## Supplemental Figures and Legends


**Figure S1: SIRT2 expression is modestly increased in $TS^{KI4}$ cells relative to $TS^{WT}$ cells. Related to Figure 1.**
(A) qRT-PCR data are shown for $TS^{WT}$ and $TS^{KI4}$ cells expressing a control shRNA (C) or $TS^{KI4}$ cells expressing two independent HDAC6 shRNAs 1 and 2 (sh1 and sh2). Data show fold-changes relative to $TS^{WT}$ cells and are the mean ± SEM of three independent experiments.
(B) Western blots of 30 µg $TS^{WT}$, $TS^{KI4}$, sh1 and sh2 whole cell lysates are representative of four independent experiments.
Student's t-test, ***p-value < 0.001.

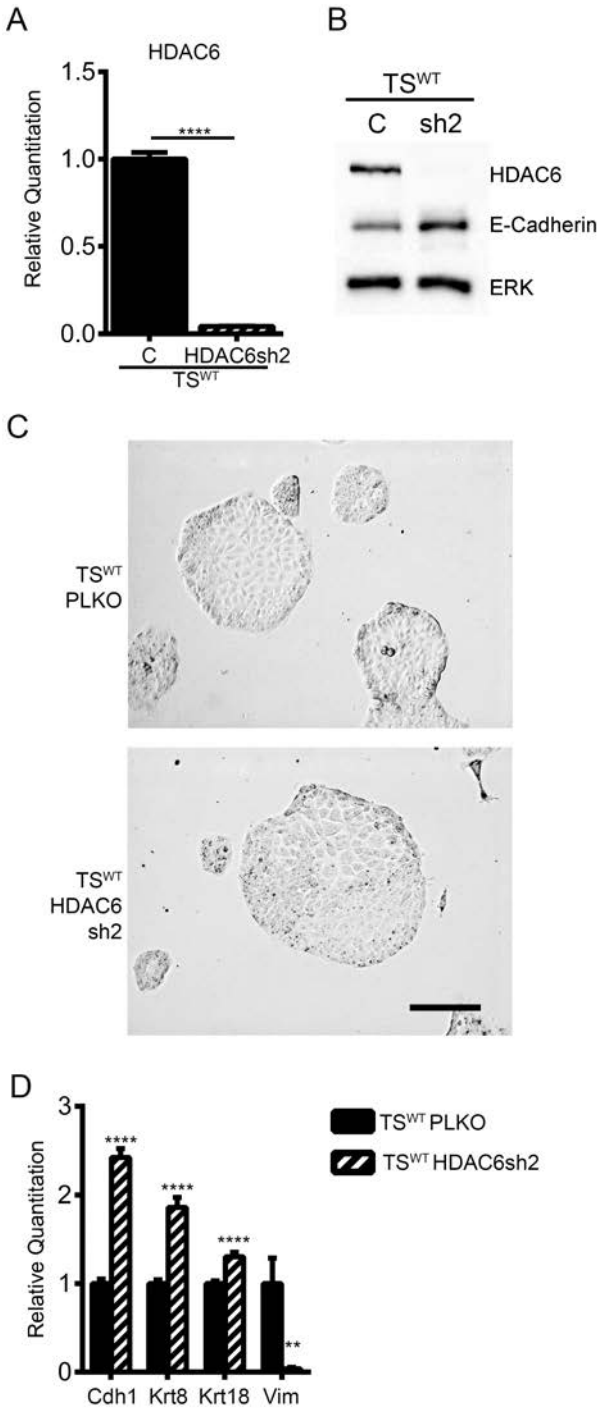


**Figure S2: TS$^{WT}$ cells with HDAC6 shRNA knockdown retain an epithelial phenotype. Related to Figure 2.**

(A) HDAC6 shRNA2 expressed in TS$^{WT}$ cells resulted in >95% knockdown of HDAC6 transcripts. qRT-PCR data show fold-change relative to TS$^{WT}$ cells expressing a control vector and are the mean ± SEM of three independent experiments.

(B) Western blots of whole cell lysates show decreased HDAC6 and increased E-cadherin protein in TS$^{WT}$ cells expressing HDAC6 shRNA relative to TS$^{WT}$ cells expressing a control vector. Data shown are representative of three independent experiments.

(C) Phase-contrast microscopy images of TS$^{WT}$ cells expressing a control vector and TS$^{WT}$ cells expressing HDAC6 shRNA2 show maintenance of an epithelial morphology. Black bar represents 200 µm. Images are representative of three independent experiments.

(D) qRT-PCR data show increased expression of epithelial markers Cdh1, Krt8, and Krt18, and decreased expression of the mesenchymal marker vimentin in TS$^{WT}$ cells expressing HDAC6 shRNA2 relative to TS$^{WT}$ cells expressing a control vector. Data shown are the mean ± SEM of three independent experiments. Student's t-test, **p-value < 0.01; ****p-value < 0.0001.

