Figure S1

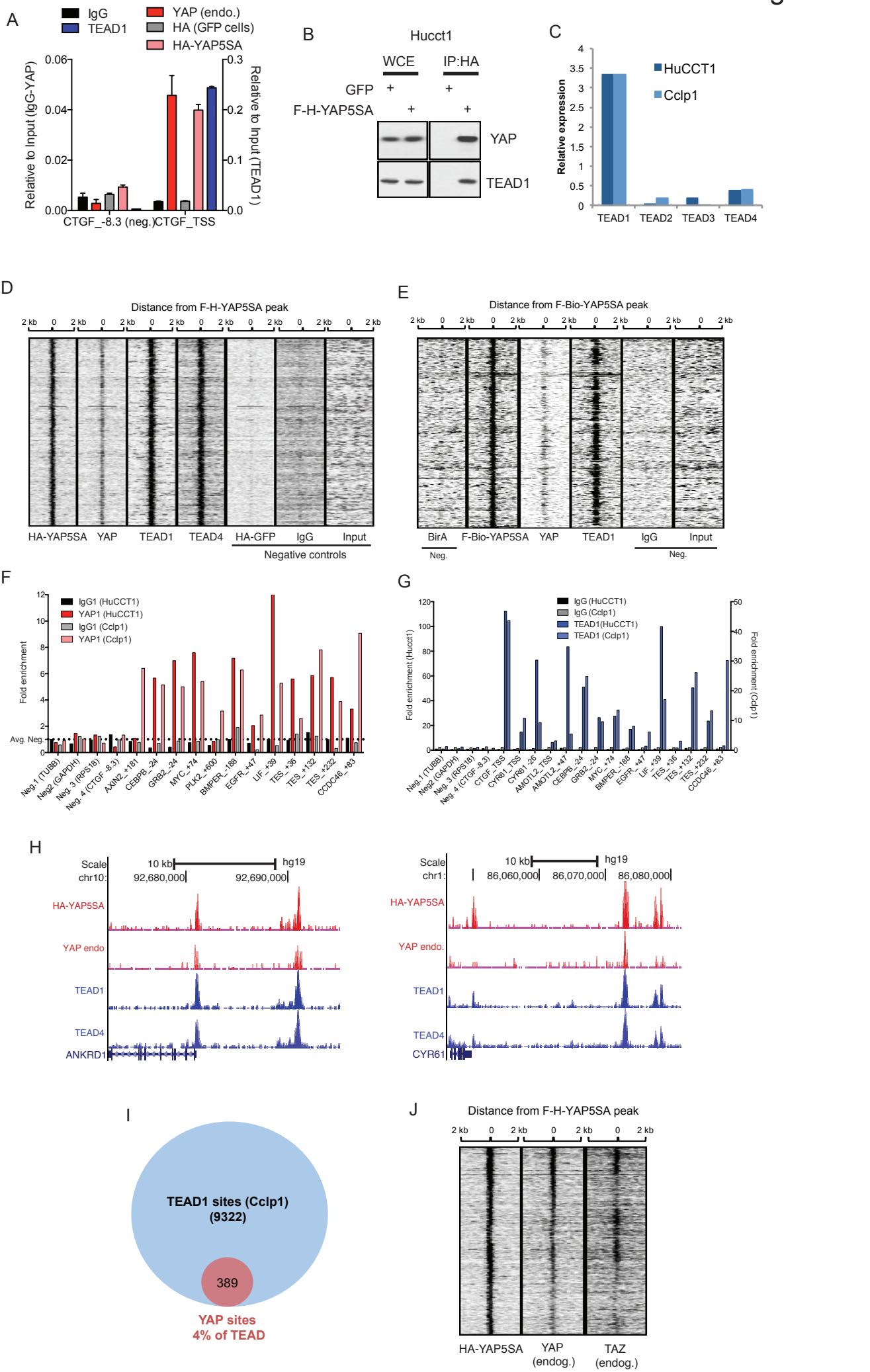

A
IgG
TEAD1
YAP (endo.)
HA (GFP cells)
HA-YAP5SA
Relative to Input (IgG-YAP)
Relative to Input (TEAD1)
CTGF_-8.3 (neg.)
CTGF_TSS
B
Hucct1
WCE
IP:HA
GFP
F-H-YAP5SA
YAP
TEAD1
C
HuCCT1
Cclp1
Relative expression
TEAD1
TEAD2
TEAD3
TEAD4
D
Distance from F-H-YAP5SA peak
HA-YAP5SA
YAP
TEAD1
TEAD4
HA-GFP
IgG
Input
Negative controls
E
Distance from F-Bio-YAP5SA peak
BirA
F-Bio-YAP5SA
YAP
TEAD1
IgG
Input
Neg.
F
IgG1 (HuCCT1)
YAP1 (HuCCT1)
IgG1 (Cclp1)
YAP1 (Cclp1)
Fold enrichment
Avg. Neg.
G
IgG (HuCCT1)
IgG (Cclp1)
TEAD1(HuCCT1)
TEAD1 (Cclp1)
Fold enrichment (Huctt1)
Fold enrichment (Cclp1)
H
Scale
chr10:
10 kb
hg19
92,680,000
92,690,000
HA-YAP5SA
YAP endo
TEAD1
TEAD4
ANKRD1
chr1:
86,060,000
86,070,000
86,080,000
YAP endo.
CYR61
I
TEAD1 sites (Cclp1)
(9322)
389
YAP sites
4% of TEAD
J
Distance from F-H-YAP5SA peak
HA-YAP5SA
YAP
(endog.)
TAZ
(endog.)

