nature research

Corresponding author(s): Maria Casanova-Acebes & Miriam Merad

Last updated by author(s): 04/01/2021

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	FACS Diva software version 7(BD), ZEN Black Imaging Software (ZEISS)
Data analysis	FlowJo v10.0.6 (Tree Star Inc.), GraphPad Prism v6.0e (GraphPad Software Inc.), Pannoramic Viewer v1.15.4 (3DHISTECH), ImageJ 2.0.0, and IncuCyte were used for flow cytometry and image analysis. DeepTools, GREAT and IGV v2.5.3 were used to generate ATACseq data. RStudio v1.2, R v3.5.3, R packages: matrixStats, Matrix.utils, mixtools, matrix. tglkmeans (https://tanaylab.bitbucket.io/tglkmeans/) were used for bulk and scRNAseq analysis. Mouse assembly version mm10/NCBI m38 was used for sequence alignments with bowtie version 2.3.4.1. Sequence duplicates in ATACseq were removed with picard-tools version 1.88. Strategies for scRNAseq clustering, cell projection, and homology analyses are similar to previously described approaches (see methods), and scripts are available upon request.

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

Mouse scRNAseq, ATACseq and bulk RNAseq (Fig. 1a-c; Fig. 2a,d-h; Fig. 3b-c; Extended Data Fig. 1e-h; Extended Data Fig. 2b-c; Extended Data Fig. 3a-b,e) data are all available on GEO (accession number: GSE147671). All human processed gene expression data will be available on GEO, and raw sequencing reads are uploaded to dbGaP, BioProject: PRJNA609924 (GEX, gene expression files on https://dataview.ncbi.nlm.nih.gov/object/PRJNA609924?reviewer=2tal0jn5fi9mm154e8dpk72bnt).

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 size. Sample sizes were estimated based on preliminary experiments, with an effort to achieve a minimum of n=3, mostly n=5 mice per treatment group, which proved sufficient to determine reproducible results.
Data exclusions	For the single-cell RNAseq data, we excluded cells based on pre-established criteria for single-cells, excluded cells with low number of detected transcripts and high mitochondria content. Cell filtering for the human data is described in Leader et al., as noted.
Replication	All experiments were reliably reproduced. All experiments (except for -omics data ie. RNA-seq, ATAC-seq and scRNA-seq) were performed independently at least two times, unless stated (see Figure legends). The ATAC/RNA-seq data from in vivo samples were reliably reproduced, as assessed by Principal Component Analysis (PCA) and hierarchical clustering methods.
Randomization	Mice with matched sex and age are randomized into different treatment groups (eg. naive, tumor). Human subjects were not randomized into experimental groups since we focused on studying macrophage heterogeneity across patients as a whole.
Blinding	Quantification of mouse tumor sizes, lung CC3, Ki67 and p27, CD4, CD8, Treg and CD3 infiltration was blinded by de-identifying samples. In all other experiments, analysis was objective and did not require blinding.

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

The following antibodies were used for flow cytometry:

CD45 (clone 30-F11) BioLegend 1:200 Catalog # 103137
SiglecF (E50-2440) BD Pharmingen 1:200 Catalog # 740956
CD11c (N418) Invitrogen 1:200 Catalog # 47-0114-82
CD24 (M1/69) Invitrogen 1:200 Catalog # 25-0242-82
CD103 (2E7) BioLegend 1:200 Catalog # 17-1031-82
CD11b (M1/70) eBioscience 1:200 Catalog # 45-0112-82
Ly6G (1A8) BioLegend 1:200 Catalog # 127621
Cx3Cr1 (SA011F11) BioLegend 1:200 Catalog # 149037
MHCII I-A/I-E (M5/114.15.2) eBioscience 1:100 Catalog # 12-5321-82
Ly6C (AL-21) eBioscience 1:200 Catalog # 553104
CD206 (Mrc1) (C068C2) BioLegend 1:200 Catalog # 141719
CD169 (3D6.112) BioLegend 1:200 Catalog # 142403
FoxP3 (FJK-16s) Invitrogen 1:100 Catalog # 12-5773-82
TCRb (H57-597) eBioscience 1:200 Catalog # 560657
CD4 (GK1.5) eBioscience 1:200 Catalog # 17-0041-82
CD8 (53-6.7) eBioscience 1:200 Catalog # 558106
IFNg (XMG1.2) eBioscience 1:200 Catalog # 25-7311-41
TNFα (MP6-XT22) eBioscience 1:200 Catalog # 17-7321-82

CTLA-4 (UC10-4B9) eBioscience 1:200 Catalog # 17-1522-82
CD73 (eBioTY/11.8) Invitrogen 1:200 Catalog # 127211
Ki67 (16A8) BioLegend 1:200 Catalog # 652409
CD2 (RM2-5) BioLegend 1:200 Catalog # 100113
Biotin Mertk (Polyclonal) R&D 1:100 Catalog # BAF591
Biotin CD64 (X54-5/7.1) BioLegend 1:100 Catalog # 139318

