nature portfolio

Peer Review File

Spatial predictors of response to chemo-immunotherapy in microsatellite stable metastatic colorectal cancer

Corresponding Author: Professor Raghav Sundar

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Choo J et al present the results of a translational study associated with the MAYA trial (published in 2022 in JCO) where aimed to elucidate the factors influencing response heterogeneity using multi-omics spatial profiling of samples from patients who participated in the MAYA trial, including baseline and on-treatment tissue and blood specimens.

What are the noteworthy results?
The fact that MSS CRC tumours are largely unresponsive to treatment emphasises the huge need to identify novel efficacious therapeutic approaches. The use of temozolomide to sensitise MSS MGMT silenced cold tumours to immunotherapies is a novel and interesting approach. In this study the authors use multi-omics spatial profiling of tumour and blood samples that show co-localisation of activated lymphocytes and macrophages are predicators of IT response. The samples analysed are unique and certainly can provide a unique insight into changes in the TME. That said, it is difficult to understand the contribution of the individual immune cell subsets to the response to TMZ and/or IT in the absence of an in vivo/pre-clinical model of MSS CRC to mimic the effects of TMZ and targeting specific immune cell subsets to definitively determine their contributions to response. The authors did not find differences in mutations, mutational signatures, copy number variations or TMB between responders versus non responders. Therefore, they assessed the TME for potential explanations as to determinants of response. Using DSP Nanostring GeoMX, Tumour and stroma areas of interest of responders had higher proportions of memory CD8 T cells and memory B cells with enriched pathways for tumour cell killing and recognition. This was confirmed by Lunaphore COMET technology where greater numbers of CD8 T cells were present in responders versus non responder. Specific dynamic changes were assessed in 3 tumours on treatment with triplet therapy using 10X Visium which showed that specific pathways such as antigen presentation and IFN gamma signatures were upregulated in CD68/Cd163 hotspots in post treatment versus baseline. This approach also validated Cd8 were increased post treatment. No significant differences were observed within the peripheral blood between responders and non-responders at all timepoints. CD8+TGIT+ cells were significantly greater in non-responders compared to responders at post triplet timepoint.

Will the work be of significance to the field and related fields? How does it compare to the established literature? If the work is not original, please provide relevant references
Progress in sensitising MSS to IT has been limited so the approach and target is of interest, and significant in this field of research. Approaches such as the NICHE Trial, using neo-adjuvant IT are reshaping ideas in the field about overcoming the challenge of non-responsiveness in MSS CRC. The approach proposed in the current study, to use TMZ in MGMT silenced tumours to increase neo-antigen load is original and the translational study, including 28 patients, aimed at identifying factors of responsiveness using state of the art approaches identify activated Lymphocytes and macrophage infiltration in responsive tumours. Although this finding is novel in relation to TMZ effects, other studies have found these cell types as indicators of response to IT in multiple tumour types including CRC.

Does the work support the conclusions and claims, or is additional evidence needed?
The data presented support the claims although the number of samples overall is quite low, but the use of multiple platforms for in depth analysis validates the findings observed.

Are there any flaws in the data analysis, interpretation and conclusions? Do these prohibit publication or require revision?
The data analysis, interpretation and conclusions are rational. Limitation in the trial and translational study is that it is difficult

to discriminate the contribution of temozolomide versus immunotherapy to tumor responses and changes in immune infiltration. Additionally, no correlative analysis is performed.

Is the methodology sound? Does the work meet the expected standards in your field?
The methodologies used are state of the art and meets the expected standards, however details on the specific reagents and conditions used in spatial imaging and transcriptomics as well as the sample preparation processes are lacking. It is not clear how cells were isolated from blood or the impact of freeze thaw on the quantity/quality of the cells isolated.

Is there enough detail provided in the methods for the work to be reproduced?
The detail on specific antibodies used, dilution, conditions for staining etc are lacking for the spatial transcriptomics, spatial proteomics and flow cytometry methodologies. Limited information is provided on statistical analysis.

Specific comments:
The data showing differential proportions of CD68+/CD163+ macrophages and interpretation of the effects of TAMs in CRC is not convincing (esp Discussion Line 426)
Figure legends need revising – esp Fig 5
Line 229 'Taken together, the data suggest that responder immune cells were more active….'. This summary statement is general and does not reflect the findings presented and also does not guide the reader as to its relevance.
Line 272' Taken together…..' The summary reference to baseline association is not clear as to what the authors mean?
Line 335 ' CD8 and CD68 cells appeared increased' – the meaning is unclear

Reviewer #2

(Remarks to the Author)
In this manuscript, Choo et al. analyzed the key spatial factors of immunotherapy response from the MAYA trial (temozolomide in combination with nivolumab and ipilimumab for pMMR/MSS MGMT-silenced mCRC). In particular, baseline CD8+KI67+ cells in tumor stromal interface (TSI) and tumor regions, and CD68+CD163+ cells in TSI, were associated with immunotherapy response. They also showed an increase in immune cell abundance after treatment by spatial transcriptomics analysis of sequential biopsy samples, and the enrichment of immune-activating pathways in responders.
Overall, the manuscript is well-written and easy to understand. This work makes full use of clinical trial samples at multiple treatment timepoints, with carefully designed experiments, and the results are relatively clear. However, the current amount of data seems relatively small, and more extensive experiments are needed to validate the main findings. In terms of clinical application, there is still a lack of analytical verification. The following points may help to improve the manuscript to meet the journal's publication criteria:

Major points:
- The authors suggested that baseline CD8+ T cells in tumor and TSI, and CD68+CD163+ macrophages in TSI, were associated with immunotherapy response. Although the compartment-specific distribution patterns are intriguing, the current analysis methods are relatively simple and the data seem inadequate. I recommend that the authors further analyze the mechanisms that influence the distribution of these cell compartments and how the above heterogeneity affects the immunotherapy response.

- Based on the study scheme (Fig1A), the spatial transcriptomics analyzed in Fig6 were from pre-treatment and post-triplet samples. Here, the authors need to further consider that the changes in immune contexture after treatment may be the confounding effect of TMZ and immunotherapy, not just immunotherapy alone. Admittedly, as mentioned in the Discussion section, biopsy samples post-TMZ and pre-immunotherapy are missing. Can peripheral blood samples post-TMZ help solve this problem?

- The clinical application of this work can be further strengthened. For example, whether the compartment-specific abundance of CD8+GzmB+/CD8+KI67+ cells correlates with clinical indicators of mCRC patients (such as tumor burden, carcino-embryonic antigen, tissue PD-L1 expression, BRAF/KRAS mutation, and progression-free survival, etc.). In addition, based on the sample size currently available, these cellular predictors may be further explored to construct quantitative biomarkers to further guide treatment strategies.

Minor points:
- Line 274-275, "a potential role for increased CD68+/CD163+ (M2-like) macrophages in the tumor-stromal interface of non-responders as a mechanism of resistance." This conclusion, at least at this stage, is not fully supported by the data presented in the corresponding results section, since there are no more CD68+CD163+ cells in NR than in R shown in Fig5A.

- The use of "determinants" in the text may be a bit overdone in my opinion.

- The presentation of the data needs to be improved. The text in the figures (such as the GSEA results, the statistical plots, and the scale bars) is too small and hard to discern.

Reviewer #3

(Remarks to the Author)
Choo et al. (2024) utilises spatial multi-omic approaches to predict an immunotherapy response in microsatellite stable (MSS) metastatic colorectal cancer (mCRC). Following up on a previously published MAYA study, during which MSS mCRC patients were treated with temozolomide and immune checkpoint inhibitors (ICI), ipilimumab, and nivolumab, the authors sought to identify factors that predict response to this therapy regimen, given than tumor mutational burden was not completely predictive. The main discovery of the study is increased proliferative CD8+ T-cells and CD68+ CD163+ tumour-associated macrophages (TAMs) located within the tumour and outside of the tumor, respectively.

Whilst sample collection across the treatment timepoints & spatial multi-omic datasets complementing one another are invaluable, the findings as presented are not novel. There are also various issues regarding sample sizes, biology, and study design. Further analyses and revision are required for such a study to be a strong candidate for Nature Communications.

Comments:
1. The major finding is the presence/absence CD8+ T cells and macrophages within or outside of the tumor proper as a predictor of ICI response in CRC. There has already been an avalanche of papers documenting this finding, so the major punchline of the manuscript is not novel. The authors spent a lot of resources to do spatial omics to discover things that we already know, whereas a couple of IHC stains would have given the same conclusions (given what is already known about ICI response).
2. The analysis gave shallow findings. Maybe the reason is that the sample size is relatively low n<10 for each group. Given that there is heterogeneity in CRC subtypes, this low sample size may not enable to query of deep mechanisms. Hence, the overall description of the results is severely lacking. What do you want the readers to concentrate on – outside of the obvious CD8+ T cells and macrophage infiltration.
3. Study design and logic challenge – the study collected invaluable longitudinal samples. However, not much insights were gained from these samples. The conclusions derived from pretreatment and posttreatment are largely similar. If that is the case, how does adding TMZ really help, since it is not via tumor mutations? What is the mechanism? Are you sampling too few patients that those that responded would have responded anyways without TMZ?
4. The assays done on various sample timepoints do not seem to be unified. Perhaps more insights can be gleaned if the same type of data are collected on all specimens.
5. Fig.1 and S.Fig.1 are identical (with additional data points) to the previously published MAYA trial paper. While the reviewer acknowledges the limited number of visualisation methods with clinical data points, different visualisations are recommended.
6. Results of Fig.2c do not describe what is being shown. What are the reasons behind your statement “no differences in mutations, mutational signatures, copy number variations or tumor mutation burden (TMB) between responders and non-responders at baseline”?
7. Abbreviations for Fig.2C are required in the figure legends. E.g. Amplification, Frameshift, etc.
8. Citation is required for MuSiCa.
9. Results section and figure legends for S.Fig.2 are required.
a. What is af?
b. What are the subtypes?
c. There are patients with alphabets. It is surmisable that they are for the treatment stages, but please define A, B, and C.
10. In Fig.3C, there are a lot of non-immune enriched pathways with higher normalised enrichment scores in the tumour. What are the non-immune pathways with the high NES in the tumour (between responders & non-responders)? These pathways could provide us with further insights into different responses.
11. Fig.3D shows epithelial-to-mesenchymal transition (EMT) as highly enriched in the tumours of non-responders compared to responders. Are subtypes of the patient tumors available?
12. In S.Fig.3, add the proportions/percentages for PC1 & PC2.
13. Results section is required for S.Fig.4. No description is available.
14. Add representative images of tumour, TSIs, and stroma selection, separated by non-responders and responders as a supplementary figure (for Fig.4).
15. Describe a strategy for how you picked TSI. 50% tumour + 50% stroma? How you choose TSI ROI would skew the results drastically.
16. In Fig.4 & S.Fig.5, cytokeratin (CK) was used to identify tumours. Did you verify the absence of normal epithelium with morphology? Or did you use the CK level threshold to define tumours (based on “image analysis confirmed increased expression of CK in the tumor regions)? What if the tumor cells lost their CK expression.
17. In Fig.4C, better representative images are required. “Tumors” are not zooming into tumors. Also need scale bars for the tumor images.
18. In Fig.5A, the proportion of TAMs in TSI between responders Vs. non-responders are not statistically different. Instead of calculating the proportion of TAMs in the given area, which can introduce technical artefacts of choosing TSI, the distance and relative frequency of TAMs from the tumour borders/edges may be of more value.
19. Inconsistent scale bars from Fig.5B.
20. For Fig.6A, choose better representative images for responders Vs. non-responders. The responder sample mainly shows stroma, whereas the non-responder sample shows mainly tumours. If it is due to a successful response to triple ICI, state it (reduced tumour size…).
21. Better visualisation and description for Fig. 6B is required.
a. How are immune scores the highest in fibroblasts in comparison to other main immune cells?
b. Continuing the above enquiry, lines 333, 334 suggest the heatmap for fibroblasts, endothelial cells, and macrophages are cell numbers, not immune cell scores. Please elaborate.

c. Inter-patient variability seems to drive the immune cell score difference. Could you reorganise to match the order of patients pre- and post-treatment?
22. S.Fig.6 shows different T-cell subset changes across the ICI treatments. Validate the finding, albeit incomplete, in Visium. Distribution? Frequency?
23. Fig.6D needs bigger labels.

Reviewer #4

(Remarks to the Author)
In this manuscript samples from the MAYA trial of refractory MGMT-silenced MSS colorectal cancer (CRC) were analyzed. Tumor and blood samples at baseline, post-cycle 2 of triplet therapy, and at disease progression to investigate how temozolomide priming and combination immune therapy reshape the tumor microenvironment (TME) and immune context were investigated. This approach aimed to identify biomarkers and uncover resistance mechanisms to improve combination therapies and personalized treatments for microsatellite-stable CRC, which is still a largely non-responding cancer entity. Overall, this study represents a very interesting and important description of this innovative trial set up. However, there are a few concerns that should be addressed before publication of this manuscript which are listed below.

Major comments
Spatial analysis
The usage of spatial data is highly relevant for heterogenic therapy responses. However, the major conclusion, based on spatial data, is missing or to weakly elaborated in the manuscript. CD8+TIGIT+ scores show association in peripheral blood but were not investigated in the spatial datasets. This should be done to generate a line of evidence that is connected to deeper spatial relevance.
TME signals
It is demonstrated that the TME compositions is altered by the treatment with TMZ and ICI. Nevertheless, factors that are potentially critical for this change in TME like TGF-beta or other means have been completely neglected. A more unbiased analysis of the probe set from e.g. the Visium platform might enhance the overall understanding and helps defining underlying mechanisms or novel targets (as state in the paper).
Minor comments
Data representation
The font and description of the figures should be harmonized to allow proper reading of all data points.
TMB
Somewhat the conclusion for the findings around the TMB is lacking. What does the increase after TMZ plus ICI mean and are they any strategies for the future?

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The reviewers have addressed my comments for the most part

Reviewer #2

(Remarks to the Author)
I thank the authors for revising this manuscript with additional analyses, and for responding to all my queries. The authors have added multiple dimensions of spatial omics analysis, newly proposing the relationship between the compartmental distribution of cancer-associated fibroblasts (CAFs) and TMZ response, and suggesting that the distance of CD8 T cells to CAFs reflects the response to TMZ. The current version of the manuscript shows improvement, particularly in the methodology and the the breadth of findings. However, several aspects in this manuscript required further attention before publication.

- In the revised manuscript, the authors present some new insights, such as the potential correlation between neighbourhood of CD8 T cell and CAFs to TMZ response. Although the authors performed many analyses and attempts, perhaps limited by the sample size, many results fail to reach statistical significance (e.g., Fig. 4F and 6C), which impacts the overall repeatability of the conclusions.

- The interpretation of some conclusions lacks sufficient depth. As the authors state in Line 582: “The interaction of CAFs with other cell types... shapes immune evasion and treatment resistance”. However, I regret to note that while the authors used spatial omics to observe the distribution features of cells at various treatment time points, they did not further explore the cellular interactions. This may improve the mechanistic depth and the novel insights that spatial omics could provide to this study.

- In the cellular neighbourhood analysis, CN9 and CN2 have similar cellular diversity but differential survival outcomes. Yet, their T cell proportions are relatively similar. What might be the underlying reason for this result? How can we understand different cellular compositions lead to divergent survival outcomes? Incorporating immune-related proteins, cytokines,

ligand-receptor, etc., into the neighbourhood analysis might help explain this finding.

- The presentation of many figures needs further improvement. For example, it is difficult to compare immune cell abundance between responders and non-responders Sup. Fig 3. It would benefit from reorganization by response group or quantitative analyses to facilitate direct comparison. Additionally, some texts in many figures remain unclear (e.g., Figs 2 and 4).

Reviewer #3

(Remarks to the Author)
We thank the authors for thoroughly addressing many of our comments specifically those that are critical to the clarity of the paper. We are impressed by the amount of details provided. Newly adopted analyses, visualisations, scopes certainly strengthened the narrative. Specifically, I think the paper now has a more clear direction.

Given that, the major issues with the paper remain unresolved and I am not sure given at this stage it is resolvable. We appreciate the study's message regarding the use of biomarker-based stratification of patients for combinatory immunotherapy and the fact that the results arose from clinical trial samples. However, the numbers of samples used to make some of the conclusions are, honestly, low, and the authors were not able to address the issues. While more analyses are done on the small number of samples already present, this does not lead to "informative and robust" analyses as the authors stated. (Doing more types of analyses on a small number of samples – giving shaky results - is not equivalent to doing a deep analysis on many more samples). Actually, for some analyses, the number of samples actually decreased from the previous submission for some reason. Because of this, some of the comments were not addressed directly but avoided (for instance, R3- comment 3- we ask about what we learn different from pretreated vs post treated samples that are different if not tumor mutations – the authors just skirted the issue and gave us a description and reference on TMB), and the unified assay comment – which was not addressed. At this point, we feel like the authors did do a lot of work to address the comments, but there does not seem to be a way to improve the small sample size issues. I would like to leave it to the editor to decide based on comments from other reviewers.

Other comments.

1. In figure 4, authors should consider showing examples of specific cellular neighbours.
2. For figure 2A, why did the number of flow cytometry & GeoMx datasets decrease from your last version?
3. For figure 2D, is Fisher's exact test appropriate? I was wondering if the test is too conservative.
4. For figure 4F, clearly state/indicate they are distances from CD8 T-cells.
5. Figure 5F-I are missing.

Reviewer #4

(Remarks to the Author)

Version 2:

Reviewer comments:

Reviewer #2

(Remarks to the Author)
The authors have addressed all my previous comments, and I have no further concerns.

Reviewer #3

(Remarks to the Author)
I went through the Choo et al. revision. As per our review, it is really up to the editor. The authors addressed all of the points we made, except for the insufficient sample size to support their conclusions. There is not a simple way to increase the sample size for a robust analysis.

Reply to Reviewer #1

Choo J et al present the results of a translational study associated with the MAYA trial (published in 2022 in JCO) where aimed to elucidate the factors influencing response heterogeneity using multi-omics spatial profiling of samples from patients who participated in the MAYA trial, including baseline and on-treatment tissue and blood specimens.

R1C1 What are the noteworthy results?

The fact that MSS CRC tumours are largely unresponsive to treatment emphasises the huge need to identify novel efficacious therapeutic approaches. The use of temozolomide to sensitise MSS MGMT silenced cold tumours to immunotherapies is a novel and interesting approach. In this study the authors use multi-omics spatial profiling of tumour and blood samples that show co-localisation of activated lymphocytes and macrophages are predicators of IT response. The samples analysed are unique and certainly can provide a unique insight into changes in the TME. That said, it is difficult to understand the contribution of the individual immune cell subsets to the response to TMZ and/or IT in the absence of an in vivo/pre-clinical model of MSS CRC to mimic the effects of TMZ and targeting specific immune cell subsets to definitively determine their contributions to response. The authors did not find differences in mutations, mutational signatures, copy number variations or TMB between responders versus non responders. Therefore, they assessed the TME for potential explanations as to determinants of response. Using DSP Nanostring GeoMX, Tumour and stroma areas of interest of responders had higher proportions of memory CD8 T cells and memory B cells with enriched pathways for tumour cell killing and recognition. This was confirmed by Lunaphore COMET technology where greater numbers of CD8 T cells were present in responders versus non responder. Specific dynamic changes were assessed in 3 tumours on treatment with triplet therapy using 10X Visium which showed that specific pathways such as antigen presentation and IFN gamma signatures were upregulated in CD68/CD163 hotspots in post treatment versus baseline. This approach also validated CD8 were increased post treatment. No significant differences were observed within the peripheral blood between

responders and non-responders at all timepoints. CD8+TGIT+ cells were significantly greater in non-responders compared to responders at post triplet timepoint.

Reply: We thank the Reviewer for underscoring the strong impetus behind our investigation and for the insightful comments. We acknowledge the importance of in vivo/pre-clinical models of MSS CRC for functional validation. Our study builds on the foundational work of Germano *et al*[1], who have already made key biological discoveries related to Temozolomide treatment. Their research demonstrated that the inactivation of MMR, driven by acquired resistance to the clinical agent temozolomide, increased mutational load, promoted continuous renewal of neoantigens in human colorectal cancers and triggered immune surveillance in mouse models[1,2].

As such, our study places a greater emphasis on a multi-omic approach using human samples, prioritizing orthogonal validation as the primary method for confirming our findings.

References

[1] Germano, G., Lamba, S., Rospo, G. et al. Inactivation of DNA repair triggers neoantigen generation and impairs tumour growth. Nature 552, 116–120 (2017). https://doi.org/10.1038/nature24673

[2] Crisafulli G, Sartore-Bianchi A, Lazzari L, et al. Temozolomide Treatment Alters Mismatch Repair and Boosts Mutational Burden in Tumor and Blood of Colorectal Cancer Patients. Cancer Discov. 2022 Jul 6;12(7):1656-1675. doi: 10.1158/2159-8290.CD-21-1434

R1C2 Will the work be of significance to the field and related fields? How does it compare to the established literature? If the work is not original, please provide relevant references Progress in sensitising MSS to IT has been limited so the approach and target is of interest, and significant in this field of research. Approaches such as the NICHE Trial, using neo-adjuvant IT

are reshaping ideas in the field about overcoming the challenge of non-responsiveness in MSS CRC. The approach proposed in the current study, to use TMZ in MGMT silenced tumours to increase neo-antigen load is original and the translational study, including 28 patients, aimed at identifying factors of responsiveness using state of the art approaches identify activated Lymphocytes and macrophage infiltration in responsive tumours. Although this finding is novel in relation to TMZ effects, other studies have found these cell types as indicators of response to IT in multiple tumour types including CRC.

Reply: We thank the Reviewer for the comment. Our study represents the first to explore the novel combination of TMZ and immune checkpoint inhibition (ICI) within the spatial context of colorectal cancer (CRC). The fact that cell types historically recognized as predictive markers of ICI responsiveness were consistently observed in our analysis serves as a validation of our findings. This recurrence reassures us that our spatially resolved approach is capturing critical elements of the tumor-immune interface. It underscores the robustness of these immune signatures and strengthens the relevance of our insights in advancing the understanding of immune modulation within CRC. This alignment with established immunological predictors further enhances the translational significance of our work and its potential impact in refining future therapeutic strategies.

In the revised version of the manuscript, we have now expanded our spatial analysis by introducing additional key cell types such as cancer-associated fibroblasts (CAFs), endothelial cells and memory T cells, just to name a few (**Supplementary Table 1)**. This provides a more comprehensive characterization of the tumor microenvironment (TME), offering greater insight into its complexity. Furthermore, to provide a more robust interrogation of the TME, we have incorporated additional spatial features, which go beyond relative abundance of cell types per compartment. We have introduced additional compartment specific spatial features including (1) Proximity analyses of adjacent cells: We now examine the spatial proximity of adjacent cells[1], analyzing how different cell types interact within the TME to better understand immune cell dynamics. (2) Relative distances of adjacent cells: Our analysis also includes compartment-

specific relative distances between cell types, allowing us to quantify how immune cells are spatially organized across distinct compartments. (3) Cellular neighborhoods: We investigate the cellular neighborhoods to understand how spatially resolved immune cell clusters orchestrate responses or resistance to immune checkpoint inhibitors (ICIs). This aspect of the analysis provides insight into the microenvironmental context of immune cell function and therapy outcomes. (**Figure R1**) These revisions enhance our spatial analysis and provide deeper insights into the TME and its impact on immune responses, particularly in the context of ICI therapy.

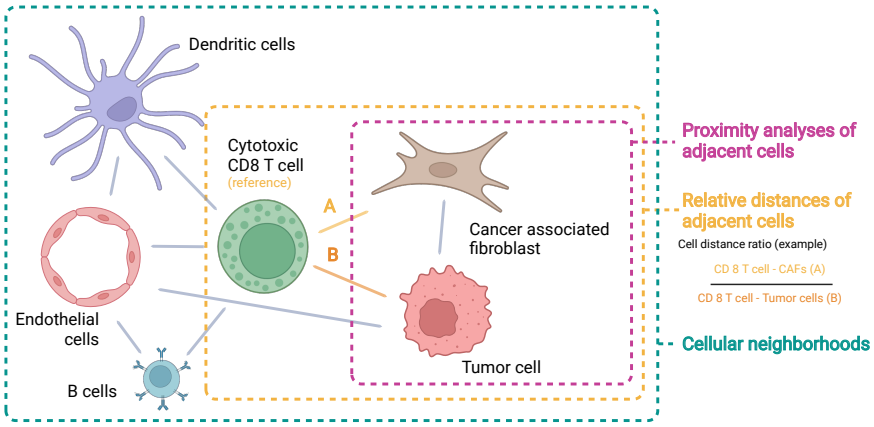


Figure R1 Illustration of additional spatial features incorporated into the revised manuscript

Remarks: IHC marker definitions per cell type are provided in **Supplementary Table 1**.

On inspection of adjacent cell abundances within the vicinity of 20μm of a tumor cell within tumor compartments of baseline samples, we observed scant numbers of effector adaptive immune cells (CD8 T cells and B cells) in contrast to immunosuppressive cells such as CAFs and memory adaptive immune cells such as memory T cells. These findings are consistent with the known phenomenon of an immune-cold microenvironment of microsatellite stable (MSS) colorectal cancer (CRC). Of interest, higher cancer associated fibroblast (CAF) infiltration within tumor

compartments is associated with poor response to TMZ alone, and subsequent ICI. In contrast, greater Memory T cell infiltration within tumor cells was associated with good response to ICI.

We also sought to understand how tumor-infiltrating cell types influenced effector CD8 T cells within tumor compartments. We evaluated relative cell distances from CD8 T cells and found that relative proximity of CD8 T cells to CAFs compared to tumor cells were associated with poor response to TMZ alone, although statistical significance was not reached. Likewise, relative proximity of CD8 T cells to endothelial cells were associated with poorer response.

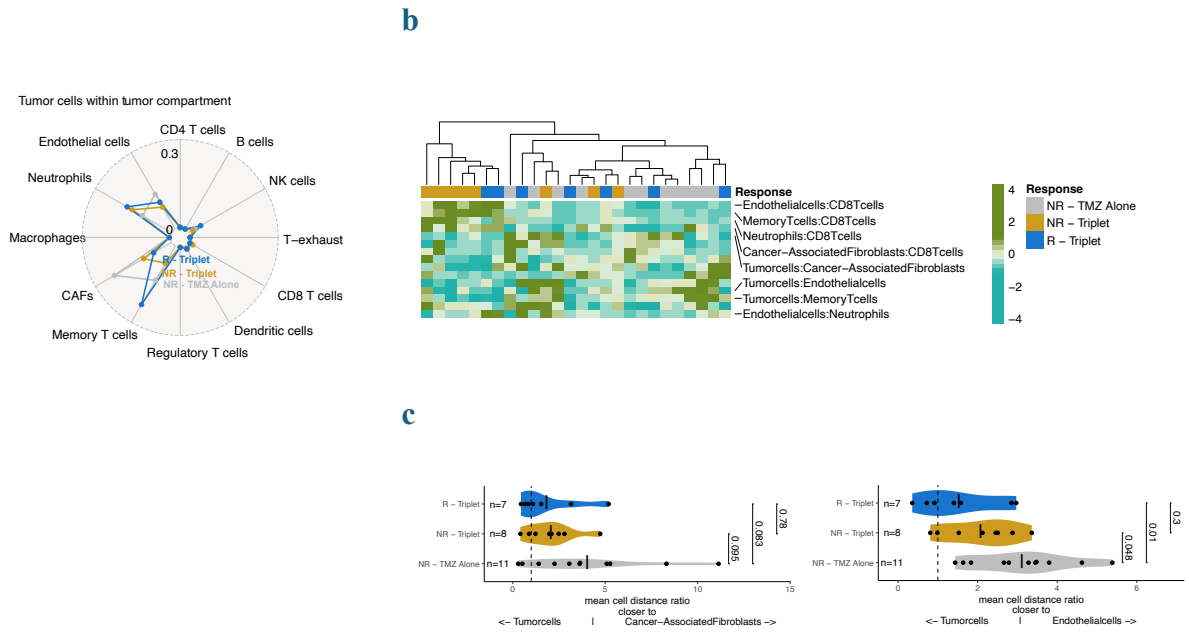


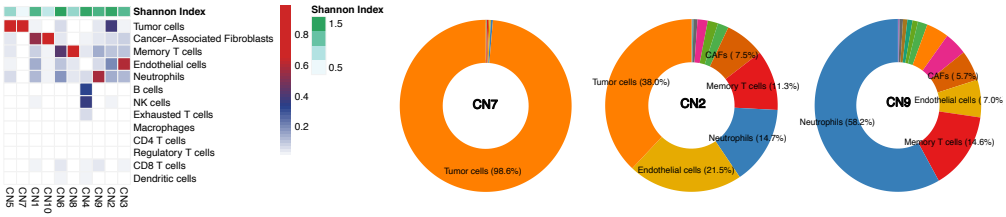
Figure R2

a) Adjacent cell abundances within 20µm of tumor cells within tumor compartments.
b) Heatmap of mean cell distance ratios from CD8 T cells per patient. Cell distance ratios flagged out had an ANOVA p-value of less than 0.15 across response group comparisons.
c) Selected violin plots comparing mean cell distance ratios per patient stratified by tumor response groups. P-values were retrieved from a two-sided T-test.

