous cancer genomic studies(Cancer Genome Atlas Network.2012a; Cancer Genome Atlas Network.2012b; Cancer Genome Atlas Network.2012c; Cancer Genome Atlas Network.2013; Cancer Genome Atlas Network.2014; Ahn et al., 2014; Cerami et al., 2012; Gao et al., 2013; Hodis et al., 2012; Imielinski et al., 2012; Kandoth et al., 2013; Krauthammer et al., 2012; Lohr et al., 2014; Peifer et al., 2012; Seshagiri et al., 2012).

Table S2. Mutual exclusive mutations of TET2 and WT1 in various cancers

Cancer types	Patient NO.	WT1 mutation	TET2 mutation	Co-occurrence	Reference
Bladder Urothelial Carcinoma	130	2	4	0	TCGA, Nature 2014
Breast Invasive Carcinoma	507	1	2	0	TCGA, Nature 2012
Colorectal Adenocarcinoma	72	3	5	1	Genentech, Nature 2012
Colorectal Adenocarcinoma	224	6	4	2	TCGA, Nature 2012
Kidney Renal Clear Cell Carcinoma	424	3	9	0	TCGA, Nature 2013
Kidney Renal Papillary Cell Carcinoma	168	1	2	0	TCGA, Provisional
Liver Hepatocellular Carcinoma	231	1	4	0	AMC, Hepatology in press
Lung Adenocarcinoma	183	5	2	0	Broad, Cell 2012
Lung Adenocarcinoma	230	8	3	0	TCGA, Nature, in press
Lung Squamous Cell Carcinoma	178	4	4	0	TCGA, Nature 2012
Small Cell Lung Cancer	29	1	2	0	CLCGP, Nature Genetics 2012
Multiple Myeloma	205	1	1	0	Broad, Cancer Cell 2014
Skin Cutaneous Melanoma	121	4	3	0	Broad, Cell 2012
Skin Cutaneous Melanoma	278	8	13	0	TCGA, Provisional
Skin Cutaneous Melanoma	91	2	2	0	Yale, Nature Genetics 2012
Stomach Adenocarcinoma	220	2	8	1	TCGA, Provisional
Uterine Corpus Endometrioid Carcinoma	248	3	13	0	TCGA, Nature 2013

Supplemental Experimental Procedures

Antibodies

Antibodies against TET2 (Abcam), WT1 (Santa Cruz), GFP (Abmart), HA (Santa Cruz), Flag (Sigma, ShanghaiGenomics), H3K4me3 (Abcam), 5mC (Eurogentec), 5hmC (Active Motif), 5fC (Active Motif), 5caC (Active Motif), mouse IgG (Santa Cruz) and rabbit IgG (Santa Cruz) were purchased commercially.

Plasmids

The cDNA encoding full-length human WT1, TET1, TET2, p53 or PEPCK1 were cloned into Flag, HA or GFP-tagged vectors (pcDNA-Flag; pcDNA3-HA; pEGFP-C3). DNA fragments of TET2 catalytic domain ($TET2^{CD}$) and truncated TET2 and WT1 were sub-cloned into pcDNA-Flag, pcDNA3-HA or pEGFP-C3 vectors. Plasmid for shRNA expression with hygromycin resistance (pMKO.1-hyg) was achieved by replacing the puromycin resistance gene with hygromycin resitance gene after introducing BamHI and NotI restriction sites into pMKO.1-puro vector. Flag-tagged human TET2 gene (full length) was subcloned into pQCXIH vector. Flag-tagged human IDH1 gene was subcloned into pBABE vector. HA-tagged CD domain of wild-type TET2 and AML-derived mutants (A1505T, P1617H, and G1861R) were subcloned into lentiviral vector pCDH-Puro. Point mutations were generated as decribed previously(Hartkamp et al., 2010) by site-directed mutagenesis using a QuickChange Site-Directed Mutagenesis kit (Stratagene). All expression constructs were verified by DNA

sequencing.

Cell culture and transfection

HEK293T and HeLa cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM) (Invitrogen, Shanghai, China) supplemented with 10% new born bovine serum (Biochrom, Germany) in the presence of penicillin, streptomycin, and 8 mM L-glutamine (Invitrogen). Plasmid transfection was carried out either by the calcium phosphate method or lipofectamine 2000 (Invitrogen). HCT116 cells were maintained in McCoy's 5A Medium (Gibco) supplemented with 10% fetal bovine serum in the presence of penicillin, streptomycin. Mouse E14TG2a ESCs were cultured on gelatin-coated tissue culture dishes in standard ES medium (DMEM supplemented with 15% FBS, 0.1mM β-mercaptoethanol (Gibco), 1X nonessential amino acids (Gibco), 2mM L-glutamine (Gibco), and with 1000 U/ml LIF (Millipore).Moreover, KG-1 and HL-60 cells were maintained in RPMI-1640 medium supplemented with 20% heat-inactivated fetal bovine serum (MP Biomedicals), 2 mmol/L L-glutamine, 100 U/mL penicillin/streptomycin.

