nature portfolio

Corresponding author(s): Vinod Balachandran

Last updated by author(s): Nov 12, 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

TCR Vβ sequencing – We used Dropsense 96 to quantify samples, dilute standard concentrations, and prepare libraries. We then generated sample data using an immunoSEQ Assay (Adaptive Biotechnologies).

Single cell RNA/TCR sequencing – We assessed the quality of PCR amplifications on an Agilent Bioanalyzer 2100. Pooled libraries were clusteredon an Illumina Novaseq pair end flow cell sequenced for 656 28-10-10-91, to obtain about 350 million clusters per sample. We processed sequencing images using Illumina's Real Time Analysis software (RTA). We used 10x Genomics Cell Ranger Single Cell Software suite v6.0.0 to demultiplex, align (hg19), filter, UMI count, and single-cell 5' end gene count samples, assemble TCRs, annotate paired VDJ, and perform quality control per manufacturer's parameters.

Flow cytometry and cell sorting – We analyzed PBMCs, transduced cells, and activated T cells on a FACS LSR Fortessa (BD Biosciences) using FACSDiva software (version 8.0.1, BD Bioscience). We sorted cells using and Aria Cell sorter (BD Biosciences).

Data analysis

All analyses were performed using GraphPad Prism (version 10.1.1) or Python (version 3.11.6).

Flow cytometry – We analyzed data using FlowJo (version 10).

Time integrated cumulative vaccine frequency – The numerical integration was done using scipy.integrate.cumtrapz.

Analysis of metaclones – We used GLIPH2.0.

Single cell analysis – We performed quality control by filtering count matrices generated by Cellranger (v6) and integrating them into a single

matrix using Scanpy. Cells were filtered by having a sequenced high confidence TRB CDR3 region called from Cellranger VDJ (v6). Phenotyping was done using the GeneVector pipeline.

Whole exome sequencing, mutation identification, and phylogeny – we aligned reads to the reference human genome (hg19) using BWA (v0.7.17), and marked duplicates by picard-2.11.0 MarkDuplicates. Then, we performed indel realignment by the Genome Analysis toolkit (GenomeAnalysisTK-3.8-1)44 using RealignerTargetCreator and IndelRealigner. We used the 1000 genome phase1 indel (1000G_phase1.indels.b37.vcf) and Mills indel calls (Mills_and_1000G_gold_standard.indels.b37.vcf) as references. We then calibrated base quality by the GATK's BaseRecalibrator using dbSNP version 138 as reference source. After a pre-processing step, we employed MuTect 1.1.745 and Strelka 1.0.1546 to identify SNVs and indels in tumor samples compared to normal. Using bam-readcounts (v0.8). We used phyloWGS (v1.0-rc2) to reconstruct the phylogeny tree and also to infer CCF on the merged set of mutations. We visualized trees with R package data.tree (v1.1.0).

Predicted HLA-I binding affinity – We used NetMHCPan4.1.

Analyses of TCR Vβ sequencing and single cell RNA/TCR sequencing data used custom code:
Code for CloneTrack is available at https://github.com/zsethna/CloneTrack
Code for PhenoTrack is available at https://github.com/zsethna/PhenoTrack
Code to identify GLIPH 2 metaclones is available at http://50.255.35.37:8080/tools

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

All single-cell sequencing data are available at GEO (accession number GSE222011). All experimental source data are provided. Per institutional data sharing policies, requests for deidentified individual participant data can be made beginning 12 months after publication and for up to 36 months post-publication. Deidentified individual participant data reported in the manuscript will be shared under the terms of a Data Use Agreement that addresses the privacy and security requirements for safeguarding the data under HIPAA and MSK policies, and may only be used for approved proposals. Requests may be made to balachav@mskcc.org.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	The sex of all patients enrolled in the trial is reported in the primary article for this dataset - Rojas et al. Nature 2023 https://www.nature.com/articles/s41586-023-06063-y
Reporting on race, ethnicity, or other socially relevant groupings	The race/ethnicity of all patients enrolled in the trial is reported in the primary article for this dataset - Rojas et al. Nature 2023 https://www.nature.com/articles/s41586-023-06063-y
Population characteristics	Patients were stratified by immune response to autogene cevumeran as described in Methods
Recruitment	Full details of inclusion/enrollment and recruitment are provided in the clinical trial protocol, and the primary article for this dataset - Rojas et al. Nature 2023 https://www.nature.com/articles/s41586-023-06063-y
Ethics oversight	The study was approved by the institutional review board (IRB) at Memorial Sloan Kettering Cancer Center (MSK), the United States Federal Drug Administration (FDA), and was registered on clinicaltrials.gov (NCT04161755). All participants provided written informed consent.