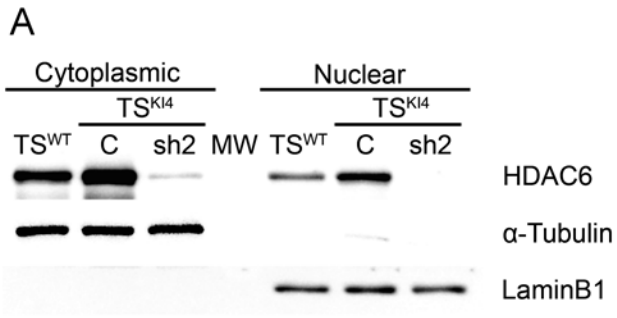


**Figure S3: Nuclear localization of HDAC6 is increased in TS$^{KI4}$ cells. Related to Figure 5.**

(A) Cytoplasmic and nuclear lysates from TS$^{WT}$, TS$^{KI4}$, and TS$^{KI4}$ HDAC6 sh2 cells probed with the indicated antibodies demonstrate the purity of cytoplasmic and nuclear fractions. Blots shown are representative of three independent experiments. Equal protein mass from cytoplasmic and nuclear fractions was run side by side on the same gel, separated by the molecular weight marker (MW, not shown).

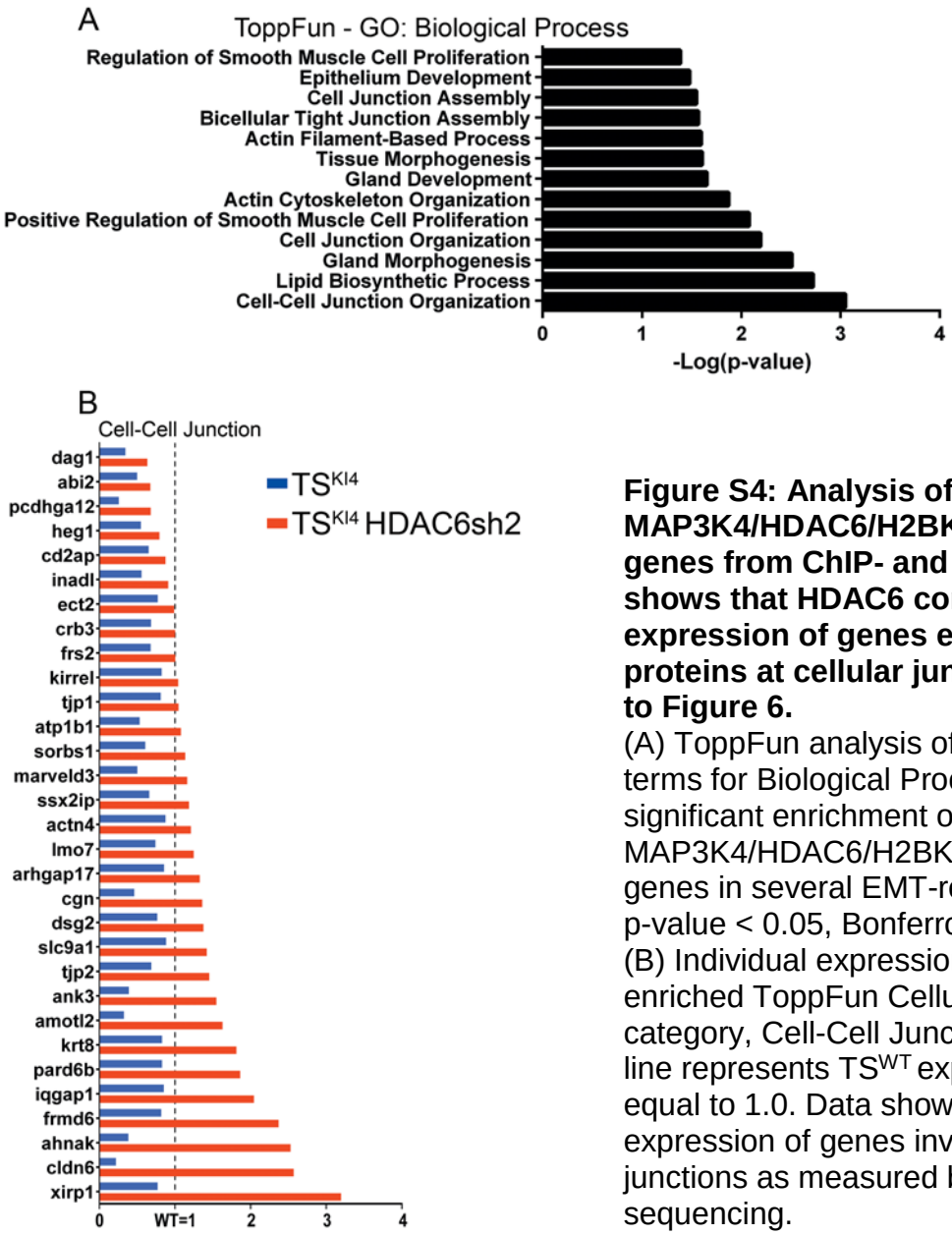


**Figure S4: Analysis of 514 MAP3K4/HDAC6/H2BK5Ac dependent genes from ChIP- and RNA-seq data shows that HDAC6 controls the expression of genes encoding proteins at cellular junctions. Related to Figure 6.**

(A) ToppFun analysis of gene ontology terms for Biological Processes shows significant enrichment of the 514 MAP3K4/HDAC6/H2BK5Ac dependent genes in several EMT-related categories. p-value < 0.05, Bonferroni corrected.

