nature portfolio

Corresponding author(s): Rongxin Zhang

Last updated by author(s): Oct 9, 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	The external data used for data analysis in this study could be accessed through the following portal: The TCGA data: https://www.cbioportal.org/study/summary?id=coadread_tcga_pan_can_atlas_2018 The external scRNA-seq data: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE205506 (CRC) https://singlecell.broadinstitute.org/single_cell/study/SCP1288/tumor-and-immune-reprogramming-during-immunotherapy-in-advanced-renal-cell-carcinoma (ccRCC) https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE115978 (melanoma)
Data analysis	This study does not involve software development. The software and their versions used in this study are listed as follow: software version resource Python 3.9.12 https://www.python.org/ R 4.1.2 https://www.r-project.org/ Scanpy 1.9.3 https://scanpy.readthedocs.io/en/stable/ Anndata 0.8.0 https://anndata.readthedocs.io/en/latest/ Umap 0.5.3 https://umap-learn.readthedocs.io/en/latest/ Scipy 1.8.0 https://scipy.org/ Numpy 1.21.2 https://numpy.org/ Pandas 1.4.2 https://pandas.pydata.org/ Scikit-learn 1.1.1 https://scikit-learn.org/stable/ Statsmodels 0.13.2 https://www.statsmodels.org/stable/index.html Python-igraph 0.10.2 https://igraph.org/python/

Pynndescent 0.5.7 https://pynndescent.readthedocs.io/en/latest/
Pyslingshot 0.0.1 https://github.com/mossjacob/pyslingshot
Monocle 2.22.0 https://github.com/cole-trapnell-lab/monocle-release
Bbknn 1.5.1 https://github.com/Teichlab/bbknn
Geopandas 0.12.2 http://geopandas.org/
Seaborn 0.11.2 https://seaborn.pydata.org/
Gseapy 1.0.0 https://github.com/zqfang/gseapy
Tidepy 1.3.8 https://jingxinfu.github.io/TIDEpy/
Diopy 0.5.5 https://github.com/JiekaiLab/diopy
Matplotlib 3.7.1 https://matplotlib.org/
Esda 2.4.3 https://github.com/pysal/esda
Libpysal 4.7.0 http://pysal.org/libpysal
Tqdm 4.64.0 https://tqdm.github.io/
Leidenalg 0.9.0 https://github.com/vtraag/leidenalg
StLearn 0.4.7 https://github.com/BiomedicalMachineLearning/stLearn
Statannot 0.2.3 https://github.com/webermarcolivier/statannot

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

The raw sequencing FASTQ files of the scRNA-seq and stereo-seq data generated in this study could be accessed on Genome Sequence Archive (accession number: PRJCA020107, https://ngdc.cncb.ac.cn/gsa-human/browse/HRA005647). The processed matrix of Stereo-seq and scRNAseq was deposited on STOmicsDB69 of China National GenBank Database (accession number: STT0000036, https://db.cngb.org/stomics/project/STT0000036). The publicly available TCGA with Z-scored gene expression matrix and the MSI scores could be downloaded through https://www.cbioportal.org/study/summary?id=coadread_tcga_pan_can_atlas_2018. The publicly available single cell RNA-seq data could be downloaded through the following links: CRC (gene expression omnibus: GSE205506): https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE20550616; RCC (single cell portal: SCP1288): https://singlecell.broadinstitute.org/single_cell/study/SCP1288/tumor-and-immune-reprogramming-during-immunotherapy-in-advanced-renal-cell-carcinoma#study-download70; melanoma (gene expression omnibus: GSE115978): https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE115978 71. Source data are provided as a Source Data file. The remaining data are available within the Article, Supplementary Information or Source Data file.