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 the primary dataset, we targeted to accrue a total of 20 evaluable patients based on estimated sample size needed to evaluate our primary endpoint of safety. Additional details can be found in the primary article – Rojas et al. Nature 2023 https://www.nature.com/articles/s41586-023-06063-y
Data exclusions	No data were excluded.
Replication	All findings were reproducible. All experimental replicates, performed as independent experiments in individual patient samples, are outlined in the figures and figure legends. Technical replicates are also indicated in the figure legends when appropriate.
Randomization	Randomization is not applicable, as the primary study source for this dataset was a single-arm, phase-I clinical trial.
Blinding	In the clinical trial underlying the primary dataset, the investigators were blinded to survival events during immunologic response determination. Other experiments: no blinding, as researches needed access to the sample information to perform the experiments.

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
☒	☐ Plants

Methods

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

Antibodies

Antibodies used

CD137 - clone 4B4-1, PE (Biolegend Cat# 309804); 3µl/sample
CD3 - clone SK-7, PE-Cy7 (Biolegend Cat# 344816); 4µl/sample
CD4 - clone OKT4, BV650 (Biolegend Cat# 317436); 2 µl/sample
CD45 - clone 2D1, Alexa Fluor 700 (Biolegend Cat# 368514); 5 µl/sample
CD56 - clone HCD56, BV605 (Biolegend Cat# 318334); 2 µl/sample
CD8 - clone SK1, APC-Cy7 (Biolegend Cat# 344714); 2 µl/sample
CD8 - clone SK1, Alexa Fluor 700 (Biolegend Cat# 344724); 2 µl/sample
IFNγ - clone 4S.B3, BV421 (Biolegend Cat# 502532); 5 µl/sample
CD107a - clone H4A3, PE (BD Biosciences Cat# 555801); 20 µl/ sample
CD56 - clone NCAM16.2, BV786 (BD Biosciences Cat# 564058); 5 µl/sample
TNFα - clone MAb11, APC (BD Biosciences Cat# 554514); 2.5 µl/sample
mTCR - clone H57-597, PE-Cy5 (Biolegend Cat #109209), 0.5 µl/sample
HLA-ABC - clone W6/32, APC (Biolegend Cat #311409), 3 µl/sample
DAPI - (BD Biosciences Cat# 564907); 0.2 µl/sample
Fixable viability Dye - eFLuor 520 (Invitrogen Cat# 65-0867-14); 1.0 µl/sample

Validation

All antibodies were validated by the manufacturer and used per their instructions. In our experiments, biological negative controls were included. Additional information on validation can be found on the manufacturer's websites, below.
CD137 - clone 4B4-1, PE (Biolegend Cat# 309804); https://www.biolegend.com/en-ie/products/pe-anti-human-cd137-4-1bb-antibody-1510?GroupID=BLG2203
CD3 - clone SK-7, PE-Cy7 (Biolegend Cat# 344816); https://www.biolegend.com/en-ie/products/pe-cyanine7-anti-human-cd3-antibody-6934
CD4 - clone OKT4, BV650 (Biolegend Cat# 317436); https://www.biolegend.com/en-ie/products/brilliant-violet-650-anti-human-cd4-antibody-7786
CD45 - clone 2D1, Alexa Fluor 700 (Biolegend Cat# 368514); https://www.biolegend.com/en-ie/products/alexa-fluor-700-anti-human-cd45-antibody-12399
CD56 - clone HCD56, BV605 (Biolegend Cat# 318334); https://www.biolegend.com/en-ie/products/brilliant-violet-605-anti-human-cd56-ncam-antibody-7668
CD8 - clone SK1, APC-Cy7 (Biolegend Cat# 344714); https://www.biolegend.com/en-ie/products/apc-cyanine7-anti-human-cd8-antibody-6391

CD8 - clone SK1, Alexa Fluor 700 (Biolegend Cat# 344724); https://www.biolegend.com/en-ie/products/alexa-fluor-700-anti-human-cd8-antibody-9062
IFNγ - clone 4S.B3, BV421 (Biolegend Cat# 502532); https://www.biolegend.com/en-ie/products/brilliant-violet-421-anti-human-ifn-gamma-antibody-7189
CD107a - clone H4A3, PE (BD Biosciences Cat# 555801); https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/pe-mouse-anti-human-cd107a.555801
CD56 - clone NCAM16.2, BV786 (BD Biosciences Cat# 564058); https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/bv786-mouse-anti-human-cd56.564058
TNFα - clone MAb11, APC (BD Biosciences Cat# 554514); https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/apc-mouse-anti-human-tnf.554514
mTCR - clone H57-597, PE-Cy5 (Biolegend Cat #109209) https://www.biolegend.com/de-at/products/pe-cyanine5-anti-mouse-tcr-beta-chain-antibody-273
HLA-ABC - clone W6/32, APC (Biolegend Cat #311409), https://www.biolegend.com/de-at/products/apc-anti-human-hla-a-b-c-antibody-1870
DAPI - (BD Biosciences Cat# 564907); https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/dapi-solution.564907
Fixable viability Dye - eFLuor 520 (Invitrogen Cat# 65-0867-14); https://www.thermofisher.com/order/catalog/product/65-0867-14

