nature research

Corresponding author(s): Joan Massague

Last updated by author(s): Nov 7, 2019

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	single cell RNAseq and ChIPseq data have been deposited on GEO under the accession numbers provided in the methods section.
Data analysis	Details of software packages used for single cell RNA sequencing and ChIP seq analysis are provided in the methods section. All single cell analyses and visualizations were performed in Python using the following open-source algorithms as described in the methods: SEQC (https://github.com/ambrosejcarr/seqc), t-SNE (https://lvdmaaten.github.io/software/), MAGIC (https://github.com/dpeerlab/magic), and the Scikit-learn implementation of a Gaussian Mixture Model. ChIPseq data analysis was performed in R and visualized using IGV.

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

scRNA-seq data have been deposited in the Sequence Read Archive under accession number SRP136919. Data are available at: https://dataview.ncbi.nlm.nih.gov/object/PRJNA448377?reviewer=cljvd9mlkg2naqlj2n47si4jmi. ChIPseq data are deposited on GEO (GSE112555).

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 clinical samples, minimum numbers of samples required for statistical power were used, in accordance with MSKCC IRB guidelines. For mouse experiments, in compliance with IACUC guidelines, a minimal number of animals for a statistically significant result was used. Age and sex matched animals were used for all experiments. Littermate Cre negative animals were used as controls. All experiments were designed to have > = 5 mice per group.
Data exclusions	Any animals that died within 24 hours of rectal or splenic injection were assumed to have died due to procedure-related complications and were excluded from further analysis.
Replication	All attempts at replication were successful. Multiple biologically independent experimets were performed wherever feasible as described in the methods and figure legends.
Randomization	Littermate controls were used for relevant animal experiments. For indicated experiments where treatments were started once tumours were established, tumours in individual animals were measured and ranked based on the BLI signal; every alternate animal was allocated to the control or treatment arm based on the measured BLI intensity.
Blinding	Investigators were blinded to patient identity or mouse group when scoring histopathological samples.

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	Antibodies used for immunohistochemistry and immunofluorescence: L1CAM (Biolegend, #826701, Clone 14.10, 1:500), Ki-67 (Roche, #790-4286, Clone 30-9, 1:1 (prediluted))) Antibodies used for flow cytometry/sorting: EpCAM-FITC (Dako, #F0860, Ber-EP4, 1:20), L1CAM-APC (Novus Biologicals, #NBP234505J, Clone UJ127, 1:10) EphB2-PE (Thermo Fisher Scientific, FAB467P, Clone 512012, 1:10), CD133-PE-Vio770 (Miltenyi Biotec, #130-101-652, Clone AC133, 1:20), CD44-BV410 (BD Biosciences, #563029, Clone G44-26, 1:20), mouse L1CAM-APC (Miltenyi Biotec, #130-102-221, Clone 555, 1:10). Goat anti-human IgG (H+L) HRP conjugate (Thermo Fisher Scientific, #A18811, 0.5 µg/mL) was used for solid-state binding assays.
Validation	Pre-validated antibodies were purchased from reputable commercial sources. All primary antibodies were tested in the relevant application to detect recombinant protein, without cross-reactivity with non-specific proteins.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	L-Wnt3a and 293T cells were purchased from ATCC. m-Rspo-Fc was a kind gift of Calvin Kuo. Organoid lines were generated de novo from patients enrolled in MSK IRB approved protocol #14-244, or from C57BL/6J mice.

Authentication	Key human colorectal cancer organoid lines generated and used for this study (MSK107Li, MSK121Li, MSK132P, MSK125P) were validated by hybridization capture and targeted exon sequencing using the MSK-IMPACT assay to validate their identity at least 6 months following original derivation. Mouse AKP organoids were genotyped via PCR using primers targeting the APC, KRAS and TP53 alleles. L-Wnt3a, 293T and m-Rspo-Fc were validated by suppliers and were re-authenticated by STR profiling.
Mycoplasma contamination	All cell lines tested negative for mycoplasma contamination
Commonly misidentified lines (See ICLAC register)	None of the cell lines used are listed in the ICLAC database