Abbreviations: CAF, cancer associated fibroblasts; T-exhaust, Exhausted T cells; NR – TMZ Alone, non-responder to temozolomide (TMZ) alone; NR - Triplet, responder to TMZ however non-responder to ICI; R - Triplet, responder to TMZ and ICI; n=, number of patients.

Remarks: IHC marker definitions are provided in **Supplementary Table 1**.

a



b

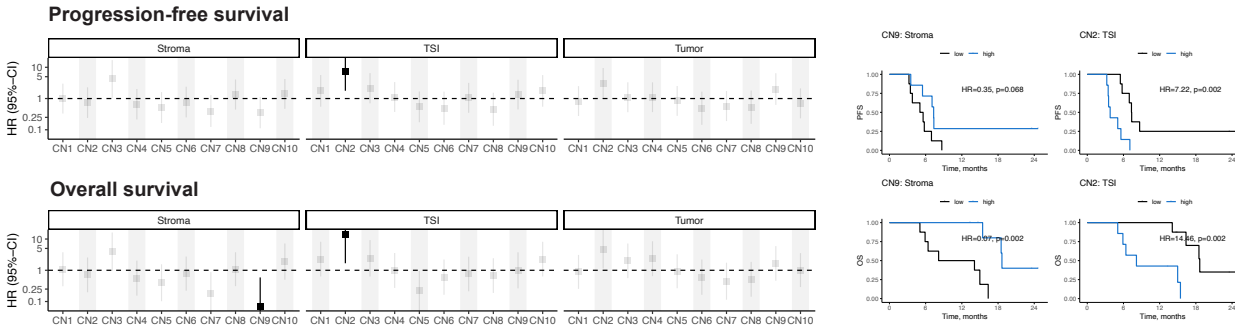


Figure R3

b) Heatmap and donut plots providing an overview of immune cell proportions per cellular neighbourhood.

b) Forest plot and selected KM plots of survival outcomes stratified by compartment specific CN cellular proportions at median split. Values were highlighted if Log-rank p-value was less than 0.01.

Abbreviations: CN, cellular neighbourhood; TSI, tumor stromal interface; HR, hazard ratio; KM, Kaplan-Meier.

The tumor microenvironment (TME) operates as an integrated entity, where distinct immune cell clusters interact to coordinate anti-tumor responses. As such, we employed spatially resolved clustering algorithms previously described by Schürch *et al*[2] to retrieve cellular neighbourhood (CN) information. Broadly, our analyses demonstrated CNs that differed in immune cell repertoire and diversity. For instance, CN7 was found to predominantly comprise tumor cells (98.6%) with low immune cell diversity (Shannon Index = 0.097). In contrast, CN2 and CN9 were found to have high immune cell diversity (CN2, Shannon Index = 1.715; CN9, Shannon Index = 1.482).

Compartment specific comparative proportions of CNs were found to be associated with survival outcomes and response. For instance, patients with high proportions of CN9 cells within stromal compartments were associated with favorable survival outcomes. Conversely, patients with high proportions of CN2 cells (heterogenous mixture of tumor cells, endothelial cells, neutrophils, memory T cells and CAFs) within the tumor stromal interface (TSI) were associated with poorer survival outcomes (**Figure R3b**). It was interesting to note that CN10, which was primarily composed of CAFs, did not modulate treatment response to the same extent and magnitude as CN2 (OS within TSI, HR=2.28, p=0.196; PFS within TSI, HR=1.76, p=0.337) (**Figure R3b**).

Our findings underscore the critical role of spatial features in reflecting the intricate, dynamic interplay of immune cells, providing valuable prognostic and therapeutic insights into tumor responses. Specifically, we demonstrate that the infiltration of immunosuppressive cells, such as CAFs, within tumor compartments—particularly their proximity to CD8+ T cells—attenuates treatment efficacy and modulates therapeutic responses. Moreover, the immunosuppressive influence of CAFs is not uniform but spatially dependent, further emphasizing the importance of cellular architecture in shaping immune evasion and resistance mechanisms.

References

[1] Jia K, Chen Y, Sun Y, et al. Multiplex immunohistochemistry defines the tumor immune microenvironment and immunotherapeutic outcome in CLDN18.2-positive gastric cancer. BMC Med. 2022 Jul 11;20(1):223. doi: 10.1186/s12916-022-02421-1. Erratum in: BMC Med. 2023 Sep 29;21(1):369. doi: 10.1186/s12916-023-03090-4. PMID: 35811317; PMCID: PMC9272556.

[2] Schürch CM, Bhate SS, Barlow GL, et al. Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front. Cell. 2020 Sep 3;182(5):1341-1359.e19. doi: 10.1016/j.cell.2020.07.005. Epub 2020 Aug 6. Erratum in: Cell. 2020 Oct 29;183(3):838. doi: 10.1016/j.cell.2020.10.021. PMID: 32763154; PMCID: PMC7479520.

R1C3 Does the work support the conclusions and claims, or is additional evidence needed? The data presented support the claims although the number of samples overall is quite low, but the use of multiple platforms for in depth analysis validates the findings observed.

Reply: We thank the Reviewer for acknowledging the use of multiple profiling approaches as orthogonal validation of the key findings in our study. This translational study was conducted on samples obtained from a clinical trial, which resulted in a positive outcome. The trial was designed with a pre-defined clinical endpoint of survival, which it successfully achieved. While the total number of samples might appear limited from a translational perspective, it is important to clarify that the study design was robust and sufficiently powered to achieve both statistically significant and clinically meaningful results.

R1C4 Are there any flaws in the data analysis, interpretation and conclusions? Do these prohibit publication or require revision?

The data analysis, interpretation and conclusions are rational. Limitation in the trial and translational study is that it is difficult to discriminate the contribution of temozolomide versus immunotherapy to tumor responses and changes in immune infiltration.

Reply: We appreciate the Reviewer's insightful comment regarding the potential confounding effects of temozolomide (TMZ) and immunotherapy on changes in immune contexture. While we fully acknowledge the value of obtaining biopsies post-TMZ and pre-immunotherapy, doing so presents practical challenges within current clinical workflows. In patients who respond well to TMZ and proceed directly to immunotherapy, additional tissue sampling is often not clinically indicated and may pose unnecessary procedural risk without influencing treatment decisions. As such, while ideal from a research standpoint, routine pre-immunotherapy biopsies remain difficult to justify in this setting.

As such, we conducted flow cytometry as an alternative to evaluate immune cell activation post-TMZ. Previous work by Chin *et al* [1] and Hwang *et al* [2] have demonstrated that peripheral blood immune cell subsets reflect changes in the tumor microenvironment. However, no significant differences in immune cell proportions were identified pre- and post-TMZ (**Supplementary Figure 8**).

However, upon accounting for tumor response, we identified divergent T cell exhaustion trajectories post-TMZ that discriminated responders from non-responders. We found that post-TMZ increase in exhausted T cell proportions predicts poor response. Fold change values (post-TMZ/baseline) of CD8+TIGIT+ (AUROC: 0.850 95%-CI: 0.605 - 1.000) and CD4+PD1+ (AUROC: 0.800 95%-CI: 0.548 - 1.000) cell proportions were predictive of response (**Figure R4a, R4b**). When analyzed with paired sample comparisons at the patient-level, increased T cell exhausted proportions post-TMZ were predictive of poorer PFS and OS (**Figure R4c**). These findings are consistent with the established phenomenon in which T cell exhaustion is associated with poorer ICI response [3]. These findings suggest that immune composition evolution post-TMZ while heterogenous, can be exploited to predict subsequent ICI response.

That being said, we concur that this approach has its inherent limitations, specifically with regard to the lack of appreciating spatial organization and distribution of immune cells like what we did

so for post-ICI analysis. We have acknowledged this, borne out of clinical workflow considerations, in our discussion.

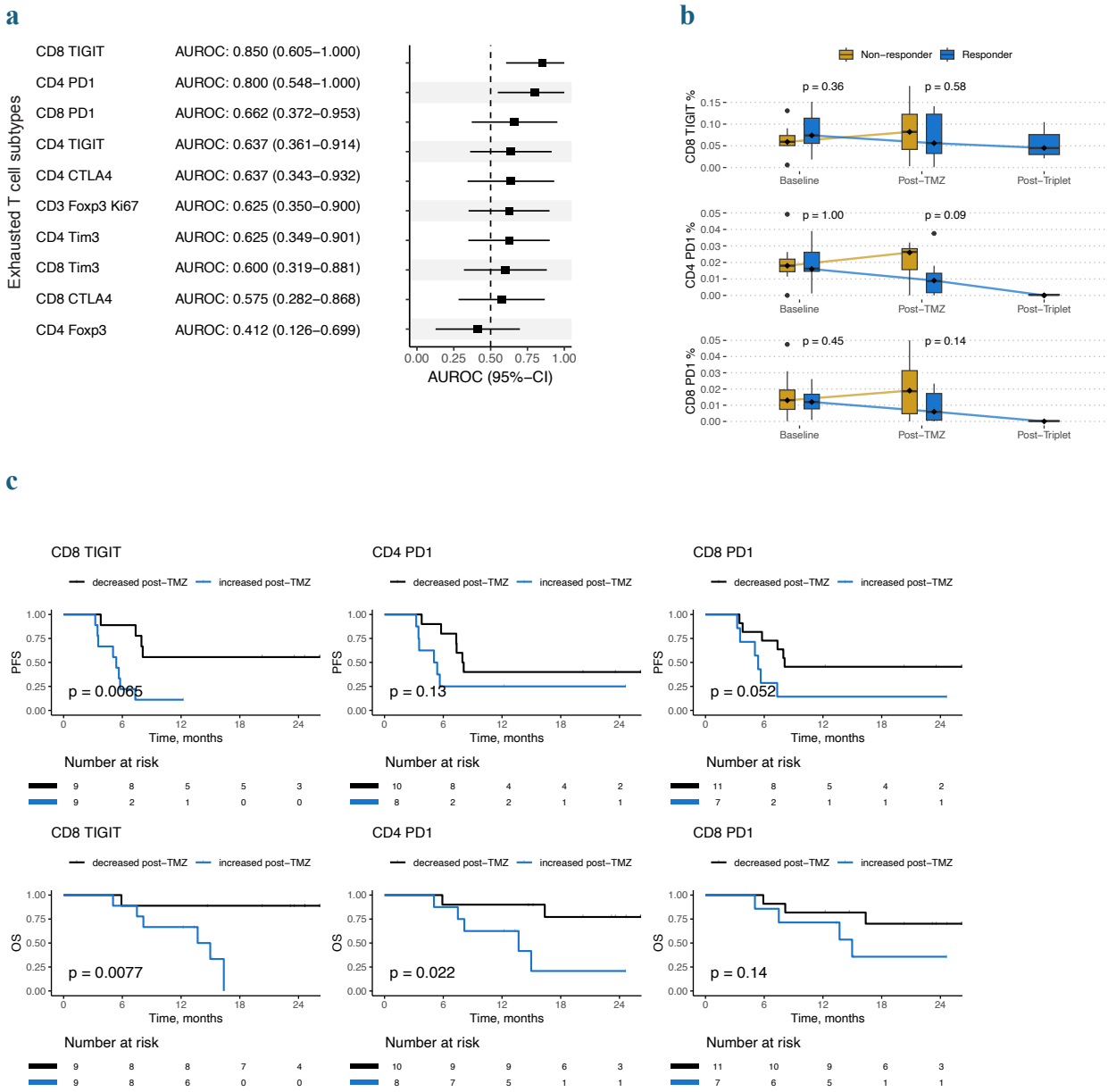


Figure R4

a) Forest plot of AUROC values of post-TMZ/baseline immune cell ratio against response.
b) Longitudinal box plots depicting immune cell type proportions evolution with treatment. P-values were retrieved from a two-sided T-test.
c) KM plots (OS, PFS) of participants stratified by dynamic changes in immune cell proportions post TMZ. P-values were retrieved from a Log-rank test.

Abbreviations: AUROC, area under receiver operating curve; TMZ, temozolomide; OS, overall survival; PFS, progression-free survival; ROI, region of interest; TSI, tumor stromal interface.

References

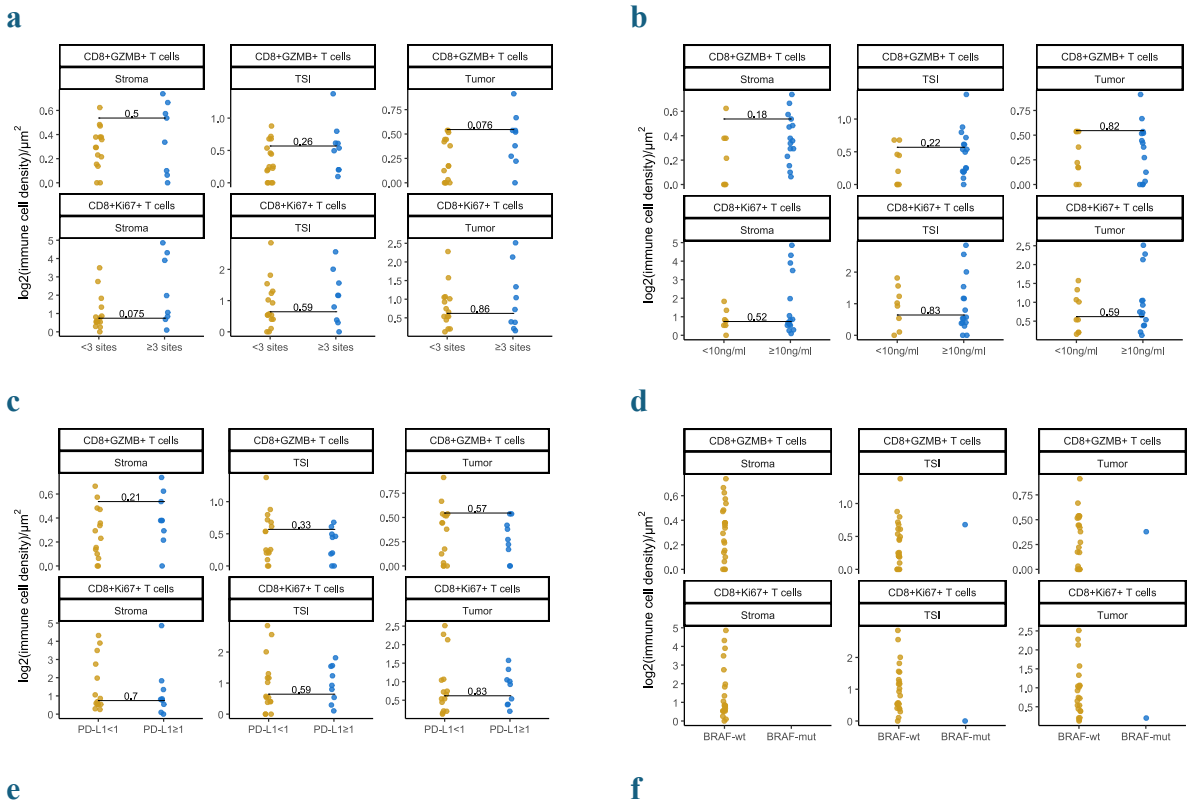
[1] Chin WL, Cook AM, Chee J, et al. Coupling of response biomarkers between tumor and peripheral blood in patients undergoing chemoimmunotherapy. Cell Rep Med. 2025 Jan 21;6(1):101882. doi: 10.1016/j.xcrm.2024.101882. Epub 2024 Dec 27. PMID: 39731918.

[2] Hwang M, Canzoniero JV, Rosner S, et al. Peripheral blood immune cell dynamics reflect antitumor immune responses and predict clinical response to immunotherapy. J Immunother Cancer. 2022 Jun;10(6):e004688. doi: 10.1136/jitc-2022-004688. PMID: 35688557; PMCID: PMC9189831.

[3] Chow A, Perica K, Klebanoff CA, Wolchok JD. Clinical implications of T cell exhaustion for cancer immunotherapy. Nat Rev Clin Oncol. 2022 Dec;19(12):775-790. doi: 10.1038/s41571-022-00689-z. Epub 2022 Oct 10. PMID: 36216928; PMCID: PMC10984554.

R1C5 Additionally, no correlative analysis is performed.

Reply: As suggested by the Reviewer, we have now conducted clinical correlative analyses of upstream identified key immune cells. We compared CD8+GZMB+ and CD8+Ki67+ T cells against clinical indicators of tumor burden as reflected by the number of metastatic sites (**Figure R5a**), CEA level (**Figure R5b**), PD-L1 combined positive score (CPS) (**Figure R5c**), BRAF and KRAS mutation (**Figure R5d, e**). While there was some suggestion for which CD8+GZMB+ was found to be higher in tumor compartments in patients with higher tumor burden, this did not reach statistical significance (p=0.076) (**Figure R5a**). All other comparisons were found to be statistically not significant.



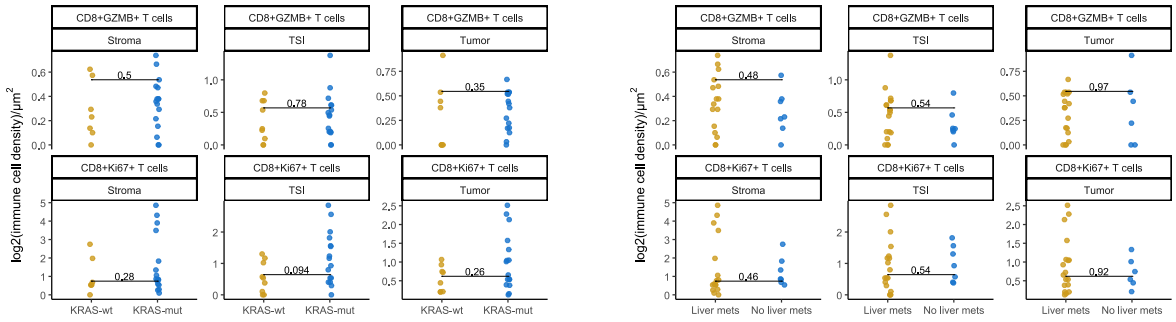


Figure R5 Dot plot comparisons for of CD8+Ki67+ and CD8+GZMB+ T cell densities stratified by

a) Number of metastatic sites
b) Baseline CEA levels
c) PD-L1 CPS
d) BRAF status
e) KRAS status
f) Presence of liver metastasis

Abbreviations: CEA, carcino-embryonic antigen; CPS, combined positive score; -wt, wild-type; -mut, mutant; PD-L1, programmed death-ligand 1.

Remarks: Progression-free survival (PFS) was not evaluated as this was already analyzed with treatment response. P-values were retrieved from a two-sided Wilcoxon-test. Statistical testing for BRAF status was not pursued as only 1 patient with available Lunaphore stain data was found to be BRAF-mut.

R1C6 Is the methodology sound? Does the work meet the expected standards in your field? The methodologies used are state of the art and meets the expected standards, however details on the specific reagents and conditions used in spatial imaging and transcriptomics as well as the

sample preparation processes are lacking. It is not clear how cells were isolated from blood or the impact of freeze thaw on the quantity/quality of the cells isolated.

Reply: As suggested by the Reviewer, we have now included the necessary methodological information in the **Methods** section to improve transparency of our study.

R1C7 Is there enough detail provided in the methods for the work to be reproduced? The detail on specific antibodies used, dilution, conditions for staining etc are lacking for the spatial transcriptomics, spatial proteomics and flow cytometry methodologies. Limited information is provided on statistical analysis.

Reply: As suggested by the Reviewer, we have now included the necessary methodological information in the **Methods** section to improve transparency of our study.

Specific comments:

R1C8 The data showing differential proportions of CD68+/CD163+ macrophages and interpretation of the effects of TAMs in CRC is not convincing (esp Discussion Line 426)

Reply: We have replaced the analysis of tumor-associated macrophages (TAMs) with a series of novel spatial analyses that provide deeper insights into the tumor microenvironment. Specifically, our revised analysis reveals that the spatial proximity between CD8+ effector T cells and cancer-associated fibroblasts (CAFs) within the tumor compartment is associated with attenuated immunotherapeutic efficacy. This finding highlights a potential mechanism of CAF-mediated immune suppression and underscores the importance of spatial context in modulating treatment response.

R1C9 Figure legends need revising – esp Fig 5

Reply: We thank the Reviewer for this comment. We have updated all figure labels and legends to ensure clear visualisations throughout the manuscript.

R1C10 Line 229 'Taken together, the data suggest that responder immune cells were more active….'. This summary statement is general and does not reflect the findings presented and also does not guide the reader as to its relevance.

Reply: As suggested by the Reviewer, we have revised the summary statement to provide greater clarity and better reflect the findings presented. The revised statement is as follows:

> *These findings highlight the role of immune activation and engagement within the tumor microenvironment as key contributors to therapeutic response, distinguishing responders from non-responders by their more robust and effective immune response.*

R1C11 Line 272' Taken together…..' The summary reference to baseline association is not clear as to what the authors mean?

> *Statement in the manuscript is: "Taken together, these findings, observed in pre treatment samples, suggest a baseline association between CD8+ T cell infiltration and response to TMZ-based immunotherapy, with a potential role for increased CD68+/CD163+ (M2-like) macrophages in the tumor-stromal interface of non-responders as a mechanism of resistance."*

Reply: As mentioned above, we have incorporated a series of novel spatial analyses that provide deeper insights into the tumor microenvironment. Specifically, our revised analysis reveals that the spatial proximity between CD8+ effector T cells and cancer-associated fibroblasts (CAFs) within the tumor compartment is associated with attenuated immunotherapeutic efficacy. This

finding highlights a potential mechanism of CAF-mediated immune suppression and underscores the importance of spatial context in modulating treatment response.

R1C12 Line 335 ‘ CD8 and CD68 cells appeared increased’ – the meaning is unclear

Reply: We have replaced the analysis of tumor-associated macrophages (TAMs) with a series of novel spatial analyses that provide deeper insights into the tumor microenvironment. Specifically, our revised analysis reveals that the spatial proximity between CD8+ effector T cells and cancer-associated fibroblasts (CAFs) within the tumor compartment is associated with attenuated immunotherapeutic efficacy. This finding highlights a potential mechanism of CAF-mediated immune suppression and underscores the importance of spatial context in modulating treatment response.

Reply to Reviewer #2

In this manuscript, Choo et al. analyzed the key spatial factors of immunotherapy response from the MAYA trial (temozolomide in combination with nivolumab and ipilimumab for pMMR/MSS MGMT-silenced mCRC). In particular, baseline CD8+KI67+ cells in tumor stromal interface (TSI) and tumor regions, and CD68+CD163+ cells in TSI, were associated with immunotherapy response. They also showed an increase in immune cell abundance after treatment by spatial transcriptomics analysis of sequential biopsy samples, and the enrichment of immune-activating pathways in responders.

Overall, the manuscript is well-written and easy to understand. This work makes full use of clinical trial samples at multiple treatment timepoints, with carefully designed experiments, and the results are relatively clear. However, the current amount of data seems relatively small, and more extensive experiments are needed to validate the main findings. In terms of clinical application, there is still a lack of analytical verification. The following points may help to improve the manuscript to meet the journal's publication criteria:

R2C1 Major points:

- The authors suggested that baseline CD8+ T cells in tumor and TSI, and CD68+CD163+ macrophages in TSI, were associated with immunotherapy response. Although the compartment-specific distribution patterns are intriguing, the current analysis methods are relatively simple and the data seem inadequate. I recommend that the authors further analyze the mechanisms that influence the distribution of these cell compartments and how the above heterogeneity affects the immunotherapy response.

Reply: We appreciate the constructive feedback and have made significant enhancements to our spatial analyses to address the points raised. We have expanded our spatial analysis by introducing additional key cell types such as cancer-associated fibroblasts (CAFs), endothelial cells and memory T cells, just to name a few (**Supplementary Table 1**). This provides a more comprehensive characterization of the tumor microenvironment (TME), offering greater insight into its complexity. Furthermore, to provide a more robust interrogation of the TME, we have

incorporated additional spatial features, which go beyond relative abundance of cell types per compartment. We have introduced additional compartment specific spatial features including (1) Proximity analyses of adjacent cells: We now examine the spatial proximity of adjacent cells [1], analyzing how different cell types interact within the TME to better understand immune cell dynamics. (2) Relative distances of adjacent cells: Our analysis also includes compartment-specific relative distances between cell types, allowing us to quantify how immune cells are spatially organized across distinct compartments. (3) Cellular neighborhoods: We investigate the cellular neighborhoods to understand how spatially resolved immune cell clusters orchestrate responses or resistance to immune checkpoint inhibitors (ICIs). This aspect of the analysis provides insight into the microenvironmental context of immune cell function and therapy outcomes. (**Figure R6**) These revisions enhance our spatial analysis and provide deeper insights into the TME and its impact on immune responses, particularly in the context of ICI therapy.

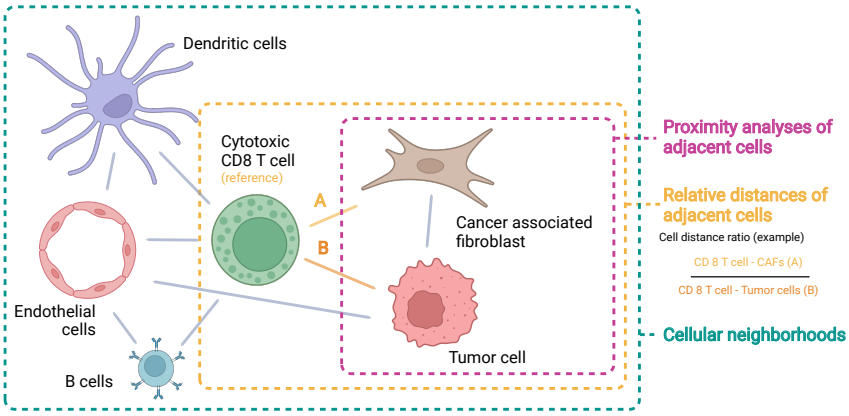


Figure R6 Illustration of additional spatial features incorporated into the revised manuscript

Remarks: IHC marker definitions per cell type are provided in **Supplementary Table 1**.

On inspection of adjacent cells abundances within the vicinity of 20µm of a tumor cell within tumor compartments of baseline samples, we observed scant numbers of effector adaptive immune cells (CD8 T cells and B cells) in contrast to immunosuppressive cells such as CAFs and memory

adaptive immune cells such as memory T cells. These findings are consistent with the known phenomenon of an immune-cold microenvironment of microsatellite stable (MSS) colorectal cancer (CRC). Of interest, higher cancer associated fibroblast (CAF) infiltration within tumor compartments is associated with poor response to TMZ alone, and subsequent ICI. In contrast, greater Memory T cell infiltration within tumor cells was associated with good response to ICI.

We also sought to understand how tumor-infiltrating cell types influenced effector CD8 T cells within tumor compartments. We evaluated relative cell distances from CD8 T cells and found that relative proximity of CD8 T cells to CAFs compared to tumor cells were associated with poor response to TMZ alone, although statistical significance was not reached. Likewise, relative proximity of CD8 T cells to endothelial cells were associated with poorer response.

