nature research

Corresponding author(s): William R. Sellers

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	All software used in the analysis was either published or is in press, and can be provided upon request. All software is described in the Supplemental Methods section of the paper.
Data analysis	We used custom R codes for most of the analysis. Source codes are available upon 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

All processed data used in this manuscript will become available on the CCLE portal (www.broadinstitute.org/CCLE) and companying R package. The raw data will be deposited to public repositories.

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	We selected the cell lines based on commercial availability and unmet medical need as was previously described in the original CCLE publication. For the in-vivo Ponatinib/pSHP2 efficacy experiment, for each model, we used 7 mice in treatment and 7 mice in control groups, to achieve statistical significance. This number was chosen based on prior experience.
Data exclusions	We used various QC metrics for RPPA, DNA methylation, microRNA as described in the supplemental methods to exclude samples with low quality.
Replication	We included 5 biological replicates and 30 technical replicates in RPPA batches as described in the supplemental methods. 28/30 technical replicates had high concordance between the two batches. In Ponatinib/pSHP2 in-vitro validation experiment, we validated RPPA pSHP2 measurements by Western blot as presented in the manuscript and described in supplemental methods.
Randomization	In Ponatinib/pSHP2 in-vivo experiment, mice were randomized to treatment/control groups.
Blinding	Genetic, transcriptomic, RPPA, DNA methylation and global chromatin profiling data collection were performed without the investigators' knowledge of cell lines identities. Investigators were not blind to cell line identities during 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	A complete list of antibodies used is given in Supplemental Table 14.
Validation	Only antibodies with a Pearson correlation coefficient between RPPA and western blotting of greater than 0.7 were used. Antibodies with a single or dominant band on western blotting were further assessed by direct comparison to RPPA using cell lines with differential protein expression or modulated with ligands/inhibitors or siRNA for phospho- or structural proteins, respectively.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	A list of the CCLE cell lines and vendors is available on the CCLE portal (www.broadinstitute.org/CCLE)
Authentication	Cell line authentication was performed using SNP-based DNA fingerprinting.
Mycoplasma contamination	Cell lines were tested for mycoplasma contamination.

Commonly misidentified lines
(See ICLAC register)

We have compared germline SNPs across CCLE cell lines and between CCLE and Sanger cell lines and annotated the cell lines that share high SNP identity including KPL-1 / MCF7 which is listed in ICLAC

Animals and other organisms

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

Laboratory animals	Each primagraft was xenotransplanted into twenty female 7-week-old NOD scid gamma (NSG) mice from Jackson Laboratory (Bar Harbor, ME).
Wild animals	N/A
Field-collected samples	N/A