nature portfolio

Corresponding author(s): Kaja Kostyrko
Alejandro Sweet-Cordero

Last updated by author(s): Jun 13, 2022

Reporting Summary

Nature Portfolio 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 Portfolio 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	Cell proliferation data was collected using Incucyte S3 software (v2019a, Essen BioScience). Images were taken on DMi8 fluorescence microscope (Leica). Flow cytometry data was collected using Accuri C6 (BD Biosciences).
Data analysis	Microscope images were analyzed with open source software: CellProfiler (version 3.1.9) and ImageJ (version 1.53c). RNA-seq FASTQ data were preprocessed using HTStream (UC Davis, github.com/s4hts/HTStream) and aligned to the GRCh38 reference genome using STAR (version 2.5.3b). Gene-level counts were obtained using the human GENCODE v37 gene annotation (gencodegenes.org/human). Differential gene expression was calculated using R packages: edgeR (version 3.34.1) and limma (version 3.48.3). Incucyte S3 software (v2019a, Essen BioScience) was used for automated microscopy analysis. Putative KRAS interacting genes included in shRNA libraries were derived using previously published algorithms described in details in Supplementary Methods. Methylation analysis. The Minfi package available in R. IDAT files was taken as input, preprocessed, and normalized. Quality control/preprocessing included removal of low-quality probes and probes on problematic positions such as probes that were on multi-mapped, SNPs, or cross-reactive sites. Following initial filtering and quality check, data were normalized using the quantile normalization method. Batch effects and cell-type heterogeneity were assessed. Both Beta and M-values were then generated. All downstream statistical analyses were generated with M-values while plots were displayed as Beta-values. Additional annotations such as promoter region and island prediction were all acquired with annotations data packages from Bioconductor. Unsupervised t-SNE was performed on the top 10000 most variable CpGs (on the basis of median absolute deviation). Single differential methylation probe (DMP) analysis was performed using the linear

modeling provided by the limma package and differential variable probe (DVP) was performed with the DiffVar package. The DMRcate packaged was then used to determine differential methylated regions (DMR). Pathway analysis. To test for over-representation we used the function goana from the limma package with custom pathway sets downloaded from MSigDB. Genes that were associated with differentially methylated probes were used as input. However, since multiple probes can be associated with the same gene, we removed any genes on probes with opposing directions followed by the removal of duplicates, resulting in a unique list of genes.

Correlation analysis. Spearman correlation and p-value follow with Benjamini-Hochberg false discovery rates were computed for UHRF1 against every gene in the tumor suppressor list.

Survival analysis. Standardized clinical survival end-points for TCGA (LUAD dataset) consisting of DSS, OSS and OS were downloaded from a dataset previously curated from Liu et al. This set contains survival end-points for overall survival (OS), disease-specific survival (DSS), disease-free interval (DFI) and progression-free interval (PFI). RNAseq expression data for LUAD was downloaded from Firehose, Broad institute and subsequently converted to log2 cpm with the edgeR package from raw read counts. Somatic mutation for the LUAD data set was acquired from the UCSC Xena public repository. Samples with non-silent KRAS mutation(s) were assigned to the KRAS-positive (KRAS+) group and all others as KRAS wildtype (KRASwt) group. Samples with KRAS silent mutations were excluded from the analysis. Kaplan-Meier and multivariate Cox hazard regression were completed using the survminer and survival packages in R. For the multivariate Cox regression analysis, we adjusted for age of diagnosis, gender, and cancer stage. We grouped samples as (Normal vs High) expression based on the quantile of the UHRF1 gene expression: normal is < 75th percentile and high is > 75th percentile.

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 Portfolio 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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The data has been deposited to NCBI GEO under a single Super Series: GSE198450. A reviewer read-only token has been generated: kvaruuswhzctjyn and can be used to access the datasets.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

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

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	For all RNAseq and methylation array experiments, experiments were performed in 2 or 3 replicates. For in vivo and in vitro experiments, sample size is indicated in the figure legend or in the methods section corresponding to each experiment. The sample size was determined based on previous experience for each experiment to detect specific effects and was not predetermined with any statistical methods.
Data exclusions	No data was excluded

Replication	Each in vivo experiment presented in the paper was repeated in multiple mice. In vitro experiments were replicated as indicted in the figure legends or Methods section.
Randomization	For all in vivo experiments mice were randomly allocated to experimental groups.
Blinding	N/A

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
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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

Antibodies

Antibodies used	UHRF1 (sc-373750, Santa Cruz Biotechnology), KRAS (Op24, Millipore), Cyclin D1 (ab134175, Abcam), actin (Sigma, A5316), Cyclin B1 (ab32053, Abcam), GAPDH (no. 9485, Abcam), Ki67 (ab15580, Abcam), biotinylated secondary horse anti-mouse IgG biotinylated antibody (Vector Laboratories, BA-2000), goat anti-rabbit AF488 secondary antibody (ThermoFisher Scientific, A-11008, 1:200), goat anti-mouse AF647 secondary antibody 1:200 (ThermoFisher Scientific, A-21235).
Validation	UHRF1, KRAS and Cyclin D1 antibodies were validated using western blot on cells treated with UHRF1, KRAS and CyclinD1 siRNAs, respectively (Fig. 5h, Extended Data Fig. 3).

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	All human cell lines were purchased from ATCC. LKR10 cells were derived from the KRAS G12D mouse model (DuPage et al, 2011).
Authentication	Human cell lines were positively authenticated using CellCheck 9 - human (9 Marker STR Profile and Inter-species Contamination Test) from Idexx Bioanalytics (https://www.idexxbioanalytics.com)
Mycoplasma contamination	All cell lines were tested for mycoplasma contamination by PCR (IDEXX BioResearch). All cell lines were negative for Mycoplasma sp.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	8-10 week old NSG mice were used for xenograft experiments. C57BL/6 mice were used for genetically engineered mouse experiments.
Wild animals	This study did not involve wild animals
Reporting on sex	Sex was not considered in this study
Field-collected samples	This study did not involve field-collected samples
Ethics oversight	All the procedures involving mice were are approved by the Institutional Animal Care and Use Committee (IACUC) at UCSF (protocol #AN15761).

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

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	Sample preparation for flow cytometry is described in the Methods section under "Apoptosis and cell cycle analysis"
Instrument	Flow cytometry data was acquired on Accuri C6 (BD Biosciences)
Software	Flow cytometry data was analyzed using FlowJo (v10).
Cell population abundance	Cell sorting was not performed in this study.
Gating strategy	FSC and SSC were used to gate out single cell populations. Negative and positive staining populations were determined based on single-stained controls.

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