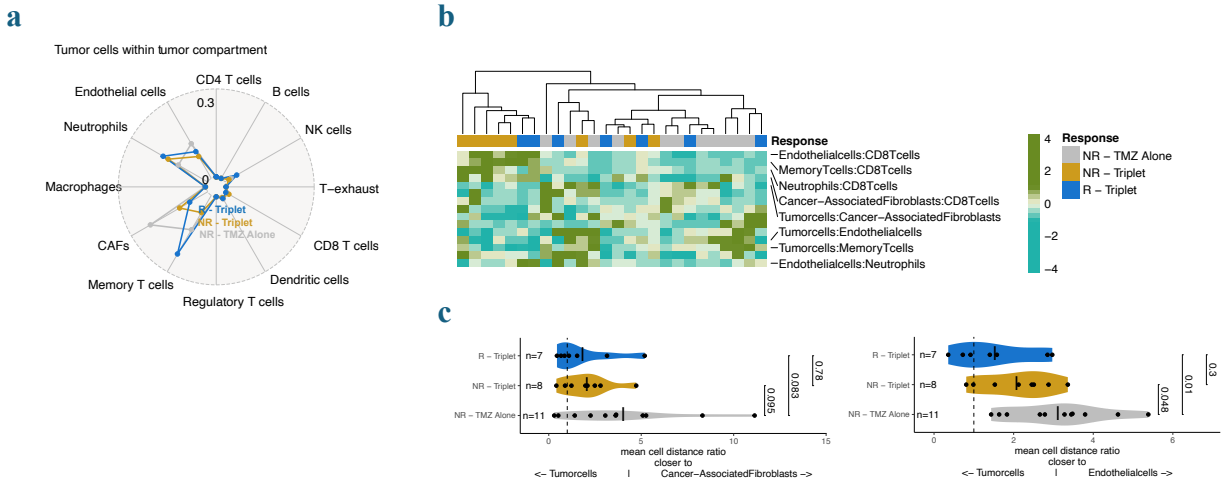


Figure R7

d) Adjacent cell abundances within 20µm of tumor cells within tumor compartments.

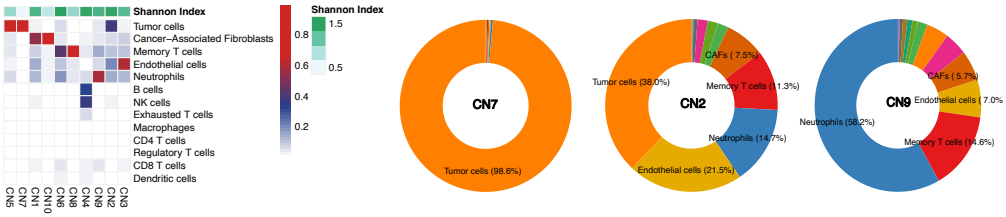
e) Heatmap of mean cell distance ratios from CD8 T cells per patient. Cell distance ratios flagged out had an ANOVA p-value of less than 0.15 across response group comparisons.

f) Selected violin plots comparing mean cell distance ratios per patient stratified by tumor response groups. P-values were retrieved from a two-sided T-test.

Abbreviations: CAF, cancer associated fibroblasts; T-exhaust, Exhausted T cells; NR – TMZ Alone, non-responder to temozolomide (TMZ) alone; NR - Triplet, responder to TMZ however non-responder to ICI; R - Triplet, responder to TMZ and ICI; n=, number of patients.

Remarks: IHC marker definitions are provided in **Supplementary Table 1**.

a



b

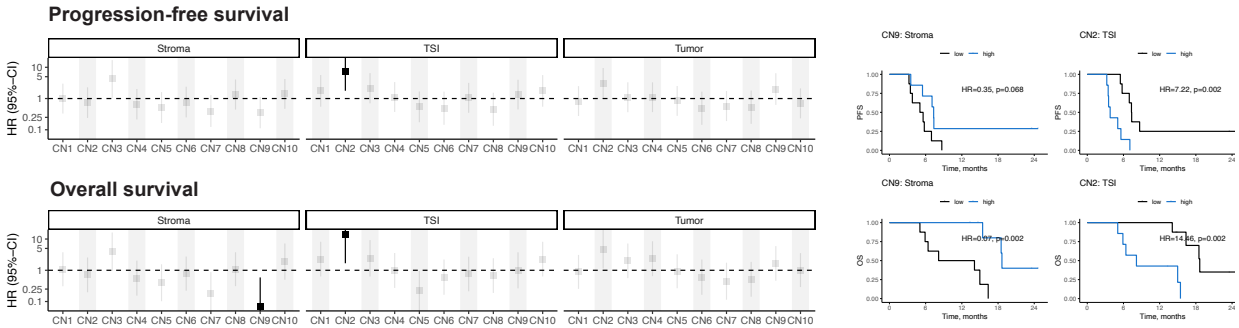


Figure R8

b) Heatmap and donut plots providing an overview of immune cell proportions per cellular neighbourhood.

b) Forest plot and selected KM plots of survival outcomes stratified by compartment specific CN cellular proportions at median split. Values were highlighted if Log-rank p-value was less than 0.01.

Abbreviations: CN, cellular neighbourhood; TSI, tumor stromal interface; HR, hazard ratio; KM, Kaplan-Meier.

The tumor microenvironment (TME) operates as an integrated entity, where distinct immune cell clusters interact to coordinate anti-tumor responses. As such, we employed spatially resolved clustering algorithms previously described by Schürch *et al*[2] to retrieve cellular neighbourhood (CN) information. Broadly, our analyses demonstrated CNs that differed in immune cell repertoire and diversity. For instance, CN7 was found to predominantly comprise tumor cells (98.6%) with low immune cell diversity (Shannon Index = 0.097). In contrast, CN2 and CN9 were found to have high immune cell diversity (CN2, Shannon Index = 1.715; CN9, Shannon Index = 1.482).

Compartment specific comparative proportions of CNs were found to be associated with survival outcomes and response. For instance, patients with high proportions of CN9 cells within stromal compartments were associated with favorable survival outcomes. Conversely, patients with high proportions of CN2 cells (heterogenous mixture of tumor cells, endothelial cells, neutrophils, memory T cells and CAFs) within the tumor stromal interface (TSI) were associated with poorer survival outcomes (**Figure R8b**). It was interesting to note that CN10, which was primarily composed of CAFs, did not modulate treatment response to the same extent and magnitude as CN2 (OS within TSI, HR=2.28, p=0.196; PFS within TSI, HR=1.76, p=0.337) (**Figure R8b**).

Our findings underscore the critical role of spatial features in reflecting the intricate, dynamic interplay of immune cells, providing valuable prognostic and therapeutic insights into tumor responses. Specifically, we demonstrate that the infiltration of immunosuppressive cells, such as CAFs, within tumor compartments—particularly their proximity to CD8+ T cells—attenuates treatment efficacy and modulates therapeutic responses. Moreover, the immunosuppressive influence of CAFs is not uniform but spatially dependent, further emphasizing the importance of cellular architecture in shaping immune evasion and resistance mechanisms.

References

[1] Jia K, Chen Y, Sun Y, et al. Multiplex immunohistochemistry defines the tumor immune microenvironment and immunotherapeutic outcome in CLDN18.2-positive gastric cancer. BMC Med. 2022 Jul 11;20(1):223. doi: 10.1186/s12916-022-02421-1. Erratum in: BMC Med. 2023 Sep 29;21(1):369. doi: 10.1186/s12916-023-03090-4. PMID: 35811317; PMCID: PMC9272556.

[2] Schürch CM, Bhate SS, Barlow GL, et al. Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front. Cell. 2020 Sep 3;182(5):1341-1359.e19. doi: 10.1016/j.cell.2020.07.005. Epub 2020 Aug 6. Erratum in: Cell. 2020 Oct 29;183(3):838. doi: 10.1016/j.cell.2020.10.021. PMID: 32763154; PMCID: PMC7479520.

R2C2 Based on the study scheme (Fig1A), the spatial transcriptomics analyzed in Fig6 were from pre-treatment and post-triplet samples. Here, the authors need to further consider that the changes in immune contexture after treatment may be the confounding effect of TMZ and immunotherapy, not just immunotherapy alone. Admittedly, as mentioned in the Discussion section, biopsy samples post-TMZ and pre-immunotherapy are missing. Can peripheral blood samples post-TMZ help solve this problem?

Reply: We appreciate the Reviewer's insightful comment regarding the potential confounding effects of TMZ and immunotherapy on changes in immune contexture. While we fully acknowledge the value of obtaining biopsies post-TMZ and pre-immunotherapy, doing so presents practical challenges within current clinical workflows. In patients who respond well to TMZ and proceed directly to immunotherapy, additional tissue sampling is often not clinically indicated and may pose unnecessary procedural risk without influencing treatment decisions. As such, while ideal from a research standpoint, routine pre-immunotherapy biopsies remain difficult to justify in this setting.

As rightfully pointed out by the Reviewer, we conducted flow cytometry as an alternative to evaluate immune cell activation post-TMZ. Previous work by Chin *et al* [1] and Hwang *et al* [2]

have demonstrated that peripheral blood immune cell subsets reflect changes in the tumor microenvironment. As advised by the Reviewer, we have expanded our analyses to take into consideration how **dynamic changes immune cell proportions** post-TMZ are associated with response.

Post-TMZ increase in exhausted T cell proportions predicts poor response

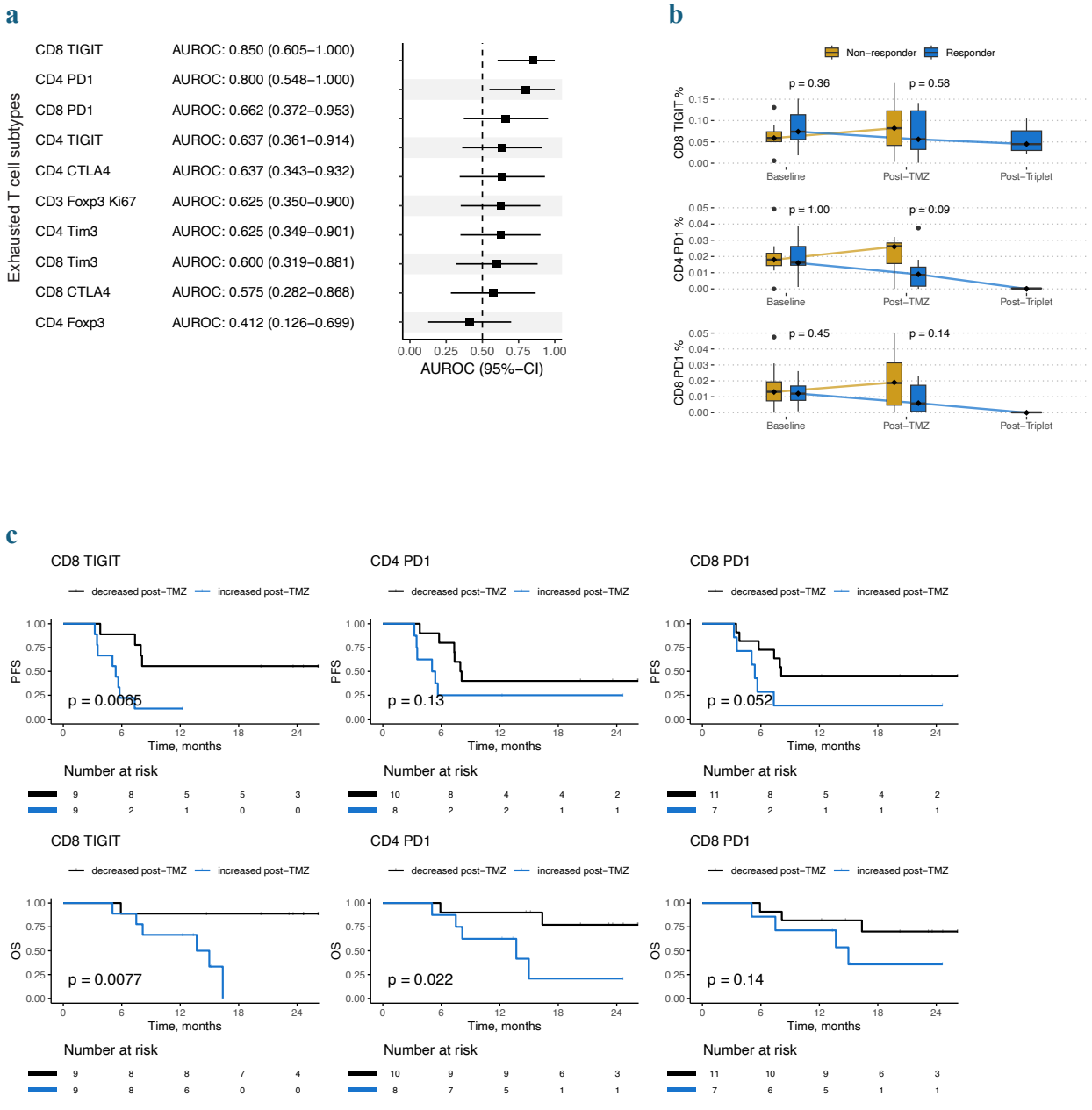
Fold change values (post-TMZ/baseline) of CD8+TIGIT+ (AUROC: 0.850 95%-CI: 0.605 - 1.000) and CD4+PD1+ (AUROC: 0.800 95%-CI: 0.548 - 1.000) cell proportions were predictive of response (**Figure R9a, R9b**). When analyzed with paired sample comparisons at the patient-level, increased T cell exhausted proportions post-TMZ were predictive of poorer PFS and OS (**Figure R9c**). These findings are consistent with the established phenomenon in which T cell exhaustion is associated with poorer ICI response [3].

Spatial distribution of Cancer Associated Fibroblasts (CAF) in baseline samples drives divergent T cell exhaustion trajectories with TMZ treatment

CAFs have been established to contribute to immunosuppression by secreting cytokines such as TGF-β and IL-6, which inhibit T cell activity. Further, CAFs can create physical barriers within the tumor microenvironment, preventing immune cells, particularly cytotoxic T cells, from effectively infiltrating the tumor. The immunosuppressive property of CAFs was confirmed through our Lunaphore dataset, in which we found that CAFs were broadly negatively correlated with other immune cell types across all compartments and especially so at the tumor stromal interface (TSI) (**Figure R9d**).

Consistently, patients that did not respond to TMZ alone were characterized by increased CAF infiltration within tumor and TSI compartments. In contrast, patients with subsequent ICI response were characterized by CAF distributions within stromal compartments of baseline samples (**Figure R9e**).

These findings, while exploratory, may collectively suggest that the CAF abundance within tumor and TSI compartments precludes effective T cell infiltration. We further posit that ineffective T cell infiltration in the context of continuous neoantigen stimulation with TMZ may drive T cell exhaustion.



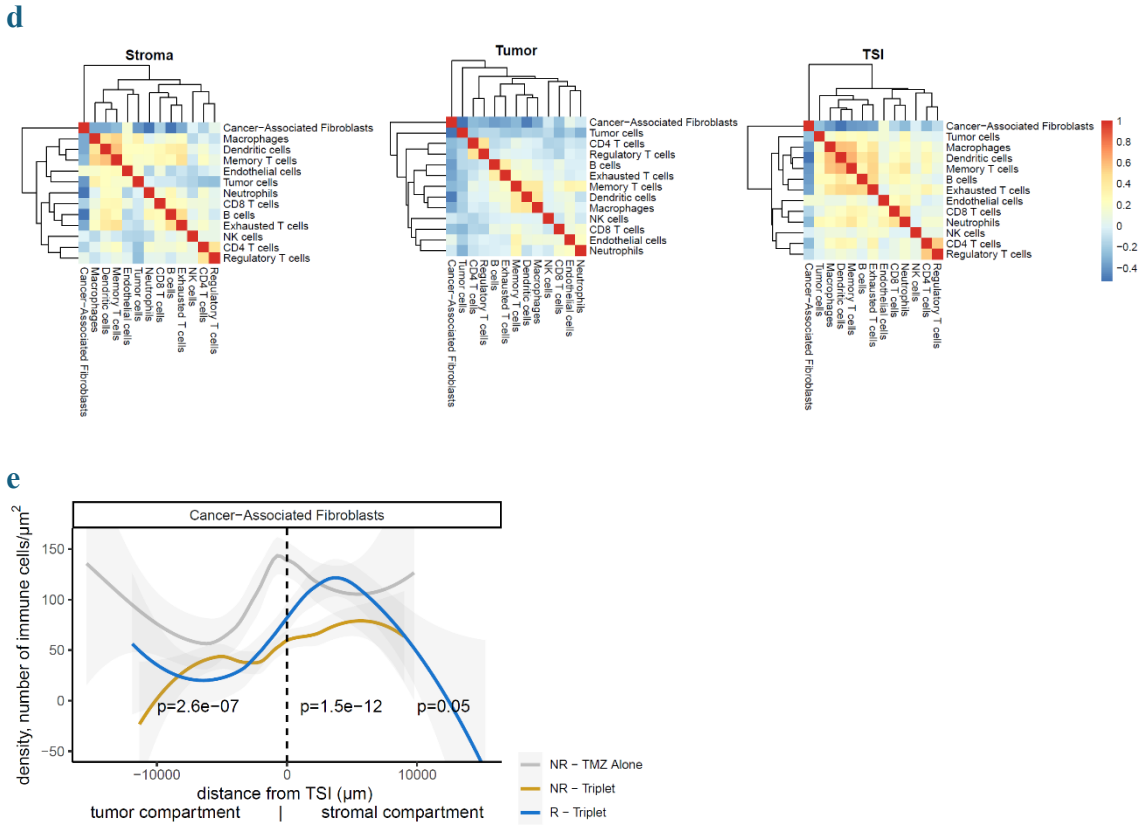


Figure R9

a) Forest plot of AUROC values of post-TMZ/baseline immune cell ratio against response.
b) Longitudinal box plots depicting immune cell type proportions evolution with treatment. P-values were retrieved from a two-sided T-test.
c) KM plots (OS, PFS) of participants stratified by dynamic changes in immune cell proportions post TMZ. P-values were retrieved from a Log-rank test.
d) Correlograms (Spearman's Rho) of immune cell densities per ROI across compartments.
e) Distribution of immune cell densities against distance from TSI. P-values were retrieved with a two-sided ANOVA test.

Abbreviations: AUROC, area under receiver operating curve; TMZ, temozolomide; OS, overall survival; PFS, progression-free survival; ROI, region of interest; TSI, tumor stromal interface.

References

[1] Chin WL, Cook AM, Chee J, et al. Coupling of response biomarkers between tumor and peripheral blood in patients undergoing chemoimmunotherapy. Cell Rep Med. 2025 Jan 21;6(1):101882. doi: 10.1016/j.xcrm.2024.101882. Epub 2024 Dec 27. PMID: 39731918.

[2] Hwang M, Canzoniero JV, Rosner S, et al. Peripheral blood immune cell dynamics reflect antitumor immune responses and predict clinical response to immunotherapy. J Immunother Cancer. 2022 Jun;10(6):e004688. doi: 10.1136/jitc-2022-004688. PMID: 35688557; PMCID: PMC9189831.

[3] Chow A, Perica K, Klebanoff CA, Wolchok JD. Clinical implications of T cell exhaustion for cancer immunotherapy. Nat Rev Clin Oncol. 2022 Dec;19(12):775-790. doi: 10.1038/s41571-022-00689-z. Epub 2022 Oct 10. PMID: 36216928; PMCID: PMC10984554.

R2C3 The clinical application of this work can be further strengthened. For example, whether the compartment-specific abundance of CD8+GzmB+/CD8+KI67+ cells correlates with clinical indicators of mCRC patients (such as tumor burden, carcino-embryonic antigen, tissue PD-L1 expression, BRAF/KRAS mutation, and progression-free survival, etc.). In addition, based on the sample size currently available, these cellular predictors may be further explored to construct quantitative biomarkers to further guide treatment strategies.

Reply: As suggested by the reviewer, we compared CD8+GZMB+ and CD8+Ki67+ T cells against clinical indicators of tumor burden as reflected by the number of metastatic sites (**Figure R10a**), CEA level (**Figure R10b**), PD-L1 combined positive score (CPS) (**Figure R10c**), BRAF and KRAS mutation (**Figure R10d, e**). While there was some suggestion for which CD8+GZMB+ was found to be higher in tumor compartments in patients with higher tumor burden, this did not reach statistical significance (p=0.76) (**Figure R10a**). All other comparisons were found to be statistically insignificant.

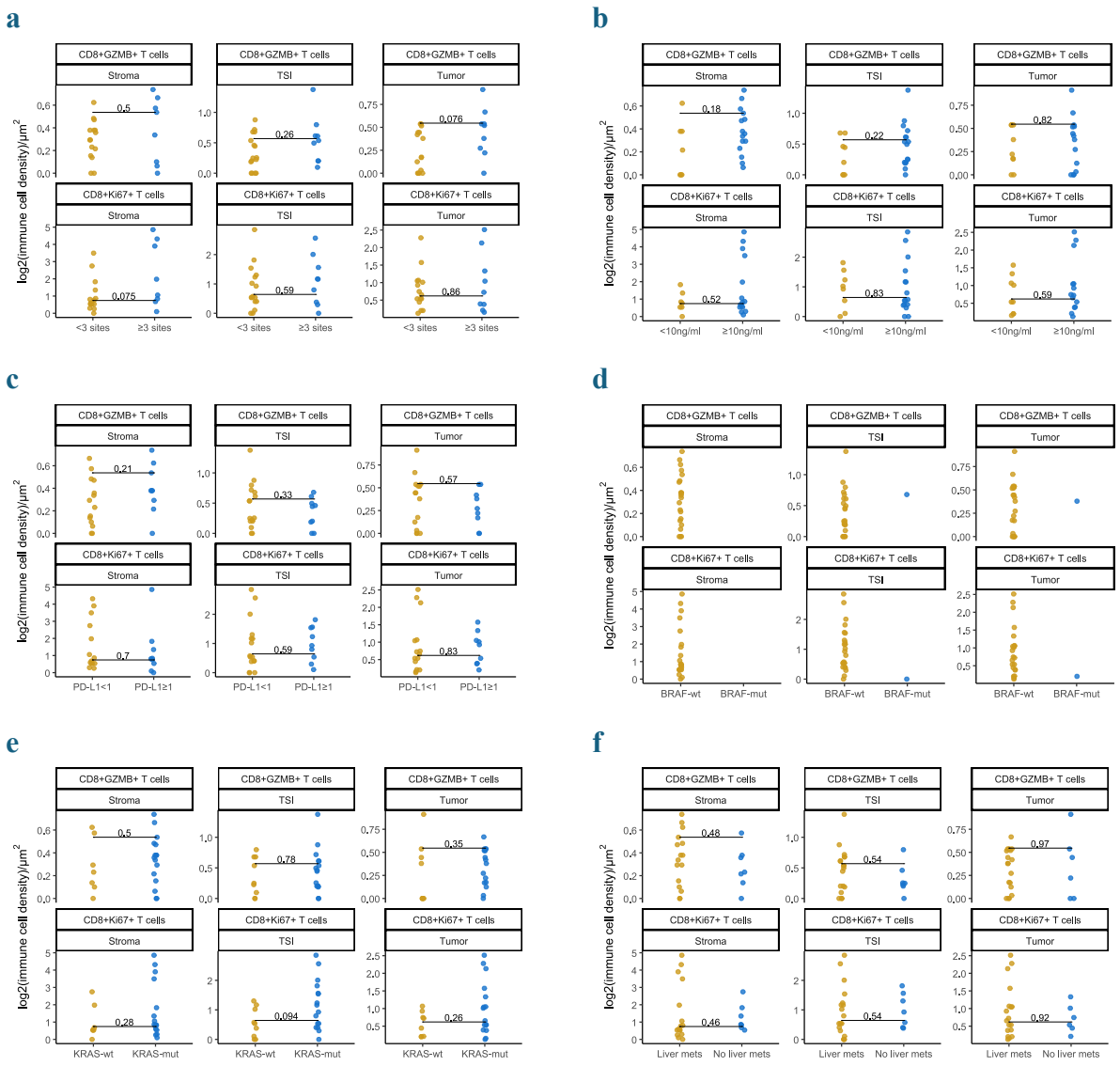


Figure R10 Dot plot comparisons for of CD8+Ki67+ and CD8+GZMB+ T cell densities stratified by

a) Number of metastatic sites

b) Baseline CEA levels
c) PD-L1 CPS
d) BRAF status
e) KRAS status
f) Liver metastasis

Abbreviations: CEA, carcino-embryonic antigen; CPS, combined positive score; -wt, wild-type; -mut, mutant; PD-L1, programmed death-ligand 1.

Remarks: Progression-free survival (PFS) was not evaluated as this was already analyzed with treatment response. P-values were retrieved from a two-sided Wilcoxon-test. Statistical testing for BRAF status was not pursued as only 1 patient with available Lunaphore stain data was found to be BRAF-mut.

Minor points:

R2C4 Line 274-275, ”a potential role for increased CD68+/CD163+ (M2-like) macrophages in the tumor-stromal interface of non-responders as a mechanism of resistance.” This conclusion, at least at this stage, is not fully supported by the data presented in the corresponding results section, since there are no more CD68+CD163+ cells in NR than in R shown in Fig5A.

Reply: We have replaced the analysis of tumor-associated macrophages (TAMs) with a series of novel spatial analyses that provide deeper insights into the tumor microenvironment. Specifically, our revised analysis reveals that the spatial proximity between CD8+ effector T cells and cancer-associated fibroblasts (CAFs) within the tumor compartment is associated with attenuated immunotherapeutic efficacy. This finding highlights a potential mechanism of CAF-mediated immune suppression and underscores the importance of spatial context in modulating treatment response.

R2C5 The use of “determinants” in the text may be a bit overdone in my opinion.

Reply: We thank the Reviewer for highlighting this and have carefully revised the phrasings in the manuscript to more accurately reflect our findings. We believe these changes better align with the intent and clarity the study’s results.

R2C6 The presentation of the data needs to be improved. The text in the figures (such as the GSEA results, the statistical plots, and the scale bars) is too small and hard to discern.

Reply: We thank the Reviewer for this comment. We have updated all figures to ensure clear visualisations throughout the manuscript.

Reply to Reviewer #3

Choo et al. (2024) utilises spatial multi-omic approaches to predict an immunotherapy response in microsatellite stable (MSS) metastatic colorectal cancer (mCRC). Following up on a previously published MAYA study, during which MSS mCRC patients were treated with temozolomide and immune checkpoint inhibitors (ICI), ipilimumab, and nivolumab, the authors sought to identify factors that predict response to this therapy regimen, given than tumor mutational burden was not completely predictive. The main discovery of the study is increased proliferative CD8+ T-cells and CD68+ CD163+ tumour-associated macrophages (TAMs) located within the tumour and outside of the tumor, respectively.

R3C1 Whilst sample collection across the treatment timepoints & spatial multi-omic datasets complementing one another are invaluable, the findings as presented are not novel.

Reply: We thank the Reviewer for the comment. Our study represents the first to explore the novel combination of TMZ and immune checkpoint inhibition (ICI) within the spatial context of colorectal cancer (CRC). The fact that cell types historically recognized as predictive markers of ICI responsiveness were consistently observed in our analysis serves not as a point of concern, but rather as a validation of our findings. This recurrence reassures us that our spatially resolved approach is capturing critical elements of the tumor-immune interface. It underscores the robustness of these immune signatures and strengthens the relevance of our insights in advancing the understanding of immune modulation within CRC. This alignment with established immunological predictors further enhances the translational significance of our work and its potential impact in refining future therapeutic strategies.

R3C2 There are also various issues regarding sample sizes, biology, and study design. Further analyses and revision are required for such a study to be a strong candidate for Nature Communications.

Reply: We sincerely thank the Reviewer for their valuable feedback. We are committed to conducting additional analyses and revisions to further strengthen our study, and we look forward to improving the manuscript accordingly.

Comments:

R3C3 1. The major finding is the presence/absence CD8+ T cells and macrophages within or outside of the tumor proper as a predictor of ICI response in CRC. There has already been an avalanche of papers documenting this finding, so the major punchline of the manuscript is not novel. The authors spent a lot of resources to do spatial omics to discover things that we already know, whereas a couple of IHC stains would have given the same conclusions (given what is already known about ICI response).

Reply: We thank the Reviewer for the valuable comment, which provides an opportunity to further clarify and emphasize the novelty of our findings. This study serves as an update to an important clinical trial published in Journal of Clinical Oncology [1] and presented at ESMO Congress 2021 [2], showcasing updated data that confirm the remarkable durability of some responses. Given that the study group is now planning first-line trials and expanding TMZ-combination strategies, and in light of complementary efforts by other research teams such as Bardelli's and Luis Diaz's, we hope that this work helps to identify specific patient populations that benefit from these treatments. This will help avoid futile enrollment in trials and demonstrate that precision medicine is achievable even in MSS CRC, where immunotherapy has typically failed.

To truly offer long-term benefits to selected patients, it is essential to implement biomarker-driven, personalized immunotherapy combinations, much like the successful treatments developed for MSI tumors. In response to the Reviewer's comment, we have revised the discussion to highlight the limited immunotherapy strategies with promising data in this setting and the need to invest in spatial transcriptomics to better understand the mechanisms of immune resistance specific to each treatment subprotocol.

