

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 our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the 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
- ☐ ☒ A statement on 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 statistical parameters 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

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

1. Picture acquisition for Western blot and DNA gels was performed in Li-Cor Fc Odyssey Imaging System (LI-COR Biosciences, Lincoln, NE) as outlined in the Supplementary Experimental procedures.
2. Cell proliferation assays were performed by Photo-micrographs taken every 6 hours using an Incucyte live cell imager (Essen Biosciences, Ann Arbor, MI), as described in the Supplementary Experimental procedures.
3. For the cell cycle/FACS analysis, Propidium Iodide stained cell cultures were analyzed on the BD FACS Calibur v2.3 (BD Biosciences, San Jose, CA, <https://www.bdbiosciences.com/en-us/instruments/clinical-instruments/clinical-cell-analyzers/facscalibur>). For cell sorting, NCI-H1299 cells after the staining with Vybrant™ DyeCycle™ Ruby Stain (Thermo Fisher, Cat. No. V10309) were sorted in BD FACS Aria III (BD Biosciences, San Jose, CA, <https://www.bdbiosciences.com/en-us/instruments/research-instruments/research-cell-sorters/facsaria-iii>). The data were analyzed using FlowJo 9.3.2.
4. All IHC/slide images were captured on the Nikon eclipse 50i microscope with attached Axiocam 506 color camera using the ZEN 2.6 blue edition software (Zeiss).

Data analysis

1. The RT-PCR quantification, qRT-PCR, ChIP-qPCR, RIP-qPCR, ChIC-qPCR, end-stage cell proliferation analysis (Fig. 6), percentage change analysis on Ki-67 Western blot (Fig. 7), Ki-67 IHC scores (Fig. 7), proliferation (Fig. S9) and G2/M phase (Fig. 6), Number of Transformed cells in Fig. S8, expression of various components of the pathway in different patient groups (Stage I/II vs III/IV, grade II vs III, M0 vs M1, Low IWS1 vs High IWS1, Relapsed vs Non-relapsed, primary Lung Adenocarcinoma vs Brain Metastatic lesions, Normal Adjacent Tissue vs Tumor Samples) in Figure 8 and S11, were analyzed by unpaired one-tailed t test using GraphPad Prism v8.4.2
2. The RT-PCR bands were quantified using the native analyzer in Li-Cor Fc Odyssey Imaging System (LI-COR Biosciences, Lincoln, NE).
3. The western blotting band intensity (Fig. 5 and Fig. 8) was measured with Image J 1.8.0.
4. The correlation of all the components of the IWS1/CDCA5 pathway in LUAD samples in Fig. 8 and Fig. S8 were performed using Pearson's linear correlation function in GraphPad Prism v8.4.2.
5. The Cancer Genome Atlas (TCGA) PanCancer Atlas LUAD cohort was used for survival analysis, clinical stage and gene expression profiling in Fig. 8 (Cell 173, 400-416 e411 (2018)). We excluded cases based on those that were masked ("Do_not_use") according to the Pan-Cancer Atlas sample quality annotations (<http://api.gdc.cancer.gov/data/1a7d7be8-675d-4e60-a105-19d4121bdebf>). Lung adenocarcinoma survival

analysis and gene expression was retrieved from two different publicly available Agilent microarray datasets present on GEO (GSE13213 (Journal of clinical oncology, 27(17), pp.2793-2799.9), GSE26939 (PLoS One 2012;7(5):e36530) by using Log Ratio values calculated by the authors and imported in the R-Studio framework (V 3.5.2) for the gene selection. Expression values from 6 dissected brain metastasis derived from lung adenocarcinoma patients were obtained from RNA-seq data provided in the publicly available dataset GSE141685 and analyzed as outlined in the Supplementary Experimental Procedures. Heat-maps generation and Kaplan-Meier Curves were performed with GraphPad Prism v8.4.2.