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	We purified patient peripheral blood mononuclear cells (PBMCs) by density centrifugation over Ficoll-Paque Plus (GE Healthcare) and cultured purified cells in RPMI medium supplemented with 10% fetal bovine serum (FBS, (Nucleus Biologics), 1 mM sodium pyruvate, 2 mM L-glutamine, non-essential amino acids and 2-mercaptoethanol, 100 U ml-1 penicillin-streptomycin (MSK medium preparation core facility), with 100 U ml-1 IL-2 (Peprotech), and 100 U ml-1 IL-15 (Peprotech) added every other day. We purchased pan negative selection purified human T cells (Precision for Medicine). We purchased T2 (174 x CEM.T2) cells (CRL-1992) and K562 cells (CCL-243) from ATCC. 293Vec-RD114 cells (Biovec Pharma) and Phoenix-AMPHO cells (ATCC CRL-3213) were kindly gifted from MSK laboratories.
Authentication	For transduction of T cells, T2 cells, and K562 cells, we authenticated the correct sequence of the retroviral vector by Sanger sequencing before transduction. Then, we validated the generated cell lines by PCR amplification of the inserted transgene and Sanger sequencing. We further validated correct HLA and TCR expression by functional assays with restricted epitopes.
Mycoplasma contamination	PBMCs and donor-derived transduced T cells were not tested for mycoplasma contamination. Cell lines were periodically tested for presence of mycoplasma using the MycoAlert-PlusTM kit (Lonza) at the Antibody and Bioresource Core Facility at MSK.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified lines were used in this study.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	NCT04161755
Study protocol	Study protocol is available (Supplementary File 1).
Data collection	Patients were enrolled from December 2019 to August 2021. Data were collected at MSK during and beyond enrollment period.
Outcomes	Primary endpoint was safety and was assessed using a predefined number of grade 3 adverse events due to atezolizumab and autogene cevumeran per number of patients enrolled (reported in primary article for this dataset – Rojas et al. Nature 2023 https://www.nature.com/articles/s41586-023-06063-y). Secondary endpoints were 18-month recurrence free survival (RFS) and overall survival (OS). We defined recurrence as new lesions by RECIST 1.1, and RFS from either the date of surgery (RFS), or from the date of the last autogene cevumeran priming dose (landmark RFS) to the date of recurrence or death, whichever occurred first. We censored patients without events at the last known date they were recurrence-free. We defined OS from the date of surgery to the date of death. As exploratory endpoints, we measured immune response. Data cut-off was December 1, 2023, extending the median follow-up to 36 months

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.

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

We purified patient peripheral blood mononuclear cells (PBMCs) by density centrifugation over Ficoll-Paque Plus (GE Healthcare) and cultured purified cells in RPMI medium supplemented with 10% fetal bovine serum (FBS, (Nucleus Biologics), 1 mM sodium pyruvate, 2 mM L-glutamine, non-essential amino acids and 2-mercaptoethanol, 100 U ml-1 penicillin-streptomycin (MSK medium preparation core facility), with 100 U ml-1 IL-2 (Peprotech), and 100 U ml-1 IL-15 (Peprotech) added every other day.

We stained cells using antibody cocktails in the dark at 4°C, washed, and analyzed. To examine expression of intracellular markers, we surface-stained, fixed, permeabilized, and stained the cells for intracellular proteins using the Fixation and Permeabilization Buffer Kit as per the manufacturer's recommendations (Invitrogen, MA, USA). Full details are provided in the Methods.

Instrument

We analyzed PBMCs, transduction efficiency, and T cell activation, on a FACS LSR Fortessa (BD Biosciences). We sorted T cells using an Aria Cell sorter (BD Biosciences).

Software

Data was analyzed using FACSDiva software (version 8.0.1, BD Bioscience) and FlowJo (version 10).

Cell population abundance

Representative cell abundance is indicated in Figure 4a and Extended Data Figure 9. Due to the scarcity of the samples, we were not able to confirm the purity of the populations within the post-sorting fractions.

Gating strategy

We gated lymphocytes based on size and complexity gating (SSC-A vs FSC-A). We considered events with high SSC-W and FSC-W and normal SSC-H and FSC-H, respectively, as doublets and excluded them from the analysis. We identified dead cells as positive for Fixable Viability Dye 520 or DAPI and excluded them from the analysis. We used the following definitions: human T cells as live, CD45+, CD56-, CD3+ cells; degranulating CD8+ T cells as live, CD56-, CD3+, CD8+, CD107a+; neopeptide-stimulated CD8+ T cells as live, CD56-, CD3+, CD8+, TNF +/IFN +; and activated transduced T cells as live, CD3+, CD8+, mTCR +, 4-1BB+. We set the gates of activation markers using mock-stimulated T cells (DMSO or irrelevant peptides).

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