nature research

Corresponding author(s): Aadel A. Chaudhuri; Aaron M. Newman

Last updated by author(s): Nov 7, 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 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	No specific software was used for data collection beyond software associated with commercially available hardware (Agilent 2100 Bioanalyzer software B.0.2.11 , Fluidigm Helios CyTOF software v7.0, MoFlo Legacy Summit software, Illumina NovaSeq control software v1.6, Illumina HiSeq 2500 control software v2.2.68).
Data analysis	Software packages used in this study are detailed in Methods, including Beckman Coulter Cytobank v8.0 or v8.1, Astrolabe Cytometry Platform v3.6 or v4, FlowJo v10, Cell Ranger v3.1.0 or v5.0.1, Loupe V(D)J Browser v3.0.0, Salmon v0.12.0, MiXCR v3.0.125, CIBERSORTx v1.0.41, Skewer v0.2.2, GSEA v4.1.0, GSEAPreranked v7.1.0, GraphPad PRISM v8 or v9, and various R v3.5.1+ packages (e.g., affy v1.6.0, Seurat [v3.1.5, v3.2.1, v4.0.1], tximport v1.10.1, vegan v2.5-6, immunarch v0.6.5, escape v1.0.1, pwr v1.3-0, dplyr v1.0.7). Custom scripts developed in this study are publicly available from https://doi.org/10.25936/f3np-k536.

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 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

All bulk and single-cell expression data generated in this work have been deposited in the Gene Expression Omnibus with accession number GSE186144. Published expression datasets analyzed in this work (Supplementary Table 20) are available from the Gene Expression Omnibus with accession numbers GSE50772, GSE61635,

GSE72509, GSE126124, GSE3365, and GSE100833. Processed CyTOF and immunoSEQ® data are publicly available from https://doi.org/10.25936/f3np-k536. Additional data supporting the findings in this work are available in the main text, figures, extended data, and supplementary files. All reasonable requests for raw data will be promptly reviewed by the corresponding authors to determine whether the request is subject to any confidentiality obligations. Any data that can be shared will be released via a material transfer agreement.

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	Baseline peripheral blood samples for a single-cell discovery cohort were obtained as a retrospective, exploratory study from 18 patients with metastatic melanoma (Fig. 1; Fig. 2a). In this cohort, the association between effector memory CD4 T cell abundance (CyTOF) and severe irAE development had an effect size of 1.99 (Fig. 2d), calculated using pwr v1.3-0 in R. Bulk cohorts 1 and 2 were designed to satisfy this effect size requirement at $\alpha = 0.05$ and $1-\beta = 0.8$ while emphasizing specificity in bulk cohort 1 (number of patients without severe irAE > number of patients with severe irAE) and balance in bulk cohort 2 (number of patients without severe irAE ≈ number of patients with severe irAE) as described in the Methods.
Data exclusions	All analyzed genomic data generated in this work satisfied quality control thresholds described in the Methods and Figure 1.
Replication	We analyzed three separate cohorts of melanoma patients, one by single-cell analysis and the other two by bulk RNA sequencing. The finding of increased activated CD4 memory T cell abundance and elevated T cell receptor diversity in patients who developed severe immune-related adverse events remained consistent across these three independent experiments.
Randomization	Sample groups were determined according to the experimental question and known or predetermined biological or clinical phenotypes. No randomization was applied as our cohorts were retrospective.
Blinding	The investigators were not blinded to group allocation during data collection or analysis, but all cohorts analyzed in this work were independently generated without prior knowledge of cell type composition or TCR diversity. CyTOF gating was performed in a blinded fashion to corroborate Astrolabe results as shown in Supplementary Fig. 2a-b and described in the Methods.