Immunoprecipitation and western blotting

Cells were lysed in ice-cold NP-40 buffer [50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 0.1% NP-40)] containing protease inhibitor cocktail (Roche). Immunoprecipitation was carried out either by incubating Flag beads (Sigma)

at 4°C with lysate for 3 hrs or by incubating appropriate antibody with cell lysate for 1 hr, followed by incubating with Protein-A beads (Upstate) for another 2 hrs at 4°C before beads were washed for three times with ice-cold NP-40 buffer. Standard western blot protocols were adopted.

RNA isolation and quantitative real-time PCR

Total RNA was isolated from cultured cells using Trizol reagent (Invitrogen) following the manufacturer's instructions. RNA was reverse transcribed with oligo-dT primers and preceded to real-time PCR with gene-specific primers in the presence of SYBR Premix Ex Taq (TaKaRa). β-actin was used as a housekeeping control. Primer sequences were present in the supplementary Table S1.

Protein concentration determination

Protein concentrations were determined by the Bradford method (Bradford, 1976) using a Bio-Rad Protein Assay Kit (Bio-Rad). SYPRO Ruby staining signal intensity was quantified by generating density histograms for each protein band and then determining the area within the corresponding peak using ImageJ.

Colony formation assay

For colony-formation assay, HL-60 cells were plated (1000 cells per 35mm

dish) in methylcellulose-based MethoCult medium (Stem Cell Technologies). Culture dishes were incubated for 6 days at 37°C in a humid atmosphere with 5% CO_2. Colonies (>50 µm diameter) were counted using Olympus fluorescence microscope IX18.

In vitro TET2 activity assay

In vitro assay of TET2 catalytic activity was performed following the procedures previously described (Zhang et al., 2014; Zhang et al., 2013) . Briefly, genomic DNA from monocyte-derived macrophage (MDM) cells was isolated using Wizard Genomic DNA Purification Kit (Promega) following the manufacturer's instructions, sonicated (Covaris Sonicator, 13 min at 4°C) and used as substrate. To prepare the TET enzyme, HA-tagged wild-type and AML-derived TET2-CD mutants were ectopically expressed in HEK293T cells and immunoprecipitated in a NP-40 buffer (50 mmol/L Tris pH 7.5, 150 mmol/L NaCl, 0.3% Nonidet P-40, 1% protease inhibitor (Pierce)) using anti-HA beads (Cell Signaling). HA-TET2 proteins were eluted by HA peptide from the agarose beads and used as the enzyme. To start the reaction, 150 ng sonicated genomic DNA and 200, 400, 800 ng immunopurified proteins were added to a 20 µl reaction buffer containing 50 mM HEPES buffer (pH 8.0), 100 mM ammonium iron (II) sulfate, 1 mM α-ketoglutarate, 2 mM ascorbic acid, 2.5 mM DTT, 100 mM NaCl, 1.2 mM ATP. The reactions were incubated at 37°C for 1 h with constant shaking on an Eppendorf ThermoMixer (700 r.p.m),

denatured in 0.4 M NaOH, 10 mM EDTA at 95°C for 10 min and then neutralized by adding an equal volume of cold 2 M ammonium acetate (pH 7.0). 4 µl reaction mixtures were spotted on nitrocellulose membrane (0.22µm, BioRad) and dried at room temperature. The membrane was then cross linked with DNA (Auto cross link, 3 min) using UV Stratalinker 2400 (STRATAGENE) and blocked with 5% non-fat milk (LabScientific) and incubated over night with antibodies recognizing 5mC (Epigentek, 1:500 dilution). The membrane was washed three times with wash buffer (1% PBS, 0.1% Tween 20), each for 10 min. HRP-conjugated secondary antibody (Millipore 1:1000) was then added, incubated for 2 h at room temperature, and washed three times with the same wash buffer described above and the binding was visualized by LumiGLO Reserve chemiluminescence (KPL). Loading controls were provided by either immunoblotting of equal amount of unreacted substrate DNA using 5mC antibody or staining with methylene blue.

In vivo TET2 activity assay

Genomic DNA was extracted using phenol chloroform method and the concentration of DNA was quantified (NanoDrop). Genomic DNA was serially diluted, and spotted on a nitrocellulose membrane (Whatman). The membrane was placed under an ultraviolet lamp for 20 min to crosslink the DNA and then blocked with 5% milk in TBS-Tween20 for 1 h, followed by incubating with the anti-5hmC (Active Motif) antibody at 4°C overnight. After incubation with a HRP-conjugated secondary antibody (GeneScript) for 1 h at room temperature,

the membrane was washed with TBS-Tween 20 for three times and then scanned by a Typhoon scanner (GE Healthcare). The quantification of 5hmC was done by Image-Quanta software (GE Healthcare).

