nature portfolio

Corresponding author(s): Raghav Sundar

Last updated by author(s): Nov 22, 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	No software was used.
Data analysis	Analyses and data visualisations were done in R (version 4.3.2). For Whole Exome Next Generation Sequencing, raw exome reads were trimmed with Trimmomatic-0.33 (parameters: LEADING: 5; TRAILING: 5; SLIDINGWINDOW: 4:20; MINLEN: 50) and aligned on the GRCh38 human genome using BWA-mem software. Multimapping reads were filtered out using SAMtools 0.1.19. Optical duplicates were marked using Picard's MarkDuplicates tool v2.20. DNA alignments were further optimized at regions around indels, and base scores were recalibrated after the optimization step using Genome Analysis Tool Kit (GATK) software v3.7. Somatic variant calling of single nucleotide variants and small indels was performed using Mutect2, VarScan 2, and Scalpel with default parameters. MuSiCa tool was used to determine the enrichment of genetic signatures in the profiles of somatic mutations detected in each patient. For GeoMx spatial analysis, all analyses were done in R (version 4.3.2) with raw counts exported from DSP Control Center (Nanostring, Seattle, WA, USA). Counts from the retained AOIs were then normalized to the 75th quartile (i.e., Q3 normalization), which were used for PCA (prcomp function from the stats R package [version 4.3.2]) and spatial deconvolution with the safeTME reference matrix (SpatialDecon package [ver. 1.12.3]). Cell type proportion comparisons were made with the propeller.ttest function of the speckle R package (version 1.2.0). Differential gene expression analysis was performed on raw counts with the DESeq2 R package (version 1.42.0). Log2(fold change) values were corrected for noise associated with genes with low counts using the apeglm estimator (42) within DESeq2's lfcShrink function. Gene set enrichment analysis was conducted with the fgsea R package (version 1.28.0) and GO:BP (Gene Ontology: Biological Processes) terms from the msigdbr R package (version 7.5.1) on genes ranked by shrunken log2(fold change) values from DESeq2. For COMET spatial proteomics analysis, we selected 70 or more fields of view (each with 2048 x 2048 pixels) from the multiplex tissue image of each case as regions of interest (ROIs) using QuPath software (v0.5.0). Cell segmentation and feature measurements were performed using an automated tissue image analysis software (cellXpress v2.4.0; https://cellXpress.org; Bioinformatics Institute, Singapore). Positive cells for each marker were identified using a web-based hyperplex tissue scoring system (ImmunoThresholdTM; ImmunoQs Pte. Ltd., Singapore) written in R (version 4.3.2). For each marker, the system employed Gaussian

mixture modeling (GMM) (mixtools v2.0.0) (46) to determine a decision threshold that can separate between cells with ("positive") or without ("negative") staining of the marker. For Visium spatial transcriptomic analysis, using Visium companion H&E-stained tissues captured at 20× magnification, a trained pathologist (JY) and histologist (YZX) developed a supervised machine learning classifier with QuPath software (version 0.3.2) to distinguish between tumor and stroma tissues, while excluding areas with damaged tissue and high concentrations of collagen or muscle. Fiducial and tissue alignment (to locate spatial barcodes in the tissue) were manually performed on the H&E images using the Loupe Browser (10x Genomics), followed by read or UMI alignment to the GRCh38 human reference genome using Space Ranger (10x Genomics). Using the Seurat R package version 4.4.0 (47), we excluded low-quality cells or empty droplets by removing Visium spots containing less than 20 unique genes. Using the sctransform function in Seurat, normalized data based on the 3,000 most variable genes were generated. For cell-type inference, we applied the Microenvironment Cell Populations (MCP) Counter method version 1.2.0 (48) to analyze sctransform-normalized Visium data, utilizing the MCPcounter.estimate function in the MCP Counter R package. Our analysis relied on the safeTME cell profile matrix, which was constructed from both PBMC and cancer tissue sample data. Gene Set Enrichment Analysis (GSEA) was conducted the GSEA function in clusterProfiler R package version 4.6.2 based on Human Molecular Signatures Database (MSigDB) Hallmark and G5:BP gene sets via msigdbr R package version 7.5.1.

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

Spatial proteomic data have been deposited at immunoatlas.org. Other data used in this study are available on request. Correspondence and requests for materials should be addressed to Filippo Pietrantonio.

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	Sex was reported as a baseline characteristic for the study population. In the first analysis of the MAYA trial clinical outcomes reported in JCO, there was no significant association between the sex of patients who enrolled into part 1 or were able to proceed onto part 2 of the study (those who did not have disease progression on part 1). The study focused on responder vs non-responder analysis.
Reporting on race, ethnicity, or other socially relevant groupings	No reporting on race or ethnicity.
Population characteristics	Trial patient baseline characteristics are described in detail in our first publication on the MAYA clinical trial outcomes (https://doi.org/10.1200/JCO.21.02583) and will be included as a supplementary table in this manuscript.
Recruitment	Patients were prescreened for trial eligibility at 12 Italian Centres.
Ethics oversight	The MAYA trial was approved by the ethical committees of all centers and by the Italian regulatory authority. All patients signed the informed consent of the MAYA trial including consent for future translational analyses and all patients signed an additional informed consent for the purpose of this work, while the overall translational study was approved by the IRB (INT 304/20).

