nature portfolio

Corresponding author(s): Raghav Sundar

Last updated by author(s): 29/3/2026

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 classification and thresholding were performed using a proprietary, web-based analysis platform (ImmunoThreshold™; ImmunoQs Pte. Ltd., Singapore). The software was pre-existing and not developed specifically for this study. 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

All data underlying the figures and analyses presented in this manuscript are provided as a Source Data file accompanying this article. The raw bulk RNA-sequencing (FASTQ) data generated in this study were produced by Auristone Pte. Ltd. and have been deposited in the European Genome-phenome Archive (EGA) under accession code EGAD50000002244 (https://ega-archive.org/datasets/EGAD50000002244). The RNA-sequencing data are available under restricted access because they originate from human clinical samples and contain genomic information that may carry a potential risk of participant re-identification. Access can be obtained by submitting a request through the EGA portal using the dataset accession number listed above. Requests from qualified researchers will be reviewed by the Data Access Committee in accordance with the Auristone Data Access Policy and applicable data-protection regulations. The committee aims to respond within 90 working days. Approved users will be granted access for up to one year, with extensions possible upon request. Processed data underlying Supplementary Fig. S3 (xCell enrichment scores derived from bulk RNA-sequencing of pre-treatment tumour samples) are provided in the Source Data file accompanying this paper.

Spatial transcriptomics data generated using the 10x Genomics Visium platform are publicly available in the Gene Expression Omnibus (GEO) under accession number GSE326101 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?&acc=GSE326101). The dataset includes spatially resolved gene expression count matrices in HDF5 format and corresponding high-resolution histology images required for spatial mapping and visualization.

Output data from spatial proteomic profiling performed using the Lunaphore platform have been uploaded and are publicly available (https://immunoatlas.org/NCIR/241129-14).

For samples profiled using the FoundationOne® clinical next-generation sequencing assay (Foundation Medicine, Inc.), the investigators received only the clinical genomic reports generated by Foundation Medicine, which include processed variant calls and copy-number alterations. The underlying raw sequencing data generated by Foundation Medicine were not returned to the investigators and are not held by the authors. Due to Health Insurance Portability and Accountability Act requirements, Foundation Medicine is not authorized to share individualized patient genomic data, which contains potentially identifying or sensitive patient information. Variant and copy-number information extracted from these clinical reports and used for the analyses presented in Fig. 2c are provided in the Source Data file accompanying this manuscript. Requests for access to the underlying raw sequencing data generated by Foundation Medicine should be directed to Foundation Medicine, Inc., which retains custody of these data. Access may be granted subject to the company's data governance policies and applicable patient-privacy regulations. For more information and mechanisms of access, please contact the Foundation Medicine, Inc. Data Governance Council at mailto:data.governance.council@foundationmedicine.com.

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 COMET (Lunaphore) (Supplementary Table 2)
No Marker Clone Company Catalogue Dilution
1 aSMA 1A4 Cell Marque 202M-96
1:2000
2 BDCA-2 10E6.1 Merck MABF94
1:100
3 CD11b EPR1344 Abcam ab133357
1:2000
4 CD11c EP1347Y Abcam ab52632
1:3000
5 CD14 EPR3653 Abcam ab133335
1:2000
6 CD15 Carb-3 Dako M363101-2
1:500
7 CD16 SP175 Cell Marque 116R-16
1:1500
8 CD163 MRQ-26 Cell Marque 163M-15
1:80