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 sample size calculation was performed. We included 2-6 independent specimens in each group for all genomics data generated in this study, which is similar to other studies in this field (e.g.,https://www.nature.com/articles/s41467-022-29366-6; https://www.nature.com/articles/s41421-021-00312-y?fromPaywallRec=false)
Data exclusions	No specimen was excluded, while the data was quality controlled before analyais. For stereo-seq, the spot size was set to bin100 (50μm x 50μm). The spots with gene numbers lower than 500 were filtered out. For scRNA-seq, cells were filtered based on gene number and mitochondria-related gene percentage. Cells with a gene count lower than 300 or higher than 5000, or with over 50% mitochondria-related genes, were excluded from analysis.
Replication	For the genomics data, the key conclusions in this study was supported using external datasets. For the wet lab results (e.g., ELISA, qRT-PCR, etc), all results were repeated performed at least 3 times before solid conclusion was drawn. The replicates could be seen in Supplementary Fig 9.
Randomization	No randomization was performed in this study. Experimental groups were defined based on mismatch repair status and microsatellite instability evaluation.
Blinding	No blinding was deemed necessary for this study, because the d/pMMR status and immunotherapeutic responses were the prerequisites to develop this study. The investigators were blinded to other information, such as gender and age. Only d/pMMR and immunotherapeutic

responses were known before the data analysis.

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
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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

Antibodies

Antibodies used

The antibodies for mIF staining was listed as follow:
antibodies; resource; identifier
Anti-PD1 rabbit mAb; Cell Signaling; CST86163; 1:200
Anti-COL1A1 rabbit mAb; Cell Signaling; CST72026S; 1:500
Anti-IL-8 rabbit mAb; Cell Signaling; CST94407; 1:200
Anti-PDL1 antibody; Abcam; Ab213524; 1:200
Anti-pan cytokeratin antibody; Abcam; Ab7753; 1:200

Validation

All antibodies used in this study was previously validated in human cells, which could be seen by checking the following literatures:
Anti-PD1 rabbit mAb; Cell Signaling; CST86163 PMID: 38040418
Anti-COL1A1 rabbit mAb; Cell Signaling; CST72026S PMID: 36093604
Anti-IL-8 rabbit mAb; Cell Signaling; CST94407 PMID: 35731343
Anti-PDL1 antibody; Abcam; Ab213524 PMID: 35139382
Anti-pan cytokeratin antibody; Abcam; Ab7753 PMID: 35275208

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

Human colorectal cancer cell lines HCT116, HCT8, HCT15, SW480, SW620 and HT29 were obtained from the American Type Culture Collection (ATCC). Human CAF cell line was obtained from Guangzhou Saliai Stem cell Science and Technology (Cat#CatiCel1-0030a). The cells were cultured in DMEM or RPMI-1640 supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin at 37°C incubator containing 5% CO2.

Authentication

All commercially available cell lines in our study were authorized by the American Type Culture Collection (ATCC) and confirmed by short tandem repeat (STR) profiling.

Mycoplasma contamination

To monitor the mycoplasma contamination status of the cell lines, we routinely perform mycoplasma detection tests (Vazyme, Cat#D101-01) every month. For cell lines found to be contaminated, we implemented eradication procedures (MCE, Cat# HY-K1059) and re-evaluated the contamination status two weeks post-treatment. Only cell lines with negative mycoplasma status were used in this study.

Commonly misidentified lines
(See ICLAC register)

We did not use any commonly misidentified cell lines in this study.

Human research participants

Policy information about studies involving human research participants

Population characteristics	No particular population characteristics was involved in this study, except for mismatch repair status and microsatellite instability.
Recruitment	All pMMR patients with colorectal cancer that was planed to undergo primary surgery and all dMMR patients that was planed to accept anti-PD1 treatement could be asked to be recruited. All participants provided written informed consent and the clinical information was collected at Sun Yat-Sen University cancer center.
Ethics oversight	This study was done in accordance with the Declaration of Helsinki (B2023-178-01). The protocol was reviewed and approved by The Institutional Review Board of BGI Ethical (BGI-IRB21083-T1).

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