nature research

Corresponding author(s): Boyi Gan, Wei Li

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Zen (Zeiss) software was used to collect confocal images; Accuri C6 (BD Bioscience) was used to collect flow cytometry data;
Data analysis	ChIP seq analysis: R package DEGseq, bowtie v1.1.0, MACS v1.4.2. ImageJ were used for confocal images analysis; Graphpad were used for bar graphs output and statistic analysis; FlowJo_V10 was used for flow data analysis. Gene expression data were generated using the UCSC Xena Browser (http://xena.ucsc.edu/). The survival impact of different signatures was analyzed in R using the Kaplan-Meier method with a Cox proportional hazards model. The software and algorithms for data analyses used in this study are all well-established from previous work. All software and custom arguments are included in Methods section. There is no unreported algorithm used in this paper. The source code for data processing are available from the corresponding author on reasonable request.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

ChIP-seq data that support the findings of this study have been deposited in the Gene Expression Omnibus (GEO) under accession codes GSE101987. Previously published ChIP-seq data that were re-analyzed here are available under accession code GSM1250375, GSM1250374, and GSM1250376. The UVM, KIRC, KIRP, PCPG and BRCA data were derived from the TCGA Research Network: http://cancergenome.nih.gov/. The data-set derived from this resource that supports the findings of this study is available in UCSC Xena Browser (http://xena.ucsc.edu/). The normal sample data were derived from the GTEx database: https://gtexportal.org/home/. The data-set derived from this resource that supports the findings of this study is available in UCSC Xena Browser (http://xena.ucsc.edu/). Source data for Fig. 1, 2 and Supplementary Fig. 1 have been provided as Supplementary Table 5 "Statistics Source Data". All other data supporting the findings of this study are available from the corresponding author on reasonable request.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculations were performed. Sample size was determined according to our experience as well as literature reporting in terms of specific experiment.
Data exclusions	No samples or animals were excluded from the analyses
Replication	Multiple independent repeats were included for related experiments. Each experiment was performed for at least twice to make sure similar results are reproducible. ChIP-seq, RNA-Seq and mass spectormetry experiments have been done once, but ChIP-qPCR, RT-PCR and immunoprecipitation have been repeated more than twice for genes or proteins of interest.
Randomization	6-8 week female nude mice were chosen as xenograft hosts, and randomly allocated into experimental groups.
Blinding	For cell-based experiments, EM, Western blotting, immunostaining and FACS, cell types were known when prepare the samples or start to treat cells at the beginning of experiments. Data measurement for cell viability and FACS or photo capture were blinded to different person who processed assay at the time. ChIP-seq, RNA-seq and mass spectormetry analysis were blinded before analysis.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	H2Aub (Cell Signaling Technology, 8240), RNA pol II CTD (Cell Signaling Technology, 2629), RNA pol II CTD phospho Ser5 antibody (Active Motif, 61085), RNA pol II CTD phospho Ser2 antibody (Active Motif, 61083), FOXK1 (Abcam, ab18196), FOXK2 (Bethyl Laboratories, A301-730A), HCFC1 (Bethyl Laboratories, A301-399A), OGT (Cell Signaling Technology, 5368), KDM1B (Abcam,

ab193080), ASXL1 (Santa Cruz, sc-293204),BAP1 (Santa Cruz, sc-28383), SLC7A11 (Cell Signaling Technology, 12691), tubulin (Cell Signaling Technology, 2144), H2Aub (Millipore, 05-678), H2Aub antibody (Cell Signaling Technology, 8240), H2A (Millipore, ABE327), vinculin (Sigma, V4505), 4-Hydroxynonenal (Abcam, ab46545).p53 (Santa Cruz, sc-126), PARP (Cell Signaling Technology, 9542). tubulin (Cell Signaling Technology, 2144), β-actin (Abcam, ab8226), and V5 (Sigma, A8012). See Methods for detailed information on dilution information.

Validation

All antibodies used in our study have been validated and detailed information could be found on the website from manufactures as listed below. Some of them have also been validated by our experiments as shown in this manuscript using either over-express, knockout or knockdown strategies.
H2Aub, https://www.cellsignal.com/products/primary-antibodies/ubiquityl-histone-h2a-lys119-d27c4-xp-rabbit-mab/8240;
RNA pol II CTD, https://www.cellsignal.com/products/primary-antibodies/rpb1-ctd-4h8-mouse-mab/2629;
RNA pol II CTD phospho Ser5 , http://www.activemotif.com/catalog/details/61085/rna-pol-ii-ctd-phospho-ser5-antibody-mab;
RNA pol II CTD phospho Ser2, http://www.activemotif.com/catalog/details/61083/rna-pol-ii-ctd-phospho-ser2-antibody-mab;
FOXK1, https://www.abcam.com/foxk1-antibody-ab18196.html;
FOXK2, https://www.bethyl.com/product/A301-730A/FOXK2+Antibody;
HCFC1, https://www.bethyl.com/product/A301-399A?utm_source=Labome2018&utm_campaign=A301-399A&utm_medium=website;
OGT, https://www.cellsignal.com/products/primary-antibodies/ogt-antibody/5368;
KDM1B, https://www.abcam.com/lsd2-aof1-antibody-epr18508-ab193080.html;
ASXL1, https://www.scbt.com/scbt/product/asxl1-antibody-6e2;
BAP1,https://www.scbt.com/scbt/product/bap1-antibody-c-4?requestFrom=search;
SLC7A11,https://www.cellsignal.com/products/primary-antibodies/xct-slc7a11-d2m7a-rabbit-mab/12691;
Tubulin,https://www.cellsignal.com/products/primary-antibodies/a-tubulin-antibody/2144;
H2Aub, http://www.emdmillipore.com/US/en/product/Anti-ubiquityl-Histone-H2A-Antibody-clone-E6C5,MM_NF-05-678;
H2Aub, https://www.cellsignal.com/products/primary-antibodies/ubiquityl-histone-h2a-lys119-d27c4-xp-rabbit-mab/8240;
H2A, https://www.cellsignal.com/products/primary-antibodies/ubiquityl-histone-h2a-lys119-d27c4-xp-rabbit-mab/8240;
Vinculin, https://www.sigmaaldrich.com/catalog/product/sigma/v9131?lang=en&region=US&gclid=CjwKCAjwspHaBRBFEiwA0eM3kVN5VwVtzO3JtQWg1edd9vj8fO4qbCsSIdzANqvZtxgz0fvdsMByQBoCeTkQAvD_BwE;
4-HNE, https://www.abcam.com/4-hydroxynonenal-antibody-ab46545.html;
Actin, https://www.abcam.com/beta-actin-antibody-mabcam-8226-loading-control-ab8226.html;
V5, https://www.sigmaaldrich.com/catalog/product/sigma/v8012?lang=en&region=US&gclid=CjwKCAjwspHaBRBFEiwA0eM3kWLItkcZWmCRw30EAW5Qw3yHylAjHN_9dCRHW4NOn0xK9qrixCOJehoCf2cQAvD_BwE