6. The percentage reduction of p-ERK levels (Fig. 5H) was analyzed by one-way ANOVA, using GraphPad Prism v8.4.2

7. The IHC Ki-67 (Fig. 7) and IHC TMA scores (Fig. S8) were performed with Image J v1.8.0, as described in details in the Supplementary Experimental procedures.

8. The tumor volume, tumor weight and Ki-67 IHC scores between the NCI-H1299, A549 shControl vs shIWS1 xenografts tumors were analyzed by paired one-tailed t test using GraphPad Prism v8.4.2

9. The RNA-Seq data alignment and DESeq, DEXSeq, GO and GSEA analysis were performed with standard pipelines as described in depth in the Supplementary Experimental Procedures.

10. The volcano plots in Fig. 1 and Fig. S1 were performed using GraphPad Prism v8.4.2, as described in the figure legends of the manuscript.

11. The ChIP-qPCR, RIP-qPCR and ChC-qPCR data were analyzed using the analysis substrate file provided online by Sigma-Aldrich, calculating the fold enrichment. (<https://www.sigmaaldrich.com/technical-documents/articles/biology/chip-qpcr-data-analysis.html>).

12. The SRSF1 binding sites on U2AF2, SLC12A2, IFT88, STXBP1, C1qTNF6 IWS1-phosphorylation target exons, were calculated using the RBPmap pipeline, with recommended parameters (Nucleic acids research 42.W1 (2014): W361-W367., <http://rbpmap.technion.ac.il/index.html>)

13. The summary graphical abstracts in Fig. 4h, Fig. 6j and Supplementary Fig. 9c were created with Biorender (Created with BioRender.com).

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 and reviewers. 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

All the raw data underlying figures 1-8 (uncropped gel images, qPCR, FACS, plates reader and proliferation data), processed counts of RNA-Seq, Junction-Seq splicing platform for visualization of the exon usage results and microscope images derived from this report have been deposited in the Mendeley Dataset in two independent publicly available datasets (<http://dx.doi.org/10.17632/y82yx8gfkj.4>, <http://dx.doi.org/10.17632/ts44bfg45k.1>). In addition, all the uncropped WB and PCR gels are provided in the source data file. Specific P values are also included in these datasets. The main uncropped full scans are provided as source data with the manuscript. The latest version of Ensembl database can be downloaded through the Ensembl project (<http://useast.ensembl.org/info/data/ftp/index.html>) or github (<https://github.com/Ensembl/ensembl-hive>). All the RNA-seq and ChIP-Seq data in this report have been deposited in the Gene Expression Omnibus (GEO), under the SuperSeries accession number GSE166955 (RNA-seq : GSE166953, ChIP-Seq : GSE165964).

Field-specific reporting

Please select the one below that is 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/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

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

Sample size	For in vitro experiments, at least three biological replicates were achieved for all the experiments. For in vivo experiments, no size calculations were performed. The sample was chosen based on the evidence of our previous work (Molecular cell, 53(4), 577-590.), on which a sample size of n = 5 mice was used per experimental group with and was sufficient on ensuring statistical power.
Data exclusions	No data were excluded in this study
Replication	All the in vitro experiments were repeated independently at least in three biological replicates Multiple mice (n = 5/group) were used for every in vivo experiment. All the attempts at replication were successful.
Randomization	For animal studies, the mice were earmarked before grouping and then were randomly separated into groups by an independent person; however, no particular method of randomization was used. For other experiments, cells/samples were randomly assigned to groups to avoid bias.
Blinding	The IHC for TMA, WB of tumor samples and their analysis were performed in a blinded fashion. For the in vitro experiments, the researchers were unaware of the abundance of the proteins/mRNA transcripts willing to detect. The tumor samples in this study were provided as unidentified samples from OSUCCC Total Cancer Care (TCC) tissue bank. The TMA samples were purchased from US Biomax and were also provided as unidentified samples to ensure blinding approach.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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