References

[1] Morano F, Raimondi A, Pagani F, et al. Temozolomide Followed by Combination With Low-Dose Ipilimumab and Nivolumab in Patients With Microsatellite-Stable, O6-Methylguanine-DNA Methyltransferase-Silenced Metastatic Colorectal Cancer: The MAYA Trial. J Clin Oncol. 2022 May 10;40(14):1562-1573. doi: 10.1200/JCO.21.02583. Epub 2022 Mar 8. PMID: 35258987; PMCID: PMC9084437.

[2] Pietrantonio, F., Morano F.,, Lonardi, S, et al. 383O MAYA trial: Temozolomide (TMZ) priming followed by combination with low-dose ipilimumab and nivolumab in patients with microsatellite stable (MSS), MGMT silenced metastatic colorectal cancer (mCRC) Annals of Oncology, Volume 32, S530 - S531

R3C4 2. The analysis gave shallow findings. Maybe the reason is that the sample size is relatively low n<10 for each group. Given that there is heterogeneity in CRC subtypes, this low sample size may not enable to query of deep mechanisms. Hence, the overall description of the results is severely lacking. What do you want the readers to concentrate on – outside of the obvious CD8+ T cells and macrophage infiltration.

Reply: We appreciate the constructive feedback and have made significant enhancements to our spatial analyses to address the points raised.

We have expanded our spatial analysis by introducing additional key cell types such as cancer-associated fibroblasts (CAFs), endothelial cells and memory T cells, just to name a few (**Supplementary Table 1**). This provides a more comprehensive characterization of the tumor microenvironment (TME), offering greater insight into its complexity. Furthermore, to provide a more robust interrogation of the TME, we have incorporated additional spatial features, which go beyond relative abundance of cell types per compartment.

We have introduced additional compartment specific spatial features including (1) Proximity analyses of adjacent cells: We now examine the spatial proximity of adjacent cells[1], analyzing how different cell types interact within the TME to better understand immune cell dynamics. (2) Relative distances of adjacent cells: Our analysis also includes compartment-specific relative distances between cell types, allowing us to quantify how immune cells are spatially organized across distinct compartments. (3) Cellular neighborhoods: We investigate the cellular neighborhoods to understand how spatially resolved immune cell clusters orchestrate responses or resistance to immune checkpoint inhibitors (ICIs). This aspect of the analysis provides insight into the microenvironmental context of immune cell function and therapy outcomes. (**Figure R11**) These revisions enhance our spatial analysis and provide deeper insights into the TME and its impact on immune responses, particularly in the context of ICI therapy.

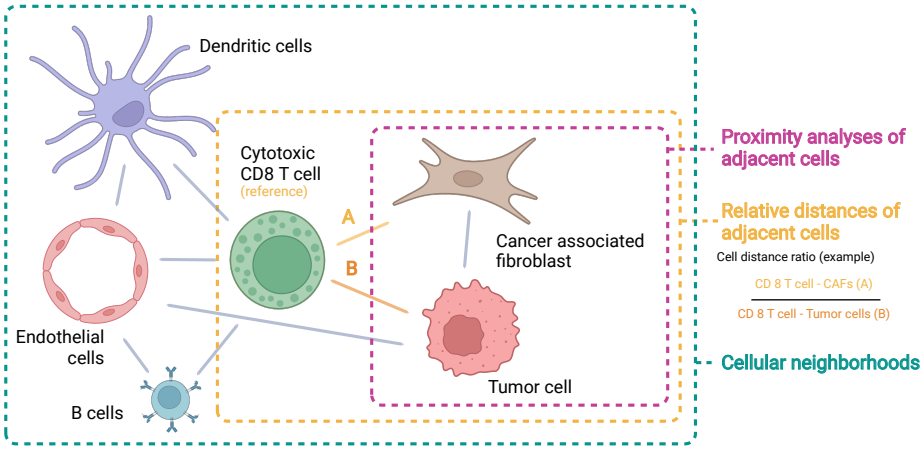


Figure R11 Illustration of additional spatial features incorporated into the revised manuscript

Remarks: IHC marker definitions per cell type are provided in **Supplementary Table 1**.

On inspection of adjacent cells abundances within the vicinity of 20µm of a tumor cell within tumor compartments of baseline samples, we observed scant numbers of effector adaptive immune cells (CD8 T cells and B cells) in contrast to immunosuppressive cells such as CAFs and memory

adaptive immune cells such as memory T cells. These findings are consistent with the known phenomenon of an immune-cold microenvironment of microsatellite stable (MSS) colorectal cancer (CRC). Of interest, higher cancer associated fibroblast (CAF) infiltration within tumor compartments is associated with poor response to TMZ alone, and subsequent ICI. In contrast, greater Memory T cell infiltration within tumor cells was associated with good response to ICI.

We also sought to understand how tumor-infiltrating cell types influenced effector CD8 T cells within tumor compartments. We evaluated relative cell distances from CD8 T cells and found that relative proximity of CD8 T cells to CAFs compared to tumor cells were associated with poor response to TMZ alone, although statistical significance was not reached. Likewise, relative proximity of CD8 T cells to endothelial cells were associated with poorer response.

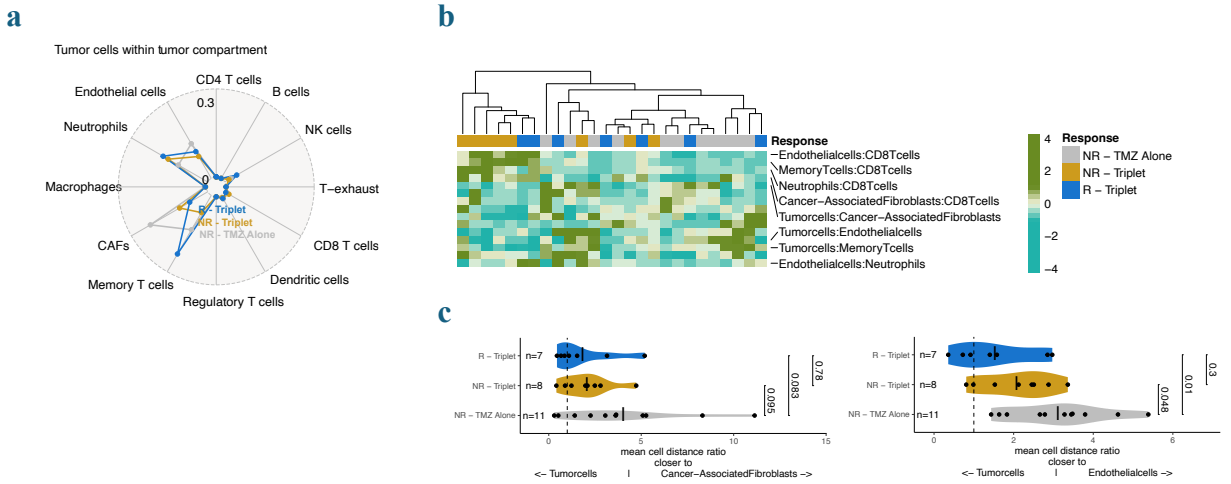


Figure R12

f) Adjacent cell abundances within 20μm of tumor cells within tumor compartments.

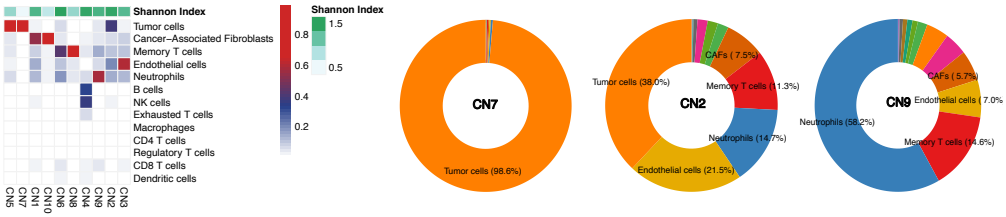
g) Heatmap of mean cell distance ratios from CD8 T cells per patient. Cell distance ratios flagged out had an ANOVA p-value of less than 0.15 across response group comparisons.

h) Selected violin plots comparing mean cell distance ratios per patient stratified by tumor response groups. P-values were retrieved from a two-sided T-test.

Abbreviations: CAF, cancer associated fibroblasts; T-exhuast, Exhausted T cells; NR – TMZ Alone, non-responder to temozolomide (TMZ) alone; NR - Triplet, responder to TMZ however non-responder to ICI; R - Triplet, responder to TMZ and ICI; n=, number of patients.

Remarks: IHC marker definitions are provided in **Supplementary Table 1**.

a



b

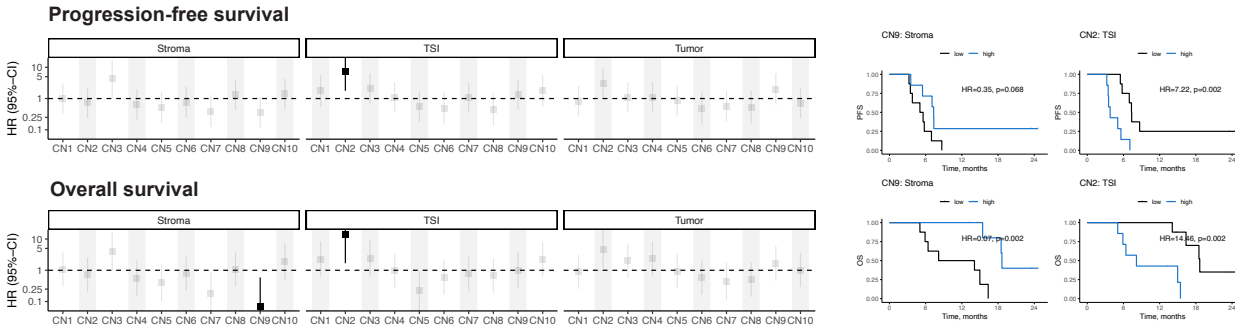


Figure R13

b) Heatmap and donut plots providing an overview of immune cell proportions per cellular neighbourhood.

b) Forest plot and selected KM plots of survival outcomes stratified by compartment specific CN cellular proportions at median split. Values were highlighted if Log-rank p-value was less than 0.01.

Abbreviations: CN, cellular neighbourhood; TSI, tumor stromal interface; HR, hazard ratio; KM, Kaplan-Meier.

The tumor microenvironment (TME) operates as an integrated entity, where distinct immune cell clusters interact to coordinate anti-tumor responses. As such, we employed spatially resolved clustering algorithms previously described by Schürch *et al*[2] to retrieve cellular neighbourhood (CN) information. Broadly, our analyses demonstrated CNs that differed in immune cell repertoire and diversity. For instance, CN7 was found to predominantly comprise tumor cells (98.6%) with low immune cell diversity (Shannon Index = 0.097). In contrast, CN2 and CN9 were found to have high immune cell diversity (CN2, Shannon Index = 1.715; CN9, Shannon Index = 1.482).

Compartment specific comparative proportions of CNs were found to be associated with survival outcomes and response. For instance, patients with high proportions of CN9 cells within stromal compartments were associated with favorable survival outcomes. Conversely, patients with high proportions of CN2 cells (heterogenous mixture of tumor cells, endothelial cells, neutrophils, memory T cells and CAFs) within the tumor stromal interface (TSI) were associated with poorer survival outcomes (**Figure R13b**). It was interesting to note that CN10, which was primarily composed of CAFs, did not modulate treatment response to the same extent and magnitude as CN2 (OS within TSI, HR=2.28, p=0.196; PFS within TSI, HR=1.76, p=0.337) (**Figure R13b**).

Our findings underscore the critical role of spatial features in reflecting the intricate, dynamic interplay of immune cells, providing valuable prognostic and therapeutic insights into tumor responses. Specifically, we demonstrate that the infiltration of immunosuppressive cells, such as CAFs, within tumor compartments—particularly their proximity to CD8+ T cells—attenuates treatment efficacy and modulates therapeutic responses. Moreover, the immunosuppressive influence of CAFs is not uniform but spatially dependent, further emphasizing the importance of cellular architecture in shaping immune evasion and resistance mechanisms.

References

[1] Jia K, Chen Y, Sun Y, et al. Multiplex immunohistochemistry defines the tumor immune microenvironment and immunotherapeutic outcome in CLDN18.2-positive gastric cancer. BMC Med. 2022 Jul 11;20(1):223. doi: 10.1186/s12916-022-02421-1. Erratum in: BMC Med. 2023 Sep 29;21(1):369. doi: 10.1186/s12916-023-03090-4. PMID: 35811317; PMCID: PMC9272556.

[2] Schürch CM, Bhate SS, Barlow GL, et al. Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front. Cell. 2020 Sep 3;182(5):1341-1359.e19. doi: 10.1016/j.cell.2020.07.005. Epub 2020 Aug 6. Erratum in: Cell. 2020 Oct 29;183(3):838. doi: 10.1016/j.cell.2020.10.021. PMID: 32763154; PMCID: PMC7479520.

R3C5 3. Study design and logic challenge – the study collected invaluable longitudinal samples. However, not much insights were gained from these samples. The conclusions derived from pre-treatment and post-treatment are largely similar. If that is the case, how does adding TMZ really help, since it is not via tumor mutations? What is the mechanism? Are you sampling too few patients that those that responded would have responded anyways without TMZ?

Reply: Our study builds on the principle that changes in TMB might have explained response to TMZ+ICI. These have been previously established by foundational work of Germano et al [1], who have already made key biological discoveries related to Temozolomide (TMZ) treatment. Their research demonstrated that the inactivation of MMR, driven by acquired resistance to the clinical agent temozolomide, increased mutational load, promoted continuous renewal of neoantigens in human colorectal cancers and triggered immune surveillance in mouse models [1,2].

Consistent with pre-clinical evidence, changes in TMB in our study was found to be associated with PFS (Pearson's r=0.863, p=0.0598 [n=5]) (**Figure R14**). However, this finding should be interpreted with caution for the following reasons. First, the sample size was notably small, with only 5 patients having paired pre- and post-treatment TMB levels, limiting the generalizability of the results. Second, post-treatment TMB was measured in samples already treated with immune checkpoint inhibition (ICI), making it challenging to disentangle the specific contribution of TMZ-

induced mutational changes from the immune-modulatory effects of the subsequent ICI therapy. Additionally, the timeframe between TMZ administration and ICI initiation may introduce confounding factors, as the mutational landscape could evolve not only due to TMZ but also in response to ongoing immune pressures [3]. Moreover, variability in the tumor microenvironment at the time of post-treatment sampling, as well as potential selection bias in the tumor subclones surviving TMZ and ICI, further complicates the interpretation of TMB changes in this setting.

Furthermore, it is also important to emphasize that while TMB has been proposed as a biomarker for predicting response to ICI, it is not the sole determinant of therapeutic efficacy. Several studies have demonstrated that TMB alone is insufficient to predict response to ICI [4]. This prompted us to further investigate the tumor microenvironment (TME) at spatial resolution, recognizing that a deeper understanding of the spatial organization of immune and non-immune cells within the tumor is crucial for unraveling the mechanisms underlying response and resistance to therapy.

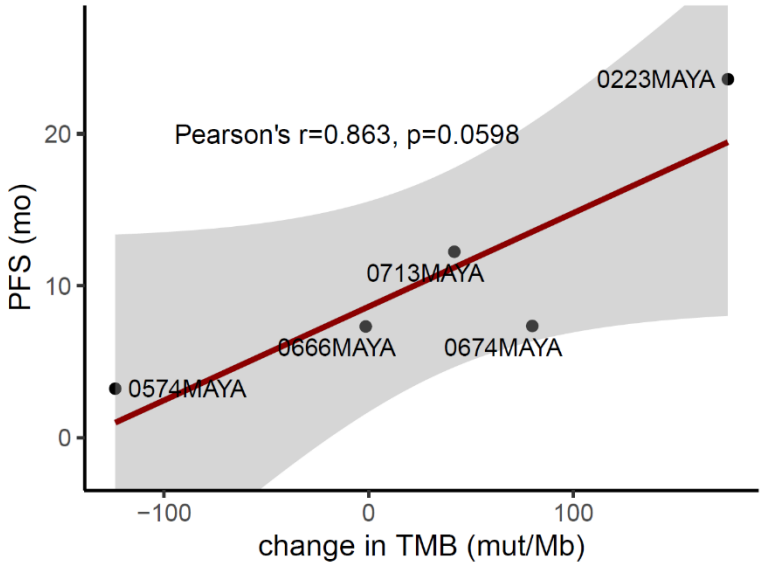


Figure R14 Correlation between TMB changes (pre vs post) and PFS

Abbreviations: TMB, tumor mutational burden; PFS, progression-free survival; mo, months.

As rightfully pointed out by the Reviewer, it would have been ideal to investigate if there was a subset of patients that would have responded to without TMZ. However, the trial was inherently not designed to answer this; an arm of patients treated with ICI only would be necessary.

References

[1] Germano, G., Lamba, S., Rospo, G. et al. Inactivation of DNA repair triggers neoantigen generation and impairs tumour growth. Nature 552, 116–120 (2017). https://doi.org/10.1038/nature24673

[2] Crisafulli G, Sartore-Bianchi A, Lazzari L, et al. Temozolomide Treatment Alters Mismatch Repair and Boosts Mutational Burden in Tumor and Blood of Colorectal Cancer Patients. Cancer Discov. 2022 Jul 6;12(7):1656-1675. doi: 10.1158/2159-8290.CD-21-1434

[3] Pham TV, Goodman AM, Sivakumar S, et al. Intra-patient stability of tumor mutational burden from tissue biopsies at different time points in advanced cancers. Genome Med. 2021 Oct 12;13(1):159. doi: 10.1186/s13073-021-00979-8. PMID: 34641956; PMCID: PMC8513181.

[4] Bortolomeazzi M, Keddar MR, Montorsi L, et al. Immunogenomics of Colorectal Cancer Response to Checkpoint Blockade: Analysis of the KEYNOTE 177 Trial and Validation Cohorts. Gastroenterology. 2021 Oct;161(4):1179-1193. doi: 10.1053/j.gastro.2021.06.064. Epub 2021 Jun 29. PMID: 34197832; PMCID: PMC8527923.

R3C6 4. The assays done on various sample timepoints do not seem to be unified. Perhaps more insights can be gleaned if the same type of data are collected on all specimens.

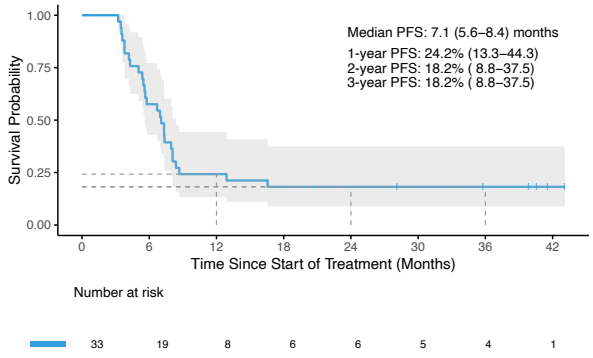
Reply: We appreciate the Reviewer's comment and agree that consistent data collection across all specimens could provide deeper insights. However, in clinical trials involving patients with refractory cancer, sample collection is often challenging. Differences in tissue availability, biopsy site, and sample size can limit what assays can be performed on each specimen. These practical

constraints, along with limited funding, required us to use multiple assay platforms and led to variability in sample timepoints. Despite this, we believe the dataset remains informative and supports robust analysis.

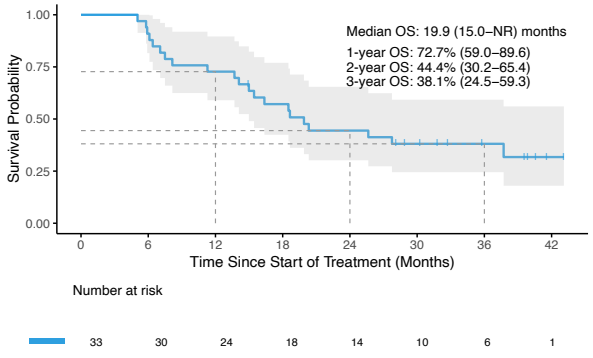
R3C7 5. Fig.1 and S.Fig.1 are identical (with additional data points) to the previously published MAYA trial paper. While the reviewer acknowledges the limited number of visualization methods with clinical data points, different visualisations are recommended.

Reply: We thank the Reviewer for this comment. The data presented in Figure 1 and Supplementary Figure 1 comprise updated survival information of the trial participants, with an updated median follow-up time of 35.8 months from 23.1 months [1]. As suggested by the Reviewer, we have also improved data visualization of our survival curves to demonstrate 95% confidence intervals, as well as survival rates at timepoints of 1-, 2- and 3-years.

a



b



c

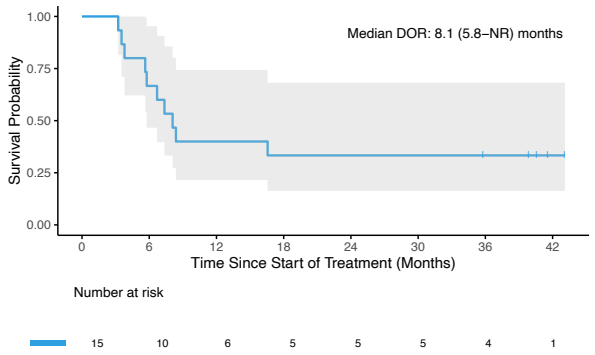


Figure R15 Revised survival curves of the MAYA trial

a) Progression-free survival
b) Overall survival
c) Duration of response

Abbreviations: mPFS, median progression-free survival; mOS, median overall survival; mDOR, median duration of response; CI, confidence interval; NR, not reached

References

[1] Morano F, Raimondi A, Pagani F, et al. Temozolomide Followed by Combination With Low-Dose Ipilimumab and Nivolumab in Patients With Microsatellite-Stable, O6-Methylguanine-DNA Methyltransferase-Silenced Metastatic Colorectal Cancer: The MAYA Trial. J Clin Oncol. 2022 May 10;40(14):1562-1573. doi: 10.1200/JCO.21.02583. Epub 2022 Mar 8. PMID: 35258987; PMCID: PMC9084437.

R3C8 6. Results of Fig.2c do not describe what is being shown. What are the reasons behind your statement “no differences in mutations, mutational signatures, copy number variations or tumor mutation burden (TMB) between responders and non-responders at baseline”?

Reply: We apologize for the lack of clarity in both the reporting of results and figures. In our revised manuscript, we now provide detailed statistical testing to support the argument that no major differences in mutational patterns (**Figure R16b**, Fisher's exact test p>0.05 for all comparisons) and tumor mutational burden (**Figure R16c**, T-test, p=0.87) were identified between responders and non-responders at baseline.

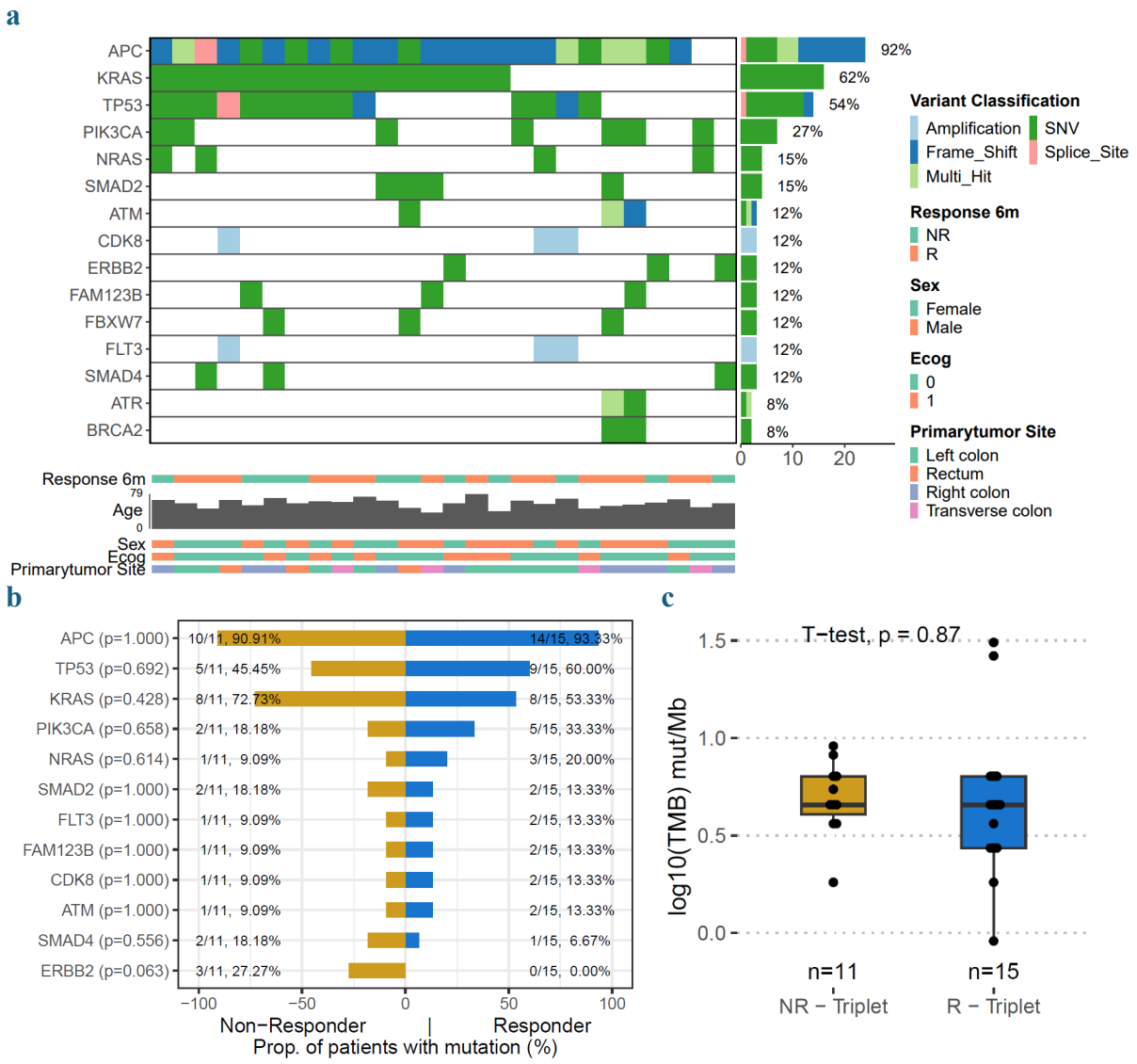


Figure R16 Comprehensive genomic profiling of baseline samples with FoundationOne CDx assay

a) Oncoplot of identified mutations on baseline samples
b) Co-barplot of mutations between Responders and Non-Responders. Only genes with more than 3 patients with mutations were shown here. P-values were retrieved from a Fisher's Exact Test.
c) Boxplot comparison of TMB between Responders and Non-Responders

Abbreviations: NR, non-responder; R, responder; TMB, tumor mutational burden; SNV, single nucleotide variant

R3C9 7. Abbreviations for Fig.2C are required in the figure legends. E.g. Amplification, Frameshift, etc.

Reply: We thank the Reviewer for this comment. All figure legends have been updated to ensure abbreviations are clarified.

R3C10 8. Citation is required for MuSiCa.

Reply: We thank the Reviewer for this comment. The citation for MuSiCa [1] has been inserted into the methods section.

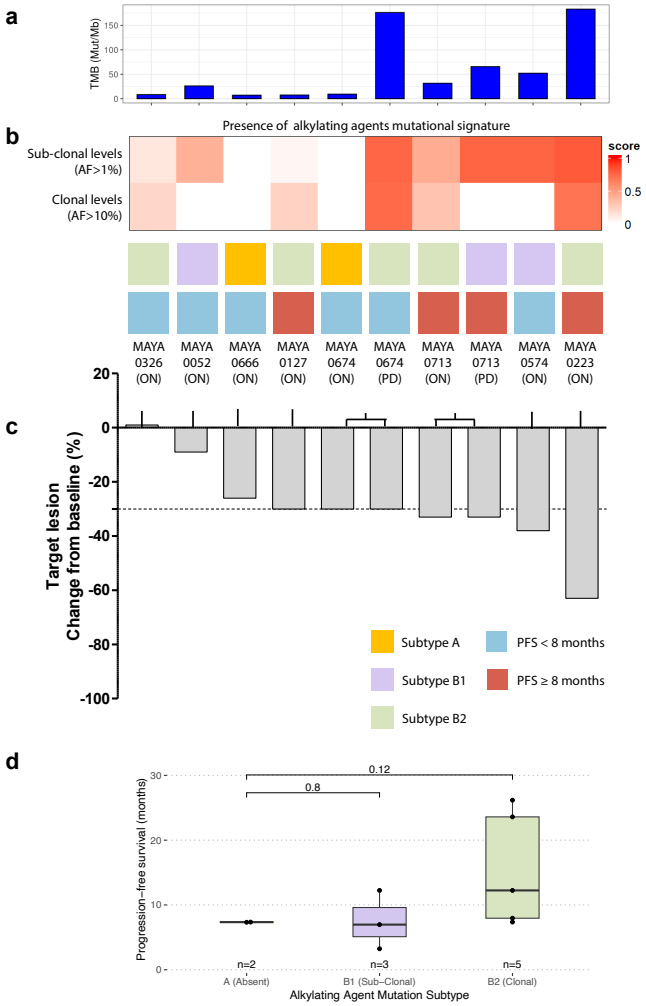
Reference

[1] Díaz-Gay M, Vila-Casadesús M, Franch-Expósito S, et al. Mutational Signatures in Cancer (MuSiCa): a web application to implement mutational signatures analysis in cancer samples. BMC Bioinformatics. 2018 Jun 14;19(1):224. doi: 10.1186/s12859-018-2234-y. PMID: 29898651; PMCID: PMC6001047.

