

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

Luminescence measurements were collected using EnVision Workstation Software (PerkinElmer); RT-qPCR measurements were collected with QuantStudio 6 and 7 Flex Software (Applied Biosystems); immunoblot scans were performed with Image Studio Software (LI-COR Biosciences); low-throughput cell counting (pertains to Fig. 1c) was performed with Vi-CELL XR Software (Beckman Coulter); high-resolution digital immunohistochemical staining images were generated using Leica ScanScope XT.

Data analysis

CRISPR-Cas9 screening data were analyzed using a reproducible GenePattern pipeline containing individual modules, all available from GParc (gparc.org); STARS (<https://portals.broadinstitute.org/gpp/public/software/stars>); and MATLAB (MathWorks, Inc.). Small-molecule sensitivity analysis and visualizations were performed with Microsoft Excel Professional Plus 2013; GraphPad Prism (v6 or v7); Pipeline Pilot (version 8.5) (Accelrys, Inc.); and MATLAB 8.4 (2014b) or MATLAB 9.4 (2018a). RNA-sequencing analysis was performed with Trim Galore! (version 0.4.1) (Babraham Bioinformatics); and STAR (version 2.4.2). Differential expression analysis was performed in R (version 3.2.1) using the DESeq2 package (version 1.10.0). Gene set enrichment analysis was performed with GSEA2 (version 2.2.0). Copy number variation analysis was performed with Picard (version 1.999) (Broad Institute Picard); BWA (version 0.7.12); GATK3 Toolkit (version 2.2) (Broad Institute GATK); and GATK4 Toolkit (Broad Institute GATK). RT-qPCR analysis was performed in R (version 3.4.4). ChIP-seq analysis was performed with Bowtie2 (version 2.2.1); Bamliquidator (version 1.0); MACS (version 1.4.2); ROSE2 (<https://github.com/BradnerLab/pipeline>); STAR; and Sailfish. ATAC-seq data analysis was performed with Bowtie2 (version 2.2.1); samtools (0.1.19); and MACS (version 1.4.2). Brachyury ChIP-seq analysis was performed with Bowtie 2 (version 2.2.1); Bedtools (version 2.26.0); MACS (version 1.4.2); and HOMER (version 4.9.1) (<http://homer.ucsd.edu/homer/>). All other statistical analyses were performed with GraphPad Prism (v6 or v7); or R (version 3.4.4).

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

CRISPR-Cas9 screening data for two chordoma cell lines (pertains to Figs. 1, 2, 4, and Extended Data Fig. 1) are available at Figshare (DOI: 10.6084/m9.figshare.7302515). CRISPR-Cas9 screening data for all other cancer cell lines (pertains to Figs. 1, 2, 4, and Extended Data Fig. 1) were generated as part of Project Achilles (Broad Institute Project Achilles; <https://depmap.org/portal/achilles/>). All RNA-sequencing data (pertains to Figs. 1 and 4) are available at Gene Expression Omnibus (GEO) (accession number: GSE121846). Small-molecule sensitivity data generated using non-chordoma cell lines and used for comparative analyses (pertains to Fig. 2) are available at the National Cancer Institute's CTD2 Data Portal (<https://ocg.cancer.gov/programs/ctd2/data-portal>) and the CTRP (www.broadinstitute.org/ctrp/). The analysis of new small-molecule primary screening data generated using chordoma cell lines (pertains to Fig. 2) was performed as described previously²², except as noted in the Methods, and the resulting AUC values are provided in Supplementary Table 2. Raw plate-reader data files and accompanying Pipeline Pilot and MATLAB scripts for small-molecule primary screening and low-throughput compound sensitivity analysis (pertains to Fig. 2, Extended Data Figs. 2a, 3a, 8c, and 10a) are available upon request. Chromatin profiling data (pertains to Figs. 3, 4, and Extended Data Figs. 4 and 6) are available at GEO (accession number: GSE109794).

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	Sample sizes were determined on the basis of previous studies that had used similar sample sizes to successfully detect a difference between groups.
Data exclusions	For primary small-molecule screening, individual microtiter wells not meeting routine quality control standards of our high-throughput screening facility were excluded from analysis. For in vivo studies, one mouse in the vehicle group did not have successful tumor engraftment and was therefore excluded from the study.
Replication	Measures taken to verify the reproducibility of the findings included some combination of the following: confirming an effect across multiple cell lines or tumor samples, replicating findings with independent experiments, using biological and technical replicates, using multiple concentrations of a compound, using multiple sgRNAs targeting a gene of interest, and confirming ex vivo results in different model systems (patient-derived tumors or mice). All attempts at replication were successful.
Randomization	Human research participants were not grouped. Animals were randomly grouped.
Blinding	Human research participants were not grouped. Blinding was not possible for animal studies, as animals were segregated into distinct, labeled cages according to the treatment group.