(B) Individual expression of genes from enriched ToppFun Cellular Component category, Cell-Cell Junction. The dashed line represents $TS^{WT}$ expression levels equal to 1.0. Data show relative expression of genes involved in cell-cell junctions as measured by RNA-sequencing.

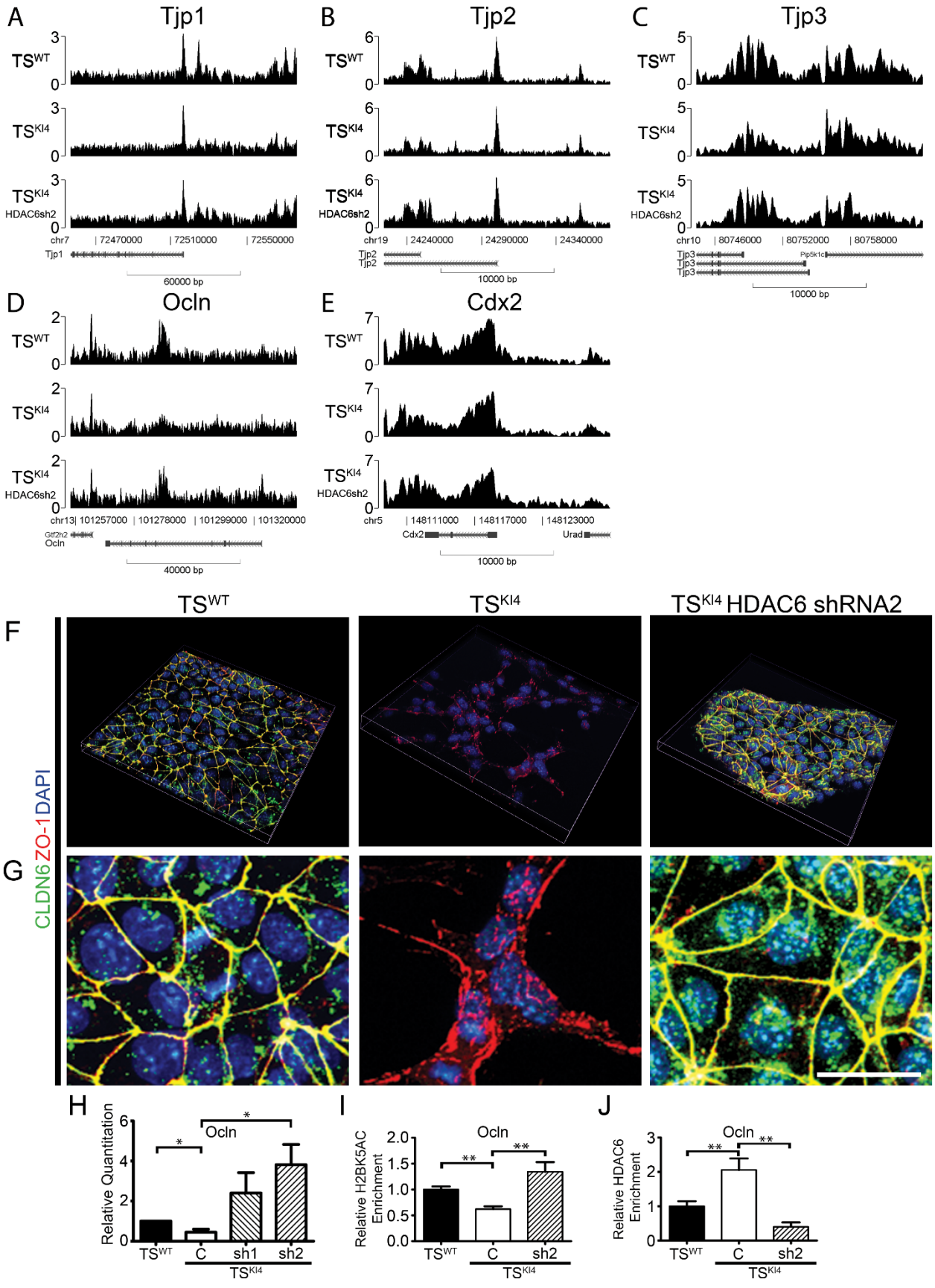
A
Tjp1
B
Tjp2
C
Tjp3
D
Ocln
E
Cdx2
TSWT
TSKI4
HDAC6sh2
chr7 | 72470000 | 72510000 | 72550000
Tjp1
60000 bp
chr19 | 24240000 | 24290000 | 24340000
Tjp2
10000 bp
chr10 | 80746000 | 80752000 | 80758000
Pip5k1c
10000 bp
chr13| 101257000 | 101278000 | 101299000 | 101320000
Gtf2h2
Ocln
40000 bp
chr5 | 148111000 | 148117000 | 148123000
Cdx2
Urad
10000 bp
F
TSWT
TSKI4
TSKI4 HDAC6 shRNA2
G
CLDN6 ZO-1 DAPI
H
Relative Quantitation
Ocln
TSWT
C
sh1
sh2
TSKI4
I
Relative H2BK5AC Enrichment
Ocln
J
Relative HDAC6 Enrichment
Ocln

