nature research

Corresponding author(s): Uri Ben-David

Last updated by author(s): Nov 9, 2020

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 *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	All software used in data collection were published , and are described in the Methods section of the paper. No commercial SW were used. Low-pass whole-genome sequencing alignment was performed using the BWQ (0.7.12) backtrack algorithm. Single cell sequencing alignment was performed using Bowtie2 (v2.2.4, duplicate reads marked with BamUtil (v1.0.3).
Data analysis	All SW used for data analysis are commercially or publicly available, and are described in the Methods section: Genomic data sets were downloaded from the Depandency Map portal (www.depmap.org/portal/). Experimantal confounders in gene expression data were corrected using ComBat (v3). Functional annotation enrichment analysis was performed using DAVID v6.8 and GSEA v4.0.0. Microscopy-based analysis of cell proliferation was performed using IncuCyte S3 Live Cell Analysis System (Essen Bioscience). Cell masking in representative images was performed using Ilastik image analysis SW. Fluorescence signal quantification was performed using the SlideBook SW. Copy number data were detected using the HMMCopy algorithm. Aneuploidy analysis of scDNAseq data was performed using AneuFinder (v1.14.0). Flow cytometry data were analyzed using the Kaluza Analysis SW. Statistical analysis and plotting were done in GraphPad Prism v8.4.3 or in Office 2016 Excel. The following packages were used: scikit-learn 'RandomForestRegressor', R packages: 'stats', 'drc', 'limma-trend', 'hclust', 'fgsea'.

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

All datasets are available within the article, its Supplementary Information, or from the corresponding authors upon request. Cell line aneuploidy profiles and scores are available at the DepMap portal (www.depmap.org/portal/). The analyzed CCLE genomic data is available in https://doi.org/10.6084/m9.figshare.11384241.v2. LP-WGS data have been deposited to SRA with BioProject accession number PRJNA672256.

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

Life sciences study design

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

Sample size	All available human cancer cell lines were used for the comparisons of the Achilles-shRNA, DRIVE-shRNA, Achilles-CRISPR, GDSC and CTD2 data sets. For validation experiments, we selected 5 highly-aneuploid and 5 near-diploid cell lines for practical reasons. For isogenic cell line systems (HCT/HPT, RPE1/RPT, RPE1/aneuploid clones), all existing clones were used.
Data exclusions	No data were excluded from the analyses.
Replication	All experiments were performed at least 3 times. The experimental findings were reliably reproduced, and all attempts were included in the presentation unless technical failures prevented the completion of the experiment.
Randomization	No randomization was done, as all available cell lines were used (therefore, randomization was not required).
Blinding	Genetic and transcriptional profiling were performed without the investigators' knowledge of each sample identity. Investigators were not blind to sample identity during the in vitro experiments because cell lines required different culture conditions.

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	The following primary antibodies were used: anti-Kif18A rabbit (1:500), affinity-purified polyclonal antibody raised against an N-terminal GST-tagged fragment (Kif18AbN), a gift from Dr. Thomas Mayer, University of Konstanz, Germany; anti-Kif18A rabbit (1:5,000), Bethyl Laboratories (catalog no. A301-080A); anti-GAPDH goat (1:1,000), Abcam (catalog no. ab9483); anti-α-Tubulin mouse (1:2,000), Sigma (catalog no. T6199).
Validation	Antibodies were selected based on their use in the literature in human cancer cell lines, and previous experience of the investigators. Full antibody information is provided in the Methods section of the paper. Positive and negative controls were used in all experiments including antibodies.

Product citations (n):
Kif18a: n=7, https://www.bethyl.com/product/A301-080A/KIF18A+Antibody
GAPDH: n=112, https://www.abcam.com/gapdh-antibody-loading-control-ab9483.html
α-Tubulin: n=1582, https://www.sigmaaldrich.com/catalog/product/sigma/t6199

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Established commonly-used human cancer cell lines were used in this study: MDAMB468, A101D, EN, VMCUB1, CAL51, SW48, SH10TC, NCIH1693, MHHNB11 and PANC0813 were purchased from the Cancer Cell Line Encyclopedia. The HCT116/HPT and RPE1/RPT lines were genetically manipulated by the Storchova lab, and the RPE1 line was also manipulated by the Santaguida lab.
Authentication	Cell line authentication was performed using SNP-based DNA fingerprinting or through copy number profiling of the cell lines.
Mycoplasma contamination	All cell lines tested negative to mycoplasma contamination using a Lonza kit.
Commonly misidentified lines (See ICLAC register)	Cell lines are not in the list of misidentified cell lines.

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	For cell cycle and cell death analyses, cells were trypsinized and incubated in cold PBS supplemented with 5% fetal calf serum (Sigma-Aldrich; PBS-FACS). DNA was stained either by propidium iodide (PI) or by Hoechst. For PI staining, cells were fixed in cold 70% ethanol, added dropwise while vortexing, and incubated on ice for 30 minutes. Cells were centrifuged and pellets were washed twice with PBS-FACS. 50 µl RNase A solution (100 µg/ml in PBS) was added to the pellet, followed by staining with 400 µl PI solution (50 µg/ml in PBS) per million cells. Cells were incubated for 10' at 25oC. For Hoechst staining, pellets were incubated in the dark with 10 mg/ml Hoechst 33358 for 15' at 4oC.
Instrument	Data acquisition was performed using the CytoFLEX flow cytometer (Beckman Coulter) or the BD FacsJAZZ cell sorter (BD Biosciences).
Software	Data analysis was performed using the Kaluza Analysis software 2.1 (Beckman Coulter).
Cell population abundance	No sorting was performed.
Gating strategy	An SSC-A/FSC-A gate was set in order to exclude cell debris, and an FSC-A/FSC-H gate was then set in order to exclude doublets. Cell cycle phases were determined manually using linear gating based on the 2N and 4N peaks of the histogram. Cell death was assessed by quantifying the fraction of cells in the subG1 population, and mitotic arrest was assessed by quantifying the fraction of cells in the G2/M population.

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