Figure S2

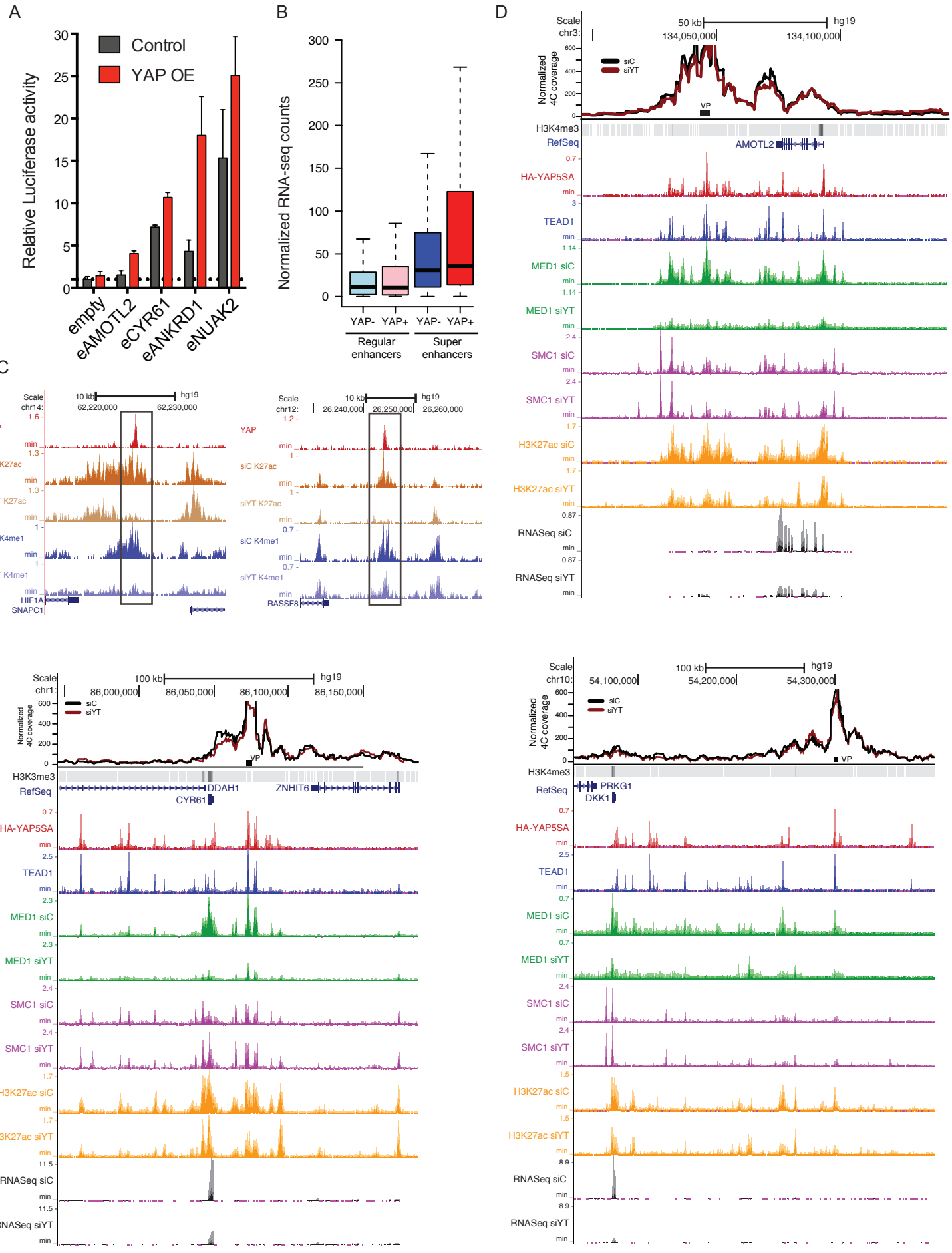