**Figure S5: MAP3K4 dependent HDAC6 expression regulates tight junctions by controlling H2BK5 promoter acetylation and expression of several tight junction genes. Related to Figure 6.**

(A-D) Read density plots for the indicated genes show reduced H2BK5Ac near the TSS of tight junction genes in $TS^{KI4}$ cells that is restored by HDAC6 knockdown. MAP3K4/HDAC6 dependent changes in H2BK5Ac were measured by H2BK5Ac ChIP-seq. X-axis represents base pairs surrounding the TSS. Y-axis expresses normalized ChIP DNA sample reads divided by normalized $TS^{WT}$ input sample reads.

(E) Acetylation of the stemness gene Cdx2 is similar in $TS^{KI4}$ and $TS^{WT}$ cells. Data shown were measured as in (A-D).

(F) 3-D z-stack confocal microscopy images of ZO-1 (red) and Cldn6 (green) show loss of colocalization in $TS^{KI4}$ cells that is restored by HDAC6 knockdown. Nuclei were stained with DAPI (blue).

(G) Enlarged insets of confocal microscopy images from Figure 6. Overlay shows colocalization of Cldn6 and ZO-1 that is absent in $TS^{KI4}$ cells. Cell nuclei were stained with DAPI. White bar represents 25 µm.

(H) Restored expression of Ocln upon knockdown of HDAC6 in $TS^{KI4}$ cells. qRT-PCR data from $TS^{WT}$, $TS^{KI4}$, and $TS^{KI4}$ cells expressing HDAC6 shRNA1 or 2 (sh1 and sh2). Data show fold changes relative to $TS^{WT}$ cells and are expressed as the mean ± SEM of three independent experiments.

(I) Acetylation of H2BK5 on the Ocln promoter is restored upon knockdown of HDAC6 in $TS^{KI4}$ cells as measured by H2BK5Ac ChIP qRT-PCR. Data shown are the mean ± SEM of three independent experiments.

(J) Enrichment of HDAC6 on the Ocln promoter in $TS^{KI4}$ cells as measured by HDAC6 ChIP qRT-PCR. Data shown are the mean ± SEM of three independent experiments. Student's t-test, *p-value < 0.05; **p-value < 0.01