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

HEK-293T, Caki1, 786-O, 769-P, ACHN and NCI-H226 cell lines were obtained from American Type Culture Collection (ATCC). RCC4, UMRC2, SLR20, and UMRC6 cell lines were obtained from Dr. William G. Kaelin at Dana-Farber Cancer Institute. TK10 cell line was obtained from Dr. Gordon Mills at MD Anderson Cancer Center.

Authentication

Cell lines were not authenticated.

Mycoplasma contamination

All cell lines tested negative for mycoplasma contamination.

Commonly misidentified lines
(See ICLAC register)

HEK-293T cells were used to for lentiviral production and luciferase reporter assay.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

NU/J nude female mice at 4-6 week old were purchased from ERO mouse facility in MD anderson cancer center. For Tumor xenograft models, tumor cells were injected subcutaneously into both flanks of 6-8 week old female nude mice. All animal experiments were approved by Department of Veterinary Medicine and Surgery, the university of Texas MD anderson cancer center.

Wild animals

No wild animals involved in this study.

Field-collected samples

This study didn't involve samples collected from field.

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?

Data access links *May remain private before publication.*	token=evavokoathgfziz acc=GSE101987
Files in database submission	C91A_H2Aub_ChIP.bw EV_H2Aub_ChIP.bw WT_H2Aub_ChIP.bw C91A_H2Aub_ChIP.fastq.gz C91A_Input.fastq.gz EV_H2Aub_ChIP.fastq.gz EV_Input.fastq.gz WT_H2Aub_ChIP.fastq.gz WT_Input.fastq.gz
Genome browser session (e.g. UCSC)	http://dldcc-web.brc.bcm.edu/lilab/jiejuns/BAP1_Gan/H2Aub_ChIP_20170515/C91A-ChIP2_subtract.bw http://dldcc-web.brc.bcm.edu/lilab/jiejuns/BAP1_Gan/H2Aub_ChIP_20170515/EV-ChIP2_subtract.bw http://dldcc-web.brc.bcm.edu/lilab/jiejuns/BAP1_Gan/H2Aub_ChIP_20170515/WT-ChIP2_subtract.bw

Methodology

Replicates	H2Aub level in the cells was verified and DNA quality was confirmed by ChIP-qPCR with appropriate control before sending for ChIP-seq. No replicates used for ChIPs-seq analysis.
Sequencing depth	Sample TotalReadsCount MappedReadsCount MappedRatio C91A-ChIP 37740545 32846643 87.03% C91A-Input 32189383 27093419 84.17% EV-ChIP 38776446 34028443 87.76% EV-Input 35778742 30024672 83.92% WT-ChIP 40504799 34956148 86.30% WT-Input 37378268 31270852 83.66%
Antibodies	H2Aub, Cell Signaling Technology, 8240
Peak calling parameters	macs14 -t alignmentA.bed -n sampleA --nomodel --nolambda -g hs --wig -S -p 1e-8
Data quality	FastQC (v0.11.2) was used for reads quality control. P < 1e-8 was used as cutoff to identify peaks with MACS (v1.4.2). sample PeakCount H2Aub_C91A 9606 H2Aub_EV 5407 H2Aub_WT 5260
Software	bowtie v1.1.0; MACS v1.4.2; deepTools v2.3.4; bedtools v2.26.0; R v3.2.3;

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Cells were incubated in 6-well plate containing 5μM BODIPYTM 581/591 C11 dye (Invitrogen, D3861). After incubation for 30 min, cells were washed with PBS and trypsinized followed by PI staining in PBS for 5 min. Then cells were subjected to flow cytometry analysis using a cytometer Accuri C6.
Instrument	Accuri C6 (BD Bioscience)
Software	Using Accuri C6 software to collect data and FlowJo_V10 software to analyze data.
Cell population abundance	At least 10000 cells were analyzed for each sample.
Gating strategy	Initial cell population gating (FSC-Area VS FSC-Height) was adopted to make sure doublet exclusion and only single cell was used for analysis. A figure exemplifying the gating strategy is provided in the Supplementary Table 5.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.