nature portfolio

Peer Review File

Glioblastoma cells increase expression of notch signaling and synaptic genes within infiltrated brain tissue

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author): Expert in spatial and single-cell transcriptomics, bioinformatics, glioma genomics, and tumour microenvironment

Harwood et al. performed image-based spatial transcriptomics on GBM samples (n=11), and revealed increase expression of notch 1 signaling and synaptic genes in the tumor periphery. Overall, these datasets are valuable to the field. I have the following concerns about the data presentation and interpretation.

1. Figure 1B-D, the authors should annotate each cell type/cell state in the same UMAP/plot, then split them by malignant vs normal or by sample ID while keeping the UMAP coordinates, instead of clustering malignant and normal cells independently.
2. Figure 2A, the authors should compare their cell composition with published dataset to backup their claim they captured increased proportions of normal cells.
3. Figure 2F, it is not clear how the authors define the boundary between CT, MVP and Necrosis.
4. The authors used two distinct spatial RNA-sequencing platforms in their study. However, the data generated from these separate methodologies are not clearly labeled or specified in the figures.
5. The authors should be consistent in their presentation of differential expression results, particularly regarding the use of p-values versus adjusted p-values. In Figure 1H, they used raw p-values, whereas in other sections they used adjusted p-values.
6. In Figure 3 and 5, the authors have identified differentially expressed genes between the core and infiltrative regions of the tumors. They should integrate these data sets to reinforce their conclusions.
7. In Figure 3F, the authors chose to illustrate their point by showing the expression of a single gene, PTGSD, as an example. What about the other genes? Do they show similar cell-type specificity?
8. The authors should provide the rationale why they only performed trajectory analysis on one sample out of 5.
9. The authors should try to integrate and compare their datasets with recently published sequencing-based spatial transcriptomics on GBM to confirm their major conclusions.
10. In Figure 4, the authors used mutant p53 protein staining to label malignant cells. However, it is well documented that even in p53-mutant tumors not all tumor cells express high level of mutant p53. Some of p53-negative/low cells may be malignant cells as well. Also, GFAP could be expressed by both astrocytes and tumor cells. Thus, the cell identities should be verified by their expression of malignant gene signatures.
11. Figure 4C, what about the expression of classical and neural subtype signature genes?

Reviewer #2 (Remarks to the Author): Expert in cancer genomics, tumour microenvironment, spatial and single-cell omics

The manuscript by Harwood et al. investigates gene expression and cellular composition of glioblastoma in the tumour core and periphery using spatial transcriptomics on 11 samples. The study reveals that the core and periphery exhibit different gene expression patterns and cellular features. The core exhibits

mesenchymal, proneural, and classical subtypes, while the periphery cells show gene expressions related to neurodevelopment and synaptic functions. This study improves our understanding of glioblastoma's invasive behaviour and provides information on pathways crucial for tumour growth and its interaction with the microenvironment. However, I have severe reservations about the claims made in the manuscript due to using a small number of samples without functional validation.

The study used cutting-edge spatial transcriptomics CosMx but only on five primary GBM tumours (Fig1). Due to interpatient heterogeneity, it is unclear how universal these observations are, especially since they are not validated in other single-cell or bulk RNA-seq datasets. For example, do these samples have very similar genetic drivers of glioblastoma?

Furthermore, the authors performed some excellent bioinformatics analysis of spatial transcriptomics data but did not provide a clear biological context for these observations. For instance, one of the essential findings is that the Notch signalling pathway is more active in the tumour periphery. However, it is not clear what is the biological relevance of this observation. Does high-notch activity in the periphery influence the immune landscape? If yes, How? Does the high-notch activity make malignant cells in the periphery more EMT-like? Can the role/function of the peripheral notch be tested by blocking notch activity?

Moreover, several published single-cell and bulk RNA-seq studies on glioblastoma exist. Still, the authors did not use their CosMx data effectively to correlate their findings with existing literature and the broader clinical relevance of their observations. This work is highly descriptive and has been obtained on very few samples, making conclusions unpredictable.

Reviewer #3 (Remarks to the Author): Expert in glioma genomics, tumour microenvironment, and spatial omics

In this study, Harwood et al utilize spatial transcriptomic approaches to compare the expression profiles of malignant cells in tumor core regions to those in the tumor periphery. They find that malignant cells in the tumor cores of mesenchymal tumors are different than those of proneural tumors, but these differences do not expand to the tumor periphery. They also show that malignant cells in the core tend to be more mesenchymal in nature, while those from the periphery tend to be more stem-like or neuronal. The analyses presented in this manuscript are interesting and make good use of the spatial data generated. However, the results are descriptive in nature and their applicability is hindered by a limited sample size, a lack of statistical testing, and ambiguity in some of the methods.