The following primary antibodies were used for immunofluorescence (IF) and immunohistochemistry (IHC) staining:
E-Cadherin (36/E-Cadherin) BD Biosciences 1:100 Catalog # 610181
Twist1 (Polyclonal) Millipore 1:100 Catalog # ABD29
Active β-catenin (8E7) Millipore 1:100 Catalog # 05-665
Ki67 (SolA15) eBioscience 1:100 Catalog # 14-5698-82
Cleaved Caspase-3 (Asp175) Cell Signaling Technology 1:100 Catalog # 9661S
p27 (D69C12) Cell Signaling Technology 1:100 Catalog # 3686S
SiglecF (1RNM44N) Invitrogen 1:200 Catalog # 14-1702-82
Pan-cytokeratin (Polyclonal) Progen 1:200 Catalog # GP11
CD45.1 (A20) Biolegend 1:100 Catalog # 110717
CD10 (200103) R&D 1:1000 Catalog # MAB1126
CD206 (Polyclonal) Abcam 1:1000 Catalog # ab64693
FABP4 (Polyclonal) R&D 1:100 Catalog # AF1443
Pan-cytokeratin (AE1/AE3) Dako 1:50 Catalog # GA053
FoxP3 (D6O8R) Cell Signaling Technology 1:100 Catalog # 12653
CD4 (EPR195) Abcam 1:1000 Catalog # ab183685
CD8 (4SM15) Invitrogen 1:200 Catalog # 14-0808-82
Pan-cytokeratin (C11) BioLegend 5ug/ml Catalog # 628602
CD206 (C068C2) BioLegend 10ug/ml Catalog # 141702
FoxP3 (FJK-16s) Invitrogen 10ug/ml Catalog # 14-5773-82
CD25 (Polyclonal) R&D 10ug/ml Catalog # AF2438

The following secondary antibodies were used for immunofluorescence (IF) staining:
AlexaFluor goat-anti-mouse 568, Invitrogen 1:1000 Catalog # A11004
AlexaFluor goat anti-rabbit 647, Invitrogen 1:1000 Catalog # A21244
AlexaFluor goat anti-rat 568 Invitrogen 1:1000 Catalog # A11077
AlexaFluor goat anti-rabbit 568 Invitrogen 1:1000 Catalog # A11011
AlexaFluor goat anti-rabbit 405 Invitrogen 1:1000 Catalog # A31556
AlexaFluor goat anti-mouse 647 Invitrogen 1:1000 Catalog # A21235
AlexaFluor goat anti-guinea pig 488 Invitrogen 1:1000 Catalog # A11073
AlexaFluor 432 Catalog # A20182
AlexaFluor 594 Catalog # A20185
AlexaFluor 647 Catalog # A20186

Validation

All the antibodies are validated for the use of immunofluorescence, immunohistochemistry and histological analyses. Data are available on the manufacturer's website. The following primary antibodies have been validated by the manufacturer and;or in this paper:
CD45: Flow cytometric analysis of CD45 expression on mouse splenocytes (website)
SiglecF (BD Pharmingen): Flow cytometric analysis of SiglecF expression on mouse bone marrow leukocytes (website)
CD11c: Flow cytometric analysis of CD11c on expression mouse splenocytes (website)
CD24: Flow cytometric analysis of CD24 expression in mouse splenocytes (website)
CD103: Flow cytometric analysis of CD103 expression on mouse splenocytes (website)
I-A/I-E: Flow cytometric analysis of I-A/I-E expression on mouse splenocytes (website)
CD11b: Flow cytometric analysis of CD11b expression on mouse bone marrow leukocytes (website)
Ly6G: Flow cytometric analysis of LY6G expression on mouse splenocytes (website)
Cx3Cr1: Flow cytometric analysis of CX3CR1 expression on mouse splenocytes (website)
Ly6C: Flow cytometric analysis of CX3CR1 expression on mouse splenocytes (website)
CD206 (BioLegend): Flow cytometric analysis of CX3CR1 expression on mouse splenocytes (website)
CD169: Flow cytometric analysis of CD169 expression on mouse bone marrow leukocytes (website)
FOXP3 (Invitrogen): Flow cytometric analysis of FOXP3 expression on mouse splenocytes (website)
CD4 (eBioscience): Flow cytometric analysis of CD4 expression on mouse splenocytes (website)
CD8 (eBioscience): Flow cytometric analysis of CD8 expression on mouse splenocytes (website)
IFNg: Flow cytometric analysis of IFNg expression on mouse splenocytes (website)
TNFa: Flow cytometric analysis of TNFa expression on mouse splenocytes (website)
CTLA-4: Flow cytometric analysis of CTLA-4 expression on mouse splenocytes (website)
CD73: Flow cytometric analysis of CD73expression on mouse splenocytes (website)
Ki67: Flow cytometric analysis of Ki67 expression on mouse splenocytes (website)
CD2: Flow cytometric analysis of CD45 expression on mouse splenocytes (website)
Biotin Mertk: Flow cytometric analysis of J774A.1 mouse reticulum cell sarcoma macrophage cell line (website)
Biotin CD64: Flow cytometric analysis of CD45 expression on mouse splenocytes (website)
E-Cadherin: Immunofluorescent staining of WIDR cells (website)
Twist1: Immunohistochemistry in ductal breast fibroadenoma and stratified squamous epithelium and cartilage of human trachea tissues (website)
Active β-catenin: Immunocytochemistry, positive staining on the membrane and cytosol (website)
Ki67: Immunohistochemistry on formalin-fixed paraffin embedded mouse skin (website)
Cleaved Caspase-3: Immunohistochemical analysis of paraffin-embedded human tonsil (website)
p27: Immunofluorescent analysis of MCF-7 cells (website)

