

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 (<i>n</i>) 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
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☐ | ☒ 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. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted
<i>Give P values as exact values whenever suitable.</i> |
| ☒ | ☐ 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 <i>d</i> , Pearson's <i>r</i>), 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	No software was used for data collection
Data analysis	Mushroom, https://github.com/ding-lab/mushroom/tree/subclone_submission Morph, https://github.com/ding-lab/morph multiplex imaging pipeline, https://github.com/estorrs/multiplex-imaging-pipeline Space Ranger v1.3.0, v 2.0.0, v2.1.0, 10X Genomics Cell Ranger v6.0.2, 10X Genomics InferCNV v1.10.1, https://github.com/broadinstitute/infercnv CalicoST, https://github.com/raphael-group/CalicoST 10Xmapping, https://github.com/ding-lab/10Xmapping samtools v1.14, https://github.com/samtools/samtools trimGalore v0.6.7, https://github.com/FelixKrueger/TrimGalore BWA-mem v0.7.17, https://github.com/lh3/bwa Somaticwrapper v2.2, https://github.com/ding-lab/somaticwrapper Picard v2.6.26, https://github.com/broadinstitute/picard Strelka v2.9.10, https://github.com/Illumina/strelka MUTECT v1.1.7, https://github.com/broadinstitute/mutect VarScan v2.3.8, https://dkoboldt.github.io/varscan Pindel v0.2.5, https://github.com/genome/pindel COCOON, https://github.com/ding-lab/COCOONS scVarScan, https://github.com/ding-lab/10Xmapping

GATK v4.1.9.0, <https://github.com/broadinstitute/gatk>
 Bedtools 2.30.0, <https://github.com/bedtools/bedtools>
 Scrublet v0.2.3, <https://github.com/swolock/scrublet>
 Seurat v4.3.0, R-package
 ComplexHeatmap v2.14.0, R-package
 tidyverse v1.3.2, R-package
 ggplot2 v3.3.5, R-package
 RColorBrewer v1.1.3, R-package
 singleR v1.8.1, R-package
 clusterProfiler v3.18.1, R-package
 msigdb v7.5.1, R-package
 ROGUE v1.0, R-package
 RCTD v1.2.0, R-package
 COMMOT, <https://github.com/zcang/COMMOT>
 PASTE2, <https://github.com/raphael-group/paste2>
 ppcor v1.1, R-package
 stats v4.1.2, R-package
 ggpubr v0.6.0, R-package
 python v3.7.12 and v.3.9.6
 sklearn v0.24.2, python-package
 matplotlib v3.4.2, py39hf3d152e_0
 numpy v1.21.1, h49503c6_1_cpython
 org.Hs.eg.db v3.12.0, R-package
 pandas v1.3.1, py39hde0f152_0
 Fiji/ImageJ, v2.14.0
 BigWarp, Fiji/ImageJ-package
 Napari v0.4.19
 Imapis v10.1.1
 DeepCell v0.12.9, <https://github.com/vanvalenlab/deepcell-tf>

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](#)

Visium, sc/snRNA-seq, WES, Xenium, and CODEX imaging data are part of Human Tumor Atlas Network (HTAN) dbGaP Study Accession: phs002371.v3.p1 (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs002371.v3.p1), and the Data can be accessed through the HTAN DCC Portal (<https://data.humantumoratlas.org/>) under the HTAN WUSTL Atlas. Individual samples can be identify using Biospecimen ID from the Sample_ID_Lookup_table.xlsx on the GitHub page (https://github.com/ding-lab/ST_subclone_publication). GRCh38 references used for sc/snRNA-seq (refdata-gex-GRCh38-2020-A) is freely available from 10X Genomics website (<https://support.10xgenomics.com>). The reference GRCh38 genome (GRCh38.d1.vd1.fa.tar.gz) used for WES reads alignment is available from GDC (<https://gdc.cancer.gov/about-data/gdc-data-processing/gdc-reference-files>). The MSigDB hallmark gene sets is available from GSEA website (<https://www.gsea-msigdb.org/gsea/msigdb/collections.jsp>). The L1000 perturbation database is available through the resource page on the Harmonizome website (<https://maayanlab.cloud/Harmonizome/resource/L1000+Connectivity+Map>) and the library page on the Enrichr website (<https://maayanlab.cloud/Enrichr/libraries>).

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	Patient sex information was collected as part of this study
Reporting on race, ethnicity, or other socially relevant groupings	Patient race or ethnicity information was collected as part of this study
Population characteristics	Our dataset comprises samples from 6 tumor types, with patients ages 37-82. The distribution of the samples and clinical information across the cohorts can be found in the Supplementary Table 1, and Extended Data Fig. 1a.
Recruitment	Patients who fit the clinical criteria and consented to the study were selected for inclusion in the genetic and molecular tumor analysis. 75 cases from 6 cancers are part of Human Tumor Atlas Network (HTAN) study, and 3 ccRCC cases are from Clinical (Clinical Proteomic Tumor Analysis Consortium). There was no self-selection bias or other biases in the recruitment of patients
Ethics oversight	All samples were collected with informed consent at the Washington University School of Medicine in St Louis. Breast cancer

Ethics oversight

(BRCA) samples, pancreatic adenocarcinoma (PDAC) samples, colorectal cancer (CRC) samples, cholangiocarcinoma (CHOL) samples, renal carcinoma (RCC) samples and uterine corpus endometrial carcinoma (UCEC) samples were collected during surgical resection and verified by standard pathology (IRB protocol 201108117, 201411135, and 202106166). Tumor samples were collected during surgical resection and verified by standard pathology.