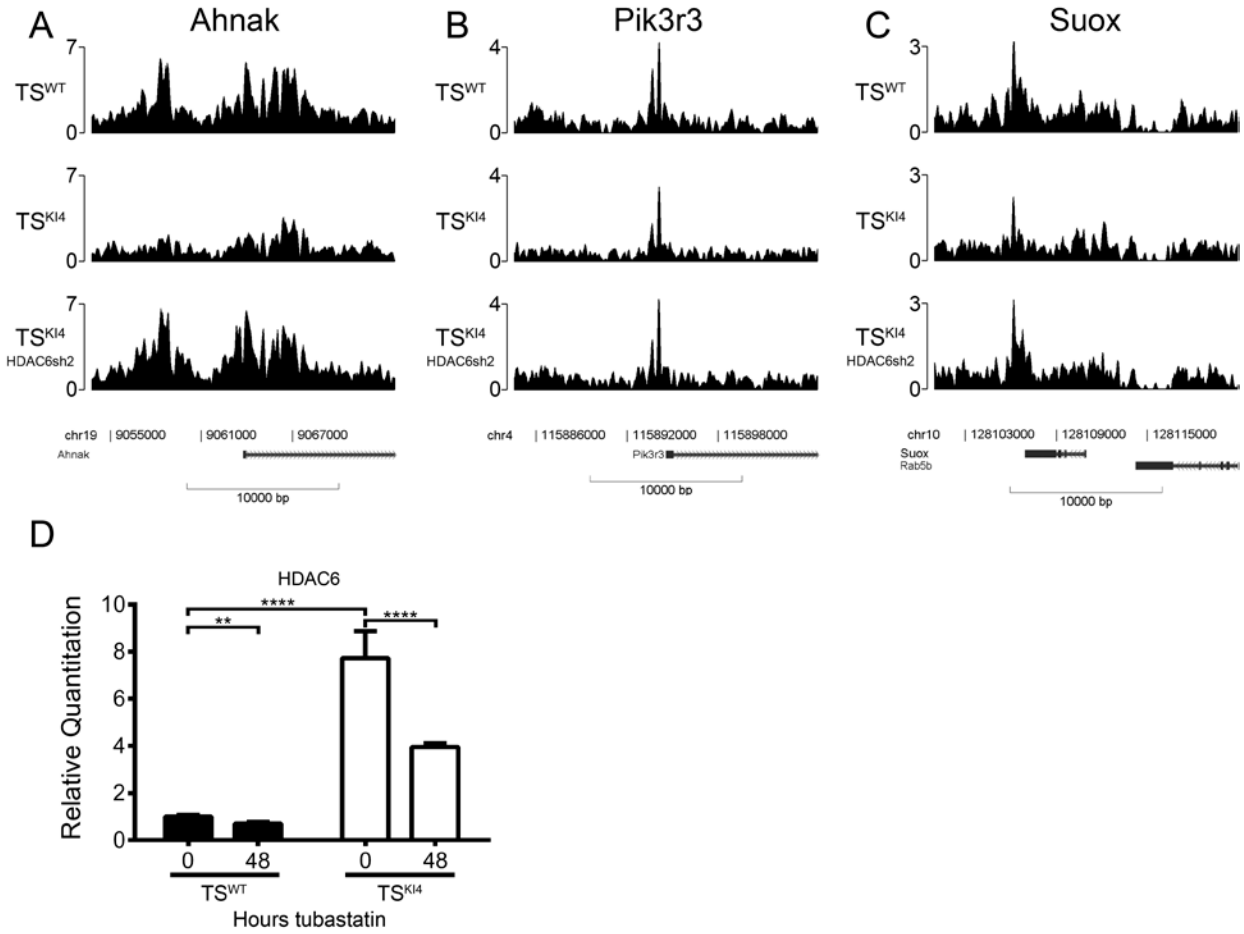


**Figure S6: MAP3K4 dependent HDAC6 expression regulates acetylation at transcription start sites of previously unidentified EMT-related genes. Related to Figure 7.**

(A-C) Read density plots for the indicated genes show MAP3K4/HDAC6 dependent changes in H2BK5Ac at previously undefined EMT-related genes as measured by H2BK5Ac ChIP-seq. Read density plots were derived as described in Figure S5.

(D) HDAC6 activity controls HDAC6 expression. Expression of HDAC6 transcript was measured by qRT-PCR. The indicated cell lines were treated for 48 hours with vehicle DMSO (0) or 10 nM Tubastatin A (48). Data show the mean ± SEM of three independent experiments.

Student's t-test, **p-value < 0.01; ****p-value < 0.0001.

## Supplemental Table Legends

**Table S1. 514 MAP3K4/HDAC6/H2BK5Ac dependent genes. Related to Figure 6.** Normalized transcripts from RNA-seq and normalized H2BK5Ac enrichment from H2BK5Ac ChIP-seq data for genes with MAP3K4/HDAC6 dependent expression and promoter H2BK5Ac.

## Supplemental Experimental Procedures

### Cell Lines, Culture Conditions, and Transfections

Cell lines with pLKO shRNA knockdown were cultured with puromycin. For differentiation, TS cells were cultured in TS media only. 293 cells were cultured in Dulbecco's modified essential medium supplemented with 10% fetal bovine serum, 1% penicillin, and 1% streptomycin. Transfections were performed using calcium chloride and BBS for 24 hours using plasmids specified in Supplemental Experimental Procedures (Chen and Okayama, 1987). Invasion assays through growth factor-reduced Matrigel were performed as previously described (Abell et al., 2011).

### Cell Lysates, Immunoprecipitation, Western Blotting, and HDAC activity

For immunoprecipitation, cells were harvested in Buffer A (20 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, and 1% Triton X) supplemented with protease inhibitors (1 mM PMSF and 17 μg/ml aprotinin) and phosphatase inhibitors (1 mM sodium vanadate and 1 mM sodium fluoride). Lysates were immunoprecipitated for 1 hour with the indicated antibodies, followed by incubation with Protein A-Sepharose (Invitrogen) for 1.5 hours. Immunoprecipitates were washed four times with cold Buffer A and boiled in Laemli buffer. Lysates and immunoprecipitates were probed with the indicated antibodies specified in Antibodies and Sources in Supplemental Experimental Procedures.

### Quantitative Real-Time PCR (qRT-PCR)

Total RNA was extracted from TS cells using the RNeasy plus minikit (Qiagen). The High-Capacity cDNA reverse transcription kit (Applied Biosystems) was used to synthesize cDNA from 3 μg of DNase treated total RNA. qRT-PCR was performed using

BioRad CFX96 Touch real-time PCR system and ABsolute Blue qPCR SYBR Green Mix (Thermo Fisher). Murine β-actin served as the endogenous control. Expression levels were calculated using the 2−ΔΔCT method and normalized to undifferentiated $TS^{WT}$ cells. Final results represent the pooling of at least two experiments in which cDNA was assayed in triplicate. ChIP-PCR was performed in duplicate on immunoprecipitated DNA and input control using BioRad SsoFast EvaGreen Supermix. Enrichment was calculated using the 2−ΔΔCT method. Primers used for all qRT-PCR are specified in Supplemental Experimental Procedures.

