

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](#).

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
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☒ | ☐ | 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
<i>Give P values as exact values whenever suitable.</i> |
| ☒ | ☐ | 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	No software was used specifically for data collection; see below for data analysis software.
Data analysis	Software usage and parameters are detailed in methods section of manuscript. Briefly, sequencing reads were aligned with BWA v. 0.7.4, Methylation data analyzed with methylTools v. 0.9.4. ChIP-seq peaks were called and analyzed via Homer v. 4.9, MEME v. 4.7, HTSeq v. 0.6.152 and DESeq2 v. 1.16.1, general data analysis and graphing was performed in Matlab v. 9.1.0.441655, R v. 3.5.3, and IGV v. 2.5.3.

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. 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

Sequencing data that support the findings of this study have been deposited in GEO with the accession code GSE107447.

Raw sequencing data generated through this project may contain identifiable human genetic information, as such it requires IRB approval to access (data has been deposited into a dbGaP dataset connected to the GEO database).

Raw data for the mouse xenograft trial (i.e. measured tumor volumes) are available as supplementary information table 3.

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 analysis of clinical tissue, no statistical methods were used to predetermine sample size; rather, all available SDH-deficient tumor specimens with validated SDH loss and enough material available for analysis were tested. For mouse studies, no specific statistical calculations were performed; rather, sample size was determined based on prior experience with similar PDX trials.
Data exclusions	No data were excluded from analyses.
Replication	For all CRISPR insulator deletions, viral introduction of CRISPR/Cas9+sgRNA vector was repeated three times with separate viral preparations and infections to generate biologically independent replicates. For epigenomic and transcriptomic characterization of clinical tissue (e.g. ChIP-seq, 4C, RNA-seq), multiple clinical specimens were analyzed, but technical replicates could not be performed on individual samples due to limited availability of material. All biological replicates (i.e. tumors of a given driver) were similar at reported sites (i.e. insulator loss/enhancer presence was consistent within driver subgroups).
Randomization	For PDX trial, xenograft-bearing mice were randomized to treatment group at ~200mm ³ tumor volume, with 8 mice per treatment group.
Blinding	For PDX trial, blinding was not possible due to preparation and delivery methods of tested drugs. This is thought to have minimal impact on the studies, as no subjective (e.g. behavioral) criteria were measured.

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
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

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

Antibodies

Antibodies used	CTCF antibody is from Cell Signaling technologies, clone D31H2, catalog number 3418. Antibody was validated by manufacturer for ChIP in human cells, has been previously utilized by the authors (Flavahan et al., Nature 2016), and was validated as part of the ENCODE project via Western Blot for CTCF, as well as motif analysis to confirm enrichment for the known CTCF motif. H3K27ac is a rabbit polyclonal antibody available from Active Motif, catalog number 39133, lot 31814008. Antibody was validated by the manufacturer for ChIP in human cells, and was validated as part of the ENCODE project.
Validation	Both antibodies were validated by manufacturers (including statements on the websites), and by the investigators and colleagues as part of the ENCODE project - including, but not limited to, western blot, immunoprecipitation, and ChIP motif finding and analysis of control cell lines and known peaks/motifs.

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	One cell line, GIST-T1, was used, and was obtained from the commercial vendor Cosmo Biosciences.
Authentication	The cell line was not authenticated via STR testing, however locus sequencing for the KIT gene confirmed the presence of the known and published KIT mutation present in GIST-T1 cells and the parental tumor the cell line was derived from.
Mycoplasma contamination	The cell line was tested for mycoplasma via PCR-based method and confirmed to be mycoplasma-free.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were utilized.

Animals and other organisms

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

Laboratory animals	For the PDX trial, female 8-week old NSG mice were utilized.
Wild animals	No wild animals were utilized in this study.
Field-collected samples	No field-collected samples were utilized in this study.
Ethics oversight	All animal experiments and study protocols were approved by the Dana Farber Cancer Institute Institutional Animal Care and Usage Committee (IACUC), and this is noted in the manuscript.

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 study involved the collection of deidentified and anonymized tumor material from patients at either Brigham and Women's Hospital, The Dana Farber Cancer Institute, or Massachusetts General Hospital. As such, no information about the patients, other than disease pathology, is known.
Recruitment	Patient tissue was obtained from tissue banks at either MGH or DFCI. Standard of care for GIST includes surgical resection of the tumor bulk - following treatment, excess surgical material from consenting patients was deposited into these banks. As the collection of these tissues occurs as part of the normal disease treatment, it is unlikely there is significant self-selection bias.
Ethics oversight	The study protocol was approved by the Massachusetts General Hospital IRB and the Dana Farber Cancer Institute IRB.

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
May remain private before publication.

GEO database accession GSE107447
<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE107447>

Files in database submission

See included file, ChIP-seq_2017-12-16052D_file_info.xlsx

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

no longer applicable

Methodology

Replicates

For epigenomic and transcriptomic characterization of clinical tissue (e.g. ChIP-seq, 4C, RNA-seq), multiple clinical specimens were analyzed, but technical replicates could not be performed on individual samples due to limited availability of material. All clinical samples within a driver subgroup (SDH-intact vs. SDH-deficient) were highly similar at tested locations (e.g. KIT/FGF insulator loss and superenhancer presence). See extended data figure 2 for more info.

Sequencing depth	See included file, <code>ChIP-seq_2017-12-16052D_file_info.xlsx</code> , which includes sequencing depth for each experiment.
Antibodies	<p>CTCF antibody is from Cell Signaling technologies, clone D31H2, catalog number 3418. Antibody was validated by manufacturer for ChIP in human cells, has been previously utilized by the authors (Flavahan et al., Nature 2016), and was validated as part of the ENCODE project via Western Blot for CTCF, as well as motif analysis to confirm enrichment for the known CTCF motif.</p> <p>H3K27ac is a rabbit polyclonal antibody available from Active Motif, catalog number 39133, lot 31814008. Antibody was validated by the manufacturer for ChIP in human cells, and was validated as part of the ENCODE project.</p>
Peak calling parameters	Peaks were called with HOMER 4.9 against input controls. To call all H3K27ac peaks, we used 'histone' settings. To call super-enhancers, we used 'super' settings and no local filtering. CTCF peaks were called with 'factor' settings.
Data quality	All reported peaks are detected with FDR < 0.1% and fold change > 4
Software	Homer 4.9, DESeq2 1.16.1, FIMO/MEME 4.7, HTSeq 0.6.1, featureCounts 1.6.2, and CNVnator 0.3