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

For immunoblotting: the goat polyclonal brachyury antibody (clone N-19, #sc-17743; 1:1,000) and the mouse monoclonal CDK7 antibody (clone C-4, #sc-7344; 1:1,000) were purchased from Santa Cruz Biotechnology. The rabbit monoclonal cofilin antibody (clone D3F9, #5175; 1:10,000), the mouse monoclonal total Rb antibody (clone 4H1, #9309; 1:2,000), the rabbit monoclonal phospho-Rb S807/S811 antibody (clone D20B12, #8516; 1:1,000), the rabbit monoclonal phospho-Rb S780 antibody (clone D59B7, #8180; 1:1,000), the mouse monoclonal total STAT1 antibody (clone 9H2, #9176; 1:1,000), the rabbit monoclonal phospho-STAT1 S727 antibody (clone D3B7, #8826; 1:1,000), and the rabbit monoclonal SOX9 (clone D8G8H, #82630; 1:1,000) were purchased from Cell Signaling Technology. The rat monoclonal phospho-POLR2A Ser2 antibody (clone 3E1, #04-1571-I; 1:1,000), the rat monoclonal phospho-POLR2A Ser5 antibody (clone 3E8, #04-1572-I; 1:1,000), and the rat monoclonal phospho-POLR2A Ser7 antibody (clone 4E12, #04-1570-I; 1:1,000) were purchased from Millipore. The rabbit polyclonal total POLR2A antibody (#A300-653A; 1:400) was purchased from Bethyl Laboratories. The mouse monoclonal V5-tag antibody (#46-0705; 1:5,000) was purchased from Invitrogen. IRDye secondary antibodies were purchased from LI-COR Biosciences (1:10,000). For ChIP-seq: the rabbit polyclonal H3K27Ac antibody (#ab4729) was purchased from abcam. For immunohistochemistry, the goat polyclonal brachyury antibody (clone N-19, #sc-17743; 1:150) was purchased from Santa Cruz Biotechnology. Lot numbers were not routinely recorded.

Validation

All antibodies are commercially available and have been validated by their manufacturers as described on the company websites. In addition, the brachyury antibody (clone N-19, #sc-17743) was validated using sgRNA-mediated T (brachyury) repression experiments as reported in Fig. 1c.

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)

UM-Chor1, MUG-Chor1, U-CH1, U-CH2, and JHC7 chordoma cell lines were provided by the Chordoma Foundation; the CH22 chordoma cell line was provided by the Chordoma Foundation and Massachusetts General Hospital; chondrosarcoma cell lines were obtained from the Broad Institute Cancer Cell Line Encyclopedia project (Barretina et al.; DOI: 10.1038/nature11003).

Authentication

All chordoma cell lines were authenticated by the Chordoma Foundation using the following criteria: each line must have a confirmed diagnosis of chordoma, be of human origin, and have a rearranged genome. Additionally, the line must be immortal, have gone through at least 50 population doublings in culture, and have a genotype that does not match any known cell line by STR analysis. Furthermore, at least two of the following criteria must be met: expression of brachyury, expression of CD24, evidence of a physaliferous morphology, or evidence of genomic rearrangement that matches the original tumor. Chondrosarcoma cell lines were authenticated by the Cancer Cell Line Encyclopedia project as described by Barretina et al. (DOI: 10.1038/nature11003).

Mycoplasma contamination

Chordoma cell lines were tested for mycoplasma contamination by the Chordoma Foundation and were determined to be mycoplasma-negative. Chondrosarcoma cell lines were tested for mycoplasma contamination by the Cancer Cell Line Encyclopedia project and were determined to be mycoplasma-negative, as described by Barretina et al. (DOI: 10.1038/nature11003).

Commonly misidentified lines
(See [ICLAC](#) register)

No commonly misidentified cell lines were used.

Animals and other organisms

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

Laboratory animals

SCID Hairless Outbred (SHO), all female, 6-8 week-old mice were purchased from Charles River Laboratories International Inc. (#474).

Wild animals

The study did not involve wild animals.

Field-collected samples

The study did not involve samples collected from the field.

Human research participants

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

Population characteristics	Relevant information on human research participants is provided in Supplementary Table 6. Covariate-relevant population characteristics of the participants provided include: age, sex, tumor location, classification as metastatic or non-metastatic disease, classification as recurrent or non-recurrent disease, and preoperative treatment.
Recruitment	Patient-derived tissue was obtained from Massachusetts General Hospital (MGH) and the Chordoma Foundation (CF) as described in the Methods. Tissue from MGH was obtained based on availability; tissue from CF was requested to be from a mixture of anatomical locations, non-metastatic, non-recurrent, and untreated.

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
May remain private before publication.

<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE109794>

Files in database submission

.wig and .fastq files are provided.

Genome browser session
(e.g. [UCSC](#))

No longer applicable.

Methodology

Replicates	ChIP-seq analysis was performed on 5 cell lines, 8 patient-derived tumor samples, and 1 patient-derived normal adjacent tissue sample (1 replicate each). ATAC-seq analysis was performed on two cell lines (2 biological replicates for U-CH2 and 1 replicate for MUG-Chor1).
Sequencing depth	All ChIP-seq datasets were sequenced with 75 bp single-end reads. All ATAC-Seq datasets were sequenced with 100 bp paired-end reads. The associated table provides read depths for each dataset in units of millions of reads.
Antibodies	ChIP-seq was performed using an anti-H3K27ac antibody (abcam, #ab4729).
Peak calling parameters	MACS 1.4.2 was applied with a p-value cutoff of 1e-9.
Data quality	Only high confidence peaks (p-value cutoff of 1e-9) were used in subsequent analysis. To ensure quality, super-enhancer analysis was aggregated across all cell lines or all patient-derived samples (see Methods). The rank of the brachyury-associated super-enhancer is also shown across multiple individual samples.
Software	ChIP-seq analysis was performed with Bowtie2 (version 2.2.1); Bamliquidator (version 1.0); MACS (version 1.4.2); ROSE2 (https://github.com/BradnerLab/pipeline); STAR; and Sailfish. ATAC-seq data analysis was performed with Bowtie2 (version 2.2.1); samtools (0.1.19); and MACS (version 1.4.2). All relevant processing scripts can be found at https://github.com/linlabcode/chordoma_code .