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	Antibodies, including supplier name and clone name, are listed in Supplementary Table 4 and in the Methods. Antibodies for mass cytometry used in this work were: CCR4 (metal 153Eu, clone L291H4, catalog no. 3153030A, vendor Fluidigm); CD11A (metal 152Sm, clone HI111, catalog no. 301233, vendor BioLegend); CD11B (metal 209Bi, clone ICRF44, catalog no. 3209003B, vendor Fluidigm); CD11C (metal 159Tb, clone Bu15, catalog no. 3159001B, vendor Fluidigm); CD127 (metal 143Nd, clone A019D5, catalog no. 3143012B, vendor Fluidigm); CD14 (metal 151Eu, clone M5E2, catalog no. 3151009B, vendor Fluidigm); CD16 (metal 148Nd, clone 3G8, catalog no. 3148004B, vendor Fluidigm); CD19 (metal 142Nd, clone HIB19, catalog no. 3142001B, vendor Fluidigm); CD20 (metal 147Sm, clone 2H7, catalog no. 3147001B, vendor Fluidigm); CD25 (metal 169Tm, clone 2A3, catalog no. 3169003B, vendor Fluidigm); CD27 (metal 155Gd, clone L128, catalog no. 3155001B, vendor Fluidigm); CD279 (metal 174Yb, clone EH12.2H7, catalog no. 3174020B, vendor Fluidigm); CD28 (metal 145Nd, clone CD28.2, catalog no. 302902, vendor BioLegend); CD3 (metal 170Er, clone UCHT1, catalog no. 3170001B, vendor Fluidigm); CD38 (metal 167Er, clone HIT2, catalog no. 3167001B, vendor Fluidigm); CD4 (metal 176Yb, clone RPA-T4, catalog no. 3176010B, vendor Fluidigm); CD45 (metal 89Y, clone HI30, catalog no. 3089003B, vendor Fluidigm);

CD45RA (metal 154Sm, clone HI100, catalog no. 304143, vendor Biolegend); CD45RO (metal 149Sm, clone UCHL1, catalog no. 3149001B, vendor Fluidigm); CD56 (metal 163Dy, clone NCAM16.2, catalog no. 3163007B, vendor Fluidigm); CD69 (metal 144Nd, clone FN50, catalog no. 3144018B, vendor Fluidigm); CD8A (metal 146Nd, clone RPA-T8, catalog no. 3146001B, vendor Fluidigm); CTLA4 (metal 161Dy, clone 14D3, catalog no. 3161004B, vendor Fluidigm); EOMES (metal 165Ho, clone WD1928, catalog no. 14-4877-80, vendor Thermo Fisher); FAS (metal 164Dy, clone DX2, catalog no. 3164008B, vendor Fluidigm); HLA-DR (metal 173Yb, clone L243, catalog no. 3173005B, vendor Fluidigm); IFNG (metal 158Gd, clone B27, catalog no. 3158017B, vendor Fluidigm); IGD (metal 141Pr, clone IA6-2, catalog no. 348235, vendor BioLegend); IGM (metal 172Yb, clone MHM-88, catalog no. 3172004B, vendor Fluidigm); LAG3 (metal 150Nd, clone 11C3C65, catalog no. 3150030B, vendor Fluidigm); MKI67 (metal 168Er, clone Ki-67, catalog no. 3168001B, vendor Fluidigm); PDL1 (metal 156Gd, clone 29E.2A3, catalog no. 3156026B, vendor Fluidigm); PRF1 (metal 175Lu, clone B-D48, catalog no. 3175004B, vendor Fluidigm); TBET (metal 160Gd, clone 4B10, catalog no. 3160010B, vendor Fluidigm); TIM3 (metal 166Er, clone D5D5R, catalog no. 45208S, vendor Cell Signaling Technology). Antibodies for flow cytometry used in this work were: FITC-conjugated anti-human CD45 (clone 2D1, catalog no. 368507, vendor BioLegend); AF700-conjugated anti-human CD3 (clone OKT3, catalog no. 31733925, vendor BioLegend); APC-conjugated anti-human CD4 (clone OKT4, catalog no. 317415, vendor BioLegend); PE/Cy7-conjugated anti-human CD8 (clone SK1, catalog no. 980910, vendor BioLegend); APC-Cy7-conjugated anti-human CD19 (clone HIB19, catalog no. 302218, vendor BioLegend); PerCp/Cy5.5-conjugated anti-human CD14 (clone HCD14, catalog no. 325621, vendor BioLegend); BV605-conjugated anti-human CD56 (clone 5.1H11, catalog no. 362537, vendor BioLegend). Each flow cytometry antibody was diluted as follows: 5uL antibody per million cells per 100uL staining mixture.

Validation

Antibodies were validated by the manufacturer, and by correlating cytometry results with orthogonal methods (i.e., single-cell RNA sequencing, CIBERSORTx deconvolution) as described in the Methods. Per the manufacturer, BioLegend tests their antibody clones in a variety of assays to see which applications they are suited for, and by cross-validating clones across orthogonal testing methods. For mass cytometry experiments, conjugated antibodies were validated and titrated using CD3-stimulated and unstimulated peripheral blood mononuclear cells with the appropriate positive and negative controls as recommended by the Fluidigm Maxpar Antibody Labeling User Guide.