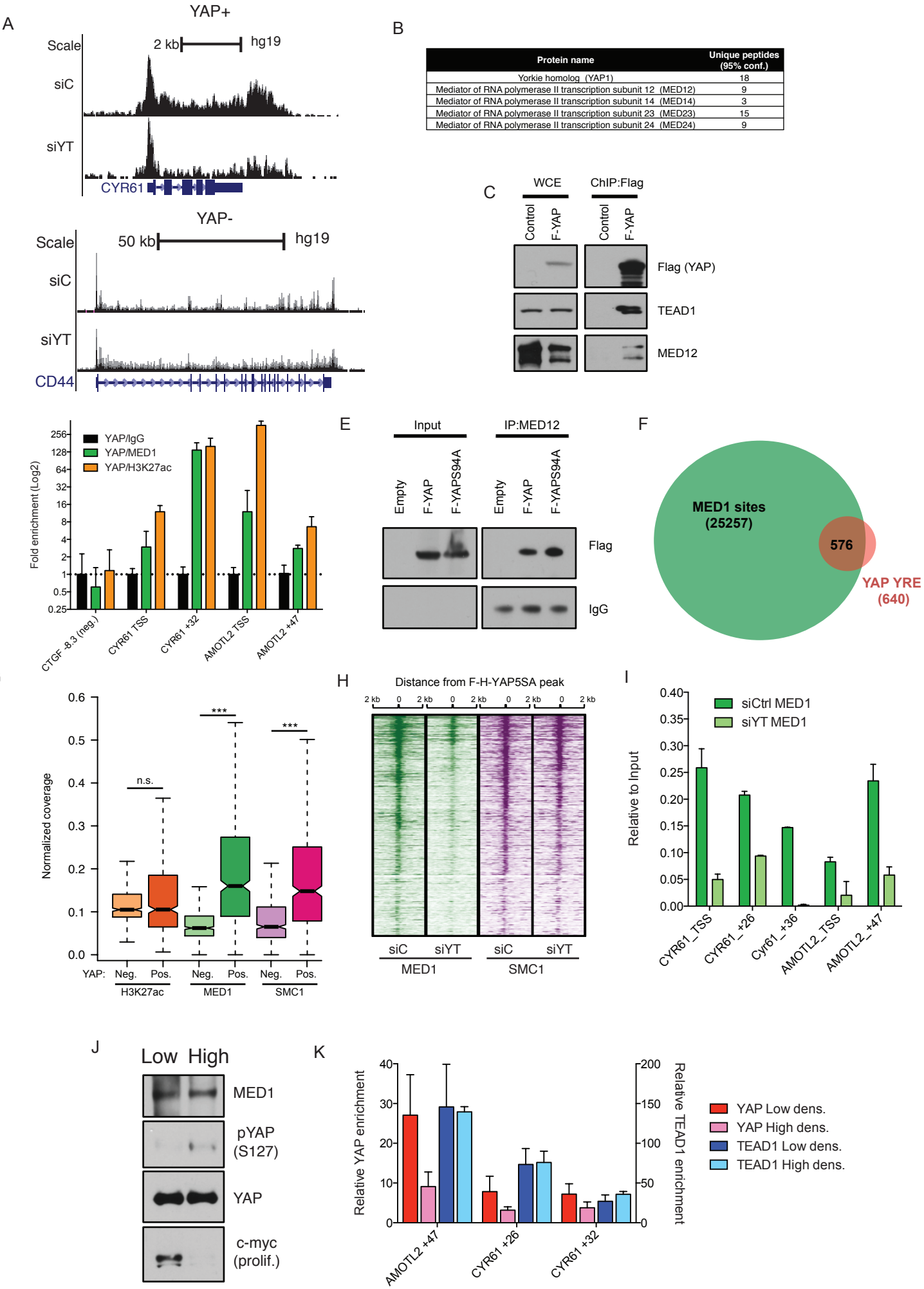
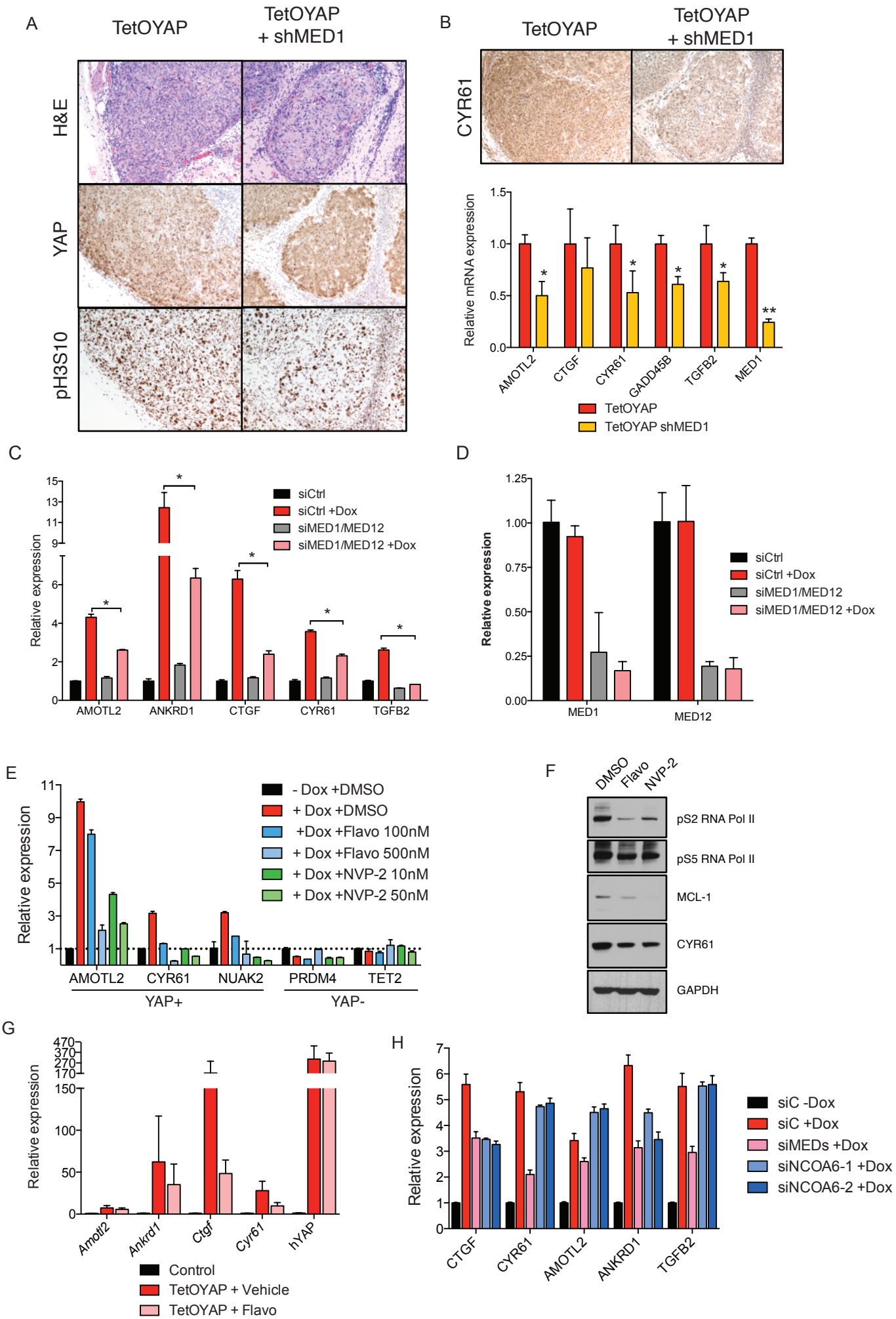

