nature portfolio

Corresponding author(s): Dr Lydia Lynch

Last updated by author(s): 2023/05/05

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	Single cell transcriptomics was performed on the 10X platform. Preliminary data quality control was performed by the BWH Single cell genomics core using Cell Ranger. Gene expression matrices were generated and provided to us for further analysis. Bulk RNA seq was performed by the DFCI transcriptomic core facilty. Data was provided as raw read counts and STAR count matrices. Flow cytometry data was acquired using BD Fortessa and CYtek Aurora.
Data analysis	Single cell RNAseq data was analysed using Seurat package in R, or Scanpy in Pyhton. Data plotting was performed using ggplot2 and viridis (0.4.2) packages. Bulk RNAseq data was analysed using VIPER data analysis pipeline and DESEQ2 for differential gene analysis. PAthway analysis was performed using GOterm and KEGG pathway analysis packages. Non-transcriptomic data was analysed in Graphpad Prism 9.Data collected in FCS files, analysed using FlowJo V10.8.

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

Sequencing data from this study has been deposited in the NCB GEO database GSE210040. Furthermore data from this study has been provided in the Supplemental Information. Publicly available data was used to generate survival curves, this available from the NCI Genetic Data Commons (https://portal.gdc.cancer.gov/). Further scRNAseq datasets were used from published sources, including GSE178341, and GSE216534.

Human research participants

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

Reporting on sex and gender	Sex and gender were not addressed in this manuscript as patient data and healthy donors were anonymised, and consent was for sharing individual level data. The sex and gender of participants was therefore not recorded.
Population characteristics	Molecular diagnosis of mismatch repair status was recorded for each patient. However, mismatch repair status was associated with different immune phenotypes.
Recruitment	Patients attending St. Vincent's University Hospital (Dublin, Ireland) for partial colectomy or the Mater Misericordiae University Hospital (Dublin, Ireland) for hysterectomy were eligible for inclusion in this study. Fifteen CRC patients and nine endometrial cancer patients were recruited prospectively. All patients provided written informed consent. Patients were recruited on a rolling basis. No selection criteria were applied to patients other than diagnosis and consent. 15 and 9 consecutive patients were used for this study who met these criteria.
Ethics oversight	This study was carried out in accordance with the declaration of Helsinki and was reviewed and approved by the ethics committees of St. Vincent's University Hospital, Mater Misericordiae University Hospital and Trinity College Dublin, and Brigham and Women's Hospital.

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	For single cell RNAseq experiments power calculations were performed using SCOPIT V1.1.4. For a bulk RNAseq a pilot study was conducted using 2 healthy donors polarised under different conditions. This data was then quantified by QPCR. For all other experiments power calculations were performed where possible, however in some instances insufficient patient samples were available for fully powered experiments. Where this occurred results were replicated in independent experiments and in larger publicily available datasets where available.
Data exclusions	No data was excluded from the analyses in this manuscript
Replication	Where possible experiments were repeated at least 3 times to confirm the reproducibility of the results. Patient samples were collected over a significant period of time, therefore samples were stored and analyzed in batches. This was carried out at least 3 times, ensuring experimental conditions were repeatable. All experiments were reproducible
Randomization	Randomization was used in adoptive transfer models. Mice were randomly assigned to PBS control group, Gen 1 or Gen 2 cell treatment groups. For other experiments no randomization was performed as all patient samples and donors were tested in each condition analysed. Covariates were not available due to anonymisation.
Blinding	Blinding was not applied in this study as no subjective data was collected which may have been influenced by observer bias. Human data was also anonymised to prevent identification of patients.

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

Target supplier clone
CD45 BioLegend H130 1:100
CD3 BioLegend SK7 1:200
CD8 BioLegend RPA-T8 1:100
CD56 BioLegend 5.1H11 1:100
VD1 Invitrogen TS8.2 1:100
VD2 BioLegend B61:100
TIGIT BD 741182 1:100
LAG3 BioLegend 11C3C65 1:100
TIM3 BioLegend F38-2F2 1:100
PD1 BioLegend EH12.2H7 1:100
TNFa Invitrogen MAb11 1:100
IL17A Invitrogen eBio64Dec7 1:100
AREG Invitrogen AREG559 1:100
IFNg BD B27 1:100
Vd1 Miltenyi Biotech REA173 1:200

Validation

All antibodies were validated by manufacturers and information is available from the suppliers website for each clone.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)

SW480, HCT116 ATCC

Authentication

Cell lines were purchased from ATCC directly. Authentication was performed at the source.

Mycoplasma contamination

Cell lines were tested for mycoplasma by PCR and returned negative results.

Commonly misidentified lines
(See ICLAC register)

No commonly misidenitifed cell lines were used in this manuscriptCell

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

NOD-scid IL2rgnull Transgenic Human IL15(NSG-Tg(HuIL15)) mice were used for adoptive transfer models.8-10 weeks old

Wild animals

No wild animals used

Reporting on sex

Adoptive transfer models were only performed in male mice. It was not possible to produce sufficient cells from humans to perform this study in both male and female mice. No notable differences in the growth of PDX has been observed in mice of different sexes. Male mice were selected for these studies, therefore no reporting of sex differences were possible.

Field-collected samples

No field collected samples used.

Ethics oversight

All animal use was in accordance with the guidelines of the Animal Care and Use Committee of the University of Massachusetts Chan

Ethics oversight	Medical School and conformed to the recommendations in the Guide for the Care and Use of Laboratory Animals (Institute of Laboratory Animal Resources, National Research Council, National Academy of Sciences, 1996).

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 suspensions were prepared from a number of sources including primary human tumors, human blood samples, patient derived xenograft tumors, and cell cultures. Surgical specimens were stored in RPMI+10% FCS for transport to the laboratory. Tumors and healthy tissue were finely minced using a dissection scissors and placed in cryostor CS10 reagent for cold storage at -80C. Sample were later thawed at 37C, cryostor was rinsed from the tissue using warm RPMI. Tissue samples were further mechanically dissociated and placed in an enzymatic digestion mixture (Collagenase type IV 0.05%, DNase in RPMI) to produce a single cell solution. Digestion solution was incubated at 37C for 45mins agitated at 180rpm in a shaking incubator. Digested tissue was strained through a 100µm filter, followed by a 70µm filter. RBC lysis was performed using ACK lysis solution for 5mins at room temperature. PBMCs were isolated from anonymized healthy blood donors attending the Kraft blood donor center. PBMCs were separated by density gradient centrifugation using Ficoll-Paque™ PLUS (GE Healthcare) and residual red blood cells were removed by adding red cell lysis solution. Cells from cell culture experiments were collected, pelleted, washed in PBS and resuspended for FACS staining
Instrument	BD Fortessa LSR 2 or BD fortessa X-20 or Cytek Aurora
Software	Data was collected using BD FACS DIVA or SpectroFlo software. Data was analysed using FlowJo V10.1
Cell population abundance	In human samples population abundance varied by cell type and patient and is graphed in the manuscript.
Gating strategy	Lymphocytes were gated using FSC/SCC profile. singlets were gated on based on FSH/FSC-A. Live (Zomebie live dead stain negative) CD45+ cells were then gated on. Lymphocytes were subsetted based on CD56 and CD3 (CD56+ CD3- NK cells, CD56 +/-, CD3+ T cells). MAIT cells were identified based on co-expression of CD161 and TCRVa7.2. After exclusion of MAIT cells, gd T cell subsets were identified based on TCR Vd1, Vd2 or Vd3 positivity. After exclusion of gd T cell subsets, CD4 and CD8 + T cells were gated respectively.

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