nature portfolio

Corresponding author(s): Dhiraj Kumar and Benjamin Izar

Last updated by author(s): Nov 22, 2023

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	University of California Santa Cruz (UCSC) Xena webserver (https://xenabrowser.net/) was used to analyze survival data according to MALAT1 expression in the breast cancer TCGA database. MALAT1 expression and correlation in different cancer databases or in primary vs metastasis tumor samples were obtained from Oncomine (https://www.oncomine.org/) or the lncAR (https://lncar.renlab.org/explorer) webserver. Kaplan-Meier plotter (https://www.kmplot.com/) was used for the survival analysis with MALAT1 and SERPINB6, expression either at transcript or protein level. Expression of SERPINB1, B6, and B9 in breast cancer samples was obtained from gepia2 (source http://gepia2.cancer-pku.cn/#analysis). The ChIRP-seq data were further analyzed using GREAT (Genomic Regions Enrichment of Annotations Tool) version 4.0.4 and visualized using the UCSC genome browser.
Data analysis	GraphPad 8 software was used for statistical analyses. FlowJo 10 was used for flow cytometry data analysis. Living Image 4.4 software was used for bioluminescence Imaging analysis. 10X Genomics Cell Ranger 6.1.1, R 4.2.1, RStudio, and Seurat v4 were used for single cell RNA-seq analyses.

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

Total-RNA-seq, single cell RNA-seq and ChIRP-seq data that support the findings of this study have been deposited in the Gene Expression Omnibus (GEO) under accession codes. GSE183562, GSE183943, and GSE183596. All data are available in the main text or the supplementary materials. Constructs and genetically engineered cell lines will be made available upon a material transfer agreement.

Human research participants

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

Reporting on sex and gender	Breast cancer primary and metastatic samples were collected from both female and male patients.
Population characteristics	Patients age 18 of over were included.
Recruitment	Samples collected from MDACC Tissue Center Access to tissue is permitted from following groups. Group 1: patients with baseline diagnostic biopsy prior to systemic therapy and surgical specimen in neoadjuvant HER2+ targeted therapy Group 2: patients with baseline diagnostic biopsy and surgery in adjuvant HER2+ targeted therapy Group 3: patients with distant recurrence, metastatic site diagnostic biopsy (sequential) in neoadjuvant/adjuvant treated HER2+ therapy Group 4: De-novo stage IV HER2 + breast cancer (concurrent primary biopsy and metastatic biopsy. Access to tissue is permitted from "MD Anderson Cancer Center Tissue and Data Donation Project" PA14-0241/ LAB03-0320.
Ethics oversight	Consent waiver was granted by the MD Anderson IRB for patients who had expired and informed consent was obtain from other patients through "front-door" tissue consent protocol, MDACC # PA 2020-0732.

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	No statistical methods were used to pre-determine sample sizes but our sample sizes are similar to those reported in previous publications. PMID: 22901808
Data exclusions	No data were excluded
Replication	For each experiments the number of biological independent animal/sample is reported in the methods. In vitro studies are represented at least 3 independent reproducible studies,and at least 3 independent experiments to enable statistical analyses. Animals studies represent at least 4 independent mice. All attempts at replication was successful.
Randomization	In immunotherapy mice experiments, groups were randomized before starting treatment. Randomization in other animal studies were not possible due to lack of treatment arms and the conditions were used in this present studies.
Blinding	The investigators were not blinded to allocation during the experimental procedures and the assessment of the outcomes. The investigators were not blinded for the allocation of groups during experiments or during in vitro or in vivo experiments as the samples or animals were marked and the same investigator set-up and performed the experiment. Fully blinded animal experiments were not possible due to personnel availability to accommodate such situations.