**HDAC Activity Assays**

HDAC assay kit (Active Motif) was used to detect HDAC activity according to the manufacturers specifications. For total HDAC activity, 5 μg of protein from cytoplasmic lysates or 2.5 μg nuclear extracts was used. Reactions were incubated at 37°C for one hour and fluorescence was measured using a Synergy H1MD plate reader (BioTek). For transfected HDAC6, Flag-HDAC6 was immunoprecipitated using an anti-Flag antibody (Rockland). For endogenous HDAC6 activity, HDAC6 was immunoprecipitated from TS cell cytoplasmic extracts with an anti-HDAC6 antibody (Cell Signaling).

**Whole Cell Immunofluorescence**

Briefly, TS were cultured on glass coverslips for two or three days. Cells were fixed with 3% paraformaldehyde in 1X PBS for 10 minutes and then washed three times with 1X PBS. Cells were permeabilized with 0.1% Triton/PBS for 3 minutes. Cells were blocked with 10% fetal bovine serum for 1 hour at RT. Coverslips were incubated with anti-ZO-1 (Thermo Fisher), Cldn6 (Santa Cruz) or E-cadherin (Biosciences) antibodies overnight at 4C. Next day the cells were washed for 30 minutes and then incubated with DAPI (50

µg/ml), anti-mouse Dy-Alexa 488 (1:500) (Thermo Fisher), anti-rabbit Alexa 594 (1:500) for one hour at RT. Coverslips were washed and mounted on slides with mounting media (90% glycerol and 10% 1 M Tris pH 7.5). Coverslips were imaged using a Nikon A1 laser scanning confocal microscope.

**Nuclear HDAC6 Immunofluorescence**

Paraformaldehyde fixed nuclei were permeabilized for 5 minutes with 0.3% Triton-X and were blocked for 1 hour at room temperature with 10% fetal bovine serum. Primary antibodies for HDAC6 and Lamin A/C (Cell Signaling) were incubated on coverslips overnight at 4C. The next day the coverslips were washed for 30 minutes and then incubated with DAPI (50 µg/ml), anti-mouse Alexa 594 (1:333) (Thermo Fisher), anti-rabbit Alexa Dylight 488 (1:500) (Thermo Fisher) for one hour at room temperature. Coverslips were washed and mounted on slides with mounting media (90% glycerol and 10% 1 M Tris pH 7.5). Coverslips were imaged using a Nikon A1 laser scanning confocal microscope.

**Confocal Microscopy**

Immunofluoroscence imaging was performed using Nikon A1 laser scanning confocal microscopy. ZO-1 and Cldn6 immunofluorescence images were obtained with 60X Plan Apochromat 1.4 numerical-aperture (NA) oil objective with lasers at 408, 488 and 594 nm. Individual z-stacks were obtained and the 2D images were displayed using maximum intensity projection (Nikon NIS Elements) for all the wavelengths. E-cadherin immunofluorescence images were obtained with 40X Plan Fluor 1.3 NA oil objective with lasers at 408 and 488 nm. Individual z-stacks were obtained and the 2D images were displayed using Extended Depth of Focus (EDF- Nikon NIS Elements) at all

wavelengths. HDAC6 immunofluorescence images were obtained using 60X Plan Apochromat 1.4 numerical-aperture (NA) oil objective with lasers at 408, 488 and 594 nm. Images were taken from a single plane on the z-axis.

**RNA Sequencing**

Libraries were constructed from 4 µg RNA using KAPA stranded mRNAseq kits and Illumina TruSeq indexed adapters following the manufacturer's protocol, except half the recommended template DNA was added to the 10 cycle PCR reaction. RNA seq Libraries were prepped by 1 x 75 bp Illumina NextSeq500, 12-plex, single-indexed sequencing. Reads were aligned to the murine mm9 reference genome using MapSplice v2.1.4 (Wang et al., 2010). Picard Tools v1.88 was used to determine the alignment profile (Broad Institute). SAMtools v1.2 was used to sort and index aligned reads, which were then translated to transcriptome coordinates and filtered for indels, large inserts, and zero mapping quality using UNC Bioinformatics utilities (UBU, https://github.co/mozack/ubu) v1.2 (Li et al., 2009). Transcript abundance for each sample was estimated using RSEM expectation-maximization algorithm (Li and Dewey, 2011).

**Chromatin Immunoprecipitation (ChIP) Coupled to High Throughput Sequencing**

ChIP was performed essentially as previously described (Abell et al., 2011; Loven et al., 2013). Cells were fixed in 1% paraformaldehyde for 10 minutes, harvested, and the nuclei were isolated and lysed. The nuclear lysate was sonicated and underwent immunoprecipitation using 50 µl of dynabeads (Invitrogen) coupled to 5 µl of Active Motif α-H2BK5Ac antibody per sample. DNA-protein crosslinks were reversed overnight at 65C, and samples were ribonuclease and proteinase treated before DNA was purified

using Min Elute columns (Qiagen). Preparation of libraries was performed using 50 ng of ChIP or $TS^{WT}$ input DNA and the KAPA HyperPrep kit with Illumina TruSeq indexed adapters. Dual SPRI size selection was performed after 18 cycles of amplification according to the manufacturer's protocol. The average library size was ~300 bp. 12-plex libraries were sequenced (1 X 75 bp) using an Illumina NextSeq500 ranging from 7.1-8.9 X $10^7$ reads per sample. Using Bowtie v1.1.2 software reads were mapped to the mouse mm9 genome using parameters -v2 -m1.