Antibodies

Antibodies used

1. IWS1, a-rabbit, Cell Signaling Technologies #5681
2. GAPDH, a-rabbit, Cell Signaling Technologies #8884
3. p-IWS1, a-rabbit, Cell Signaling Technologies Gift (Molecular cell, 53(4), 577-590.)
4. Akt3, a-rabbit, Cell Signaling Technologies #14982
5. pan-Akt, a-rabbit, Cell Signaling Technologies #4691
6. p-Akt (T308), a-rabbit, Cell Signaling Technologies #13038
7. p-Akt (S473), a-rabbit, Cell Signaling Technologies #4060
8. SETD2, a-rabbit, Cell Signaling Technologies #23486
9. H3K36me3, a-rabbit, Cell Signaling Technologies #4909
10. Vincullin, a-rabbit, Cell Signaling Technologies #13901
11. HA-tag, a-rabbit, Cell Signaling Technologies #3724
12. NSD1, a-rabbit, AbCam #Ab70732
13. NSD2, a-rabbit, Cell Signaling Technologies #65127
14. NSD3, a-rabbit, Cell Signaling Technologies #92056
15. SMYD2, a-rabbit, AbCam #Ab108217
16. ASH1L, a-mouse, AbCam #Ab50981
17. V5-Tag, a-rabbit, Cell Signaling Technologies #13202
18. V5-tag, a-mouse, Cell Signaling Technologies #80076
19. MRG15, a-rabbit, Cell Signaling Technologies #14098
20. Actin, a-rabbit, Cell Signaling Technologies #4970
21. PHF1, a-rabbit, AbCam #Ab184951
22. BRPF1, a-mouse, AbCam #Ab168673
23. GLYR1, a-rabbit, AbCam #Ab167155
24. MSH6, a-rabbit, AbCam #Ab92471
25. LEDGF, a-rabbit, Cell Signaling Technologies #2088
26. a-tubulin, a-mouse, Sigma #T5168
27. PTB, a-rabbit, Cell Signaling Technologies #57246
28. SRSF1, a-rabbit, Cell Signaling Technologies #14908
29. Prp19, a-rabbit, GeneTex #GTX111233
30. U2AF65, a-mouse, Sigma #U4758
31. U2AF65, a-rabbit, Sigma #HPA043562
32. Sororin, a-rabbit, Thermofisher #PA544877
33. p-ERK (T202/Y204), a-rabbit, Cell Signaling Technologies #4370
34. ERK, a-rabbit, Cell Signaling Technologies #4695
35. p-CDK1 (Y15), a-mouse, Cell Signaling Technologies #4539
36. CDK1, a-mouse, Cell Signaling Technologies #9116
37. Cyclin B1, a-rabbit, Cell Signaling Technologies #12231
38. Lamin A/C, a-rabbit, Cell Signaling Technologies #2032
39. PCNA, a-rabbit, Cell Signaling Technologies #13110
40. Ki-67, a-rabbit, US biological #C100511453
41. Histone H3, a-rabbit, Cell Signaling Technologies #4499
42. Cyclin E1, a-rabbit, Cell Signaling Technologies #20808
43. CDC25A, a-rabbit, Cell Signaling Technologies #3652
44. phospho-Histone H3 (S10), a-rabbit, Cell Signaling Technologies #53348
45. RAS (G12V Mutant Specific), a-rabbit, Cell Signaling Technologies #14412
46. RAS (G12D Mutant Specific), a-rabbit, Cell Signaling Technologies #14429
47. EGF Receptor (L858R Mutant Specific), a-rabbit, Cell Signaling Technologies #3197
48. EGF Receptor (E746-A750del Specific), a-rabbit, Cell Signaling Technologies #2085
49. Flag-tag, a-rabbit, Cell Signaling Technologies #14793
50. PTBP2, a-rabbit, AbCam #ab154787
51. Akt1, a-rabbit, Cell Signaling Technologies #2938

Validation

52. Akt2, a-rabbit, Cell Signaling Technologies #3063
53. Akt3, a-rabbit, Cell Signaling Technologies #14982
54. α -actinin, a-rabbit, Cell Signaling Technologies #6487
55. SF3B1, a-rabbit, Cell Signaling Technologies #14434