SiglecF (Invitrogen): Immunohistochemistry of frozen mouse lung (website)
Pan-cytokeratin (Progen): Immunofluorescence staining of mouse CDβGeo cells (website)
CD10: Immunohistochemistry of paraffin-embedded sections of human kidney (website)
CD206 (Abcam): Immunohistochemistry of human lung formalin fixed paraffin embedded tissue section (website)
FABP4: Immunohistochemistry of frozen sections of mouse intestine (website)
Pan-cytokeratin (Dako): Immunohistochemistry of adenocarcinoma (website)
FoxP3 (Cell Signaling Technology): Immunohistochemical analysis of paraffin-embedded 4T1 metastatic tumor in mouse lung (website)
CD4 (Abcam): Immunohistochemistry of mouse colon tissue (website)
CD8 (Invitrogen): Immunohistochemistry of formalin-fixed paraffin embedded mouse spleen (website)
CD25: Immunohistochemistry of immersion fixed frozen sections of mouse spleen (website)

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	KP cells were received from Tyler Jacks and B16-BFP/OVA were kindly donated by John Finnigan.
Authentication	Cells were functionally authenticated: KP and B16 cells were injected intravenously into mice and lung tumor progression at different timepoints was followed. No other authentication was performed.
Mycoplasma contamination	All cell lines tested negative for mycoplasma.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used in the study

Animals and other organisms

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

Laboratory animals	All mice included in this study were analyzed between 8 and 12 weeks of age. Males and females were used in different experiments, each of them sex-controlled. The following mice strains were also included: CD169-Cre R26-LSL-tdTomato, Ms4a3TdT reporter, CD169-DTR, Pdzk1ip1-CreER R26Tom/Tom, C57BL/6, Cx3CR1CreER/+ R26-LSL-YFP, KP-GEMM (SPC-Cre;Trp53fl/fl;LSLKrasG12D), OT-I Tg, OTII-Tg.
Wild animals	The study did not involve wild animals.
Field-collected samples	The study did not involve field-collected samples.
Ethics oversight	Ethical approval for mouse experiments was obtained by the Internal Animal Care and Use Committee at Mount Sinai Hospital.

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	Characteristics of the patients analyzed in this study are fully described in Leader et al. Patients were undergoing resection with curative intent of early-stage non-small-cell lung cancer lesions.
Recruitment	Patients over the age of 45 years old were identified and recruited for the study by the surgeons performing the above procedure in collaboration with clinical research coordinators. No self-selection biased was anticipated.
Ethics oversight	The study protocol was approved by the Mount Sinai IRB (study protocols HS# 10-00472 and HS#-00135)

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	Single cell suspension was generated from mouse lungs and lymph nodes. Both tissues were minced with scissors prior to digestion with Collagenase IV for 30min (lung) or 15min (lymph node and spleen). The cells were filtered and subjected to red blood cell lysis.
Instrument	BD LSRFortessa
Software	FACS Diva software version 7 (BD)
Cell population abundance	Purity of sorted populations for ATACseq was ~99% in all samples and was assessed by flow cytometry.
Gating strategy	Gating for TRMs: 1. Gate on fsc-a vs. ssc-a was set to include all cell populations, but excluding debris. 2. Gate on fsc-a vs. fsc-w was set to exclude doublets. 3. gate on ssc-a vs. ssc-w was set to exclude doublets. 4. gate on fsc-a vs. live/dead was set to exclude dead cells (live/dead+). 5. gate on CD45+ vs. ssc-a was set to exclude CD45- cells. 6. gate on CD11b+ Ly6G+ was set to exclude neutrophils. 7. gate on CD64-Mertk double positive cells was set to enrich for macrophages, which were further gated as CD11blowCD2+, CD169+, Mrc1+, SiglecF+. Gating for KP-GFP in spheroids: 1. Gate on fsc-a vs. ssc-a was set to include all cell populations, but excluding debris. 2. Gate on ssc-a vs. ssc-h was set to exclude doublets. 3. gate on fsc-a vs. DAPI was set to exclude dead cells (DAPI+). 4. gate on CD45+ vs. GFP+ was set to sort KP-GFP cells (CD45-GFP+).

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