Supplemental References

Cancer Genome Atlas Network.(2012a). Comprehensive genomic characterization of squamous cell lung cancers. Nature *489*, 519-525.

Cancer Genome Atlas Network. (2012b). Comprehensive molecular characterization of human colon and rectal cancer. Nature *487*, 330-337.

Cancer Genome Atlas Network. (2012c). Comprehensive molecular portraits of human breast tumours. Nature *490*, 61-70.

Cancer Genome Atlas Network. (2013). Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature *499*, 43-49.

Cancer Genome Atlas Network. (2014). Comprehensive molecular characterization of urothelial bladder carcinoma. Nature *507*, 315-322.

Ahn, S. M., Jang, S. J., Shim, J. H., Kim, D., Hong, S. M., Sung, C. O., Baek, D., Haq, F., Ansari, A. A., Lee, S. Y.*, et al.* (2014). A genomic portrait of resectable hepatocellular carcinomas: Implications of RB1 and FGF19 aberrations for patient stratification. Hepatology.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical biochemistry *72*, 248-254.

Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., Jacobsen, A., Byrne, C. J., Heuer, M. L., Larsson, E.*, et al.* (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov *2*, 401-404.

Deplus, R., Delatte, B., Schwinn, M. K., Defrance, M., Mendez, J., Murphy, N., Dawson, M. A., Volkmar, M., Putmans, P., Calonne, E.*, et al.* (2013). TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. The EMBO journal *32*, 645-655.

Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E.*, et al.* (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal *6*, pl1.

Hartkamp, J., Carpenter, B., and Roberts, S. G. (2010). The Wilms' tumor suppressor protein WT1 is processed by the serine protease HtrA2/Omi. Molecular cell *37*, 159-171.

Hodis, E., Watson, I. R., Kryukov, G. V., Arold, S. T., Imielinski, M., Theurillat, J. P., Nickerson, E., Auclair, D., Li, L., Place, C.*, et al.* (2012). A landscape of driver mutations in melanoma. Cell *150*, 251-263.

Imielinski, M., Berger, A. H., Hammerman, P. S., Hernandez, B., Pugh, T. J., Hodis, E., Cho, J., Suh, J., Capelletti, M., Sivachenko, A.*, et al.* (2012). Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. Cell *150*, 1107-1120.

Kandoth, C., Schultz, N., Cherniack, A. D., Akbani, R., Liu, Y., Shen, H., Robertson, A. G., Pashtan, I., Shen, R., Benz, C. C.*, et al.* (2013). Integrated genomic characterization of endometrial carcinoma. Nature *497*, 67-73.

Krauthammer, M., Kong, Y., Ha, B. H., Evans, P., Bacchiocchi, A., McCusker, J. P., Cheng, E., Davis, M. J., Goh, G., Choi, M.*, et al.* (2012). Exome sequencing identifies recurrent somatic RAC1 mutations in melanoma. Nat Genet *44*, 1006-1014.

Lohr, J. G., Stojanov, P., Carter, S. L., Cruz-Gordillo, P., Lawrence, M. S., Auclair, D., Sougnez, C., Knoechel, B., Gould, J., Saksena, G.*, et al.* (2014). Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. Cancer Cell *25*, 91-101.

Peifer, M., Fernandez-Cuesta, L., Sos, M. L., George, J., Seidel, D., Kasper, L. H., Plenker, D., Leenders, F., Sun, R., Zander, T.*, et al.* (2012). Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. Nat Genet *44*, 1104-1110.

Seshagiri, S., Stawiski, E. W., Durinck, S., Modrusan, Z., Storm, E. E., Conboy, C. B., Chaudhuri, S., Guan, Y., Janakiraman, V., Jaiswal, B. S.*, et al.* (2012). Recurrent R-spondin fusions in colon cancer. Nature *488*, 660-664.

Song, C. X., Szulwach, K. E., Dai, Q., Fu, Y., Mao, S. Q., Lin, L., Street, C., Li, Y., Poidevin, M., Wu, H.*, et al.* (2013). Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. Cell *153*, 678-691.

Zhang, L., Chen, W., Iyer, L. M., Hu, J., Wang, G., Fu, Y., Yu, M., Dai, Q., Aravind, L., and He, C. (2014). A TET homologue protein from Coprinopsis cinerea (CcTET) that biochemically converts 5-methylcytosine to 5-hydroxymethylcytosine, 5-formylcytosine, and 5-carboxylcytosine. J Am Chem Soc *136*, 4801-4804.