All Cell Signaling antibodies were validated by the manufacturer using various approaches (<https://www.cellsignal.com/about-us/our-approach-process/cst-antibody-validation-principles>). All Abcam antibodies used in this report, were validated by the manufacturer as well (<https://www.abcam.com/primary-antibodies/how-we-validate-our-antibodies>). In the case of abcam and Genetex antibodies, we also include the links from the commercial website, where the validation is shown. Specifically, the applications of the antibodies listed in this report can be found below :

1. IWS1, a-rabbit, Cell Signaling Technologies #5681, WB (1:1000), ChIP (1:50), ChIC (1:50)
2. GAPDH, a-rabbit, Cell Signaling Technologies #8884, WB (1:1000)
3. p-IWS1, a-rabbit, Cell Signaling Technologies Gift (Molecular cell, 53(4), 577-590.), WB (1:1000), IHC (1:50)
4. Akt3, a-rabbit, Cell Signaling Technologies #14982, WB (1:1000)
5. pan-Akt, a-rabbit, Cell Signaling Technologies #4691, WB (1:1000)
6. p-Akt (T308), a-rabbit, Cell Signaling Technologies #13038, WB (1:1000)
7. p-Akt (S473), a-rabbit, Cell Signaling Technologies #4060, WB (1:1000)
8. SETD2, a-rabbit, Cell Signaling Technologies #23486, WB (1:1000), ChIC (1:50)
9. H3K36me3, a-rabbit, Cell Signaling Technologies #4909, WB (1:1000), ChIP (1:50), ChIC (1:50)
10. Vincullin, a-rabbit, Cell Signaling Technologies #13901, WB (1:1000)
11. HA-tag, a-rabbit, Cell Signaling Technologies #3724, WB (1:1000), ChIP (1:50)
12. NSD1, a-rabbit, AbCam #Ab70732, WB (1:1000), (<https://www.abcam.com/kmt3b-nsd1-antibody-2658c1a-ab70732.html>)
13. NSD2, a-rabbit, Cell Signaling Technologies #65127, WB (1:1000)
14. NSD3, a-rabbit, Cell Signaling Technologies #92056, WB (1:1000)
15. SMYD2, a-rabbit, AbCam #Ab108217, WB (1:1000), (<https://www.abcam.com/kmt3c-smyd2-antibody-ab108217.html>)
16. ASH1L, a-mouse, AbCam #Ab50981, WB (1:1000), (<https://www.abcam.com/ash1l-antibody-ash5h03-ab50981.html>)
17. V5-Tag, a-rabbit, Cell Signaling Technologies #13202, WB (1:1000), ChIP (1:50), RIP (1:100)
18. V5-tag, a-mouse, Cell Signaling Technologies #80076, IP (1:50)
19. MRG15, a-rabbit, Cell Signaling Technologies #14098, WB (1:1000)
20. Actin, a-rabbit, Cell Signaling Technologies #4970, WB (1:1000)
21. PHF1, a-rabbit, AbCam #Ab184951, WB (1:1000), (<https://www.abcam.com/phf1-antibody-epr14222-ab184951.html>)
22. BRPF1, a-mouse, AbCam #Ab168673, WB (1:1000), (<https://www.abcam.com/peregrinbrpf1-antibody-ab168673.html>)
23. GLYR1, a-rabbit, AbCam #Ab167155, WB (1:1000), (<https://www.abcam.com/glyr1-antibody-epr10077b-ab167155.html>)
24. MSH6, a-rabbit, AbCam #Ab92471, WB (1:1000), (<https://www.abcam.com/msh6-antibody-epr3945-ab92471.html>)