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	C57BL/6J (strain 000664), Villin-Cre (B6.Cg-Tg(Vil1-cre)997Gum/J; 004586), APCfl/fl (C57BL6-APCtm1Tyj/J; 009045), Lgr5-EGFP-IRES-creERT2 (B6.129P2-Lgr5tm(cre/ERT2)/Cle/J; stock 008875), KRASLSL-G12D (B6.129S4-Krastm4Tyj/J; stock 008179), P53fl/fl (B6.129P2-Trp53tm1Brn/J; stock 008462), athymic nude (Foxn1nu; stock 002019) and NSG (NOD.Cg-Prkdcscid IL2rgtm1Wjl/SzJ; stock 005557) mouse strains were obtained from the Jackson laboratory. KRASLSL-G12D, P53fl/fl and Lgr5-EGFP-IRES-creERT2 mice were backcrossed onto a C57BL/6J background for at least 6 generations in the Lowe laboratory. L1CAMfl/fl mice were a gift of Mellitta Schachner. Mice were inbred on a C57BL/6J background to generate the necessary crosses. Since L1cam is on the X chromosome, male mice were used for experiments involving this allele to minimize the number of breeding generations and mice needed. 8-10 week old male C57Bl/6J mice were used for colitis experiments. Mice in tumor initiation experiments were euthanized at 3 months. Xenograft injections were performed in athymic nude (Foxn1nu; stock 002019) and NSG (NOD.Cg-Prkdcscid IL2rgtm1Wjl/SzJ; stock 005557) mice matched to the gender of the organoid. Mice were obtained from the Jackson laboratory and transplanted at 6 weeks of age
Wild animals	The study did not involve wild animals
Field-collected samples	The study did not involve samples collected in the field
Ethics oversight	All animal studies were conducted under the supervision of the MSKCC Institutional Animal Care and Use Committee

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

Human research participants

Policy information about studies involving human research participants

Population characteristics	Clinical and genetic characteristics of the patients from whom MSK107Li, MSK121Li, MSK125P and MSK132P organoids were derived are listed in Supplementary Figure 2a. Patients were both male and female, with an age range of 47-65 years at recruitment. 3/4 patients had previously received FOLFOX chemotherapy. All patients had TP53-mutant Stage IV colorectal cancer with liver metastases. Human subjects involvement was restricted to biospecimen collection and no interventions were performed directly on human subjects.
Recruitment	All human tissues were obtained under Memorial Sloan Kettering Institutional Review Board approved biospecimen research protocols 14-244 and 15-101. Suitable patients with metastatic colorectal cancer undergoing surgery were identified by detailed chart review by a medical oncologist (K.G.) and all patients provided pre-procedure informed consent to tissue procurement. Recruitment was based solely on tissue availability, potentially biasing against early-stage and smaller tumors. No other patient-characteristic related bias is likely to be prevalent.
Ethics oversight	All human biospecimens were obtained from patients who had provided informed consent for tissue usage under MSK Institutional Review Board approved protocol #06-107 and distributed under #14-244.

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

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE112555 The following secure token has been created to allow review of record GSE112555 while it remains in private status: ctebcooipncxtez
Files in database submission	PROCESSED DATA FILES CRC107LishCDH1_Input_IGO_08204_2_S14_R1_001_val.hg19.sorted.RmDup.10mNorm.bw bigWig CRC107LishCDH1_REST_IGO_08204_14_S24_R1_001_val.hg19.sorted.RmDup.10mNorm.bw bigWig CRC107LishCTRL_D_Input_IGO_08204_6_S18_R1_001_val.hg19.sorted.RmDup.10mNorm.bw bigWig CRC107LishCTRL_D_REST_IGO_08204_18_S28_R1_001_val.hg19.sorted.RmDup.10mNorm.bw bigWig

CRC107LishCTRL_I_Input_IGO_08204_3_S15_R1_001_val.hg19.sorted.RmDup.10mNorm.bw bigWig
CRC107LishCTRL_I_REST_IGO_08204_15_S25_R1_001_val.hg19.sorted.RmDup.10mNorm.bw bigWig
CRC121Li_D_input_IGO_08204_B_1_S2_R1_001_val.hg19.sorted.RmDup.10mNorm.bw bigWig
CRC121Li_D_REST_IGO_08204_B_2_S3_R1_001_val.hg19.sorted.RmDup.10mNorm.bw bigWig
CRC121Li_I_input_IGO_08204_B_3_S4_R1_001_val.hg19.sorted.RmDup.10mNorm.bw bigWig
CRC121Li_I_REST_IGO_08204_B_4_S5_R1_001_val.hg19.sorted.RmDup.10mNorm.bw bigWig