**Analysis of ChIP-seq Data**

Reads aligned to the mm9 genome were analyzed using Easeq (Lerdrup et al., 2016). Aligned ChIP-seq data files were entered as datasets and annotated using the mouse mm9 genome geneset. Peaks were called with a p-value ≤ of $1.0x10^{-5}$ in each immunoprecipitated sample relative $TS^{WT}$ input using the Peaks utility. Quantified comparisons of peak regionsets for each sample were performed in Easeq using the quantify tool set to measure the TSS ± 10 kbp of each gene. Normalized H2BK5Ac signal at each TSS was compared, and genes showing >10% decrease in $TS^{KI4}$ relative to $TS^{WT}$, and >10% increase in $TS^{KI4}$ HDAC6sh2 relative to $TS^{KI4}$ were designated as having HDAC6/MAP3K4 dependent H2BK5Ac.

**Visualization of Heatmaps and Read-density Plots**

In Easeq, ChIP-seq read-density heatmaps of the 1817 genes displaying MAP3K4/HDAC6 dependent H2BK5Ac (sorted by descending $TS^{WT}$ H2BK5Ac) were generated to visualize the TSSs ± 10 kbp. The 20 kbp regions were divided into 1.0 x $10^5$ bins. For the dataset used as the numerator ($TS^{WT}$, $TS^{KI4}$, and $TS^{KI4\ HDAC6sh2}$) and the dataset used as the denominator ($TS^{WT}$ input control), the intensities were

normalized to the number of reads multiplied by one million, the fragment size, and the binsize (2 bp), so that the density represents DNA fragments per million reads per 1 kbp.

Individual gene ChIP-seq read-density plots were generated in Easeq using the FillTrack utility. After using the Peaks utility as described above, we used the FillTrack utility to display the read density of the peak regionsets that were called relative to the $TS^{WT}$ input negative control. The units of the y-axis were normalized and binned read counts calculated as specified in Lerdrup et al.

The gene expression heatmap in Figure 6D was generated using the Easeq utility Parmap. Of the 1817 genes displaying MAP3K4/HDAC6 dependent H2BK5Ac, 514 showed >10% decrease in expression by RNA-seq in $TS^{KI4}$ cells relative to $TS^{WT}$ cells and >10% increase in expression in $TS^{KI4}$ HDAC6sh2 cells relative to $TS^{KI4}$ cells. The RNA-sequencing data parameters (relative reads for each gene in each sample) were added to the ChIP-seq regionset of the 514 MAP3K4/HDAC6/H2BK5Ac dependent genes. From this regionset, the relative expression parameters were used to generate the heatmaps.

**Analyses of RNAseq and ChIP-seq Data Using EaSeq and TopFunn**

Gene list functional enrichment analysis was performed using Toppfun (Chen et al., 2009). The 514 MAP3K4/HDAC6/H2BK5Ac dependent genes were input as the training set, of which 493 had annotation. Gene Ontology (GO) terms of Biological Process and Cellular component were searched for enrichment with FDR < 0.05. Annotations shown were selected that had enrichment Bonferroni p-value < 0.05.

**Plasmids and Sources**

| Plasmid | Source/Reference |
|---|---|
| HA-MAP3K4 WT | (Gerwins et al., 1997) |
| HA-MAP3K4 KI | (Gerwins et al., 1997) |
| FLAG-MAP3K4 WT | (Abell and Johnson, 2005) |
| FLAG-MAP3K4 KI | (Abell and Johnson, 2005) |
| FLAG-HDAC6 | (Addgene, #30482) |
| FLAG-HDAC7 | (Addgene, #13824) |
| HA-MAP3K4 kinase domain KD | (Gerwins et al., 1997) |
| EGFP-HDAC6 | (Addgene, #36188) |

**Antibodies and Sources:**

| Antibody specificity | Company | Catalogue number |
|---|---|---|
| Ac-Lysine | Cell Signaling | 9441 |
| Ac-α-Tubulin | Sigma | T6793 |
| Actin | Sigma | A4700 |
| CBP | Cell Signaling | 7425 |
| Claudin-6 | Santa Cruz | sc-393671 |
| E-cadherin | BD | 610181 |
| ERK2 | Santa Cruz | sc-154 |
| FLAG mouse | Thermo Scientific | MA1-91878 |
| FLAG rabbit | Rockland | 600-401-383 |
| GFP | Santa Cruz | sc-8334 |

| | | |
|---|---|---|
| H2AK5Ac | Cell Signaling | 2576 |
| H2AK5Ac | Active Motif | 39123 |
| H2A total | Cell Signaling | 2578 |
| H2BK5Ac | Cell Signaling | 2574 |
| H3K9Ac | Cell Signaling | 9649 |
| H4K8Ac | Cell Signaling | 2594 |
| HA | Santa Cruz | sc-805 |
| Hdac6 | Abcam | ab12173 |
| Hdac6 | Cell Signaling | 7612 |
| Lamin A/C | Cell Signaling | 4777 |
| Lamin B1 | Cell Signaling | 13435 |
| c-Myc mouse | Santa Cruz | sc-40 |
| SIRT2 | Santa Cruz | sc-20966 |
| α Tubulin | Santa Cruz | sc-53646 |
| α Tubulin | Sigma Aldrich | T6793 |
| ZO-1 | Thermo Fisher | 40 2200 |