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	The sample size was calculated on the basis of the primary end point of 8-month PFS rate. According to our previously published results, the PFS of patients with MGMT-silenced mCRC and clinical benefit from single-agent temozolomide is almost always <8 months. Therefore, we aimed to increase the 8-month PFS rate from 5%to 20% with the combination of temozolomide, nivolumab, and ipilimumab. According to a single-stage design and selecting p0 (8-month PFS in the null hypothesis) = 0.05, and p1 (8-month PFS in the alternative hypothesis) 5 0.20, with 1- sided a- and b-error of 5% and 20%, respectively, a total of 27 patients were required in the second treatment part. The null hypothesis would have been rejected with ≥4 patients progression-free and alive by the 8-month time point.
Data exclusions	No
Replication	NA
Randomization	No randomisation. Single arm phase II study.
Blinding	No blinding, single arm phase II study.

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

Antibody panel for peripheral flow cytometry
S/N Fluorophore Marker Company
1 BB630 CD56 BD Biosciences
2 BB660 CD19 BD Biosciences
3 BB700 TIGIT BD Biosciences
4 BB790 CD45 BD Biosciences
5 BUV395 CD3 BD Biosciences
6 BUV496 Live Dead Thermofisher
7 BUV563 CD16 BD Biosciences
8 BUV615 TIM3 BD Biosciences
9 BUV661 PD-L1 BD Biosciences
10 BUV737 CD123 BD Biosciences
11 BUV805 CD14 BD Biosciences
12 APC-R700 CD8 BD Biosciences
13 R718 CD206 BD Biosciences
14 APC-Vio770 CD172a Miltenyi Biotec
15 Pe-CF594 PD1 BD Biosciences
16 Pe-Cy7 FceR1a Biolegend
17 Pe-Fire810 HLA-DR Biolegend
18 BV510 CD15 BD Biosciences
19 BV570 CD11b Biolegend
20 BV605 CD141 BD Biosciences
21 BV750 CD4 BD Biosciences
22 BV786 CD33 BD Biosciences
Intracellular markers
23 FITC IDO eBioscience
24 APC CD68 Miltenyi Biotec

25 PE Foxp3 BD Biosciences
26 Pe-Cy5 CTLA4 BD Biosciences
27 BV421 GZMB BD Biosciences
28 BV480 Ki67 BD Biosciences
29 BV711 Perforin Biolegend

Validation

Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJEguidelines for publication of clinical research and a completedCONSORT checklist must be included with all submissions.

Clinical trial registration	NCT03832621
Study protocol	https://ascopubs.org/action/downloadSupplement?doi=10.1200%2FJCO.21.02583&file=protocol_JCO.21.02583.pdf
Data collection	Between March 22, 2019, and November 1, 2020, 716 patients with treatment-refractory MSS colorectal cancer were pre-screened, 204 (29%) were molecularly eligible (MGMT-silenced) and a total of 135 patients were enrolled and started single-agent TMZ. The primary analysis was based on a data cutoff date of December 17, 2021 and has been previously published (https://doi.org/10.1200/JCO.21.02583). Here we present the final analysis of the study at the data cutoff date of May 1, 2023.
Outcomes	The primary end point of the trial was investigator-assessed 8-month progression-free survival (PFS) rate in patients who started the second treatment part and was defined as the proportion of patients alive and progression-free by the 8-month time point from the start of the first treatment part. The secondary end points were PFS, defined as the interval from the date of enrollment in the first treatment part to the date of PD by RECIST1.1 and ir-RECIST criteria, or death from any cause; overall survival (OS), defined as the interval from the date of enrollment in the first treatment part to the date of death from any cause or censored to the last followup for alive patients; ORR, defined as the proportion of patients achieving an objective response (complete response or PR) by RECIST1.1 and ir-RECIST criteria using the scan obtained before temozolomide monotherapy as baseline; duration of response (DoR); PFS, ORR, and DoR according to BICR; safety profile and AEs in each Treatment Part according to NCI CTCAE v4.0.

Plants

Seed stocks	NA
Novel plant genotypes	NA
Authentication	NA

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 samples were obtained from patients at 4 timepoints in the study for analysis: (I) Baseline: before temozolomide, (II) Post-TMZ: after 2 cycles of temozolomide, (III) Post-Triplet: after 2 cycles of temozolomide and 2 additional cycles of temozolomide plus nivolumab and ipilimumab, (IV) PD: at time of disease progression. Cells were quick thawed in a 37oC water bath and washed with RPMI + 10% FBS (Gibco). A cell count was performed, and the cells were washed with stain buffer (PBS + 5% FBS + 2mM EDTA). Antibodies against TIGIT, TIM3, PD-L1 and PD1 were added to the cells

	and stained for 10 minutes at room temperature. The remaining surface marker antibodies were then added to the cell suspension and incubated for 20 minutes at 4oC. The cells were washed twice with stain buffer and subsequently fixed and permeabilized with Foxp3/ transcription factor staining buffer set (eBioscience) according to manufacturer's instructions. Antibodies against intracellular markers were then added to the cells and incubated for 30 minutes at room temperature.
Instrument	Samples were analysed on a BD FACSymphony (BD Biosciences).
Software	Data was analysed using FlowJo software (Flowjo LLC).
Cell population abundance	We will upload the raw flow cytometry data detailing the cell population abundances as a supplementary file.
Gating strategy	*Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.*

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