nature portfolio

Corresponding author(s): Alex K. Shalek

Last updated by author(s): 08/21/2024

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	Single-cell RNA sequencing data were generated using the Illumina NextSeq2000 and Illumina NovaSeqX.
Data analysis	We used publicly available software in R(4.3.0; Seurat (4.3.0), Azimuth (0.4.6)) and Python (3.10.14; Scanpy (1.9.8), decoupler (1.5.1), KneeLocator (function is part of "kneed" package version 0.8.5), sklearn (1.4.2)) and other open source tools (CellRanger (v6.0.1), starsolo (2.0.0), bcl2fastq (v2.2), mkfastq (part of CellRanger v6.0.1), CellphoneDB (v3.1), cNMF (custom code)) to analyze sequencing data and further analyses. Kaplan-Meier curves were generated and analyzed using the package lifelines (V0.27.7) from python. CellProfiler (v.1.2) was used to process CellPainting data. In-house scripts and pipelines are available upon request. Code is available on Github at https://github.com/ShalekLab/compressed_screening.

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

Genomic Data Commons (GDC) TCGA Pancreatic Cancer (PAAD) gene expression RNAseq and phenotype (survival) data was retrieved from GDC Data Portal: (https://xenabrowser.net/datapages/?cohort=GDC%20TCGA%20Pancreatic%20Cancer%20(PAAD)&removeHub=https%3A%2F%2Fxena.treehouse.gi.ucsc.edu%3A443). Primary PDAC genomic and transcriptomic data (TCGA, Pancreatic Ductal Adenocarcinoma) was retrieved from (https://portal.gdc.cancer.gov/projects/). A Summarized and deidentified single-cell RNA-sequencing PDAC data is available via the Broad Institute's Single Cell Portal: https://singlecell.broadinstitute.org/single_cell/study/SCP2621/compressed-phenotypic-scree[...]discovery-pdac-compressed-screen-and-validation-screen.Summarized and deidentified single-cell RNA-sequencing PBMC data is available via the Broad Institute's Single Cell Portal: https://singlecell.broadinstitute.org/single_cell/study/SCP2622/compressed-phenotypic-screening-empowers-scalable-biological-discovery-pbmc-screens. Raw sequencing data for the PDAC screen can be accessed on GEO with accession number GSE267243 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE267243). Raw sequencing data for the PBMC screen will be available upon request. Links and accession numbers for the publicly available data analyzed in this study are listed in the key resources table. Cell Painting images can be accessed through BioImage Archive (BIA) with accession number S-BIAD1319 (https://www.ebi.ac.uk/biostudies/bioimages/studies/S-BIAD1319). Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Human research participants

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

Reporting on sex and gender	Pancreatic cancer samples used in this study were previously reported in (Raghavan et al., Cell, 2021). For this study, sex and gender were not explicitly used for analyses. PBMCs were obtained from an anonymous donor of unknown gender.
Population characteristics	Pancreatic cancer organoid samples used in this study were previously reported in (Raghavan et al., Cell, 2021) and derived from one patient. The PANFR0562 patient-derived organoid model was generated from a liver metastatic biopsy of a 76-year-old female (unknown ethnicity) with PDAC. 5 cores were collected and 2 were allocated for scRNA-seq and organoids. No prior treatment was administered for primary or metastatic disease. Patient survived for 111 days from time of diagnosis.
Recruitment	As described in (Raghavan et al., Cell, 2021), Patients with pancreatic cancer were screened for upcoming biopsies. If they were deemed appropriate candidates from a clinical and safety standpoint, they were consented to research tissue collection protocols. Samples were collected and analyzed based on tissue availability rather than by recruitment of more specific patient subsets.
Ethics oversight	Patients at the Dana-Farber Cancer Institute gave their consent to Dana-Farber/Harvard Cancer Center Institutional Review Board (IRB)-approved protocols. PBMC samples were derived from MGB under IRB-approved protocols.