R3C11 9. Results section and figure legends for S.Fig.2 are required.

Reply: We thank the Reviewer for the comment. As suggested, we have incorporated the results of findings for **Supplementary Figure 2** into the Results section of the revised manuscript.

> *In our previous work, we described a distinct alkylating agent mutational signature in tumor samples already treated with TMZ, which correlates with disease stabilization upon immune checkpoint inhibition [1]. Albeit a small sample size, we note a non-significant trend towards longer progression-free survival in patients with the signature at clonal levels (B2 subtype) compared to patients without the signature (Wilcoxon Test, p=0.12) (**Supplementary Figure 2a-d**).*



Supplementary Figure 2

A, Barplot showing the tumor mutational burden estimated as number of non-synonymous mutations/Mb in tumor samples obtained from patients after treatment with triplet therapy.

B, Heatmap showing the enrichment of the alkylating agents mutational signature (single base substitution 11 [SBS11]) occurring at sub-clonal levels (AF>1%) and clonal levels (AF>10%) in tumor samples obtained from patients after treatment with triplet therapy.

Colors in heatmap represent the score output of MuSiCa[2]. Patients were classified into three subtypes - A, B1, and B2 where subtype A corresponds to lack of the alkylating agents mutational signature, while B1 and B2 correspond to presence of the alkylating agents mutational signature at sub-clonal and clonal levels respectively.

C, Barplot representing the changes (%) in the target lesions from baseline. Colored boxes indicate patients who had durable response (progression free survival >8 months [red]) or non-durable response (progression-free survival <8 months [blue]).

D, Boxplot comparisons of PFS between patient subtypes.

Abbreviations: AF, allele frequency; ON, on-treatment tumor sample; PD, tumor sample obtained at time of progressive disease; TMZ, temozolomide; PFS , progression-free survival

Remarks: The alkylating agents mutational signature in samples treated with TMZ was previously described by Crisafulli et al[1]. The mutational signature correlates with disease stabilization upon immune checkpoint inhibition.

References

[1] Crisafulli G, Sartore-Bianchi A, Lazzari L, et al. Temozolomide Treatment Alters Mismatch Repair and Boosts Mutational Burden in Tumor and Blood of Colorectal Cancer Patients. Cancer Discov. 2022 Jul 6;12(7):1656-1675. doi: 10.1158/2159-8290.CD-21-1434. PMID: 35522273; PMCID: PMC9394384.

[2] Díaz-Gay M, Vila-Casadesús M, Franch-Expósito S, et al. Mutational Signatures in Cancer (MuSiCa): a web application to implement mutational signatures analysis in cancer samples. BMC Bioinformatics. 2018 Jun 14;19(1):224. doi: 10.1186/s12859-018-2234-y. PMID: 29898651; PMCID: PMC6001047.

R3C12 a. What is af?

Reply: We have revised the figure legend to clarify that af corresponds to “allele frequency”.

R3C13 b. What are the subtypes?

Reply: The alkylating agents mutational signature in samples treated with TMZ was previously described by Crisafulli et al[1]. The mutational signature correlates with disease stabilization upon immune checkpoint inhibition. Supplementary Fig 2B is a heatmap showing the enrichment of the alkylating agents mutational signature (SBS11) occurring at sub-clonal levels (AF>1%) and clonal levels (AF>10%) in tumor samples obtained from patients after treatment with triplet therapy. Colors in heatmap represent the score output of MuSiCa[2].

Patients were classified into three subtypes - A, B1, and B2 where subtype A corresponds to lack of the alkylating agents mutational signature, while B1 and B2 correspond to presence of the alkylating agents mutational signature at sub-clonal and clonal levels respectively.

References

[1] Crisafulli G, Sartore-Bianchi A, Lazzari L, et al. Temozolomide Treatment Alters Mismatch Repair and Boosts Mutational Burden in Tumor and Blood of Colorectal Cancer Patients. Cancer Discov. 2022 Jul 6;12(7):1656-1675. doi: 10.1158/2159-8290.CD-21-1434. PMID: 35522273; PMCID: PMC9394384.

[2] Díaz-Gay M, Vila-Casadesús M, Franch-Expósito S, et al. Mutational Signatures in Cancer (MuSiCa): a web application to implement mutational signatures analysis in cancer samples. BMC

Bioinformatics. 2018 Jun 14;19(1):224. doi: 10.1186/s12859-018-2234-y. PMID: 29898651; PMCID: PMC6001047.

R3C14 c. There are patients with alphabets. It is surmisable that they are for the treatment stages, but please define A, B, and C.

Reply: We have edited the figure to clarify the corresponding timepoints the biopsies were obtained and improved the clarity of the figure legend.

R3C15 10. In Fig.3C, there are a lot of non-immune enriched pathways with higher normalised enrichment scores in the tumour. What are the non-immune pathways with the high NES in the tumour (between responders & non-responders)? These pathways could provide us with further insights into different responses.

Reply: As suggested, analysis of baseline non-immune enriched pathways between responders and non-responders within the tumor compartment highlighted distinct biological processes that may provide insights into the tumor's pre-treatment characteristics. In responders, pathways such as aerobic respiration, cellular respiration, and oxidative phosphorylation were enriched, suggesting that tumors with higher baseline metabolic activity may be more susceptible to Temozolomide (TMZ) + immune checkpoint inhibition (ICI). These metabolic processes are crucial for sustaining the energy demands of rapidly dividing cells, and their enrichment could indicate a tumor microenvironment more conducive to treatment-induced stress and subsequent immune-mediated regression.

Conversely, in non-responders, increased enrichment of pathways related to endocardial cushion formation, fibrinolysis, and regulation of vascular endothelial cellular proliferation suggests a different baseline tumor biology. These pathways may reflect a tumor environment favoring tissue remodeling, angiogenesis, and fibrinolytic activity, which could contribute to tumor persistence

and resistance to treatment. Such pathways might promote structural and vascular adaptations that enable tumor survival despite the therapeutic pressure from TMZ and immunotherapy.

That being said, immune-related pathways were the primary focus of our analyses, given their direct relevance to the combination therapy's mechanism of action. TMZ, in combination with immunotherapy, exerts its therapeutic effect through immune modulation, and thus, understanding immune cell dynamics, infiltration, and activation within the tumor microenvironment is critical to evaluating treatment efficacy. While the non-immune pathways offer important contextual information, immune-related processes are more pertinent to understanding the response to this specific therapy, justifying our decision to focus on these pathways in our analyses.

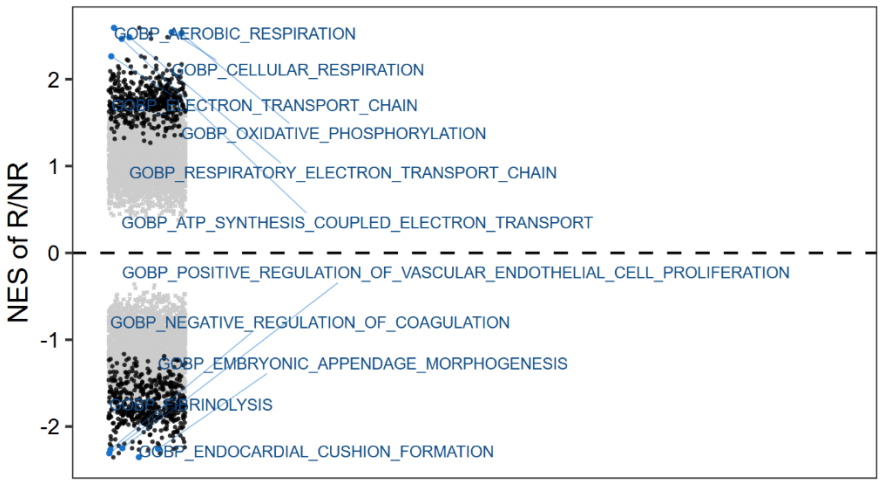


Figure R17 GO:BP Pathway analysis comparisons between R and NR within tumor compartments.

Remarks: Pathways highlighted had an adjusted p-value of <0.001 & |NES|>2.25.

Abbreviations: NES, normalized enrichment score; R, responder; NR, non-responder; GO, gene ontology; BP, biological processes.

R3C16 11. Fig.3D shows epithelial-to-mesenchymal transition (EMT) as highly enriched in the tumours of non-responders compared to responders. Are subtypes of the patient tumors available?

Reply: As suggested by the Reviewer, we have retrieved the Consensus Molecular Subtypes (CMS)[1] in colorectal cancer through whole transcriptome sequencing of baseline samples. Consistent with our cohort's profile, no samples were classified as CMS1 (characterized by MSI). Consistent with our upstream findings, samples with an EMT profile (CMS4) had a higher proportion of Non-Responders, although statistical significance was not reached (Fisher's, p=0.801); The proportion of Non-Responders was highest amongst CMS4 tumors at 40.0%(characterized by EMT), followed by CMS3 tumors at 33.3% and finally CMS2 tumors at 16.7% (**Figure R18**).

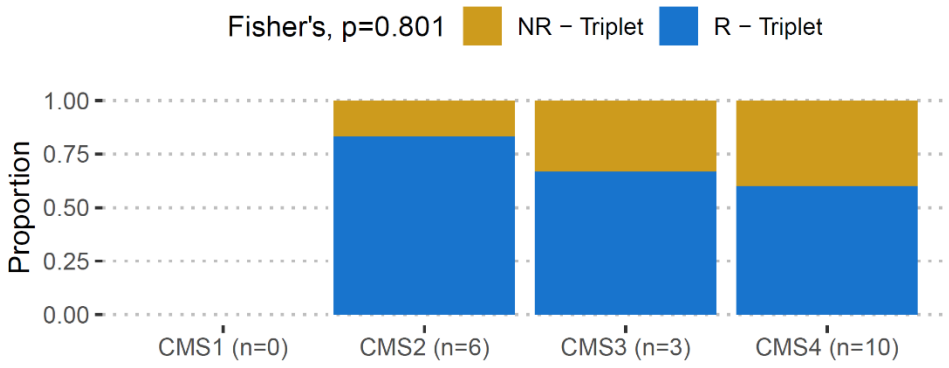


Figure R18 Stacked barchart comparing CMS subtype of baseline samples against response outcomes at 6 months

Abbreviations: CMS, consensus molecular subtype; NR, non-responder; R, responder

Remarks: CMS subtypes were retrieved through the CMSCaller package implemented in R.

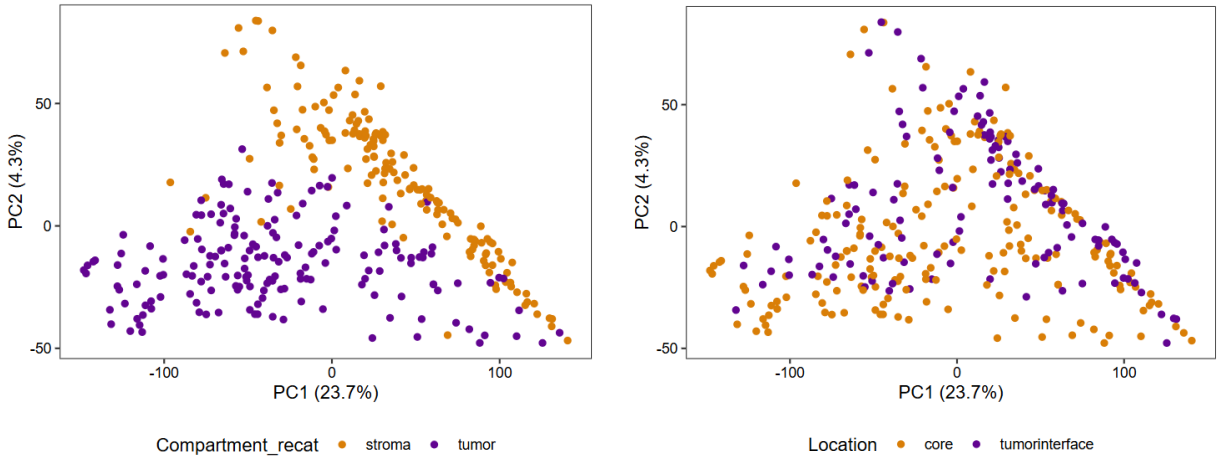
References

[1] Guinney, J., Dienstmann, R., Wang, X. et al. The consensus molecular subtypes of colorectal cancer. Nat Med 21, 1350–1356 (2015). https://doi.org/10.1038/nm.3967

[2] Eide, P.W., Bruun, J., Lothe, R.A. et al. CMScaller: an R package for consensus molecular subtyping of colorectal cancer pre-clinical models. Sci Rep 7, 16618 (2017). https://doi.org/10.1038/s41598-017-16747-x

R3C17 12. In S.Fig.3, add the proportions/percentages for PC1 & PC2.

Reply: Thank you for the suggestion. We have incorporated the percentage of variance explained by principal component (PC) 1 and 2.



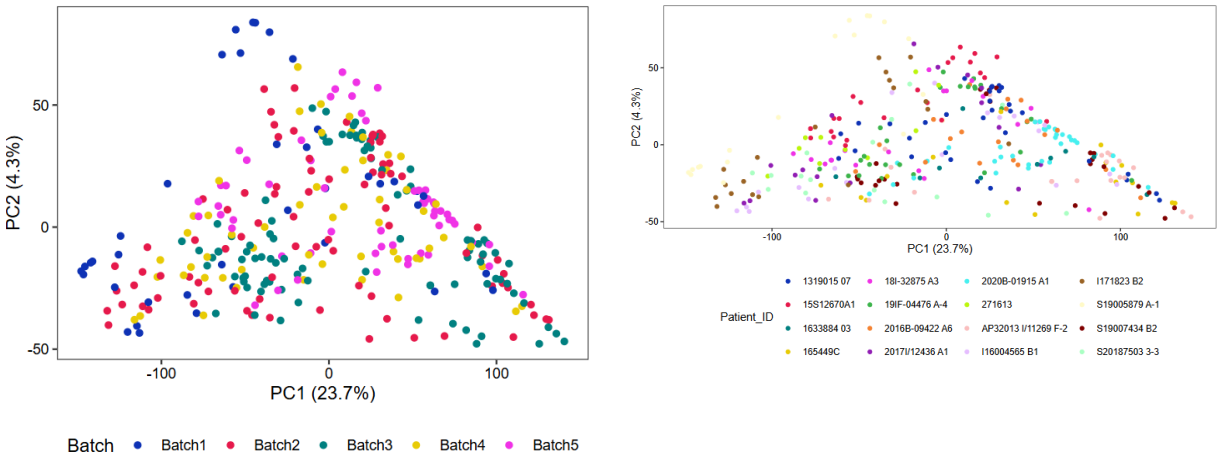


Figure R19 Revised PCA dimension reduction of gene expression in AOIs, colored according to labels shown

Abbreviations: PCA, Principal Component Analysis; AOI, areas of illumination

R3C18 13. Results section is required for S.Fig.4. No description is available.

Reply: As suggested by the Reviewer, we have provided further clarification regarding the differential gene expression analysis between responders and non-responders in both tumor and stromal compartments, as illustrated in our volcano plots in **Supplementary Figure 4**.

> *Differential gene expression yielded findings suggestive of biological processes that may potentially explain increased immunogenicity induced by temozolomide (TMZ) and enhanced response to immune checkpoint inhibitors (ICIs) in responders.*
>
> *In the tumor compartment, the upregulation of genes such as LCN2[1] and PLA2G2A[2] suggests that responders may have a pre-existing inflammatory or stressed tumor environment, potentially making these tumors more susceptible to further stress and immunogenicity induced by TMZ. Tumors exhibiting higher baseline inflammation or immune activation may be more likely to generate neoantigens or expose tumor-specific*

antigens upon DNA damage caused by TMZ, leading to enhanced immune recognition when ICIs are introduced.

In the stromal compartment, the upregulation of immune-related genes like CXCL9, CXCL8, and HLA class II molecules suggests a baseline immune-activated state. CXCL9 and CXCL8 are chemokines critical for recruiting effector immune cells such as T cells and neutrophils, and their higher expression may indicate that responders already have an immune-favorable microenvironment[3]. The presence of HLA-DQA1 and HLA-DQB1, which are involved in antigen presentation, hints that responders may have more effective immune surveillance capabilities at baseline, potentially leading to a more robust immune response to TMZ-induced neoantigens and subsequent ICI treatment.

These baseline characteristics suggest that responders have a tumor microenvironment more primed for immune engagement, which may explain their better outcomes with TMZ plus ICI therapy. The inherent presence of inflammatory, immune recruitment, and antigen-presentation pathways in responders' tumors likely facilitates a stronger response to TMZ-induced immunogenicity and the subsequent immune-activating effects of ICIs.

References

[1] Santiago-Sánchez GS, Pita-Grisanti V, Quiñones-Díaz B, et al. Biological Functions and Therapeutic Potential of Lipocalin 2 in Cancer. Int J Mol Sci. 2020 Jun 19;21(12):4365. doi: 10.3390/ijms21124365. PMID: 32575507; PMCID: PMC7352275.

[2] Hidalgo I, Sorolla MA, Sorolla A, et al. Secreted Phospholipases A2: Drivers of Inflammation and Cancer. Int J Mol Sci. 2024 Nov 19;25(22):12408. doi: 10.3390/ijms252212408. PMID: 39596471; PMCID: PMC11594849.

[3] Nagarsheth, N., Wicha, M. & Zou, W. Chemokines in the cancer microenvironment and their relevance in cancer immunotherapy. Nat Rev Immunol 17, 559–572 (2017). https://doi.org/10.1038/nri.2017.49

R3C19 14. Add representative images of tumour, TSIs, and stroma selection, separated by non-responders and responders as a supplementary figure (for Fig.4).

Reply: We thank the Reviewer for this suggestion. We have included representative images of tumor, TSI and stromal compartments between Responders and Non-Responders into the **Supplement** (**Figure R20**).

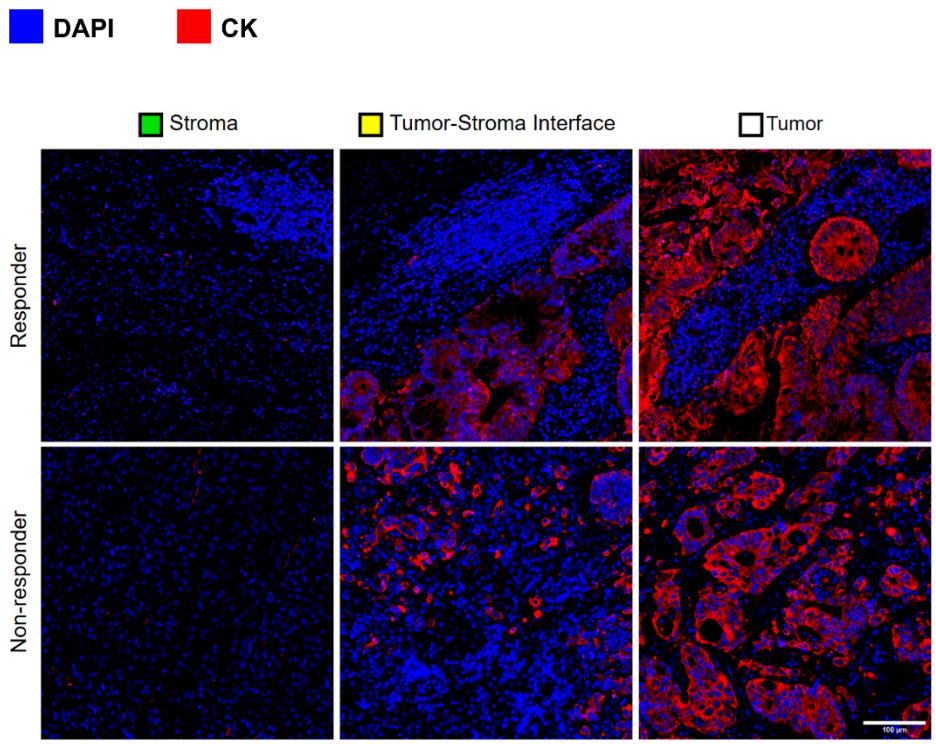


Figure R20 Representative DAPI/CK-stained images across tumor, TSI, and stromal compartments in responders and non-responders. Scale bar in white corresponds to 100 μm.

Abbreviations: TSI, tumor stromal interface

R3C20 15. Describe a strategy for how you picked TSI. 50% tumour + 50% stroma? How you choose TSI ROI would skew the results drastically.

Reply: The selection of the tumor-stromal interface (TSI) in our study was performed intentionally and manually by our pathologist using hematoxylin and eosin (H&E) stained slides. This manual approach was chosen to ensure the accurate identification of the tumor-stromal boundary, given the complexity and variability in histological appearance within the colon.

While computational methods have been demonstrated by others to automate the retrieval of the tumor-stromal interface with deep learning models[1], these approaches were not applicable in our context due to the presence of normal epithelium in the colon, which can also express cytokeratin (CK). This overlap would confound automated algorithms reliant on CK staining to differentiate between tumor and non-tumor tissues. In contrast, manual selection based on the gross histological characteristics of cancer—including irregular glandular architecture, nuclear atypia, and stromal desmoplasia—allows for a more precise demarcation of the interface, minimizing the risk of including normal CK-positive epithelium in our analysis.

Given the need for accuracy in defining the regions of interest, particularly for spatial profiling, we prioritized this manual approach to ensure the reliability of the results and to align with the histopathological complexity of the colon tumor microenvironment.

References

[1] Jiang J, Tekin B, Yuan L, et al. Computational tumor stroma reaction evaluation led to novel prognosis-associated fibrosis and molecular signature discoveries in high-grade serous ovarian

carcinoma. Front Med (Lausanne). 2022 Sep 7;9:994467. doi: 10.3389/fmed.2022.994467. PMID: 36160147; PMCID: PMC9490262.

R3C21 16. In Fig.4 & S.Fig.5, cytokeratin (CK) was used to identify tumours. Did you verify the absence of normal epithelium with morphology? Or did you use the CK level threshold to define tumours (based on "image analysis confirmed increased expression of CK in the tumor regions)? What if the tumor cells lost their CK expression.

Reply: As rightfully pointed out by Reviewer, CK is also positive in normal epithelium. As such, the selection of tumor compartments was done manually by our pathologist using hematoxylin and eosin (H&E) stained slides. Manual selection based on the gross histological characteristics of cancer—including irregular glandular architecture, nuclear atypia, and stromal desmoplasia—allows for a more precise demarcation of the interface, minimizing the risk of including normal CK-positive epithelium in our analysis.

Given the need for accuracy in defining the regions of interest, particularly for spatial profiling, we prioritized this manual approach to ensure the reliability of the results and to align with the histopathological complexity of the colon tumor microenvironment.

R3C22 17. In Fig.4C, better representative images are required. "Tumors" are not zooming into tumors. Also need scale bars for the tumor images.

Reply: We thank the Reviewer for the comment and have revised the selection of images, ensuring that they are representative of the regions of interest (ROI) depicted (**Figure R21**).

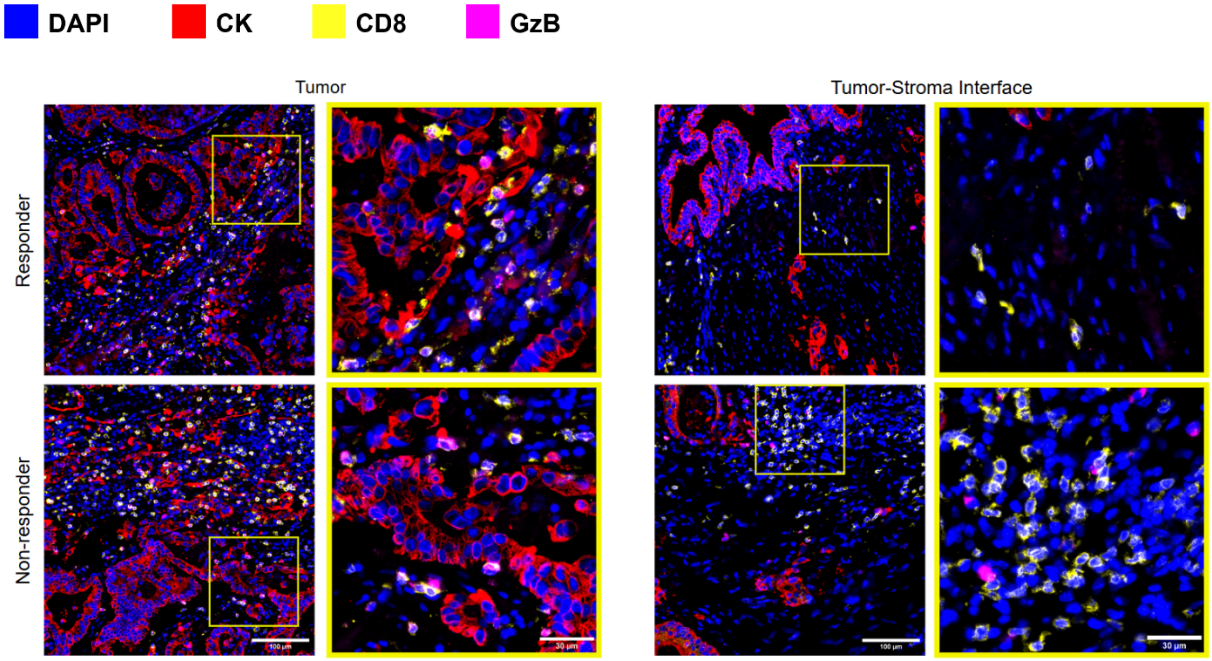


Figure R21 Revised selection of Lunaphore images

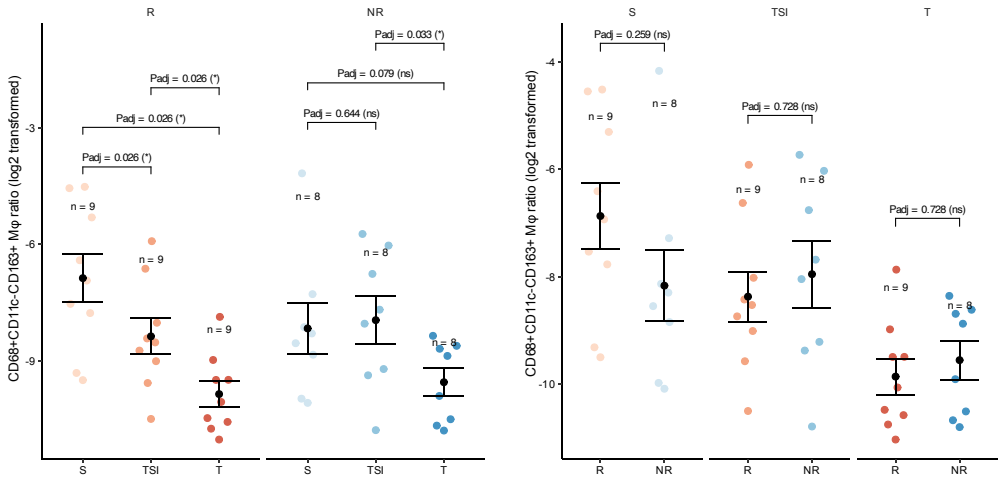
Multiplex immunofluorescence images show paired ROIs from tumor and TSI regions with corresponding high-magnification insets (yellow boxes). Stains include DAPI (blue), CK (red), CD8 (yellow), and Granzyme B (magenta). Scale bars in white correspond to 100 µm for the main panels and 30 µm for high-magnification insets.

Abbreviations: TSI, tumor stromal interface

R3C23 18. In Fig.5A, the proportion of TAMs in TSI between responders Vs. non-responders are not statistically different. Instead of calculating the proportion of TAMs in the given area, which can introduce technical artefacts of choosing TSI, the distance and relative frequency of TAMs from the tumour borders/edges may be of more value.