A
Relative Luciferase activity
Control
YAP OE
empty
eAMOTL2
eCYR61
eANKRD1
eNUAK2
B
Normalized RNA-seq counts
YAP-
YAP+
YAP-
YAP+
Regular enhancers
Super enhancers
C
YAP
siC K27ac
siYT K27ac
siC K4me1
siYT K4me1
HIF1A
SNAPC1
RASSF8
D
Normalized 4C coverage
siC
siYT
VP
H3K4me3
RefSeq
AMOTL2
HA-YAP5SA
TEAD1
MED1 siC
MED1 siYT
SMC1 siC
SMC1 siYT
H3K27ac siC
H3K27ac siYT
RNASeq siC
RNASeq siYT
H3K3me3
DDAH1
ZNHIT6
CYR61
PRKG1
DKK1

Figure S3


A
YAP+
Scale
2 kb
hg19
siC
siYT
CYR61
YAP-
Scale
50 kb
hg19
siC
siYT
CD44
B
Protein name
Unique peptides (95% conf.)
Yorkie homolog (YAP1) 18
Mediator of RNA polymerase II transcription subunit 12 (MED12) 9
Mediator of RNA polymerase II transcription subunit 14 (MED14) 3
Mediator of RNA polymerase II transcription subunit 23 (MED23) 15
Mediator of RNA polymerase II transcription subunit 24 (MED24) 9
C
WCE
ChIP:Flag
Control
F-YAP
Flag (YAP)
TEAD1
MED12
D
YAP/IgG
YAP/MED1
YAP/H3K27ac
Fold enrichment (Log2)
CTGF -8.3 (neg.)
CYR61 TSS
CYR61 +32
AMOTL2 TSS
AMOTL2 +47
E
Input
IP:MED12
Empty
F-YAP
F-YAPS94A
Flag
IgG
F
MED1 sites (25257)
576
YAP YRE (640)
G
Normalized coverage
n.s.
***
***
YAP:
Neg.
Pos.
H3K27ac
MED1
SMC1
H
Distance from F-H-YAP5SA peak
2 kb
siC
siYT
MED1
SMC1
I
siCtrl MED1
siYT MED1
Relative to Input
CYR61_TSS
CYR61_+26
Cyr61_+36
AMOTL2_TSS
AMOTL2_+47
J
Low
High
MED1
pYAP (S127)
YAP
c-myc (prolif.)
K
Relative YAP enrichment
Relative TEAD1 enrichment
AMOTL2 +47
CYR61 +26
CYR61 +32
YAP Low dens.
YAP High dens.
TEAD1 Low dens.
TEAD1 High dens.


A
TetOYAP
TetOYAP + shMED1
H&E
YAP
pH3S10
B
TetOYAP
TetOYAP + shMED1
CYR61
Relative mRNA expression
AMOTL2
CTGF
CYR61
GADD45B
TGFB2
MED1
TetOYAP
TetOYAP shMED1
C
Relative expression
siCtrl
siCtrl +Dox
siMED1/MED12
siMED1/MED12 +Dox
AMOTL2
ANKRD1
CTGF
CYR61
TGFB2
D
Relative expression
siCtrl
siCtrl +Dox
siMED1/MED12
siMED1/MED12 +Dox
MED1
MED12
E
Relative expression
- Dox +DMSO
+ Dox +DMSO
+Dox +Flavo 100nM
+ Dox +Flavo 500nM
+ Dox +NVP-2 10nM
+ Dox +NVP-2 50nM
AMOTL2
CYR61
NUAK2
PRDM4
TET2
YAP+
YAP-
F
DMSO
Flavo
NVP-2
pS2 RNA Pol II
pS5 RNA Pol II
MCL-1
CYR61
GAPDH
G
Relative expression
Amotl2
Ankrd1
Ctgf
Cyr61
hYAP
Control
TetOYAP + Vehicle
TetOYAP + Flavo
H
Relative expression
CTGF
CYR61
AMOTL2
ANKRD1
TGFB2
siC -Dox
siC +Dox
siMEDs +Dox
siNCOA6-1 +Dox
siNCOA6-2 +Dox