**Primers for for ChIP coupled to Real-time quantitative PCR:**

| Gene Name | Forward Primer | Reverse Primer |
|---|---|---|
| ahnak | CCT GGC AAG GCG TTT GC | GCA CCA TTC CAC TTA CTA ACC TTT G |
| cldn6 | GGG AAC TTT GGG AAA CGC AC | AAT GAC GTC ACT GGA CCA CC |
| ocln | CAA ACC TGG CAG CAG AAC AC | TTC CGG GAG GAG AGT CAG AG |
| pik3r3 | TTT CCC CCT TCC TCT CCA TT | TCA GCG GAA GCC ACT TTT G |
| suox | GAG AGG CCC GCT GTG TTG | TTT GCA AAA CCA GGC CTT GT |

**Primers for Real-time quantitative PCR for mRNA:**

| Gene Name | Forward Primer | Reverse Primer |
|---|---|---|
| actb | AGC CAT GTA CGT AGC CAT CC | CTC TCA GCT GTG GTG GTG AA |
| ahnak | CCA CCC CAA CTG GGA CTT TG | CAC TCC CCT GTA ACT TGC CTG |
| cdh1 | CAG GTC TCC TCA TGG CTT TGC | CTT CCG AAA AGA AGG CTG TCC |
| cdh2 | AGC GCA GTC TTA CCG AAG G | TCG CTG CTT TCA TAC TGA ACT TT |
| cldn6 | ATG GCC TCT ACT GGT CTG CAA | GCC AAC AGT GAG TCA TAC ACC TT |
| ets2 | CCT GTC GCC AAC AGT TTT CG | TGG AGT GTC TGA TCT TCA CTG A |
| hdac6 | TCC ACC GGC CAA GAT TCT TC | CAG CAC ACT TCT TTC CAC CAC |
| ocln | TTG AAA GTC CAC CTC CTT ACA GA | CCG GAT AAA AAG AGT ACG CTG G |
| pik3r3 | TAC AAT ACG GTG TGG AGT ATG GA | GAG TCA TTG GCT TAG GTG GCT |
| sirt2 | GCC TGG GTT CCC AAA AGG AG | GAG CGG AAG TCA GGG ATA CC |
| snai1 | CAC ACG CTG CCT TGT GTC T | GGT CAG CAA AAG CAC GGT T |
| snai2 | TGG TCA AGA AAC ATT TCA ACG CC | GGT GAG GAT CTC TGG TTT TGG TA |
| twist1 | GGA CAA GCT GAG CAA GAT TCA | CGG AGA AGG CGT AGC TGA G |

| | | |
|---|---|---|
| vim | TCC ACA CGC ACC TAC AGT CT | CCG AGG ACC GGG TCA CAT A |

**Supplemental References**


Abell, A.N., and Johnson, G.L. (2005). MEKK4 is an effector of the embryonic TRAF4 for JNK activation. The Journal of biological chemistry *280*, 35793-35796.

Abell, A.N., Jordan, N.V., Huang, W., Prat, A., Midland, A.A., Johnson, N.L., Granger, D.A., Mieczkowski, P.A., Perou, C.M., Gomez, S.M.*, et al.* (2011). MAP3K4/CBP-regulated H2B acetylation controls epithelial-mesenchymal transition in trophoblast stem cells. Cell stem cell *8*, 525-537.

Chen, C., and Okayama, H. (1987). High-efficiency transformation of mammalian cells by plasmid DNA. Molecular and cellular biology *7*, 2745-2752.

Chen, J., Bardes, E.E., Aronow, B.J., and Jegga, A.G. (2009). ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. Nucleic acids research *37*, W305-311.

Gerwins, P., Blank, J.L., and Johnson, G.L. (1997). Cloning of a novel mitogen-activated protein kinase kinase kinase, MEKK4, that selectively regulates the c-Jun amino terminal kinase pathway. The Journal of biological chemistry *272*, 8288-8295.

Lerdrup, M., Johansen, J.V., Agrawal-Singh, S., and Hansen, K. (2016). An interactive environment for agile analysis and visualization of ChIP-sequencing data. Nat Struct Mol Biol *23*, 349-357.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics *12*, 323.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Loven, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I., and Young, R.A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell *153*, 320-334.

Wang, K., Singh, D., Zeng, Z., Coleman, S.J., Huang, Y., Savich, G.L., He, X., Mieczkowski, P., Grimm, S.A., Perou, C.M.*, et al.* (2010). MapSplice: accurate mapping of RNA-seq reads for splice junction discovery. Nucleic acids research *38*, e178.