Reply: We thank the Reviewer for the suggestion and comment. We have incorporated distance from TSI into the analyses. Consistent with our previous approach with dichotomization of regions of interest (ROI) (**Figure R22a**), there was only a slight difference in M2-polarised macrophage distribution, with a minor shift of M2-polarised macrophages towards the stromal compartments in Responders compared to Non-Responders (**Figure R22b**). In light of these findings, we have replaced the analysis of tumor-associated macrophages (TAMs) with a series of novel spatial analyses that provide deeper insights into the tumor microenvironment. Specifically, our revised analysis reveals that the spatial proximity between CD8+ effector T cells and cancer-associated fibroblasts (CAFs) within the tumor compartment is associated with attenuated immunotherapeutic efficacy. This finding highlights a potential mechanism of CAF-mediated immune suppression and underscores the importance of spatial context in modulating treatment response.

a



b

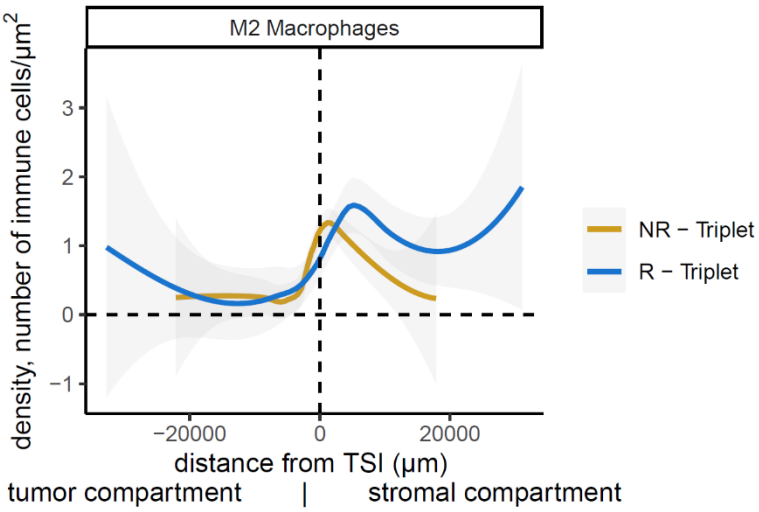


Figure R22

a) Comparisons of M2-polarised cell density across response status and compartments
b) M2-polarised cell density curves against distance from TSI stratified by response patterns

Abbreviations: R, responder; NR, non-responder; TSI, tumor stromal interface

Remarks: Tumor compartments were taken to be at negative distance while stromal compartments were taken to be at a positive distance from the TSI. Smoothed conditional means density curves were retrieved with the geom_smooth() function in ggplot2 in R.

R3C24 19. Inconsistent scale bars from Fig.5B.

Reply: We apologize for this error. We have updated all figures including those with scale bars to ensure consistency and ensure clear visualisations throughout the manuscript.

R3C25 20. For Fig.6A, choose better representative images for responders Vs. non-responders.

The responder sample mainly shows stroma, whereas the non-responder sample shows mainly tumours. If it is due to a successful response to triple ICI, state it (reduced tumour size…).

Reply: We thank the Reviewer for this feedback. The aim of this figure was to demonstrate how tumor and stromal regions were annotated using a machine learning approach and then confirmed when upregulation of EPCAM transcripts in tumor regions were visualised upon pathology review. We have replaced this with **Figure R23** which more closely aligns with our intention; it demonstrates examples of tumor and stromal segment identification via pathologist-guided supervised machine learning approach using QuPath (**Figure R23a)**. This method of tissue category assignment provided superior segmentation accuracy compared to unsupervised clustering approaches such as BayesSpace (**Figure R23b)**. Compartment validation was supported by tumor-specific enrichment of epithelial transcripts (e.g., EPCAM) (**Figure R23c**).

A B C

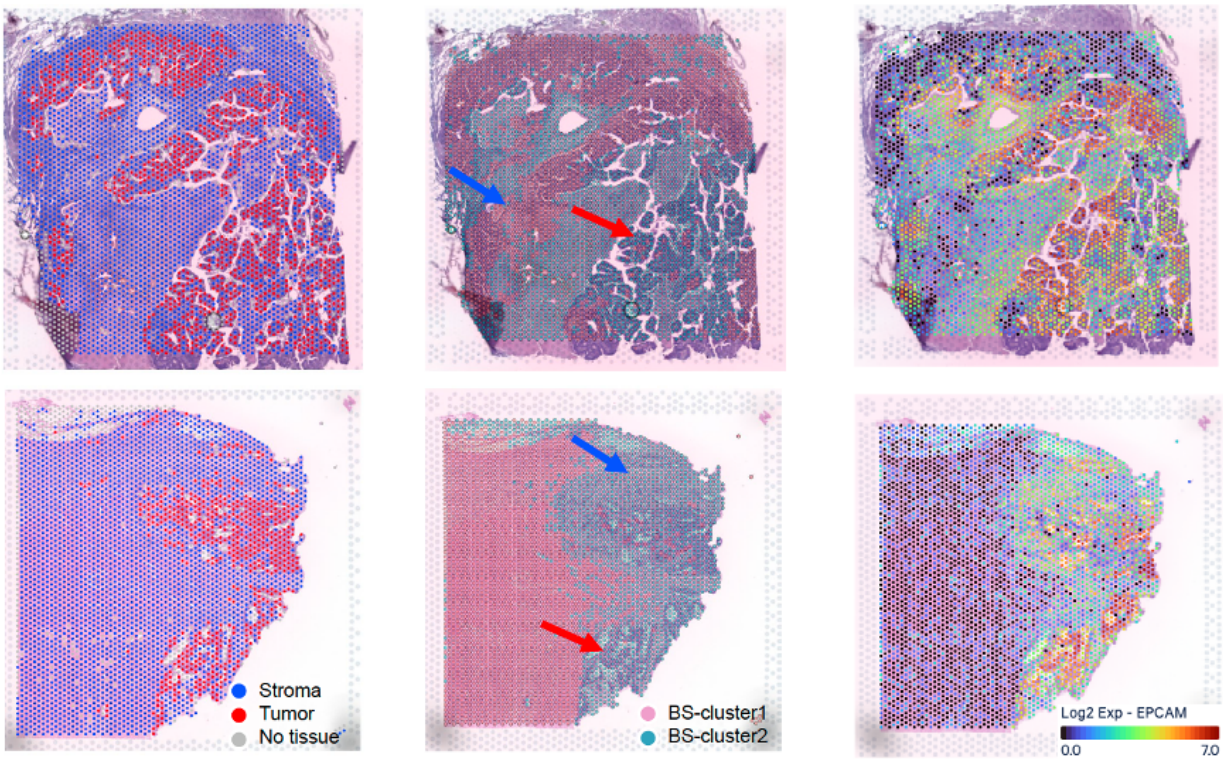


Figure R23

a) Tumor and stromal segment identification via pathologist-guided supervised machine learning approach using QuPath
b) Tumor and stroma segment identification via BayesSpace clustering. Blue arrows indicate pathologist-annotated stroma spots that were clustered as tumor, while red arrows indicate pathologist-annotated tumor spots that were clustered as stroma.
c) EPCAM gene transcripts are upregulated in tumor regions

R3C26 21. Better visualisation and description for Fig. 6B is required.

Reply: We thank the Reviewer for this feedback. We have edited the visualisation for the heatmap in accordance with the suggestions and improved the description to clarify that the heatmap represents cell-type abundance scores related to immune cell subsets, fibroblasts and endothelial cells.

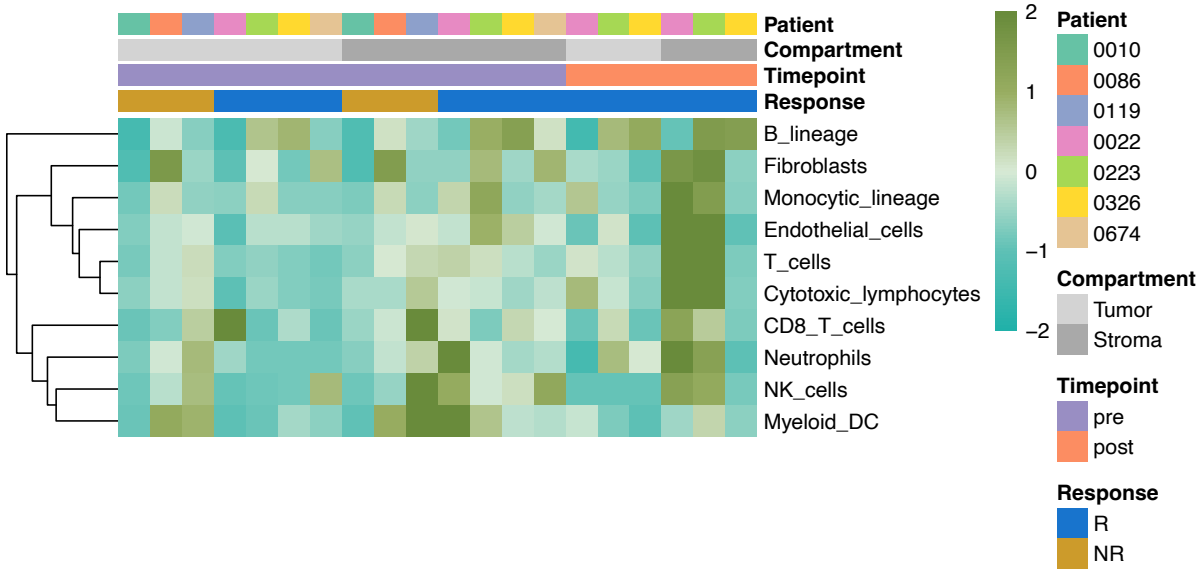


Figure R24 Revised **Figure 5d**, Spatial cell type enrichment was estimated using UCell scoring applied to log-normalized gene expression data. Gene sets were derived from the MCP-counter framework to capture fibroblasts, endothelial cells, and key immune subsets across spatial transcriptomic spots. To facilitate comparative visualization, UCell scores were Z-score normalized within each cell type, highlighting relative enrichment patterns across clinical response groups (Responder vs. Non-Responder), timepoints (pre- vs. post-treatment) and tissue compartments (tumor vs. stroma).

Abbreviations: R, Responder; NR, Non-Responder; NK, Natural Killer; DC, Dendritic Cell

R3C27 a. How are immune scores the highest in fibroblasts in comparison to other main immune cells?

Reply: We thank the Reviewer for this important comment and apologize for the confusion caused by the previous heatmap labeling. The original version displayed unscaled cell abundance scores, which reflected the absolute relative abundance of each cell type across the dataset. This presentation emphasized that fibroblasts exhibited a higher overall signal compared to other cell types. We also acknowledge that the original figure legend was incorrectly labeled as representing immune scores, when in fact it encompassed stromal (e.g., fibroblasts, endothelial cells) and immune populations.

In the revised heatmap, we now apply Z-score normalization within each cell type to facilitate the comparison of relative enrichment patterns across tissue compartments, response groups, and timepoints. While this normalization no longer permits direct comparison of absolute abundance between cell types (e.g., fibroblasts vs. T cells), it enables more meaningful interpretation of within-cell-type variability across biological conditions—an approach that we believe better supports the objective of understanding treatment-associated changes in the tumor microenvironment.

We have updated the figure legend and Methods section to clarify this normalization step and to ensure the intended interpretation of the heatmap is accurately conveyed.

R3C28 b. Continuing the above enquiry, lines 333, 334 suggest the heatmap for fibroblasts, endothelial cells, and macrophages are cell numbers, not immune cell scores. Please elaborate.

Reply: We thank the Reviewer for this important observation and appreciate the opportunity to clarify. The heatmap presented in the updated Figure 5d (originally Figure 6B) does not represent absolute cell counts, but relative cell type enrichment scores derived from gene expression data.

We used predefined gene sets from the MCP-counter framework to estimate the enrichment of fibroblasts, endothelial cells, macrophages, and other stromal and immune cell subsets across spatial transcriptomic spots. These scores capture the degree to which the expression profile of a given spot aligns with the signature of each cell type and should be interpreted as a proxy for relative abundance based on transcriptional activity, not a direct quantification of cell numbers.

We have revised the text to clarify this distinction and ensured consistent terminology with the updated figure and methods description.

R3C29 c. Inter-patient variability seems to drive the immune cell score difference. Could you reorganise to match the order of patients pre- and post-treatment?

Reply: We thank the Reviewer for this feedback. We have edited the visualisation for the heatmap in accordance with the suggestions, reorganised to match the order of patients pre- and post-treatment (**Figure R24).**

R3C30 22. S.Fig.6 shows different T-cell subset changes across the ICI treatments. Validate the finding, albeit incomplete, in Visium. Distribution? Frequency?

Reply: We thank the reviewer for this important point. In Supplementary Fig. 6, which presents peripheral flow cytometry data, we observed an increase in $CD8^+TIGIT^+$ T cells in non-responders (NR) compared to responders (R) following treatment. Apart from a reduction in circulating $CD8^+PD1^+$ cells, no other significant differences were observed at static timepoints or relative to baseline in either group. To approximate changes in intratumoral $CD8^+$ T cell subsets in the Visium dataset, we assessed transcriptional correlates of T cell dysfunction using the TIDE-derived Dysfunction score[1]. This model incorporates multiple canonical exhaustion markers (e.g., PDCD1, TIGIT, CTLA4, LAG3, IFNG) and has been validated as a proxy for $CD8^+$ T cell dysfunction in the tumor microenvironment. Within tumor regions enriched for $CD8^+$ T cells

(defined as the top quartile of $CD8^+$ abundance), we compared Dysfunction scores pre- and post-treatment. Although scores were numerically higher post-treatment in responders, the difference was not statistically significant. Notably, post-treatment samples were available for only three patients, all of whom were responders, limiting our ability to compare differences between responders and non-responders (**Figure R25a**).

In parallel, we examined spatial gradients of CD8A, PDCD1, TIGIT, CTLA4 and LAG3 expression across the tumor–stroma axis. Following treatment, we observed diffuse increases in the expression of these markers, particularly within the stromal compartment. While these transcripts are often associated with T cell exhaustion, they are also induced during T cell activation—supported here by concomitant increases in GZMB and MKI67 expression. Collectively, these patterns are consistent with immune re-engagement and post-ICI remodeling of the $CD8^+$ T cell compartment. Given the limited number of paired samples (n=3), these findings are hypothesis-generating and warrant further validation in larger cohorts (**Figure R25b**).

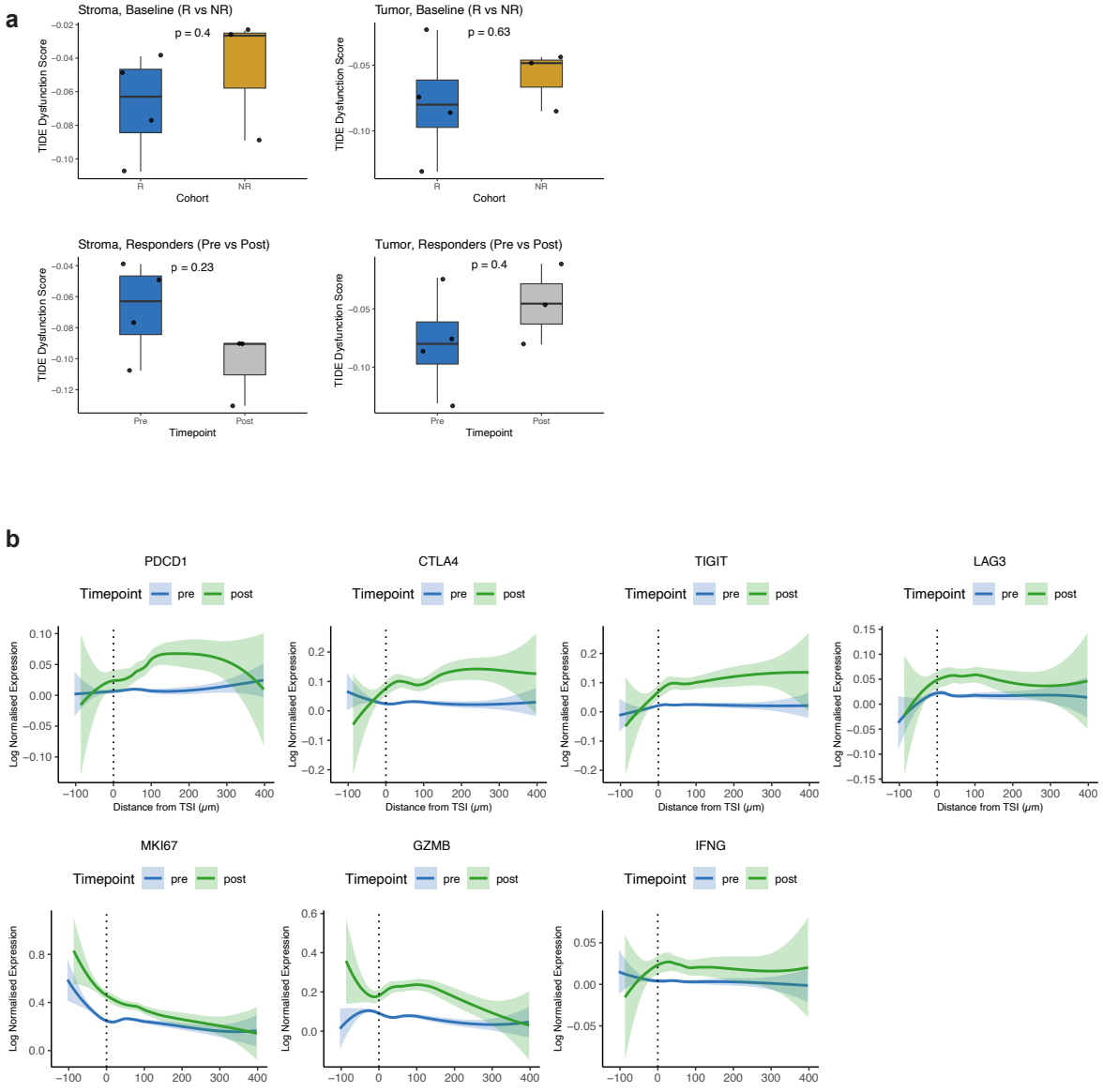


Figure R25

a) CD8⁺ T cell dysfunction scores in responders and non-responders, pre- and post-treatment. Post-treatment samples were only available for three patients, all of whom were responders. T cell dysfunction was estimated using the TIDE-derived Dysfunction score.

b) Spatial gene expression gradients of $CD8^+$ T cell exhaustion and activation markers across the tumor–stroma axis.

References:

[1] Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nature medicine. 2018;24(10):1550-8.

R3C31 23. Fig.6D needs bigger labels.

Reply: We thank the Reviewer for this comment. We have updated all figure labels and legends to ensure clear visualisations throughout the manuscript.

Reply to Reviewer #4

In this manuscript samples from the MAYA trial of refractory MGMT-silenced MSS colorectal cancer (CRC) were analyzed. Tumor and blood samples at baseline, post-cycle 2 of triplet therapy, and at disease progression to investigate how temozolomide priming and combination immune therapy reshape the tumor microenvironment (TME) and immune context were investigated. This approach aimed to identify biomarkers and uncover resistance mechanisms to improve combination therapies and personalized treatments for microsatellite-stable CRC, which is still a largely non-responding cancer entity. Overall, this study represents a very interesting and important description of this innovative trial set up. However, there are a few concerns that should be addressed before publication of this manuscript which are listed below.

Major comments

R4C1 Spatial analysis. The usage of spatial data is highly relevant for heterogenic therapy responses. However, the major conclusion, based on spatial data, is missing or to weakly elaborated in the manuscript. CD8+TIGIT+ scores show association in peripheral blood but were not investigated in the spatial datasets. This should be done to generate a line of evidence that is connected to deeper spatial relevance.

Reply: We appreciate the constructive feedback and have made significant enhancements to our spatial analyses to address the points raised. We have expanded our spatial analysis by introducing additional key cell types such as cancer-associated fibroblasts (CAFs), endothelial cells and memory T cells, just to name a few (**Supplementary Table 1**). This provides a more comprehensive characterization of the tumor microenvironment (TME), offering greater insight into its complexity. Furthermore, to provide a more robust interrogation of the TME, we have incorporated additional spatial features, which go beyond relative abundance of cell types per compartment. We have introduced additional compartment specific spatial features including (1) Proximity analyses of adjacent cells: We now examine the spatial proximity of adjacent cells [1], analyzing how different cell types interact within the TME to better understand immune cell dynamics. (2) Relative distances of adjacent cells: Our analysis also includes compartment-

specific relative distances between cell types, allowing us to quantify how immune cells are spatially organized across distinct compartments. (3) Cellular neighborhoods: We investigate the cellular neighborhoods to understand how spatially resolved immune cell clusters orchestrate responses or resistance to immune checkpoint inhibitors (ICIs). This aspect of the analysis provides insight into the microenvironmental context of immune cell function and therapy outcomes. (**Figure R26**) These revisions enhance our spatial analysis and provide deeper insights into the TME and its impact on immune responses, particularly in the context of ICI therapy.

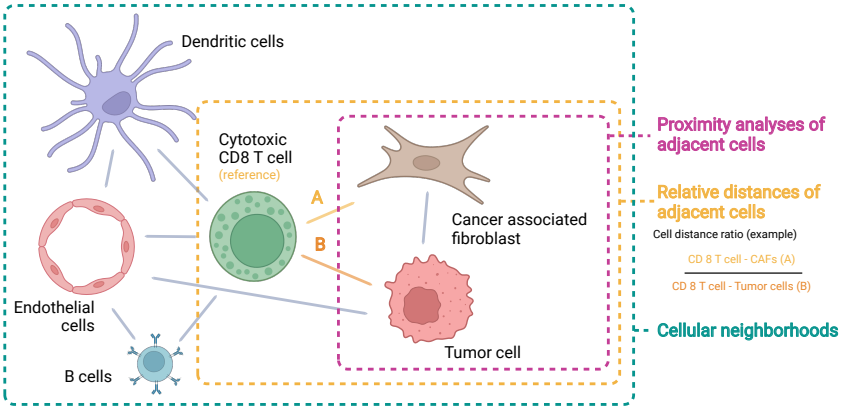


Figure R26 Illustration of additional spatial features incorporated into the revised manuscript

Remarks: IHC marker definitions per cell type are provided in **Supplementary Table 1**.

On inspection of adjacent cells abundances within the vicinity of 20μm of a tumor cell within tumor compartments of baseline samples, we observed scant numbers of effector adaptive immune cells (CD8 T cells and B cells) in contrast to immunosuppressive cells such as CAFs and memory adaptive immune cells such as memory T cells. These findings are consistent with the known phenomenon of an immune-cold microenvironment of microsatellite stable (MSS) colorectal cancer (CRC). Of interest, higher cancer associated fibroblast (CAF) infiltration within tumor compartments is associated with poor response to TMZ alone, and subsequent ICI. In contrast, greater Memory T cell infiltration within tumor cells was associated with good response to ICI.

We also sought to understand how tumor-infiltrating cell types influenced effector CD8 T cells within tumor compartments. We evaluated relative cell distances from CD8 T cells and found that relative proximity of CD8 T cells to CAFs compared to tumor cells were associated with poor response to TMZ alone, although statistical significance was not reached. Likewise, relative proximity of CD8 T cells to endothelial cells were associated with poorer response.

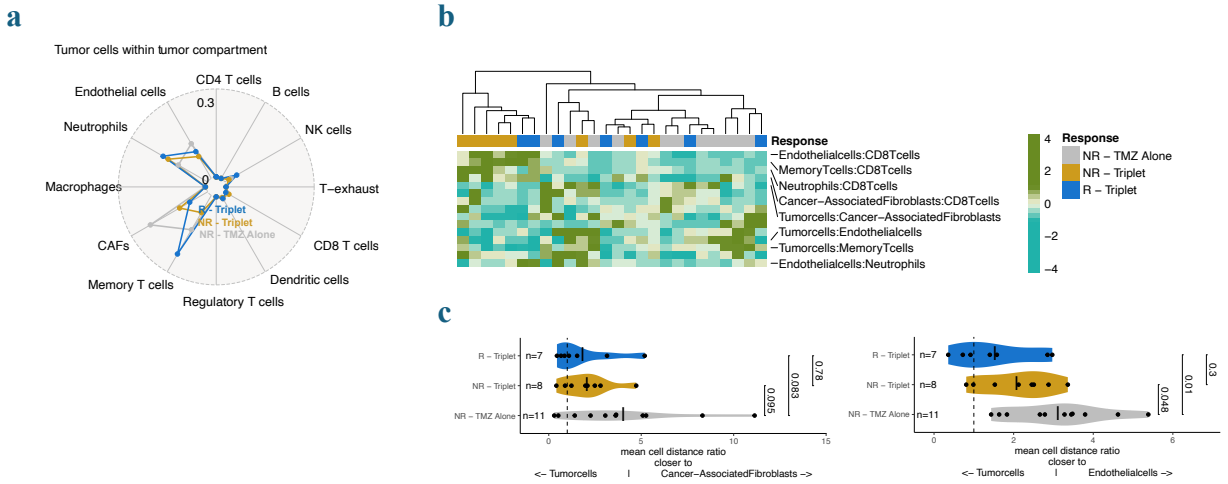


Figure R27

g) Adjacent cell abundances within 20µm of tumor cells within tumor compartments.

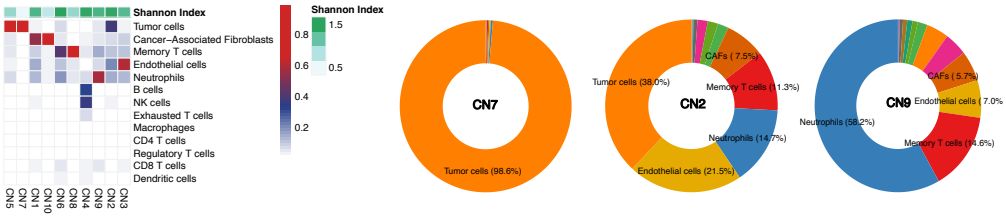
h) Heatmap of mean cell distance ratios from CD8 T cells per patient. Cell distance ratios flagged out had an ANOVA p-value of less than 0.15 across response group comparisons.

i) Selected violin plots comparing mean cell distance ratios per patient stratified by tumor response groups. P-values were retrieved from a two-sided T-test.

Abbreviations: CAF, cancer associated fibroblasts; T-exhaust, Exhausted T cells; NR – TMZ Alone, non-responder to temozolomide (TMZ) alone; NR - Triplet, responder to TMZ however non-responder to ICI; R - Triplet, responder to TMZ and ICI; n=, number of patients.

Remarks: IHC marker definitions are provided in **Supplementary Table 1**.

a



b

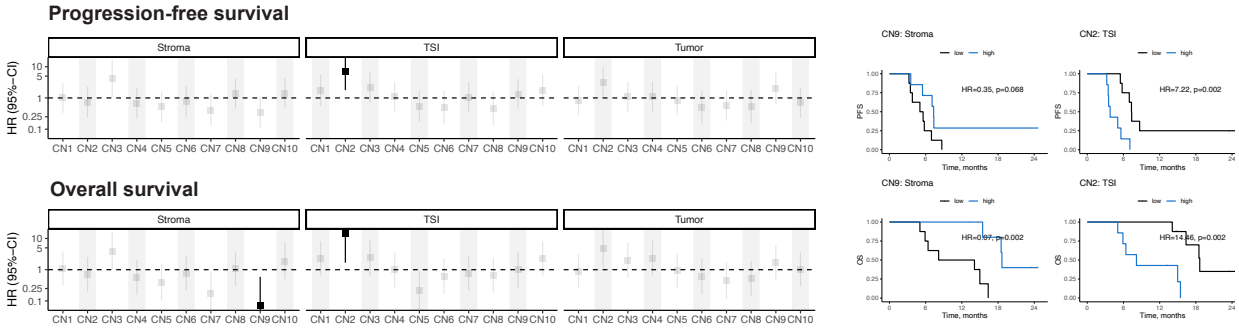


Figure R28

b) Heatmap and donut plots providing an overview of immune cell proportions per cellular neighbourhood.

b) Forest plot and selected KM plots of survival outcomes stratified by compartment specific CN cellular proportions at median split. Values were highlighted if Log-rank p-value was less than 0.01.

Abbreviations: CN, cellular neighbourhood; TSI, tumor stromal interface; HR, hazard ratio; KM, Kaplan-Meier.