SUPPLEMENTAL FIGURE LEGENDS

Figure S1 – Related to Figure 1. YAP binds to distal regulatory elements

A) Representative ChIP-qPCR for controls and YAP (left Y axis) and TEAD1 (right Y axis) on a negative (CTGF -8.3) and positive (CTGF TSS) control regions. B) Immunoprecipitation with HA antibody and whole cell extract (WCE) in HuCCT1-F-H-YAP5SA cells and immunoblotting with YAP and TEAD1 antibody. C) qPCR for TEAD genes in HuCCT1 and Cclp1 cells. D) Heatmap of ChIP-seq signal for the indicated samples +/- 2 Kb from the centre of YAP peaks in HuCCT1 cells. E) Heatmap of ChIP-seq signal for the indicated samples +/- 2 Kb from the center of YAP peaks in Cclp1 cells. F-G) ChIP-qPCR validation of YAP (F) and TEAD1 (G) binding to a set of randomly selected distal elements. Four negative control regions are indicated and data are normalized on the average of the signal in the negative controls. H) Examples of ChIP-seq signal for the indicated antibodies in HuCCT1 cells around the identified peaks. I) Venn diagram of the intersection between TEAD1 peaks and YAP peaks in Cclp1. J) Heatmap of YAP (tagged and endogenous) and TAZ (endogenous) signal +/- 2Kb centered on YAP binding sites in HuCCT1 cells.

Figure S2 – Related to figure 2. YAP controls high transcription rate fromm enhancer elements

A) Luciferase assay for the indicated enhancers in control cells (grey bars) or cells overexpressing YAP (red bars). B) Normalized RNA-seq counts for genes associated with the indicated type of genomic elements. C) Representative genome browser snapshot for regions around YAP target genomic loci (two panels). Colored tracks represent ChIP-seq signal for the indicated factors in control and YAP/TAZ knockdown cells. Grey box highlights regions bearing H3K27ac and/or H3K4me1 affected by knockdown of YAP and TAZ. D) Representative genome browser snapshot for regions around three YAP target genomic loci (three panels). Top track

lines represent normalized 4C coverage from the indicated viewpoint (VP = black block, siC = black line, siYAP/TAZ = red line). Density track represents H3K4me3 signal as a reference for active promoters. Colored tracks represent normalized ChIP-seq / RNA-seq signal for the indicated factors in control and YAP/TAZ knockdown cells.

Figure S3. Related to Figure 3. YAP recruits Mediator to regulate transcription

A) ChIP-seq tracks for RNA Pol II around genomic loci encoding a YRE-associated gene (upper panel) and a YAP- gene (lower) in HuCCT1 cells treated with siControl or siYAP/TAZ. B) Number of unique peptides for each Mediator subunit identified specifically in YAP IP-MS experiments. C) ChIP-WB experiment demonstrating interaction between Flag-YAP and MED12 onto chromatin. D) Sequential ChIP/re-ChIP between YAP and IgG, MED1 or H3K27ac in HuCCT1 cells. CTGF – 8.3 is used as a negative control region. E) co-IP between Flag-YAP bearing S94A mutation and MED12. F) Venn diagram representing the fraction of MED1 peaks positive for YREs in HuCCT1 cells. G) Boxplot representing H3K27ac, MED1 and SMC1 coverage between YAP+ YRE and a size-matched set of YAP- enhancers which display similar mean levels of H3K27ac. H) Heatmap of ChIP-seq signal for indicated antibodies in control and siYT cells. I) ChIP-qPCR validation for MED1 occupancy in cells treated with control or siYT. J) Western blot analysis for YAP phosphorylation and MED1 in low and high density H69 cells. C-MYC is used as a proliferation marker. K) ChIP-qPCR for YAP and TEAD in H69 cells grown at low and high density.

Figure S4 – Related to figure 4. Mediator is downstream of YAP and allows CDK9 dependent elongation

A) Histological and immunohistochemical examination of harvested xenografts at 7 weeks. B) Immunohistochemistry (top panel) and qPCR (bottom panel) for YAP target genes in harvested tumors from xenografts at 7 weeks. C) qPCR for YAP target genes in HuCCT1 cells overexpressing Dox inducible YAPS127A (TetOYAP) treated with control or siMED1+siMED12

oligos. D) qPCR for MED1 and MED12 validating knockdown efficiency of the experiment in C. E) qPCR for four YAP+ and two YAP- target genes in HuCCT1-TetOYAP cells treated with Doxycycline for 24 hours and/or treated with CDK9 inhibitors at the indicated concentration for 3 hours. F) Western blot analysis for RNA Pol II CTD phosphorylation in cells treated with CDK9 inhibitors for 3 hours. Flavopiridol was used at 500nM while NVP-2 at 50nM. pS2 and Mcl-1 are used as CDK9 activity markers. G) qPCR for YAP target genes and human YAP in TetOYAP mice treated with vehicle or Flavopiridol for 6 days. H) Knockdown of Ncoa6 affects transcription of a subset of YAP target genes.