25. LEDGF, a-rabbit, Cell Signaling Technologies #2088, WB (1:1000)
26. α -tubulin, a-mouse, Sigma #T5168, WB (1:1000), (<https://www.sigmaaldrich.com/catalog/product/sigma/t5168>)
27. PTB, a-rabbit, Cell Signaling Technologies #57246, WB (1:1000)
28. SRSF1, a-rabbit, Cell Signaling Technologies #14908, WB (1:1000)
29. Prp19, a-rabbit, GeneTex #GTX111233, WB (1:1000), RIP (1:100), (<https://www.genetex.com/Product/Detail/PRP19-antibody-N2C2-Internal/GTX111233>)
30. U2AF65, a-mouse, Sigma #U4758, WB (1:1000), IP (1:100), RIP (1:100), (<https://www.sigmaaldrich.com/catalog/product/sigma/u4758>), the RIP validation has been demonstrated in Genes & development 25.7 (2011): 701-716.
31. U2AF65, a-rabbit, Sigma #HPA043562, WB (1:1000), (<https://www.sigmaaldrich.com/catalog/product/sigma/hpa043562>)
32. Sororin, a-rabbit, Thermofisher #PA544877, WB (1:1000), IHC (1:50), (<https://www.thermofisher.com/antibody/product/CDCA5-Antibody-Polyclonal/PA5-44877>)
33. p-ERK (T202/Y204), a-rabbit, Cell Signaling Technologies #4370, WB (1:1000), IHC (1:50)
34. ERK, a-rabbit, Cell Signaling Technologies #4695, WB (1:1000)
35. p-CDK1 (Y15), a-mouse, Cell Signaling Technologies #4539, WB (1:1000), IHC (1:50)
36. CDK1, a-mouse, Cell Signaling Technologies #9116, WB (1:1000)
37. Cyclin B1, a-rabbit, Cell Signaling Technologies #12231, WB (1:1000)
38. Lamin A/C, a-rabbit, Cell Signaling Technologies #2032, WB (1:1000)
39. PCNA, a-rabbit, Cell Signaling Technologies #13110, WB (1:1000)
40. Ki-67, a-rabbit, US biological #C100511453, IHC (1:100)
41. Histone H3, a-rabbit, Cell Signaling Technologies #4499, WB (1:1000)
42. Cyclin E1, a-rabbit, Cell Signaling Technologies #20808, WB (1:1000)
43. CDC25A, a-rabbit, Cell Signaling Technologies #3652, WB (1:1000)
44. phospho-Histone H3 (S10), a-rabbit, Cell Signaling Technologies #53348, WB (1:1000)
45. RAS (G12V Mutant Specific), a-rabbit, Cell Signaling Technologies #14412, WB (1:1000)
46. RAS (G12D Mutant Specific), a-rabbit, Cell Signaling Technologies #14429, WB (1:1000)
47. EGF Receptor (L858R Mutant Specific), a-rabbit, Cell Signaling Technologies #3197, WB (1:1000), IHC (1:50)
48. EGF Receptor (E746-A750del Specific), a-rabbit, Cell Signaling Technologies #2085, WB (1:1000)
49. Flag-tag, a-rabbit, Cell Signaling Technologies #14793, ChIP (1:50)
50. PTBP2, a-rabbit, AbCam #ab154787, WB (1:1000)
51. Akt1, a-rabbit, Cell Signaling Technologies #2938, WB (1:1000)
52. Akt2, a-rabbit, Cell Signaling Technologies #3063, WB (1:1000)
53. α -actinin, a-rabbit, Cell Signaling Technologies #6487, WB (1:1000)
54. SF3B1, a-rabbit, Cell Signaling Technologies #14434, WB (1:1000)