The tumor microenvironment (TME) operates as an integrated entity, where distinct immune cell clusters interact to coordinate anti-tumor responses. As such, we employed spatially resolved clustering algorithms previously described by Schürch *et al*[2] to retrieve cellular neighbourhood (CN) information. Broadly, our analyses demonstrated CNs that differed in immune cell repertoire and diversity. For instance, CN7 was found to predominantly comprise tumor cells (98.6%) with low immune cell diversity (Shannon Index = 0.097). In contrast, CN2 and CN9 were found to have high immune cell diversity (CN2, Shannon Index = 1.715; CN9, Shannon Index = 1.482).

Compartment specific comparative proportions of CNs were found to be associated with survival outcomes and response. For instance, patients with high proportions of CN9 cells within stromal compartments were associated with favorable survival outcomes. Conversely, patients with high proportions of CN2 cells (heterogenous mixture of tumor cells, endothelial cells, neutrophils, memory T cells and CAFs) within the tumor stromal interface (TSI) were associated with poorer survival outcomes (**Figure R28b**). It was interesting to note that CN10, which was primarily composed of CAFs, did not modulate treatment response to the same extent and magnitude as CN2 (OS within TSI, HR=2.28, p=0.196; PFS within TSI, HR=1.76, p=0.337) (**Figure R28b**).

Our findings underscore the critical role of spatial features in reflecting the intricate, dynamic interplay of immune cells, providing valuable prognostic and therapeutic insights into tumor responses. Specifically, we demonstrate that the infiltration of immunosuppressive cells, such as CAFs, within tumor compartments—particularly their proximity to CD8+ T cells—attenuates treatment efficacy and modulates therapeutic responses. Moreover, the immunosuppressive influence of CAFs is not uniform but spatially dependent, further emphasizing the importance of cellular architecture in shaping immune evasion and resistance mechanisms.

With respect to **Supplementary Fig. 6** (revised to **Supplementary Fig. 8**), which presents peripheral flow cytometry data, we observed an increase in $CD8^{+}TIGIT^{+}$ T cells in non-responders (NR) compared to responders (R) following treatment. Apart from a reduction in circulating $CD8^{+}PD1^{+}$ cells, no other significant differences were observed at static timepoints or relative to

baseline in either group. To approximate changes in intratumoral $CD8^+$ T cell subsets in the Visium dataset, we assessed transcriptional correlates of T cell dysfunction using the TIDE-derived Dysfunction score. This model incorporates multiple canonical exhaustion markers (e.g., PDCD1, TIGIT, CTLA4, LAG3, IFNG) and has been validated as a proxy for $CD8^+$ T cell dysfunction in the tumor microenvironment. Within tumor regions enriched for $CD8^+$ T cells (defined as the top quartile of $CD8^+$ abundance), we compared Dysfunction scores pre- and post-treatment. Although scores were numerically higher post-treatment in responders, the difference was not statistically significant. Notably, post-treatment samples were available for only three patients, all of whom were responders, limiting our ability to compare differences between responders and non-responders (**Figure R29a**).

In parallel, we examined spatial gradients of CD8A, PDCD1, TIGIT, CTLA4 and LAG3 expression across the tumor–stroma axis. Following treatment, we observed diffuse increases in the expression of these markers, particularly within the stromal compartment. While these transcripts are often associated with T cell exhaustion, they are also induced during T cell activation—supported here by concomitant increases in GZMB and MKI67 expression. Collectively, these patterns are consistent with immune re-engagement and post-ICI remodeling of the $CD8^+$ T cell compartment. Given the limited number of paired samples, these findings are hypothesis-generating and warrant further validation in larger cohorts (**Figure R29b**).

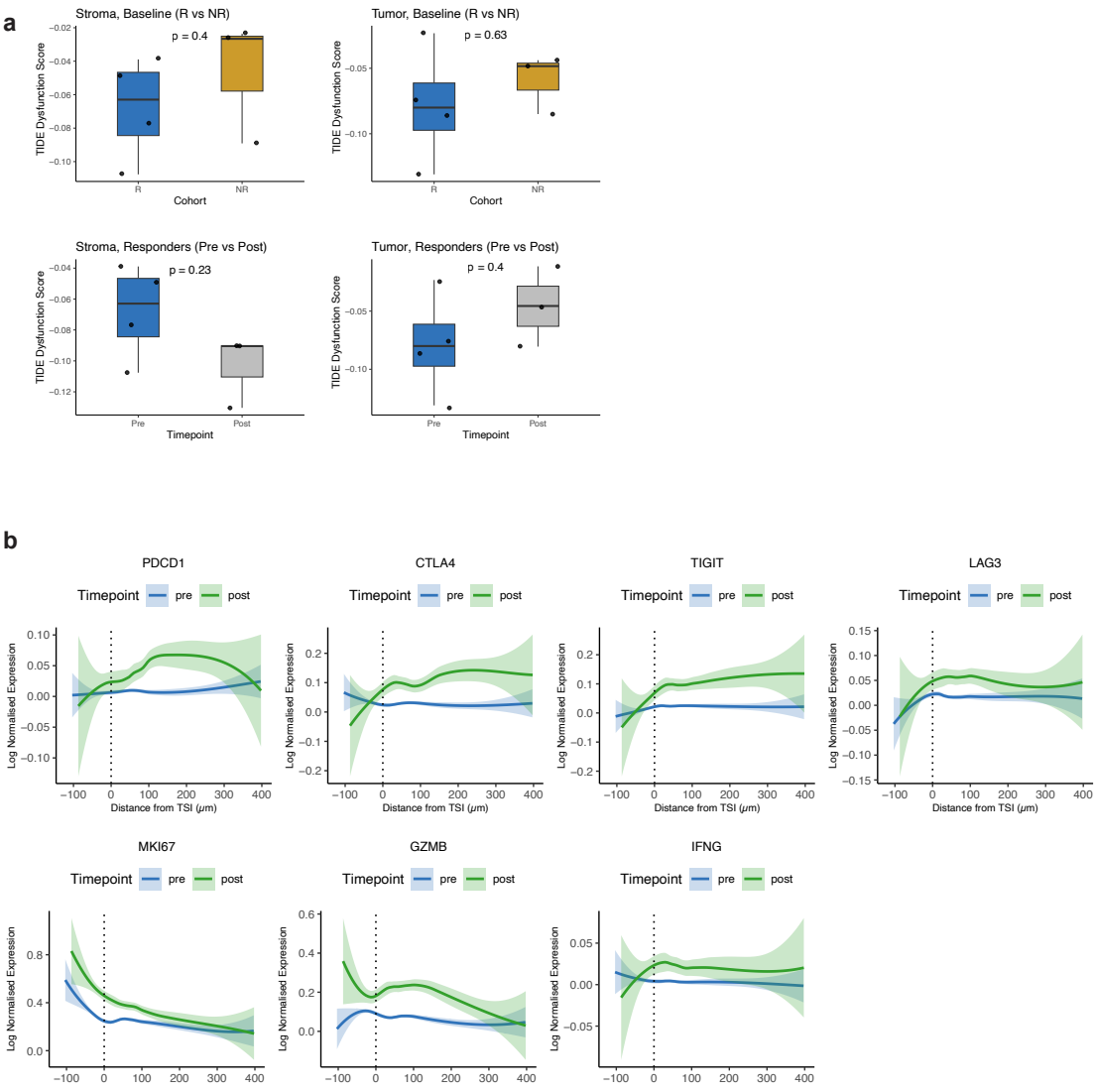


Figure R29

a) CD8⁺ T cell dysfunction scores in CD8-enriched tumor and stromal regions in responders and non-responders, pre- and post-treatment. Post-treatment samples were only available for three patients, all of whom were responders. T cell dysfunction was estimated using the TIDE-derived Dysfunction score. Tumor regions enriched for CD8⁺ T cells were defined as spatial transcriptomic spots within the top quartile of CD8⁺ UCell scores.

b) Spatial gene expression gradients of $CD8^+$ T cell exhaustion and activation markers across the tumor–stroma axis.

References

[1] Jia K, Chen Y, Sun Y, et al. Multiplex immunohistochemistry defines the tumor immune microenvironment and immunotherapeutic outcome in CLDN18.2-positive gastric cancer. BMC Med. 2022 Jul 11;20(1):223. doi: 10.1186/s12916-022-02421-1. Erratum in: BMC Med. 2023 Sep 29;21(1):369. doi: 10.1186/s12916-023-03090-4. PMID: 35811317; PMCID: PMC9272556.

[2] Schürch CM, Bhate SS, Barlow GL, et al. Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front. Cell. 2020 Sep 3;182(5):1341-1359.e19. doi: 10.1016/j.cell.2020.07.005. Epub 2020 Aug 6. Erratum in: Cell. 2020 Oct 29;183(3):838. doi: 10.1016/j.cell.2020.10.021. PMID: 32763154; PMCID: PMC7479520.

R4C2 TME signals. It is demonstrated that the TME compositions is altered by the treatment with TMZ and ICI. Nevertheless, factors that are potentially critical for this change in TME like TGF-beta or other means have been completely neglected. A more unbiased analysis of the probe set from e.g. the Visium platform might enhance the overall understanding and helps defining underlying mechanisms or novel targets (as state in the paper).

Reply: For our Nanostring DSP analyses, analyses of oncological programmes (HALLMARK pathways) were conducted and previously reported in **Figure 3d**. This exercise was well aligned with the Reviewer's suggestion to conduct an unbiased analyses of important oncological pathways that may have implications on underlying mechanisms driving response.

However, for reporting purposes, we only included pathways that were found to be significantly enriched in responders or non-responders (adjusted p-value<0.05). Our current revised manuscript includes results of all 50 HALLMARK pathways (**Figure 3d**). Specifically, we show that no significant differences in TGF-beta enrichment scores were found between responders and non-responders in both tumor (NES=-1.35, adjusted p-value=0.101) and stromal (NES=-1.14, adjusted p-value=0.344) compartments (**Figure R30b-c**).

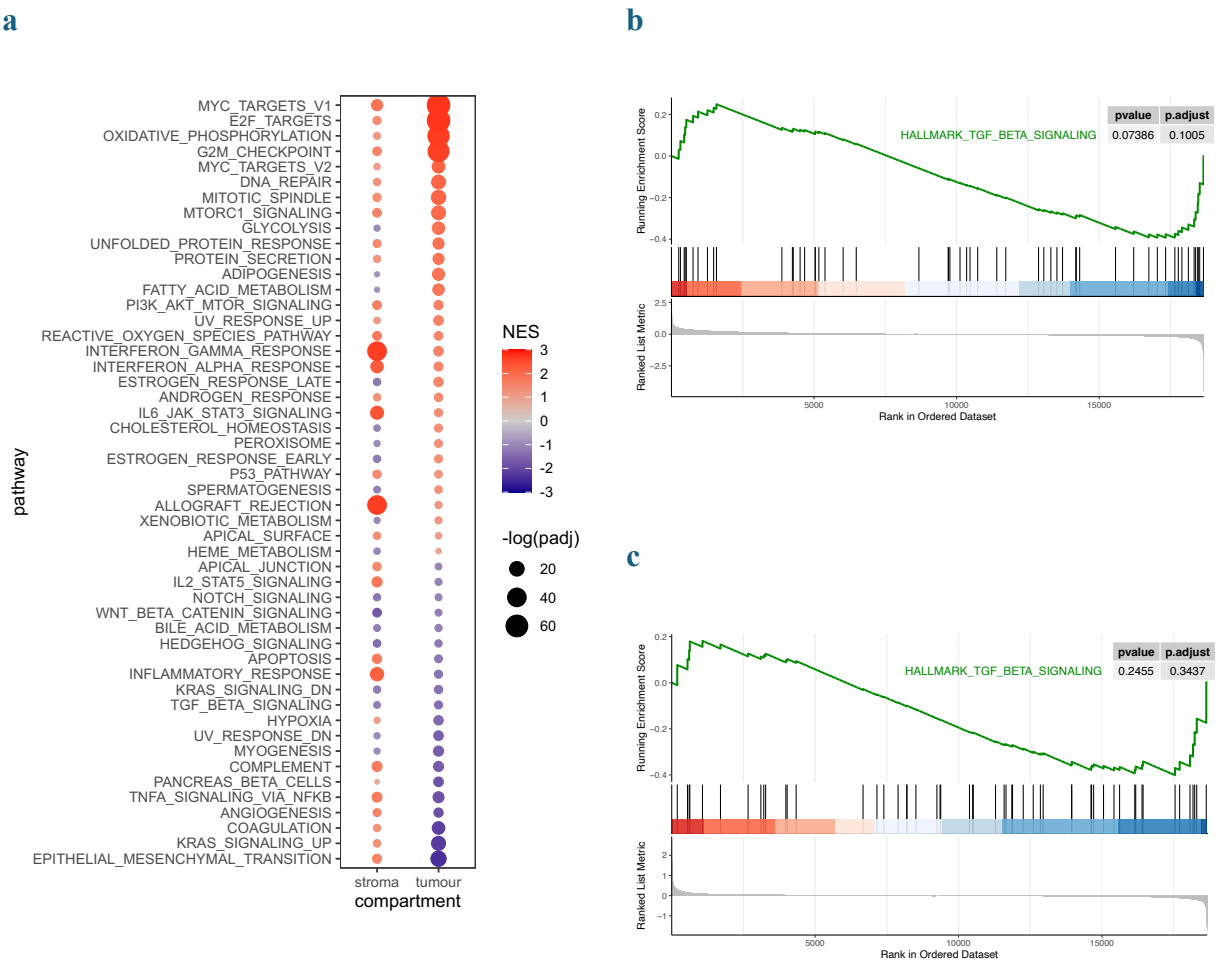


Figure R30

a) Complete results of pathway analyses between responders and non-responders within tumor and stromal compartments
b) GSEA plot of TGF-beta pathway within tumor compartment
c) GSEA plot of TGF beta pathway within stromal compartment

Abbreviations: NES, normalized enrichment score; padj, adjusted p value

Remarks: A positive NES value denotes enrichment in responders while a negative NES value denotes enrichment in non-responders.

Minor comments

R4C3 Data representation. The font and description of the figures should be harmonized to allow proper reading of all data points.

Reply: We thank the Reviewer for the comment. We have harmonized the fonts and descriptions of the figures in our revised manuscript.

R4C4 TMB. Somewhat the conclusion for the findings around the TMB is lacking. What does the increase after TMZ plus ICI mean and are they any strategies for the future?

Reply: We apologize for the lack of clarity. Our study builds on the principle that changes in tumor mutational burden (TMB) might have explained response to TMZ+ICI. These have been previously established by foundational work of Germano *et al* [1], who have already made key biological discoveries related to Temozolomide (TMZ) treatment. Their research demonstrated that the inactivation of MMR, driven by acquired resistance to the clinical agent temozolomide, increased mutational load, promoted continuous renewal of neoantigens in human colorectal cancers and triggered immune surveillance in mouse models [1,2].

In our study, we demonstrated that TMB was associated with PFS (Pearson's r=0.863, p=0.0598 [n=5]) (**Figure R31**), albeit a small sample size. Changes in TMB could offer a promising avenue for predicting responses to TMZ and ICI. However, to achieve meaningful clinical impact, TMB measurement should ideally be conducted after TMZ administration and before ICI therapy. A potential method for this assessment would be non-invasive monitoring of TMB through circulating tumor DNA (ctDNA). However, some studies have suggested that TMB measured from ctDNA (bTMB) may not be directly concordant with tissue-based TMB (tTMB); discordant associations between bTMB and clinical outcomes were seen [3], potentially reflecting tumor heterogeneity rather than pure mutational load [4]. Thus, while ctDNA offers a minimally invasive alternative for tracking treatment-induced hypermutation, the interpretation of bTMB should consider potential confounding by tumor burden, subclonality, and assay-specific technical limitations.

We have discussed this approach in the discussion, suggesting its potential utility and caveats in interpretation of results for guiding treatment strategies.

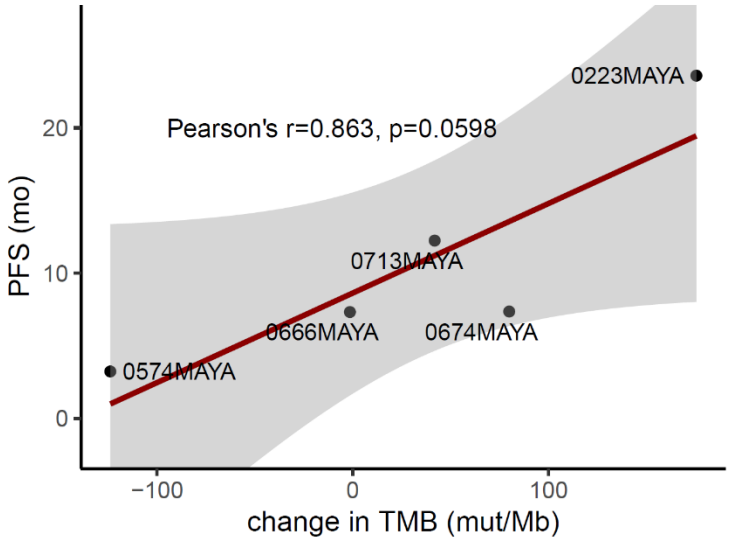


Figure R31 Correlation between TMB changes (pre vs post) and PFS

Abbreviations: TMB, tumor mutational burden; PFS, progression-free survival; mo, months.

References

[1] Germano, G., Lamba, S., Rospo, G. et al. Inactivation of DNA repair triggers neoantigen generation and impairs tumour growth. Nature 552, 116–120 (2017). https://doi.org/10.1038/nature24673

[2] Crisafulli G, Sartore-Bianchi A, Lazzari L, et al. Temozolomide Treatment Alters Mismatch Repair and Boosts Mutational Burden in Tumor and Blood of Colorectal Cancer Patients. Cancer Discov. 2022 Jul 6;12(7):1656-1675. doi: 10.1158/2159-8290.CD-21-1434

[3] Chae YK, Davis AA, Agte S, Pan A, Simon NI, Iams WT, et al. Clinical Implications of Circulating Tumor DNA Tumor Mutational Burden (ctDNA TMB) in Non-Small Cell Lung Cancer. The oncologist. 2019;24(6):820-8.

[4] Wang Z, Duan J, Wang G, Zhao J, Xu J, Han J, et al. Allele Frequency-Adjusted Blood-Based Tumor Mutational Burden as a Predictor of Overall Survival for Patients With NSCLC Treated With PD-(L)1 Inhibitors. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer. 2020;15(4):556-67.

Reply to Reviewer #1

R1C1 The reviewers have addressed my comments for the most part

Reply: We thank the Reviewer for the previous comments, which have significantly improved the quality of the manuscript.

Reply to Reviewer #2

R2C1 I thank the authors for revising this manuscript with additional analyses, and for responding to all my queries. The authors have added multiple dimensions of spatial omics analysis, newly proposing the relationship between the compartmental distribution of cancer-associated fibroblasts (CAFs) and TMZ response, and suggesting that the distance of CD8 T cells to CAFs reflects the response to TMZ. The current version of the manuscript shows improvement, particularly in the methodology and the breadth of findings. However, several aspects in this manuscript required further attention before publication.

Reply: We thank the reviewer for acknowledging our efforts in revising the manuscript and for their thoughtful feedback. We appreciate your recognition of the improvements in our methodology and findings. We remain committed to further improving the quality of the manuscript and have carefully addressed the additional points raised in this round of review.

R2C2 In the revised manuscript, the authors present some new insights, such as the potential correlation between neighbourhood of CD8 T cell and CAFs to TMZ response. Although the authors performed many analyses and attempts, perhaps limited by the sample size, many results fail to reach statistical significance (e.g., Fig. 4F and 6C), which impacts the overall repeatability of the conclusions.

Reply: We thank the reviewer for highlighting the impact of sample size on the statistical significance and repeatability of our findings. We acknowledge that sample size is a core limitation of our study, which may have affected our ability to reach statistical significance in some analyses (e.g., **Figures 4F** and **6C**). However, we would like to underscore that, despite this limitation, we observed a clear and consistent trend in our data. These trends are reassuringly in line with established biological knowledge, which increases our confidence in the robustness of our findings.

Specifically regarding **Figure 6C**, while we did not observe statistically significant differences in exhausted CD4/8 T cells between individual checkpoints in responders versus non-responders when assessed through a t-test, we found that the dynamic changes in these populations across time translated into significant survival differences as shown in the survival analysis (**Figure 6D**). This suggests that temporal dynamics, rather than static differences at individual timepoints, are important for understanding the relationship between T cell exhaustion and response and this is consistent with the dynamic nature of the strategy investigated in the trial. Importantly, this finding is immediately clinically relevant, as it indicates the potential to predict patient response based on serial monitoring of peripheral markers.

Additionally, rather than restricting our report to statistically significant findings, we included the full spectrum of results to ensure balance and transparency.

R2C3 The interpretation of some conclusions lacks sufficient depth. As the authors state in Line 582: “The interaction of CAFs with other cell types... shapes immune evasion and treatment resistance”. However, I regret to note that while

the authors used spatial omics to observe the distribution features of cells at various treatment time points, they did not further explore the cellular interactions. This may improve the mechanistic depth and the novel insights that spatial omics could provide to this study.

> *Full line in manuscript: "Moreover, the immunosuppressive influence of CAFs is not uniform but modulated by their spatial context; cellular neighborhood analysis showed that CAF-enriched neighborhoods alone did not significantly impact survival outcomes, suggesting that* ***the interaction of CAFs with other cell types—rather than their presence in isolation—shapes immune evasion and treatment resistance****".*

Reply: We thank the Reviewer for this constructive comment. In line with this suggestion, we extended our analysis beyond spatial distribution to examine CAF-centric ligand–receptor interactions using Visium-derived CellChat. By collapsing inferred interactions into pathway families and relating them to clinical outcomes, we identified distinct signaling programs. Both responders and non-responders demonstrated stromal/tumor CAF signaling, with elevated scores for FN1, laminin, agrin, vitronectin, HSPG, PDGF, LIGHT, and tenascin in tumor, fibroblast, and endothelial compartments, encompassing pathways previously implicated in adhesion, angiogenesis, and immune exclusion. In contrast, enrichment of anti-tumor CAF–immune signaling was observed primarily in responders, with higher scores for complement, APP, and ICAM across $CD8^+$ T cells, B lineage cells, and dendritic cells (**Figure R1**). These findings provide additional mechanistic depth by indicating that the prognostic influence of CAFs reflects not only their abundance or spatial context, but also the qualitative nature of their signaling with immune and tumor compartments. We have also revised the Discussion section to integrate these results, highlighting that CAF influence is shaped by their functional interactions with neighboring cell types rather than abundance alone.

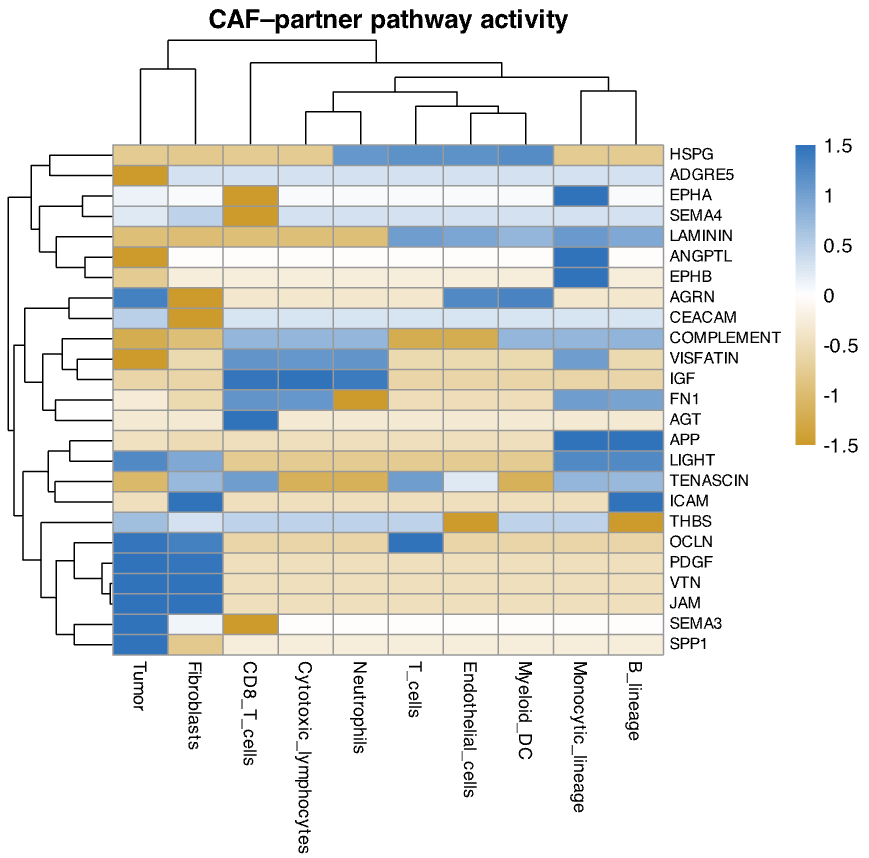


Figure R1. CAF–partner pathway differences between responders and non-responders. Heatmap of $\log_2$ fold-change (R/NR) in CAF–partner ligand–receptor activity collapsed to pathway families. Each row represents a pathway, and each column represents a partner cell type. Blue indicates relative enrichment in responders, while gold indicates relative enrichment in non-responders.

R2C4 In the cellular neighbourhood analysis, CN9 and CN2 have similar cellular diversity but differential survival outcomes. Yet, their T cell proportions are relatively similar. What might be the underlying reason for this result? How can we understand different cellular compositions lead to divergent survival outcomes? Incorporating immune-related proteins, cytokines, ligand-receptor, etc., into the neighbourhood analysis might help explain this finding.

Reply: We apologize for the lack of clarity in our previous response. To address the reviewer's question, we have conducted dedicated analyses to further inspect differences in T cell populations between CN2 (pro-tumor) and CN9 (anti-tumor). We observe an increased abundance of T cells within the CN9 neighborhoods, particularly within the Stroma and TSI ROI (**Figure R2**). This enrichment corresponds well with the observed improved responses to checkpoint inhibition in CN9.

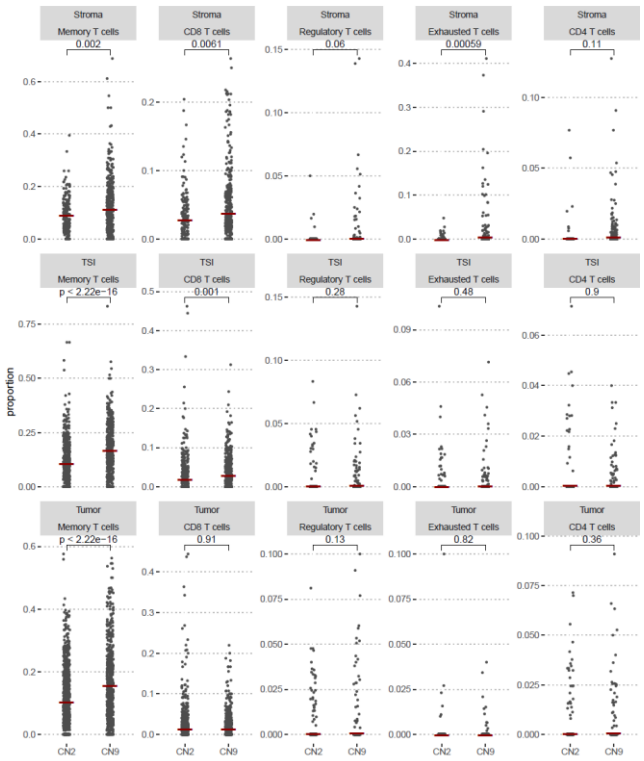


Figure R2 Comparisons of T cell subtype proportions between CN2 and CN9 stratified by region of interest

Abbreviations: CN, cellular neighborhood

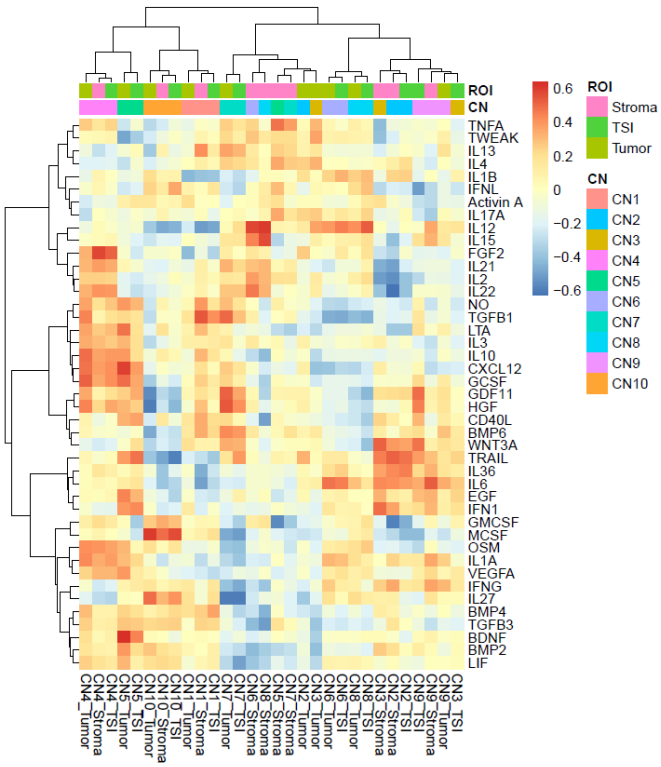


Figure R3 Heatmap of Pearson correlation between CN proportions per ROI against cytokine z-scores (CytoSig)
Abbreviations: CN, cellular neighborhood; TSI, tumor stromal interface; ROI, region of interest

As suggested by the reviewer, we have also investigated cytokine expression profiles using deconvolution analysis via CytoSig[1]. We correlated the inferred cytokine levels against CN proportions, stratified by ROI type, to elucidate immune activity within these microenvironments. For CN2 (TSI, pro-tumor), we noted negative correlations with GM-CSF (r = -0.46) and M-CSF (r = -0.41), and positive correlations with IL-36 (r = 0.45) and TRAIL (r = 0.48). Conversely, for CN9 (Stroma, anti-tumor), IL-6 showed a strong positive correlation (r = 0.56) with CN9 abundance (**Figure R3**). These distinct cytokine profiles may help explain the divergent survival outcomes. In CN2, increased TRAIL and IL-36 may reflect pro-inflammatory yet potentially immune-dampening environments, while negative correlations with growth factors like GM-CSF and M-CSF could suggest impaired recruitment or maintenance of effective immune cell types. In contrast, the strong association of CN9 with IL-6 in the stroma could indicate a more robust immune activation and effector function, supporting anti-tumor immunity and improved response to checkpoint therapy.