SUPPLEMENTARY TABLES

Table S1

Antibodies used in different applications

Table S2

Primers used in different applications

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Cell culture experiments

HuCCT1 (parental and Flag-HA-YAP5SA expressing) cells were maintained in RPMI containing 10% FBS, 1X Hepes, 1X L-glutamine. Cclp1 (parental and Flag-Bio-YAP5SA expressing) were maintained in DMEM containing 10% FBS. H69 were maintained in DMEM : DMEM/F12 containing 10%FBS and supplemented with Adenine, Insulin, Epinephrine, T3-T, EGF and Hydrocortisone as previously reported (Li et al., 2014). HuCCT1-GFP or –FH-YAP5SA cells were obtained by retroviral transduction of pBABE viruses expressing GFP or FH-YAP5SA. HuCCT1 TetOYAP cells were obtained by lentiviral transduction of a plasmid expressing M2-rtTA construct and a pNL-TRE-YAPS127A plasmid. Cclp1-BirA or –Fbio-YAP5SA cells were obtained by lentiviral transduction of previously described BirA system plasmids, known to express at sub-endogenous levels (Kim et al., 2009). MED1 shRNAs were obtained by OpenBiosystems (RHS4696-200768597). siRNAs were obtained by Ambion (YAP s20367, TAZ s24787, MED1 s10889, MED12 s19362, NF2 s224112, LATS2 s25503). All plasmid transfections were performed with Lipofectamine 2000 while siRNA transfections were performed with Lipofectamine RNAiMAX according to manufacturer recommendations.

Flavopiridol was purchased from Sigma while CDK9-specific inhibitor NVP-2 was a kind gift of Nathaniel Gray (Dana Farber Cancer Institute). For in-vitro experiments, compounds were resuspended in DMSO and cells were treated at indicated concentrations for 3 hours.

Mouse models

Tetracycline-inducible YAPS127A expression mice were previously described (Camargo et al., 2007). AdGFP, AAV8-TBG-LacZ, AAV8-TBG-Cre were obtained by University of Pennsylvania Vector Core. $1x10^{11}$ pfu of AAV viruses are injected retro-orbitally. For TetOYAP overexpression, 3 days after AAV-Cre delivery, mice were administered doxycycline (1 mg/ml)

ad libitum in their water. For Flavopiridol experiments, mice were injected IP with 3mg/kg of Flavopiridol resuspended in 10% DMSO in 10% hydroxypropyl beta cyclodextrin (Sigma) once a day for 6 days.

For xenografts assays, 2×10^6 HuCCT1 cancer cells resuspended in 100 µL of Matrigel (BD) and 100 µL PBS were injected s.c. in the flank of Nude mice (Jackson labs). After 2 weeks, tumor volume was measured and mice were divided in groups for randomization with equal tumor size. Mice were then given doxycycline (1 mg/ml) ad libitum in their water. Tumor size was evaluated once weekly by caliper measurements using the following formula: tumor volume = [length × width^2] / 2, as reported previously (Jimeno et al., 2007). A minimum of three mice was examined per experiment.

Gene expression analyses.

For immunoblotting cells were harvested and lysed in RIPA buffer supplemented with protease inhibitor cocktail (Roche). Protein samples were resolved on SDS-page, transferred onto nitrocellulose membranes and probed with the indicated antibodies (Table S1).

For quantitative PCR, total RNA was extracted from cell pellets using the Trizol (Life Technologies) according to the manufacturer's instructions. cDNA was synthetized using iScript Reagents (Biorad). qRT-PCR was performed with the One Step plus Sequence Detection System (Applied Biosystems) using Fast SYBR green master mix reagent (Applied Biosystems). Gene expression levels were normalized to the average of at least two housekeeping genes or to input levels for ChIP-enrichment. qPCR primers sequences are listed in Supplementary table S2.

For microarray analysis, HuCCT1 cells, treated for 3 hours with 500nM Flavopiridol or 50nM NVP-2, were harvested in Trizol and their RNA isolated according to standard procedure. Retrotranscription, cRNA labeling and hybridization to Human Gene 2.0ST arrays (Affymetrix)

were performed according to standard protocols at the IDDRC Molecular Genetics core facility (Boston Children's Hospital, Boston, USA).

For immunohistochemistry, tissue was fixed overnight in 10% formalin (Sigma-Aldrich, St. Louis, MO) and embedded in paraffin for sectioning. 5 uM tissue sections were rehydrated followed by antigen retrieval using low pH Antigen Unmasking Solution (Vector Labs, Burlingame, CA). Quenching of endogenous peroxidase and protein block were performed prior to overnight antibody incubation. For immunohistochemistry, Vectastain Elite ABC kit and secondary antibody (Vector Labs) were used to detect primary antibody binding. Slides were developed using the Vectastain Elite kit used as directed by the manufacturer and mounted, after dehydration with Vectamount (Vector labs). Antibodies used in this study are noted in Supplementary Table S1.

4C-sequencing

4C templates were prepared as described previously (Splinter et al., 2012). DpnII digestion was used as the first restriction enzyme to generate high resolution 3C template, which was further trimmed with Csp6I, NlaIII or BfaI. 4C primers were design following the general consideration as described (Splinter et al., 2012, van de Werken et al., 2012). The primers carried additional 5′ overhangs composed of adaptor sequences for Illumina single-read sequencing. Samples were sequenced on a Hi-seq 2000 machine (Illumina).

ChIP-seq and RNA-seq

Cells were cross-linked in 1% formaldehyde for 10 min at room temperature after which the reaction was stopped by addition of 0.125M glycine. Cells were lysed and harvested in ChIP buffer (100 mM Tris at pH 8.6, 0.3% SDS, 1.7% Triton X-100, and 5 mM EDTA) and the chromatin disrupted by sonication using a Diagenode Bioruptor sonicator UCD-200 to obtain fragments of average 200-500 bp in size. Suitable amounts of chromatin were incubated with

specific antibodies overnight. Immunoprecipitated complexes were recovered on Protein-A/G agarose beads (Pierce) and, after extensive washes, DNA was recovered by reverse crosslinking and purification using QIAquick PCR purification kit (QIAGEN). Med1 and Smc1 ChIP were performed as previously described (Whyte et al., 2013).