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)

The human NCI-H522, A549, NCI-H460 non-small cell lung cancer adenocarcinoma cell line originated from an NCI-60 cell line panel obtained from Daniel Haber at Massachusetts General Hospital. The NCI-H1299, NCI-H1975, PC-9, NCI-H1650 non-small cell lung cancer adenocarcinoma cell line and the immortalized Human Bronchial Epithelial Cells (HBEC-hTERT) originated

from Dr Carbone's Lab at the Ohio State University. The HEK-293T human embryonic kidney cell line originated from Richard Van Etten's laboratory at Tufts Molecular Oncology Research Institute and purchased from ATCC (<https://www.atcc.org/products/all/crl-3216.aspx>)

Authentication	Cell line authentication was performed by the originating cell line repositories and then used immediately upon receipt. After 1 year the original stock was expanded to insure appropriate short tandem repeat profiling.
Mycoplasma contamination	All cell lines were negative for mycoplasma contamination. Cell lines were checked monthly for mycoplasma, using the PCR mycoplasma detection kit (ABM, Cat No. G238). All experiments were carried out in mycoplasma-free cultures. Evidence of mycoplasma-free cultures can be found in the electronic lab book in the Mendeley dataset.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell line was used in the study.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	Immunocompromised 6 week old NSG (NOD-scid-IL2R- γ null) mice were obtained from the Jackson Laboratory. The mice were housed in accordance to the University of Michigan Institutional Animal Care & Use Committee (IACUC) rules and guidelines. The Ohio State University Laboratory Animal Resources (ULAR) ensured that the housing temperature was kept at 74 degrees F, relative humidity between 30%-70%, and was on a light/dark cycle of 12 hours on/12 hours off.
Wild animals	No wild animals were used in this study.
Field-collected samples	No field collected samples were used in the study.
Ethics oversight	All experimental procedures were approved by the Institutional Animal Care and Use Committee (IACUC) of the Ohio State University. IACUC protocol number 2018A00000134; P.I.: Lalit Sehgal, IACUC protocol number 2018A00000126; P.I.: Philip Tschlis.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	The tumor samples in this study were provided as unidentified samples from OSUCCC Total Cancer Care (TCC) tissue bank. The TMA samples were purchased from US Biomax and were also provided as unidentified samples to ensure blinding approach.
Recruitment	Not specific patients were recruited for this study. All the cases were lung adenocarcinoma patients treated at The Ohio State University Comprehensive Cancer Center.
Ethics oversight	Thirty Lung Adenocarcinoma samples with matching normal adjacent tissue were obtained from the Tissue Bank of The Ohio State University, under the universal consenting and biobanking protocol, Total Cancer Care (TCC). TCC is the single protocol used by the Oncology Research Information Exchange Network (ORIEN), which was formed through a partnership between the OSUCCC – James and Moffitt Cancer Center in Tampa, FL. For more information, please advise the Biospecimen Core Services facility of The Ohio State university Comprehensive Cancer Center (https://cancer.osu.edu/for-cancer-researchers/resources-for-cancer-researchers/shared-resources/biospecimen-services) and ORIEN project (https://cancer.osu.edu/for-cancer-researchers/resources-for-cancer-researchers/orien).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

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 <i>May remain private before publication.</i>	All the RNA-seq and ChIP-Seq data in this report have been deposited in the Gene Expression Omnibus (GEO), under the Superseries accession number GSE166955 (RNA-seq : GSE166953, ChIP-Seq : GSE165964).
Files in database submission	The RNA-Seq and ChIP-Seq raw (fastq) and processed files (raw counts and BigWig respectively) can be found in the Gene Expression Omnibus (GEO), under the Superseries accession number GSE166955 (RNA-seq : GSE166953, ChIP-Seq : GSE165964).
Genome browser session (e.g. UCSC)	The Bigwig files can be imported in IGV for peak calls.