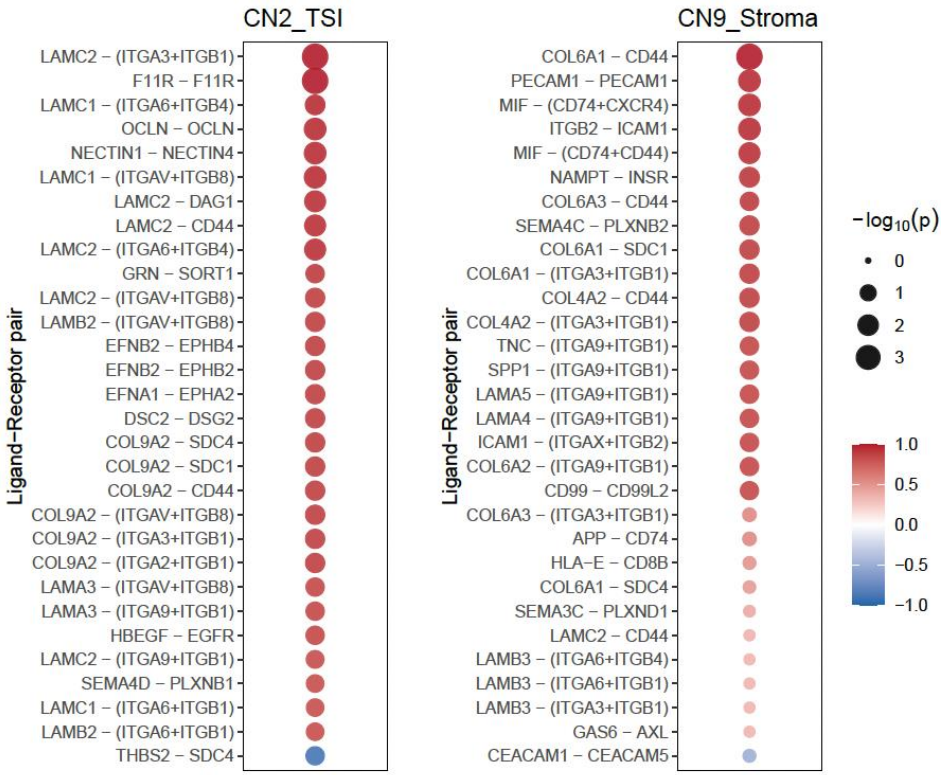


Figure R4. Ligand–receptor correlations with CN2 and CN9 across compartments. Bubble plot showing Spearman correlations between CellChat-inferred ligand–receptor interaction scores and CN2 or CN9 proportions in the stroma and tumor–stroma interface (TSI). Red indicates positive and blue negative correlations. Bubble size reflects the nominal statistical significance (−log10 p).

In addition, we extended this analysis to CellChat-inferred ligand–receptor interactions. CN9 (Stroma, anti-tumor) correlated positively with both ECM-related signaling (e.g., collagen–CD44, laminin–integrins, fibronectin–syndecans) and immune-associated signaling pathways (ICAM1–ITGB2, PECAM1–PECAM1). Meanwhile, CN2 (TSI, pro-tumor), was enriched for adhesion and guidance interactions such as laminin–integrins, Eph–Ephrin, semaphorin–plexin, and PDGF–PDGFR, consistent with a tumor–stroma dominant and exclusionary interface lacking compensatory immune signaling (**Figure R4**).

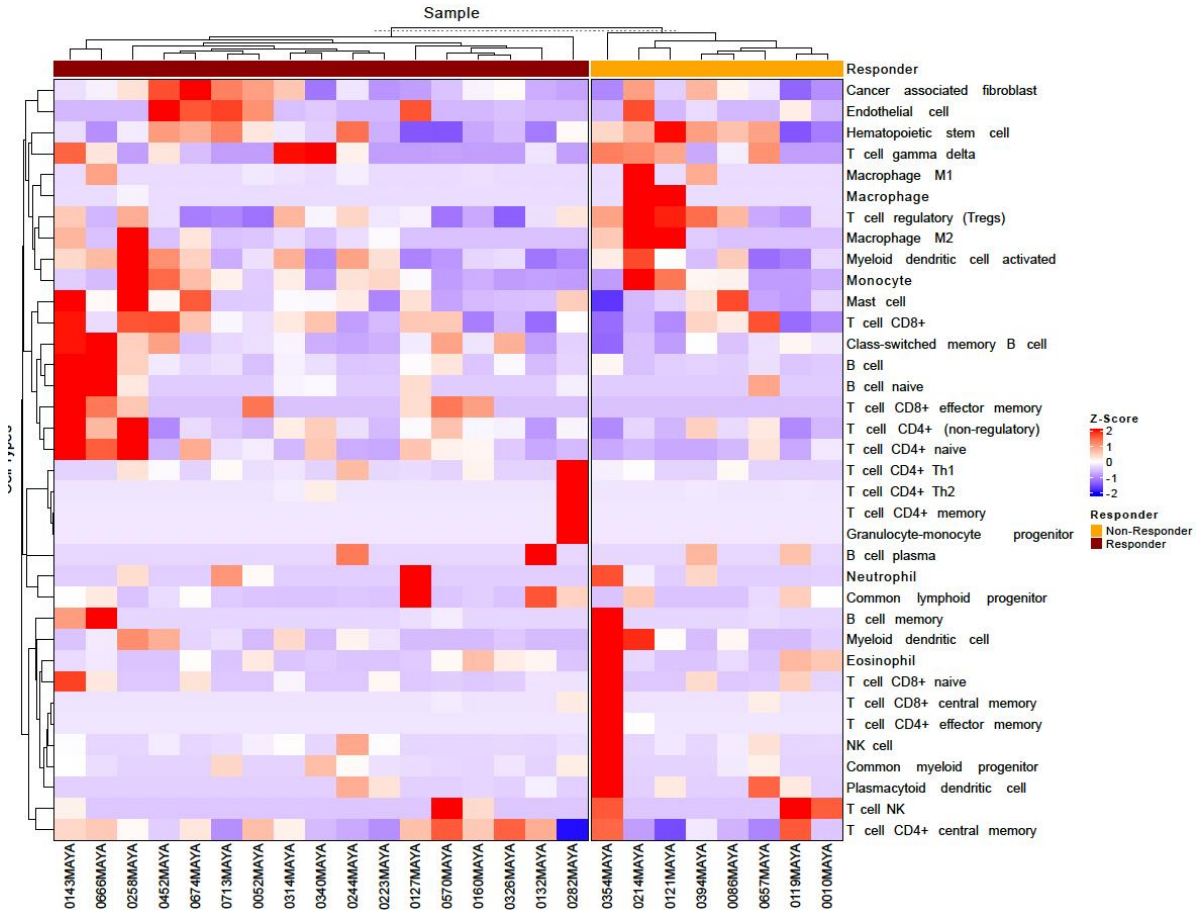
Together, these data suggest that CN2 and CN9 differ not only in their T cell content, but also by distinct immune signaling milieus and stromal–immune communication patterns. We hope this clarifies the underlying biology and supports the reviewer's suggestion to incorporate additional immune-related features in neighbourhood analyses.

References

[1] Jiang P, Zhang Y, Ru B, et al. Systematic investigation of cytokine signaling activity at the tissue and single-cell levels. Nat Methods. 2021 Oct;18(10):1181-1191. doi: 10.1038/s41592-021-01274-5. Epub 2021 Sep 30. PMID: 34594031; PMCID: PMC8493809.

R2C5 The presentation of many figures needs further improvement. For example, it is difficult to compare immune cell abundance between responders and non-responders Sup. Fig 3. It would benefit from reorganization by response group or quantitative analyses to facilitate direct comparison. Additionally, some texts in many figures remain unclear (e.g., Figs 2 and 4).

Reply: We apologize for the lack of clarity in the presentation of figures. As suggested by the reviewer, we have reorganized the heatmap for Sup. Fig. 3 by Responder status to facilitate visual comparisons. We have also edited the text size in Fig 2 and 4 to improve the clarity.



Revised Supplementary Figure 3. Immune cell composition of baseline tumors estimated by xCell deconvolution.

Bulk RNA sequencing data from pre-treatment tumor samples were analyzed using the xCell algorithm to infer the relative enrichment of immune and stromal cell types. The heatmap displays xCell scores across individual patients, representing inferred abundance of immune subsets including T cells, B cells, macrophages, dendritic cells, and NK cells, as well as stromal components such as fibroblasts and endothelial cells. Scores are relative enrichment values rather than absolute proportions and are normalized across samples to enable inter-individual comparison. Patients are grouped by response category [responder vs non-responder] to assess baseline differences in immune landscape.

Reply to Reviewer #3

R3C1 We thank the authors for thoroughly addressing many of our comments specifically those that are critical to the clarity of the paper. We are impressed by the amount of details provided. Newly adopted analyses, visualisations, scopes certainly strengthened the narrative. Specifically, I think the paper now has a more clear direction.

Reply: We thank the reviewer for their positive feedback and for recognizing our efforts to address the previous comments, especially those crucial for improving the clarity of the paper. We appreciate your acknowledgment of the additional analyses, visualizations, and expanded scope, and are pleased that these have helped to strengthen the overall narrative and provide a clearer direction for the manuscript.

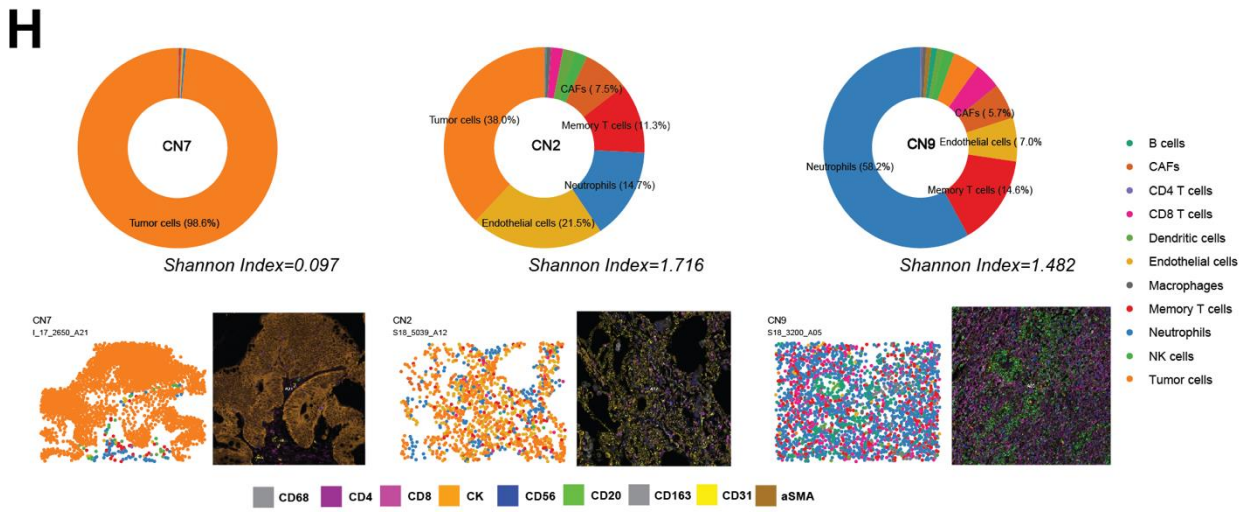
R3C2 Given that, the major issues with the paper remain unresolved and I am not sure given at this stage it is resolvable. We appreciate the study's message regarding the use of biomarker-based stratification of patients for combinatory immunotherapy and the fact that the results arose from clinical trial samples. However, the numbers of samples used to make some of the conclusions are, honestly, low, and the authors were not able to address the issues. While more analyses are done on the small number of samples already present, this does not lead to "informative and robust" analyses as the authors stated. (Doing more types of analyses on a small number of samples – giving shaky results - is not equivalent to doing a deep analysis on many more samples). Actually, for some analyses, the number of samples actually decreased from the previous submission for some reason. Because of this, some of the comments were not addressed directly but avoided (for instance, R3- comment 3- we ask about what we learn different from pretreated vs post treated samples that are different if not tumor mutations – the authors just skirted the issue and gave us a description and reference on TMB), and the unified assay comment – which was not addressed. At this point, we feel like the authors did do a lot of work to address the comments, but there does not seem to be a way to improve the small sample size issues. I would like to leave it to the editor to decide based on comments from other reviewers.

Reply: We thank the reviewer for this important observation. The reduction in sample numbers in this revision is due to stricter methodological criteria: for the flow cytometry dataset, we excluded progressive disease (PD) timepoints, as these were not informative for evaluating treatment-associated immune changes between responders and non-responders. For the GeoMx analysis, we retained the 16 available baseline samples and excluded the 2 post-treatment samples, as the limited number of post-treatment cases precluded meaningful interpretation.

Unfortunately, as the clinical trial has already concluded, it is not possible to retrieve additional samples. For TMZ-treated patients, paired pre- and post-treatment biopsies were not collected due to ethical considerations surrounding repeat sampling in the absence of progression. We have clarified these points in the revised manuscript and highlighted the limitations related to sample size and assay consistency. Despite these constraints, we believe the study nevertheless provides important insights, while recognizing the need for validation in larger, independent cohorts.

R3C3 In figure 4, authors should consider showing examples of specific cellular neighbours.

Reply: We the thank the Reviewer for the suggestion. We have revised **Figure 4h** as shown below to include both the original Lunaphore images with selected markers, as well as the cell type phenotypes retrieved. We also provide now in the **Supplement** an overview representative regions of interests for spatially identified cellular neighborhoods (**Figure R5**).



Revised Figure 4h Donut plots providing an overview of immune cell proportions per cellular neighborhood with selected representative regions of interests.
Abbreviations: CN, cellular neighborhood.

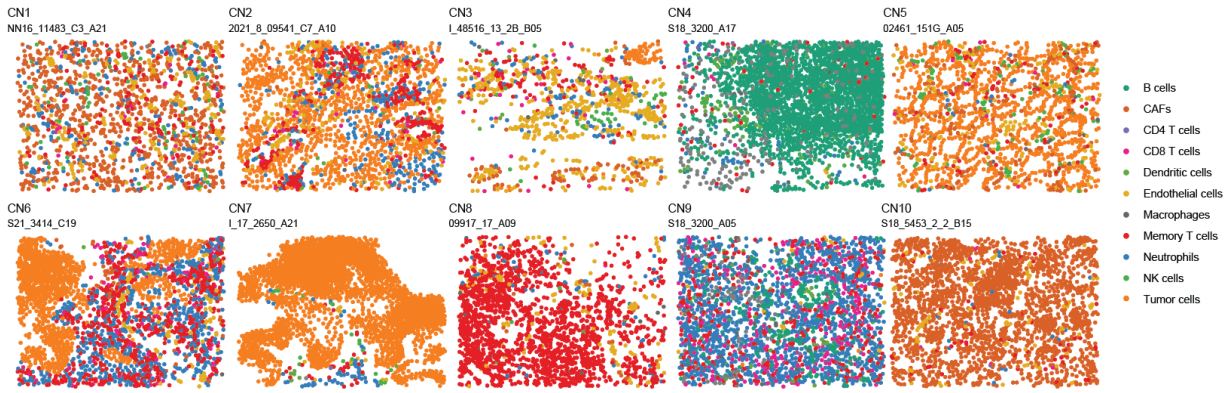


Figure R5 Overview representative regions of interests for spatially identified cellular neighborhoods
Abbreviations: CN, cellular neighborhood.

R3C4 For figure 2A, why did the number of flow cytometry & GeoMx datasets decrease from your last version?

Reply: We thank the reviewer for this observation. In the current revision, we excluded the progressive disease (PD) timepoints from the flow cytometry dataset, as they were not informative for evaluating treatment-associated immune changes between responders and non-responders. For the GeoMx analysis, we retained the 16 available baseline

samples and removed the 2 post-treatment samples, as the limited post-treatment sampling precluded meaningful comparative interpretation.

R3C5 For figure 2D, is Fisher's exact test appropriate? I was wondering if the test is too conservative.

Reply: We apologize for the earlier error in the figure legends and would like to clarify that not all comparisons in our analysis employed Fisher's exact test. For contingency tables where all expected cell counts were greater than 5, we used the chi-squared test. When any expected cell count was less than 5, we employed Fisher's exact test instead. We have made the necessary revision to the figure legend, which now reads "*Co-barplot of mutations between Responders and Non-Responders. Only genes with more than 3 patients with mutations were shown here. The chi-squared test was used when all expected cell counts exceeded 5; otherwise, the Fisher's exact test was applied.*"

The statistical rationale for this approach is that the chi-squared test provides reliable results only when expected frequencies in each cell are sufficiently large (generally at least 5). If this condition is not met, the approximation underlying the chi-squared test may yield inaccurate p-values and increase the risk of statistical errors. In contrast, Fisher's exact test calculates the exact probability of observing the data under the null hypothesis, and is therefore more robust and appropriate when cell counts are small [1].

In response to the reviewer's comment, here we include comparisons employing the less conservative chi-square test for all comparisons, which likewise demonstrate no significant difference across all genes of interest (**Figure R6**).

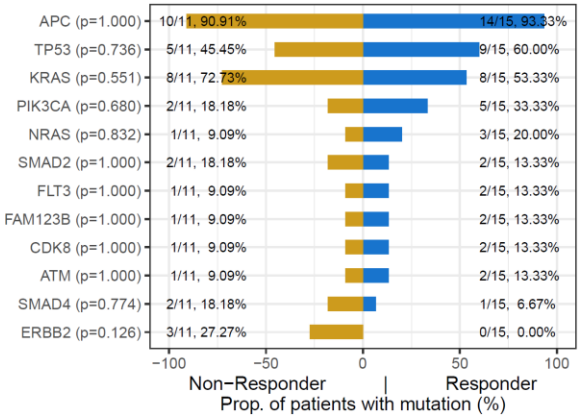


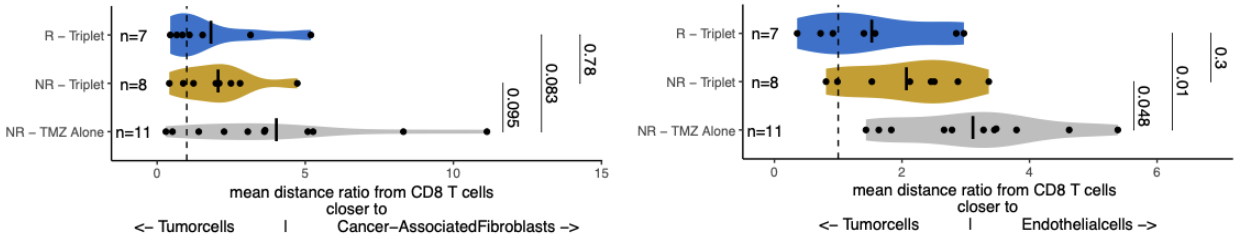
Figure R6 Co-barplot of mutations between Responders and Non-Responders. Only genes with more than 3 patients with mutations were shown here. P-values here were retrieved exclusively from a Chi-square test.

References

[1] Fisher, R. A. "On the Interpretation of χ2 from Contingency Tables, and the Calculation of P." Journal of the Royal Statistical Society 85, no. 1 (1922): 87–94. https://doi.org/10.2307/2340521.

R3C6 For figure 4F, clearly state/indicate they are distances from CD8 T-cells.

Reply: We apologize for the lack of clarity. We have edited the figure and legend to reflect clearly that CD8 T cells are the reference point and that the analysis compares the ratio of distances from CD8 T cells to tumor cells versus distances to other cell types, specifically endothelial cells and cancer-associated fibroblasts.



Revised Figure 4f. Selected violin plots showing mean distance ratios from CD8 T cells to tumor cells versus endothelial cells or cancer-associated fibroblasts, stratified by tumor response groups. P-values were calculated using a two-sided t-test.

R3C7 Figure 5F-I are missing.

Reply: We thank the reviewer for pointing this out. The panels (**Figure 5F–I**) were included in an interim revision draft but were removed from the final resubmission, as they did not add mechanistic insight beyond existing spatial and immune analyses and introduced unnecessary visual complexity. The figure and legend have been updated accordingly.

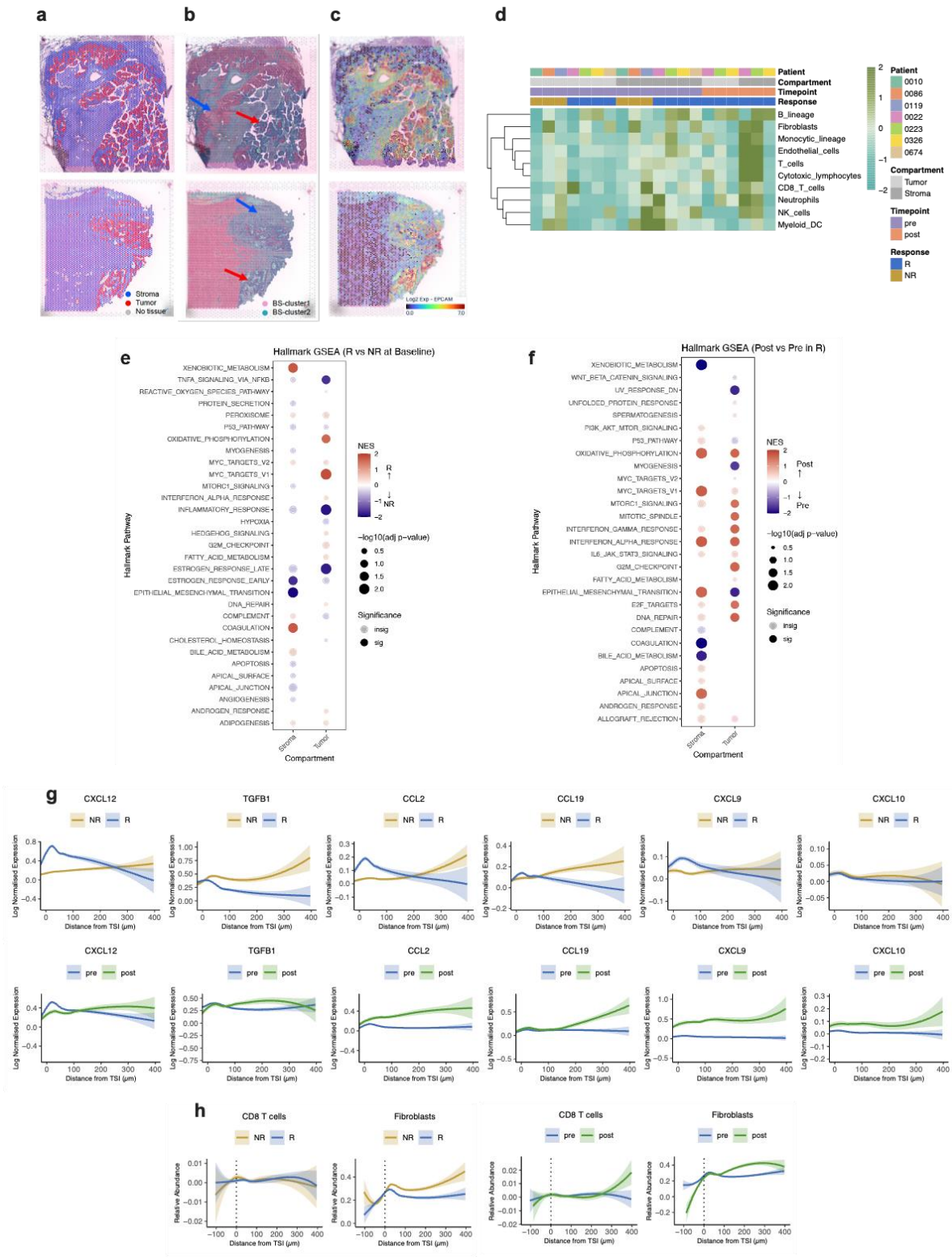


Fig. 5. Spatial transcriptomics of baseline and post-treatment samples on Visium

a, Tumor and stromal segment identification via pathologist-guided supervised machine learning approach using QuPath. **b,** Tumor and stroma segment identification via BayesSpace clustering. Blue arrows indicate pathologist-

annotated stroma spots that were inaccurately clustered as tumor, while red arrows indicate pathologist-annotated tumor spots that were clustered as stroma. **c,** EPCAM gene transcripts are upregulated in tumor regions. **d,** Relative abundance of stromal and immune cell types by response, timepoint and compartment. Spatial cell type enrichment was estimated using UCell scoring applied to log-normalized gene expression data. Gene sets were derived from the MCP-counter framework to capture fibroblasts, endothelial cells, and key immune subsets across spatial transcriptomic spots. To facilitate comparative visualization, UCell scores were Z-score normalized within each cell type, highlighting relative enrichment patterns across tissue compartments (tumor vs. stroma), clinical response groups (Responder vs. Non-Responder), and timepoints (pre- vs. post-treatment). **e,** GSEA of MSigDB Hallmark gene sets enriched in tumor and stroma of R vs. NR at baseline. **f,** GSEA of MSigDB Hallmark gene sets enriched in tumor and stroma of post-treatment vs pre-treatment samples in R. **g,** Expression profiles for cytokines and **h,** $CD8^+$ T cell and fibroblast scores along a continuous tumor–stroma gradient. This gradient was defined by the shortest Euclidean distance from each spot to the nearest spot in the opposite compartment using a nearest-neighbor approach with positive distances assigned to stromal spots and negative distances to tumor spots. Cytokine gradients were analyzed exclusively within the stromal compartment, where spatial heterogeneity is more pronounced and cytokine signaling is more likely to influence immune cell positioning. Smoothed conditional means density curves were retrieved with the geom_smooth() function in ggplot2 in R.

Abbreviations: MCP, Microenvironment Cell Populations; pDCs, plasmocytic Dendritic Cells; mDCs, myeloid dendritic cells; R, Responder; NR, Non-Responder; NK, Natural Killer; DC, Dendritic Cell; GSEA, Gene set enrichment analysis (GSEA).

Response to Reviewer #2

R2C1 The authors have addressed all my previous comments, and I have no further concerns.

Reply: We thank the reviewer for the positive re-evaluation of the revised manuscript and for confirming that all prior concerns have been addressed.

Response to Reviewer #3

R3C1 I went through the Choo et al. revision. As per our review, it is really up to the editor. The authors addressed all of the points we made, except for the insufficient sample size to support their conclusions. There is not a simple way to increase the sample size for a robust analysis.

Reply: We thank the reviewer for the careful reassessment of the revision. We acknowledge the limitation of sample size. This translational analysis included all available tumor biopsy and peripheral blood samples from patients enrolled in the parent clinical study; biospecimen availability was determined by trial participation, tissue adequacy, and patient consent. The resulting cohort therefore represents a complete series of eligible samples, which is standard practice for embedded translational studies. This limitation is acknowledged in the manuscript, and conclusions are framed conservatively and as hypothesis-generating.