RAW FILES
CRC107LishCDH1_Input_IGO_08204_2_S14_R1_001.fastq.gz fastq
CRC107LishCDH1_Input_IGO_08204_2_S14_R2_001.fastq.gz fastq
CRC107LishCDH1_REST_IGO_08204_14_S24_R1_001.fastq.gz fastq
CRC107LishCDH1_REST_IGO_08204_14_S24_R2_001.fastq.gz fastq
CRC107LishCTRL_D_Input_IGO_08204_6_S18_R1_001.fastq.gz fastq
CRC107LishCTRL_D_Input_IGO_08204_6_S18_R2_001.fastq.gz fastq
CRC107LishCTRL_D_REST_IGO_08204_18_S28_R1_001.fastq.gz fastq
CRC107LishCTRL_D_REST_IGO_08204_18_S28_R2_001.fastq.gz fastq
CRC107LishCTRL_I_Input_IGO_08204_3_S15_R1_001.fastq.gz fastq
CRC107LishCTRL_I_Input_IGO_08204_3_S15_R2_001.fastq.gz fastq
CRC107LishCTRL_I_REST_IGO_08204_15_S25_R1_001.fastq.gz fastq
CRC107LishCTRL_I_REST_IGO_08204_15_S25_R2_001.fastq.gz fastq
CRC121Li_D_input_IGO_08204_B_1_S2_R1_001.fastq.gz fastq
CRC121Li_D_input_IGO_08204_B_1_S2_R2_001.fastq.gz fastq
CRC121Li_D_REST_IGO_08204_B_2_S3_R1_001.fastq.gz fastq
CRC121Li_D_REST_IGO_08204_B_2_S3_R2_001.fastq.gz fastq
CRC121Li_I_input_IGO_08204_B_3_S4_R1_001.fastq.gz fastq
CRC121Li_I_input_IGO_08204_B_3_S4_R2_001.fastq.gz fastq
CRC121Li_I_REST_IGO_08204_B_4_S5_R1_001.fastq.gz fastq
CRC121Li_I_REST_IGO_08204_B_4_S5_R2_001.fastq.gz fastq

Genome browser session
(e.g. UCSC)

Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.

Methodology

Replicates	Two biological replicates were used per integrity condition
Sequencing depth	50 bp paired end sequencing was performed to obtain 30 million read depth
Antibodies	REST (Millipore Sigma, 07-579, lot 2843200)
Peak calling parameters	Peaks in ChIP-seq read density were called with version 2.1.1.20160309 of MACS2. 0.001 was the setting for the p-value option, with "BAM" as the format setting and the "--nomdel" option activated.
Data quality	Details of data analysis and quality assurance are in the methods section
Software	Details of the software packages used for ChIP-seq data analysis are in the methods section

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

Human tumour samples were grossly dissected out from colon mucosa (primary tumours) or liver/peritoneum (metastases). Visible normal tissue at the tumour margin was trimmed away, and the sample was washed with PBS-Abx (PBS containing penicillin 500 U/mL; Streptomycin 500 µg/mL; gentamicin 100 µg/mL; amphotericin B 12.5 µg/mL) prior to finely chopping the tumour into 5mm fragments using a pair of razor blades (Fisher Scientific). Fragments were incubated in Dissociation Buffer (DMEM with 5% fetal bovine serum (FBS; Gibco), 2mM L-glutamine (Fisher Scientific), penicillin/streptomycin (Fisher Scientific), 40 µg/mL gentamicin (Thermo Fisher Scientific), 250 U/mL Type III collagenase (Worthington), 1U/mL dispase (Sigma-Aldrich)) on a shaker for 30 min at 37°C, filtered through a 70 µm cell strainer (Greiner Bio-One), centrifuged at 600xg for 5 min and washed once with ADF (Advanced DMEM/F12 (Thermo Fisher Scientific)).

Organoids were harvested from matrigel by aspirating media, replacing with Cell Recovery Solution (Corning), incubating at 4°C for 30 min, mechanically disrupting matrigel by pipetting, washing with ADF and centrifuging at 600xg for 5 min to pellet

	organoids. To generate dissociated single cell suspensions, the organoid pellet was treated with TrypLE (Thermo Fisher Scientific) for 10 min at 37°C, washed with ADF and filtered through a 70 µm cell strainer. Dissociated fresh tumours or organoids were passed through a 70 µm filter to generate single cells and stained with the antibodies on ice for 30 min in PBS containing 0.1% FBS according to manufacturer's instructions.
Instrument	Flow sorting was performed using a BD FACSAria II cell sorter fitted with a 100 µM nozzle. Analysis was performed using a BD LSR/Fortessa.
Software	The manufacturer's FACSDiva (BD) software was used for data collection. Flow cytometry data were analyzed using FlowJo software (Treestar).
Cell population abundance	Post-sort reanalysis was typically not possible due to limiting numbers of human tumor cells or organoid cells. When performed, we verified a post-sort purity of at least 90%.
Gating strategy	The gating strategies used are defined in the figures or supplementary figures

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