Libraries for ChIP-sequencing were generated by using NEBNext Ultra DNA Library Prep Kit for Illumina (NEB) and barcoded added using NEBNext Multiplex Oligos for Illumina (Index Primers Set 1) (NEB) according to manufacturer's recommendation.

RNA-seq libraries were generated using TruSeq RNA Sample Prep Kit v2 (Illumina) according to manufacturer's recommendation. All the High-throughput sequencing experiments were run on a Hi-seq2000 (Illumina) sequencer at the Center for Cancer Computational Biology (Dana Farber Cancer Institute, Boston, USA).

Bioinformatic analyses.

ChIP-seq. Reads were aligned to human Hg19 reference using Bowtie2 software (version 2.0.6) (Langmead and Salzberg, 2012) set to report only the best alignment with default effort parameters. Duplicate reads were removed using Picard Tools suite (version 1.82) (http://picard.sourceforge.net).

In order to isolate regions with significant signal enrichment (defined as coverage or number of reads correctly aligned on said region), a peak-calling analysis was run using MACS software (version 1.3.7.1) (Zhang et al., 2008) setting IgG signal (or control-engineered cell lines) for the background noise definition and a lower fold change cutoff for model building to 5. An additional significance cutoff was set to p-value less or equal 10^{-8}. In order to obtain "YAP binding sites", the obtained peaks from tagged YAP5SA ChIP were additionally filtered by the presence of increase reads number in "endogenous YAP" ChIP versus the signal in the three described negative control samples.

Available peaks were annotated using multiple database sources (UCSC KnownGene, ENSEMBL, Entrez) and software (GREAT (McLean et al., 2010)).

Available peak sets were interrogated for overlapping when interested in co-localization or positivity of a peak for a specific signal. Overlaps were performed on the basis of the absolute position of the genome, so that a co-localization of at least 1 bp resulted in a positive hit. Alternatively given a peak dataset, overlap was based on the presence of signal above the background centered on each peak based on reads density calculated by Seqminer (Ye et al., 2011).

Available peaks were interrogated for the amount of signal of a number of factors. The amount of signal was calculated as the maximum number of reads covering a single position within the region, normalized by millions of reads available in the library. Strandness was handled using the mean value of the two strands when applicable.

Cumulative signal trends for peak sets were calculated to understand the preferential localization patterns. For each peak available in the set, a range of 10Kb around the peak center (5Kb upstream and downstream) was taken into consideration and the sum of the coverage for each position was calculated and plotted as a function of the position itself. All values have been normalized by library size. Heatmaps were generated using Seqminer (Ye et al., 2011) and peaks were clustered by K-means algorithm, where applicable.

TEAD1 human consensus matrix was downloaded from JASPAR website (Sandelin et al., 2004) and was used to confirm if peaks inside peak sets of interest were on regions enriched of or containing sequences compatible with a TEAD1 binding site. The matrix was used to scan the region using Bioconductor package Biostrings, setting a similarity cutoff between consensus matrix and sequence to 80%. Results were reported as absolute number of hits within the region, without normalization by peak width.

Where not otherwise specified, all analyses were performed using custom scripts for the R Statistical Program, aided by packages from Bioconductor repositories (Gentleman et al., 2004).

Enhancers were defined as regions showing positivity and overlap for both H3K27Ac and H3K4Me1 signals (based on peak calling vs. IgG control). Superenhancers were defined as clusters of enhancers not farther than 12.5Kb one from another and searched for using ROSE software, as previously described (Hnisz et al., 2013). $H3K27Ac^{+}$/$H3K4Me1^{+}$ peaks (putative enhancers) from MACS were used as input constituents with a ranking BAM file containing H3K27ac reads (HuCCT1 cells, siRNA control) and a control BAM file containing IgG reads (HuCCT1 cells).

Pausing index was calculated as the ratio between the normalized coverage in the gene body of genes (from TSS to TTS) and the normalized coverage in the promoter of genes (defined as +/- 300 bp from the TSS). Metagene profiles were generated by retrieving the signal for RNA PolII pS2 or pS5 in siC and siYT cells. The signal for genes belonging to each category (YAP+ super enhancers and YAP- superenhancer) was then normalized in bins to achieve equal gene size, averaged and normalized by signal across the whole genome.

RNA-seq. Reads were aligned to human Hg19 reference using TopHat software (version 2.0.9) (Kim et al., 2013), setting an average mate inner distance of 200 bp. Counts on HG19 UCSC KnownGene transcriptome reference were calculated using the scripts part of the HTSeq framework (version 0.5.3p9) (Anders et al., 2014). Gene-level differential expression analysis was performed using the package DESeq (Anders and Huber, 2010) for the R statistical program available in Bioconductor's repositories. A cutoff to 10% FDR and 1 logarithm in base 2 of fold change variation (positive or negative) was set to define significant differences in expression. Expression levels were calculated based on RNA-seq counts normalized on library size and unit length as FPKM.

Boxplots were generated using boxplotR (Spitzer et al., 2014). Center lines show the medians. Box limits indicate the 25th and 75th percentiles and whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles.