Methodology

Replicates	Two replicates for each of the ChIP-Seq and RNA-Seq experiments. Only common targets were selected for further analysis.
Sequencing depth	1. Paired End sequencing (2x100 bp) of Poly-A+ selected RNAs was realized for shControl, shIWS1, shIWS1/WT-IWS1 rescue and shIWS1/phosphorylation site mutant IWS1 rescue NCI-H522 cells. All RNA-Seq experiments were performed in duplicates, and average depth of sequenced samples was 37,5 M (± 5 M fragments). The whole procedure was performed in Tufts University Core Genomic Facility (TUCF- http://tucf-genomics.tufts.edu). 2. All ChIP-Seq experiments were performed in duplicates, and average depth of sequenced samples was 49M (± 5 M fragments). Paired-End sequencing (2x100bp) was used. All the procedures were performed in the DNA Sequencing Center of Brigham Young University. (Provo, UT) (https://biology.byu.edu/dnasc)
Antibodies	1. H3K36me3, a-rabbit, Cell Signaling Technologies #4909, 1:50 2. HA-tag, a-rabbit, Cell Signaling Technologies #3724, 1:50 3. Flag-tag, a-rabbit, Cell Signaling Technologies #14793, 1:50
Peak calling parameters	ChIP-Seq libraries were quality checked using FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adapters and sequence contaminants were detected and removed using Cutadapt (EMBnet. J. 17, 10–12, 2011). Paired end reads were aligned against the human reference genome (GRCh38/hg38) using Bowtie (version 2.2.6) (with default parameters). Peak discovery was performed with HOMER (version 4.6) (Mol. Cell 38, 576–589, 2010). Sonicated input was used as a control for peak discovery. Snapshots of genomic areas were created using the Integrative Genomic Viewer (IGV) of the Broad Institute. (https://software.broadinstitute.org/software/igv/home)
Data quality	For both the RNA-seq and ChIP-seq analysis, QC took place. Quality metrics used included base calling quality (MAPQ>30). Specifically for ChIP-seq, the plot Fingerprint function from Deeptools (http://deeptools.readthedocs.io) was used to visualize sufficient enrichment over background.
Software	1. Bowtie is available here : https://sourceforge.net/projects/bowtie-bio/files/bowtie2/2.2.6/ 2. HOMER is available here: http://homer.ucsd.edu/homer/ (Heinz S, Benner C, Spann N, Bertolino E et al. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. Mol Cell 2010 May 28;38(4):576-589) 3. The RNA-Seq was performed using DESeq (Genome Biol. 2010;11:R106.) and DEXSeq (Nature Precedings (2012): 1-1.)

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	1. For cell cycle analysis in Fig. 6, cells were harvested and fixed with pre-cooled 70% ethanol overnight, followed by staining with Propidium Iodide/RNAase A satining buffer in 37oC for 30' 2. For the sorting of NCI-H1299 cells in different phases of the cell cycle in Fig. 6, cells were harvested and were stained by adding 2uL/mL Vybrant™ DyeCycle™ Ruby Stain (Thermo Fisher, Cat. No. V10309) and incubated in the dark at 37oC for 30 minutes
Instrument	1. For FACS : BD FACS Calibur (BD Biosciences, San Jose, CA) 2. For Sorting : BD FACS Aria III (BD Biosciences, San Jose, CA) The analysis was performed in the Flow Cytometry Shared Resource (FCSR) of the Ohio State University (https://cancer.osu.edu/for-cancer-researchers/resources-for-cancer-researchers/shared-resources/flow-cytometry)
Software	FlowJo 9.3.2
Cell population abundance	The Vybrant® DyeCycle™ Ruby stain is cell membrane- permeant, DNA-selective, and essentially nonfluorescent until bound to double-stranded DNA. This stain takes advantage of the commonly available 488 nm and 633/5 nm excitation sources with emission >670 nm, placing cell cycle studies within reach of all flow cytometrists. Vybrant® DyeCycle™ Ruby stain may also be used with any excitation source from 488 nm through to 690 nm wavelengths. The Vybrant® DyeCycle™ Ruby stain allows simultaneous co-staining of the cell population for other parameters and offers the possibility of cell sorting based on DNA content. Based on the DNA content of cells being on the different phase of the cell cycle, cells were sorted in G1, S and G2/M phase pools.

The gating strategy for FACS and sorting analysis can be found in the supplementary Information file in supplementary figure S13.

Gating strategy

For the gating we used linear amplification for DNA content and low flow rate for acquisition. Analyze without washing cells on a flow cytometer using 488 nm or 633/5 nm excitation and >670 nm emission. The Vybrant® DyeCycle™ Ruby stain may be excited with any excitation source from 488 nm through 690 nm wavelengths, as suggested by the manufacturer's protocol.

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