Human research participants

Policy information about studies involving human research participants

Population characteristics

Eligible patients for this study were age >18 years with metastatic melanoma treated with standard immune checkpoint inhibitor treatment comprising either anti-PD1 blockade (nivolumab or pembrolizumab) or combination immune checkpoint blockade (anti-PD1 [nivolumab] and anti-CTLA-4 [ipilimumab]; Fig. 1). Ninety percent of patients were naïve to any prior immune checkpoint blockade at the time of pretreatment blood collection (Supplementary Table 1). The median age of the study population was 65 years, and the gender distribution was 72% male and 28% female. Population characteristics are described in detail in the Methods, Figure 1, Figure 2a, and in Supplementary Tables 1, 2, 3, 7, 9 and 18.

Recruitment

This study utilized a retrospective design using banked clinical samples. We identified 78 patients with metastatic melanoma, 71 of whom were evaluable following exclusion criteria (Fig. 1; Supplementary Table 1). Among these patients, 33 were treated with anti-PD-1 monotherapy, 38 were treated with anti-PD-1 plus anti-CTLA-4 combination therapy, and 90% had no prior immune checkpoint inhibitor (ICI) history (Supplementary Table 1). All patients were monitored closely during and after ICI treatment for immune-related adverse event (irAE) development (median follow-up time of 14.9 months; median time to grade 3+ irAE of 1.5 months). We stratified the 71 patients into three non-overlapping cohorts, a single-cell discovery cohort and a larger bulk cohort divided into training and validation sets (Fig. 1; Supplementary Table 1). We observed no significant differences in the fraction of activated CD4 memory T cells (as a function of total peripheral blood mononuclear cells) across the three different modalities utilized in this study (CyTOF, scRNA-seq, CIBERSORTx) (Extended Data Fig. 5f). In line with this, pooling CIBERSORTx and CyTOF estimates of activated CD4 memory T cells across all evaluable patients in this work (n=71) confirmed the significant association between baseline levels of activated CD4 memory T cells and severe irAE development.

Ethics oversight

Samples analyzed in this study were collected with informed consent for research use and were approved by the Yale University Institutional Review Board, or the Washington University School of Medicine Institutional Review Board, in accordance with the Declaration of Helsinki as part of observational registry studies focusing on melanoma.

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

Peripheral blood specimens were collected in K2EDTA Vacutainer tubes (Becton Dickinson, Franklin Lakes, NJ) and processed within 1 hour of phlebotomy. Peripheral blood mononuclear cell (PBMC) extraction was by the Lymphoprep (Stemcell Technologies, Vancouver, Canada) protocol, which was applied according to the manufacturer's instructions. PBMC samples were cryopreserved in 10% DMSO / 90% fetal bovine serum. Cryovials were placed in Nalgene Mr. Frosty containers (Thermo Fisher Scientific, Waltham, MA) for 24 hours, then stored in liquid nitrogen until thawing for flow cytometry. Cryopreserved

	PBMCs were thawed by holding cryovials in a 37-degree Celsius water bath for 1-2 minutes without submerging the cap. Subsequently, 2-5 million PBMCs were treated with TruStain FcX block (BioLegend, San Diego, CA) for 10 minutes at room temperature to block Fc receptors, and then stained with fluorophore-tagged surface antibodies for 30 minutes at room temperature. See Methods for details.
Instrument	MoFlo Legacy instrument (Beckman Coulter, Pasadena, CA).
Software	MoFlo Legacy Summit software (Beckman Coulter, Pasadena, CA), FlowJo version 10 (BD Biosciences, San Jose, CA).
Cell population abundance	Major lymphocyte population abundances, including B cells (mean 13% of total lymphocytes), CD4 T cells (mean 53% of total lymphocytes), CD8 T cells (mean 24% of total lymphocytes), and NK cells (mean 10% of total lymphocytes) were enumerated from healthy donor peripheral blood mononuclear cells as a percentage of total lymphocytes, and are shown in Supplementary Figure 3 and in Extended Data Figure 5e. No cell sorting was performed.
Gating strategy	Following exclusion of DAPI-positive cells and putative doublets based on forward and side scatter analysis, major lymphocyte populations including B cells, CD4 T cells, CD8 T cells and NK cells were enumerated as a percentage of total lymphocytes using FlowJo version 10 (BD Biosciences, San Jose, CA) as shown in Supplementary Figure 3.

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