9 CD19 EPR5906 Abcam ab134114
1:1500
10 CD20 SP32 Thermo MA1-39416
1:100
11 CD21 EP3093 Cell Marque 121R-16
1:1000
12 CD3 MRQ-39 Cell Marque 103R-96
1:200
13 CD31 EP3095 Abcam ab134168
1:4000
14 CD34 EP88 Cell Marque 134R-16
1:4000
15 CD38 SP149 Cell Marque 118R-15-RUO
1:100
16 CD39 EPR20627 Abcam ab223842
1:3000
17 CD4 EPR6855 Abcam ab133616
1:1000
18 CD45 2B11 + PD7/26 Dako M070101-2
1:150
19 CD45RA HI100 BioLegend 304102
1:3000
20 CD45RO UCHL1 Dako M074201-2
1:1000
21 CD56 MRQ-42 Cell Marque 156R-95
1:5000
22 CD68 KP1 Thermo MA5-13324
1:500
23 CD8 4B11 BioRad MCA1817T
1:250
24 CK AE1/AE3 Dako M351501-2
1:150
25 E-Cadherin 36/E-Cadherin (RUO) BD Bioscience 610182
1:1000
26 Eomes WD1928 Invitrogen 14-4877-82
1:200
27 FoxP3 SP97 Thermo MA5-16365
1:200
28 GZMB EPR20129-217 Abcam ab208586
1:2000
29 HLA-DR TAL 1B5 Santa cruz sc-53319
1:2000
30 ICOS EPR20560 Abcam ab224644
1:100
31 IDO-1 V1NC3IDO Invitrogen 14-9750-82
1:3000
32 Ki67 MIB-1 Dako M724001-2
1:1500
33 LAG-3 17B4 NovusBio NBP1-97657-100
1:8000
34 Lamin A 133A2 Thermo MA1-06101
1:200
35 PD1 EPR4877(2) Abcam ab137132
1:6000
36 PD-L1 (73-10) 73-10 Abcam ab228415
1:2000
37 Podoplanin D2-40 Dako M361929-2
1:50
38 S100 4C4.9 Thermo MA5-12969
1:100
39 Tryptase AA1 BioLegend 369402
1:10000
40 VISTA D1L2G Cell Signalling 64953
1:150

Antibody panel for peripheral flow cytometry (Supplementary Table 3)
S/N Fluorophore Marker Company
S/N Fluorophore Marker Company Dilution
Surface markers

1 BB630 CD56 BD Biosciences 1:100
2 BB660 CD19 BD Biosciences 1:100
3 BB700 TIGIT BD Biosciences 1:50
4 BB790 CD45 BD Biosciences 1:200
5 BUV395 CD3 BD Biosciences 1:50
6 BUV496 Live Dead Thermofisher 1:100
7 BUV563 CD16 BD Biosciences 1:50
8 BUV615 TIM3 BD Biosciences 3:100
9 BUV661 PD-L1 BD Biosciences 1:50
10 BUV737 CD123 BD Biosciences 1:50
11 BUV805 CD14 BD Biosciences 1:50
12 APC-R700 CD8 BD Biosciences 1:50
13 R718 CD206 BD Biosciences 1:50
14 APC-Vio770 CD172a Miltenyi Biotec 3:100
15 Pe-CF594 PD1 BD Biosciences 1:50
16 Pe-Cy7 FceR1a Biolegend 1:50
17 Pe-Fire810 HLA-DR Biolegend 1:50
18 BV510 CD15 BD Biosciences 1:50
19 BV570 CD11b Biolegend 1:25
20 BV605 CD141 BD Biosciences 1:50
21 BV750 CD4 BD Biosciences 1:200
22 BV786 CD33 BD Biosciences 1:50
Intracellular markers
23 FITC IDO eBioscience 1:20
24 APC CD68 Miltenyi Biotec 1:50
25 PE Foxp3 BD Biosciences 1:5
26 Pe-Cy5 CTLA4 BD Biosciences 1:50
27 BV421 GZMB BD Biosciences 1:100
28 BV480 Ki67 BD Biosciences 1:50
29 BV711 Perforin Biolegend 1:100

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 ICMJE guidelines for publication of clinical research and a completed CONSORT 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 3 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. 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	Detailed flow cytometry statistics, including immune cell subset abundances at each timepoint, are provided in the Supplementary Data.
Gating strategy	Time gate was set based on the region of stable PE MFI versus time to eliminate aberrant signal caused by any stream instability. FSC versus SSC gate was set at the expected cells size and granularity as per previous experiments using cells fixed and permeabilized with a commercial Foxp3 staining buffer set to eliminate debris from analysis. Singlet gating via FSC-H versus FSC-A and Live-Dead viability exclusion was then gated to further eliminate aberrant fluorescent signals. The majority of positive and negative population gates were defined on pseudocolour plots by segmenting at the area of lowest density or divided in the white space between the distinct positive and negative populations. These gates were also set in a fashion to be generally applicable to multiple samples in the cohort in order to reduce gating variability between samples, whilst taking into account the expected human biological variability of antigen density. For populations with indistinct resolution, antigen positivity was defined based on internal controls of other non-marker expressing cell subsets, or based on the sample with a distinct lowest negative population and applied to all samples as a means of standardization. Detailed flow cytometry statistics, including immune cell subset abundances at each timepoint, are provided in the Supplementary Data (Sup. Fig. 11).

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