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

SRSF1 ThermoFisher Scientific Cat.#32-4500 (1:1000 for WB, 1:100 for IF)
SRSF2 ThermoFisher Scientific Cat.#PA5-12402 (1:1000 for WB)
SRSF4 ThermoFisher Scientific Cat.#PA5-36366 (1:1000 for WB)
Cas9 Cell Signaling Cat.#65832S (1:2000 for WB)
NR2F1 Cell Signaling Cat.#6364S (1:1000 for WB, 1:100 for IF)
Dec-2 Novus Biologicals Cat.#NBP1-19613 (1:1000 for WB, 1:50 for IF)
p-p38 Cell Signaling Cat.#4511S (1:1000 for WB, 1:100 for IF)
p-ERK1/2 Cell Signaling Cat.#4377S (1:1000 for WB, 1:100 for IF)
p-ERK1/2 Cell Signaling Cat.#4370T (1:100 for IHC)
Ki67 Cell Signaling Cat.#9129S (1:200 for IHC)
GFP Aves Labs Inc Cat.#GFP-1020 (1:100 for IF)
Cl. Cap3 Cell Signaling Cat.#9664S (1:100 for IHC)
MPO R&D Systems Cat.#AF3667 (1:200 for IHC)
CD4 Cell Signaling Cat.#25229 (1:100 for IHC)
CD8 Cell Signaling Cat.#98941 (1:100 for IHC)
CD31 Cell Signaling Cat.# 77699 (1:100 for IHC)
CD4 Abcam Cat.#ab183685 (1:100 for IF)
CD8 Cell Signaling Cat.#98941 (1:100 for IF)
GsdmD Abcam Cat.#ab219800 (1:2000 for WB)
GsdmE Abcam Cat.#ab215191 (1:2000 for WB)
Cathepsin G Abcam Cat.#ab197354 (1:1000 for WB)
p-Yap Cell Signaling Cat.#4911S (1:2000 for WB)
p-Taz Cell Signaling Cat.#59971S (1:2000 for WB)
CTGF Santa Cruz Cat.#sc-365970 (1:2000 for WB)
Cyr61 Cell Signaling Cat.#39382S (1:2000 for WB)
Yap Cell Signaling Cat.#14074S (1:100 for IF)
Porcn Abcam Cat.#ab105543 (1:2000 for WB)
Wnt3 Abcam Cat.#ab32249 (1:1000 for WB, 1:50 for IF)
Wnt6 Abcam Cat.#ab154144 (1:1000 for WB)
β-Cat Cell Signaling Cat.#8480S (1:1000 for WB)
HA ThermoFisher Scientific Cat.#26183 (1:2000 for WB)
Flag Sigma-Aldrich Cat.#F1804 (1:2000 for WB)
RhoGDI Santa Cruz Cat.#sc-360 (1:5000 for WB)
Actin Santa Cruz Cat.#sc-47778 (1:5000 for WB)
Granzyme A-PE-Cyanine7 eBioscience™ Cat.#25-5831-82 (1:100 for FC)
CD45-APC/Cy7 Biolegend Cat.#103116 (1:50 for FC)
CD3-Brilliant Violet 510 Biolegend Cat.#100234 (1:50 for FC)
CD11b-Brilliant Violet 421 Biolegend Cat.#101251 (1:100 for FC)
CD4-PerCP/Cyanine5.5 Biolegend Cat.#100434 (1:200 for FC)
CD8-Brilliant Violet 650 Biolegend Cat.#100742 (1:200 for FC)
Ly6G-PE/Cy7 Biolegend Cat.#127618 (1:100 for FC)
Ly6C-APC Biolegend Cat.#128018 (1:100 for FC)
F4/80-PE Biolegend Cat.#123110 (1:50 for FC)
CD49b-PE/Dazzle Biolegend Cat.#108924 (1:100 for FC)
Perforin-APC Biolegend Cat.#154304 (1:100 for FC)
Granzyme B-PE/Dazzle™ 594 Biolegend Cat.#372215 (1:200 for FC)
β2-microglobulin-APC Biolegend Cat.#154505 (1:100 for FC)
H-2Kd-APC Biolegend Cat.#116619 (1:100 for FC)
H-2Dd-PE Biolegend Cat.#110607 (1:100 for FC)
CD8 Alexa Fluor 647 MBL Int Cop Cat.#K0227-A64 (1:100 for FC)
Goat anti-Chicken IgY (H+L) Secondary Antibody, Alexa Fluor™ 555 Invitrogen Cat.#A-21437 (1:200 for FC)
Donkey anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 647 Invitrogen Cat.#A-31573 (1:200 for FC)

Validation

All antibodies used were commercially available and validated by the manufacturers and have been extensively used in previous publications.

Company validation statement:
Abcam: 'How we validate Antibodies: a closer look at stander we use in the validation of antibodies. Here you will find mare information on what is carried out during our application-specific validation processes' ... (see https:// www.abcam.com/primary-antibodies/how-we-validate-our-antibodies).

Invitrogen (ThermoScientific, eBioSciences): 'Invitrogen antibodies are currently undergoing a rigorous two-part testing approach Part 1-Target specificity verification - This helps ensure the antibody will bind to the correct target. Our antibodies are being tested using at least one of the following methods to ensure proper functionality in researcher's experiments' ... (see https:// www.thermofisher.com/uk/en/home/life-science/antibodies/invitrogen-antibody-validation.html).

Cell Signaling: 'To ensure product performance, we validate all of our antibodies, in-house, in multiple research applications' (see https ://www.cellsignal.ca.uk/about-us/our-approaeh-process/est -antibody-performance-guarantee).