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	The sample size was determined from prior experience (Raghavan et al., Cell, 2021) and pilot experiments, and at least three biological replicates were used for each experiment except for the single-cell RNA-sequencing. Major considerations for sample size selection included the assurance of obtaining the amount of output sample required for robust and confident analyses as sample loss can be up to 50-60% during sample processing. Refer to Methods section for more detailed discussion of sample size determination for each experiment.
Data exclusions	For the Cell Painting experiment, wells with fewer than 50 segmented cells were excluded from analyses for lack of confidence on signal gathered. Otherwise, remaining data analysis followed standard QC protocols.
Replication	All experiments were performed in at least three independent biological replicates, all of which were successful. Biological replicates were not performed in single-cell RNA sequencing experiments, since a single sample contained single cells derived from thousands of organoids

and PBMCs of identical condition. Cell line and organoid screens were run in replicate as annotated in the manuscript. Across all wells in each screen, each compound tested appeared at least 3 times.

Randomization

For a given model, the allocation of cells into wells and replicates was random. However, we only selected one PDAC organoid model (derived from one patient) and PBMCs were derived from one healthy donor's blood sample; so allocation of samples was not random in each experiment. However, for all compressed screening experiments, distinct drug pools were randomly generated. For a more detailed explanation on parameters and criteria of our randomized perturbation assignment, refer to "Compressed screen design principles" in the Methods section.

Blinding

The samples were not blinded as the study used one model in each instance (one patient derived PDAC model and one donor derived PBMC model) and relied on objective, quantifiable data (genetic sequencing)

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

MitoTracker Deep Red (ThermoFisher #M22425), Wheat Germ Agglutinin – AlexaFluor 594, (ThermoFisher #W11262), ConcanavalinA-AlexaFluor 488 (ThermoFisher, #C11252), Phalloidin-AlexaFluor 594 (ThermoFisher #A12381), SYTO 14 green fluorescent nucleic acid stain (ThermoFisher #S7576), Hoechst 33342 (ThermoFisher #H3570), BD AbSeq Ab-Oligos (BD Biosciences #633781), BioLegend's TotalSeq-B with 10x Feature Barcoding Technology antibodies (BioLegend 394631, 394633, 394635, 394637, 394639, 394641, 394643, 394645).

Validation

The following antibodies were validated some previous studies such as (Bray et al., Na.t Protoc., 2016) and by their manufacturer: MitoTracker Deep Red (ThermoFisher #M22425), Wheat Germ Agglutinin – AlexaFluor 594, (ThermoFisher #W11262), ConcanavalinA-AlexaFluor 488 (ThermoFisher, #C11252), Phalloidin-AlexaFluor 594 (ThermoFisher #A12381), SYTO 14 green fluorescent nucleic acid stain (ThermoFisher #S7576), Hoechst 33342 (ThermoFisher #H3570). BD AbSeq Ab-Oligos (BD Biosciences #633781), were validated in some previous studies such as (Malkani et al.Biology, 2020) and by their manufacturer https://www.bdbiosciences.com/en-us/products/reagents/single-cell-multiomics-reagents/bd-single-cell-multiplexing-kits/hu-single-cell-sample-multiplexing-kit.633781. The following antibodies were validated some previous studies such as (Mylka et al., Genome Biol., 2022) and by their manufacturer: TotalSeq-B with 10x Feature Barcoding Technology antibodies (BioLegend 394631, 394633, 394635, 394637, 394639, 394641, 394643, 394645).

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)

ATCC U2OS cell line (HTB-96). Patient-derived organoid models from biopsy specimens.

Authentication

Cell lines were used as supplied from ATCC. For PDAC organoids, we compared inferred CNV profiles of organoid models with those of malignant cells from matched fresh tumor specimens. The latter inferred CNV profiles were compared to and were concordant with copy-number estimates from DNA sequencing of patient tumor samples.

Mycoplasma contamination

Cell lines tested negative for mycoplasma. Mycoplasma testing is routinely performed in batches for established organoid cultures, and was negative for organoids used in this study.

Commonly misidentified lines
(See ICLAC register)

No commonly misidentified cell lines were used in this study.