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](https://www.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

Sample sizes were chosen based on the availability of samples collected across 6 cancer types. No specific statistical methods were used to predetermine sample size. This cohort collected a total of 132 sections from 79 cases samples for Visium ST, along with several samples for paired snRNA-seq, WES, and CODEX. This sample size aligns with or exceeds that of many previously published Visium spatial transcriptomics data on the date of submission.

Data exclusions

No data were excluded from the analysis

Replication

To ensure reproducibility, we used multiple bioinformatics tools for the same analyses and compared results across technologies (including bulk, single-cell, spatial sequencing, and multiplexed imaging) from the same tumor samples. To confirm subclonal copy number alterations in tumor spatial regions, we used two independent copy number inference methods, CalicoST (github.com/raphael-group/CalicoST) and InferCNV (<https://github.com/broadinstitute/infercnv>) (Supplementary Fig. 5a). We further validated spatial CNV inference with matching WES-based CNV calling and observed 0.43-0.82 agreement in Jaccard similarity (methods) between the dominant spatial subclone and WES (Supplementary Fig. 5c). Further, spatial subclonal mutations were validated with matching WES-based mutation calling (with 4%-33% WES mutations detected in spatial transcriptomics, Supplementary Fig. 5d).

To support our findings in the tumor center and periphery regions, we used matching snRNA-seq from the same tumor samples and demonstrated expression of most center-enriched and periphery-enriched genes in malignant cells (Supplementary Fig. 8a). Subclonal tumor microenvironment compositions were confirmed with multiplexed immunostaining with co-detection by indexing (CODEX) (Fig. 5c).

Randomization

The study design didn't involve the allocation of patients into treatment groups, therefore the randomization procedure was not relevant.

Blinding

The study design didn't involve the allocation of patients into treatment groups, therefore blinding procedure was not relevant.

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 |
| ☒ | ☐ Plants |

Methods

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

Antibodies

Antibodies used

CD11b (Abcam #ab187537;EP1345Y),CD20 (Abcam #ab236434;SP32),CD31 (Abcam #ab226157;EP3095),CD68 (Abcam #ab233172;KP1),CK19 (Abcam #ab195872;EP1580Y),GLUT1 (Abcam #ab196357;EPR3915),CD163 (Akoya #4250079;AKYP0114),CD3e (Akoya #4550125;EP449E),CD4 (Akoya #4550112;EPR6855),CD8 (Akoya #4250012;C8/144B),FOXP3 (Akoya #4550071;AKYP0102),HLA-DR (Akoya #4450095;EPR3692),Ki67 (Akoya #4450096;B56),Pan-Cytokeratin (Akoya #4450093;AE-1/AE-3),Podoplanin (BioLegend #337002;NC-08),Vimentin (BioLegend #677802;O91D3),SMA (Invitrogen #MA1-06110;1A4),CD45 (Novus Biologicals #NBP2-34528;2B11+PD7/26),E-cadherin (ThermoFisher #33-4000;4A2C7),AQP1 (Abcam #ab178352;EPR11588(B)),AQP2 (Abcam #ab230170;EPR21080),CALB1 (Abcam #ab233018;EP3478),CK17 (Abcam

#ab239986;EP1623), cKit (Abcam #ab216450;YR145), COX6c (Abcam #ab243916;EPR9938), GATA3 (Abcam #ab214804;EPR16651), Her2 (Abcam #ab194979;EP1045Y), LRP2 (Abcam #ab76969;Polyclonal), MGP (Abcam #ab273657;OTI8D6), MUC2 (Abcam #ab272706;EPR23479-47), P21 (Abcam #ab212247;CIP1/823), PAI1 (Abcam #ab237780;EPR21850-82), PLAT/tPA (Abcam #ab240146;EPR7232(2)), PR (Abcam #ab239793;SP2), SOX9 (Abcam #ab225541;EPR14335-78), UCHL1 (Abcam #ab220823;EPR4118), UMOD (Abcam #ab223540;EPR20071), CK5 (Akoya #4450090;AKYP0121), ER (Akoya #4250074;AKYP0105), CP (Bethyl Laboratories #A80-124A;Polyclonal), CK14 (BioLegend #905304;Polyclonal), MLPH (Invitrogen #PA5-118065;Polyclonal), CA9 (Novus Biologicals #NB100-417;Polyclonal), CCL2 (Novus Biologicals #NBP2-22115;2D8), CK7 (Novus Biologicals #NBP-47940;OV-TL12/30), Hep-Par-1 (Novus Biologicals #NBP2-45272;HepPar1), CD36 (Cell Signaling Technology #39914;D8L9T), TFF1 (Abcam #ab239908;EPR3972)

Validation

Antibodies used in this application have already been used in our laboratory and confirmed by one or more of the following methods. First: Western blotting produces the correct banding pattern (correct molecular weight, absence in non-transfected cells or protein extracts from mice deficient for the gene but present in wild-type samples or samples with transfection of an expression construct). Second, immunostaining shows the correct pattern (correct tissue specificity, intracellular location, absent/low in non-transfected cells, present/high in transfected cells, colocalization with the tags for fusion proteins expressed in cells). Third, 2 or more independent antibodies against the same antigen show the same expected pattern. In addition, we have included detail antibody information in Supplementary Table_04_CODEX_Antibody.xlsx, further validation images in 2_Response_letter.pdf and 2b_Respons_Ab_IF_testing_0610.pdf

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.