Santa Cruz: ' Primary antibodies directed to mammalian target proteins have been characterized for reactivity against mouse, rat and human proteins' (see https://www.scbt.com/p/endomucin-antibody-v-7c7?gclid=CjOKCQiAiJSeBhCCARIsAHnAzT8RJNnzQgHgQKPuiyzeqBbJPcvyGTMmxijlmVp7Qvic7Xu8VGRWtOwaAnNXEALw_wcB)

BioLegend: 'All of our products undergo industry-leading rigorous quality control (QC) testing to ensure the highest level of performance and reproducible results. Each lot is com pared to an internally established "gold standard" to maintain lot-to-lot consistency. We also conduct wide-scale stability studies to guarantee an accurate shelf-life for our products. Additionally, we test the majority of our products on endogenous cells rather than transfected or immortal cells that may overexpress the analyte ... ' (see https://www.biolegend.com/en-us/quality/quality-control).

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	4T1 ATCC CRL-2539 LLC1 ATCC CRL-1642 HEK-293 ATCC CRL-1573 67NR, 168FARN, and 4T07 generously provided by Dr. Fred R. Miller (Wayne State University, Detroit, MI) CT26 From Dr. Kunal Rai (Original source ATCC) B16F10 ATCC CRL-6475 D2A1-m From Dr. R. Weinberg, MIT (Derived from D2A1 which was obtained from Dr. Fred R. Miller) D2A1-d From Dr. R. Weinberg, MIT (Derived from D2A1 which was obtained from Dr. Fred R. Miller) 293FT Thermo Fisher R70007
Authentication	We confirmed the authenticity of all cells by analyzing short tandem repeat (STR), whereas cell lines were confirmed through genotyping.
Mycoplasma contamination	Cell lines were tested for mycoplasm negative. Every month, a routine examination of cell lines in culture for any possible mycoplasma contamination was performed using PCR method.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified lines were used in the study.

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	Six- to eight-week-old female mice were used for all experiments. Female Balb/c, C57BL/6, Nude, and FVB-Tg(MMTV-Erbb2)NK1Mul/J mice were purchased from The Jackson Laboratory. Animals were housed at 20-22°(with 12h:12h light:dark cycles at 50-60% humidity. Mice tumor volume 2000 mm3 was considered as a maximal tumor size by our Institutional Animal Care and Use Committee of Columbia University and MD Anderson Cancer Center. In this study maximal tumor size was not exceeded.
Wild animals	No wild animals were used in the study.
Reporting on sex	For breast cancer studies we purchased female mice from The Jackson Laboratory. For melanoma, lung and colon cancer we also used female mice to keep the sex consistency without affecting the outcomes.
Field-collected samples	No field-collected samples were used in the study.
Ethics oversight	Mouse studies were conducted in accordance with protocols approved by the Institutional Animal Care and Use Committee of Columbia University and MD Anderson Cancer Center. Our research complies with all relevant ethical regulations of the Columbia University and MD Anderson Cancer Center. The animal experiments that were conducted as part of this research were completed in accordance with the guidelines provided by the Columbia University and MD Anderson Cancer Center Animal Care and Use Committee. Mice were monitored weekly for signs of ill health or overt tumors; once mice displayed signs of hydrocephalus (domed head) or neurological duress, they were humanely killed as defined by IACUC. IACUC guidelines recommend limiting solid tumors to 10% of the host's body weight. This criteria was not exceeded in this study. Animal studies were approved by the internal ethics committee and the local governmental committee.

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	For leukocytes analysis, lungs were digested in collagenase D (Roche Life Science) and single cells suspensions were prepared. Cells were spun down at 400 g for 5 min and resuspended in 2 ml fluorescence-activated cell sorting (FACS) buffer (D-PBS without calcium and magnesium containing 0.5% FBS, 2 mM EDTA, and 0.09% sodium azide). The cell suspension was passed through a 30 µm cell strainer and used for flow cytometric analysis. For systemic immune profiling, blood was collected by either tail vein or retro-orbital puncture into 1.0 ml of PBS containing 2 mM EDTA. Red blood cells were lysed with lysing buffer (Miltenyi Biotec) and cells were washed twice with PBS. For blood or tissue flow cytometry analysis, cells were blocked with rat monoclonal anti-CD16/CD32 (Fc block antibody, BD Biosciences) in PBS for 30 min at 4°C. Cells were then stained with antibodies in the dark. For cell line study single cells suspension were prepared using Accutase, stained with antibodies and resuspended in the FSCS buffer for analysis or sorting.
Instrument	Samples were acquired on LSRFortessa X-20 or FACSCelesta for acquisition and FACSAria Fusion for sorting. Data were analyzed using FlowJo software.
Software	We used concatenated files in FlowJo at the recommended setting.
Cell population abundance	For flow studies of leukocytes analysis 50k cells were acquired. For enrichment of GFP+ cells we sorted alomoste 90% of the cells. After CRISPR_Cas9 gRNAs virus infection and selection single cell was sorted in the 96-well plates. More details the methods are provided in the Methods section.
Gating strategy	For flow based sorting of the luciferase-GFP labeled cells, parental cells without transfection of luciferase-GFP were used to define negative cell populations and set gates for analysis. Unstained and compensation beads were used for compensation in the multicolor antibodies analysis and gating strategies.

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