4C-Seq. 4C-sequencing analysis was performed as previously described (van de Werken et al., 2012) with some adaptations. After trimming primer sequences, reads were mapped against a restricted genome (*Homo Sapiens, hg19*) consisting of sequences flanking 4C restriction sites using custom perl scripts. Non-unique sequences (repeats) that flank a restriction site were removed from the analysis. To account for differences in sequencing depth, all reads were scaled by multiplying with a constant, such that the total number of reads is equal in all experiments considered (after discarding the top 3 most frequently sequenced captures). Normalized 4C profiles were generated by computing the mean fragment coverage (observed number of mapped reads) in a running window of 21 4C fragment-ends.

SUPPLEMENTAL REFERENCES

ANDERS, S. & HUBER, W. 2010. Differential expression analysis for sequence count data. *Genome Biol,* 11**,** R106.

ANDERS, S., PYL, P. T. & HUBER, W. 2014. HTSeq, A Python framework to work with high-throughput sequencing data. *bioRxiv preprint*.

CAMARGO, F. D., GOKHALE, S., JOHNNIDIS, J. B., FU, D., BELL, G. W., JAENISCH, R. & BRUMMELKAMP, T. R. 2007. YAP1 increases organ size and expands undifferentiated progenitor cells. *Curr Biol,* 17**,** 2054-60.

GENTLEMAN, R. C., CAREY, V. J., BATES, D. M., BOLSTAD, B., DETTLING, M., DUDOIT, S., ELLIS, B., GAUTIER, L., GE, Y., GENTRY, J., HORNIK, K., HOTHORN, T., HUBER, W., IACUS, S., IRIZARRY, R., LEISCH, F., LI, C., MAECHLER, M., ROSSINI, A. J., SAWITZKI, G., SMITH, C., SMYTH, G., TIERNEY, L., YANG, J. Y. & ZHANG, J. 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol,* 5**,** R80.

HNISZ, D., ABRAHAM, B. J., LEE, T. I., LAU, A., SAINT-ANDRE, V., SIGOVA, A. A., HOKE, H. A. & YOUNG, R. A. 2013. Super-enhancers in the control of cell identity and disease. *Cell,* 155**,** 934-47.

JIMENO, A., RUBIO-VIQUEIRA, B., AMADOR, M. L., GRUNWALD, V., MAITRA, A., IACOBUZIO-DONAHUE, C. & HIDALGO, M. 2007. Dual mitogen-activated protein kinase and epidermal growth factor receptor inhibition in biliary and pancreatic cancer. *Mol Cancer Ther,* 6**,** 1079-88.

KIM, D., PERTEA, G., TRAPNELL, C., PIMENTEL, H., KELLEY, R. & SALZBERG, S. L. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol,* 14**,** R36.

KIM, J., CANTOR, A. B., ORKIN, S. H. & WANG, J. 2009. Use of in vivo biotinylation to study protein-protein and protein-DNA interactions in mouse embryonic stem cells. *Nat Protoc,* 4**,** 506-17.

LANGMEAD, B. & SALZBERG, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods,* 9**,** 357-9.

LI, J., RAZUMILAVA, N., GORES, G. J., WALTERS, S., MIZUOCHI, T., MOURYA, R., BESSHO, K., WANG, Y. H., GLASER, S. S., SHIVAKUMAR, P. & BEZERRA, J. A. 2014. Biliary repair and carcinogenesis are mediated by IL-33-dependent cholangiocyte proliferation. *J Clin Invest,* 124**,** 3241-51.

MCLEAN, C. Y., BRISTOR, D., HILLER, M., CLARKE, S. L., SCHAAR, B. T., LOWE, C. B., WENGER, A. M. & BEJERANO, G. 2010. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol,* 28**,** 495-501.

SANDELIN, A., ALKEMA, W., ENGSTROM, P., WASSERMAN, W. W. & LENHARD, B. 2004. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res,* 32**,** D91-4.

SPITZER, M., WILDENHAIN, J., RAPPSILBER, J. & TYERS, M. 2014. BoxPlotR: a web tool for generation of box plots. *Nat Methods,* 11**,** 121-2.

SPLINTER, E., DE WIT, E., VAN DE WERKEN, H. J., KLOUS, P. & DE LAAT, W. 2012. Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: from fixation to computation. *Methods,* 58**,** 221-30.

VAN DE WERKEN, H. J., LANDAN, G., HOLWERDA, S. J., HOICHMAN, M., KLOUS, P., CHACHIK, R., SPLINTER, E., VALDES-QUEZADA, C., OZ, Y., BOUWMAN, B. A., VERSTEGEN, M. J., DE WIT, E., TANAY, A. & DE LAAT, W. 2012. Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat Methods,* 9**,** 969-72.

WHYTE, W. A., ORLANDO, D. A., HNISZ, D., ABRAHAM, B. J., LIN, C. Y., KAGEY, M. H., RAHL, P. B., LEE, T. I. & YOUNG, R. A. 2013. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell,* 153**,** 307-19.

YE, T., KREBS, A. R., CHOUKRALLAH, M. A., KEIME, C., PLEWNIAK, F., DAVIDSON, I. & TORA, L. 2011. seqMINER: an integrated ChIP-seq data interpretation platform. *Nucleic Acids Res,* 39**,** e35.

ZHANG, Y., LIU, T., MEYER, C. A., EECKHOUTE, J., JOHNSON, D. S., BERNSTEIN, B. E., NUSBAUM, C., MYERS, R. M., BROWN, M., LI, W. & LIU, X. S. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol,* 9**,** R137.