Major Comments

1. The analyses described in Figure 2 (lines 125-136) need statistics in order to be properly evaluated. While the increase in MES-like 2 near necrosis is visually striking in the presented image, proper metrics that compare the distribution of MES-like 2 cells to other cell types will allow the readers to determine whether this behavior is cell type specific. Similarly, the differential expression analyses presented in this

section should be presented with p-values adjusted for multiple testing as opposed to the unadjusted p-values they are currently presented with.

2. It's unclear how "histological neighborhoods" were defined. The methods describes cellular neighborhoods, which are defined as the 100 cells closest to a given cell type but does not describe how this approach was extended to histological entities. The histological neighborhoods depicted in Figure 2F do not seem to correspond to the H&Es. For instance, regions of MVP in the H&E slide seem to frequently be in neighborhoods of cells that are labelled as cellular tumor. Collectively, this makes it difficult to interpret the conclusions reached in Lines 137-153.

3. Similar to comment 1, the analyses described in Figure 3 (lines 161-197) are interesting and produce intriguing qualitative results. However, without proper statistics, including p-values that compare cell abundance and gene expression across regions, it is difficult to evaluate how reproducible and meaningful these results are.

4. The comparisons in Figure 4F need p-values in order to be interpreted properly.

5. It is difficult to understand the analyses described in Figure 5. Which single-nucleus RNA dataset did the authors use? The dataset provided in the methods is a single-cell RNAseq dataset of cultured glioma stem cells rather than a single-nucleus RNAseq dataset containing a mix of malignant and resident brain cells. The authors refer to this approach as deconvolution, but it is not clear what deconvolution is being done. Instead, the authors seem to have clustered genes based on single-cell expression and then selected the gene clusters that were enriched in tumor cells. Once this is done, they then use these clusters to interpret their results from a differential expression analysis of periphery vs core-associated malignant cells, though its not clear in the text what this annotation adds. The authors should provide additional details on what hypotheses they were testing here and how it relates to their previous spatial analyses.

Minor Comments

1. The color scheme in Figure 2B is confusing to read as it is centered at R ~ 0.3. The authors should rescale this figure so that white, the center of the color scheme, corresponds to R = 0 and the min and max values are -1 and 1, respectively.

2. Line 139: The method for defining histological neighborhoods is important to understanding this section and should briefly be outlined in the results section.

3. The authors should provide more information in the results section on the GeoMX and CosMX methods, as it may not be clear to readers what these assays are measuring and how they differ.

Answer summary

We thank all reviewers for taking the time to review our manuscript, and for their comments and concerns which have all helped us improve our manuscript. We have made substantial revisions in all sections to address the concerns of the reviewers, and we are very grateful for their insights which have, in our opinion, helped us lift the manuscript to a higher level.

Reviewer #1 (Remarks to the Author): Expert in spatial and single-cell transcriptomics, bioinformatics, glioma genomics, and tumour microenvironment

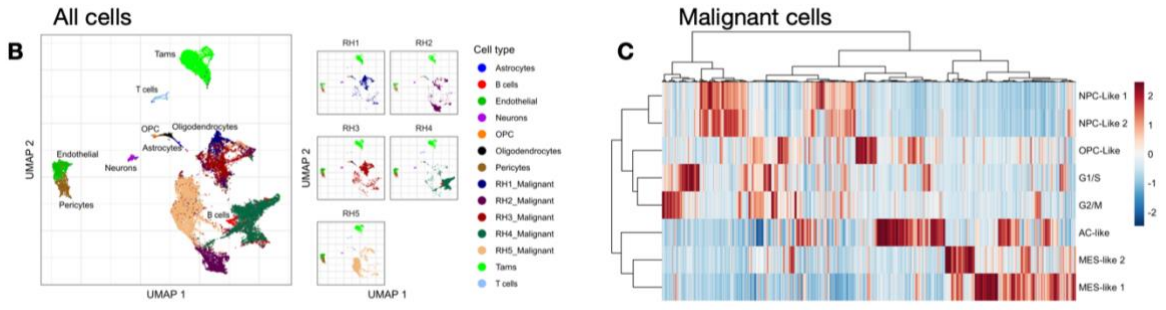
Harwood et al. performed image-based spatial transcriptomics on GBM samples (n=11), and revealed increase expression of notch 1 signaling and synaptic genes in the tumor periphery. Overall, these datasets are valuable to the field. I have the following concerns about the data presentation and interpretation.

We thank the reviewer for stating that our datasets are valuable to the field. We agree with the reviewer on all the mentioned points and have responded to them individually below.

1. Figure 1B-D, the authors should annotate each cell type/cell state in the same UMAP/plot, then split them by malignant vs normal or by sample ID while keeping the UMAP coordinates, instead of clustering malignant and normal cells independently.

We thank the reviewer for this comment. We have now used the same UMAP coordinates for all plots and have chosen to display the patient-specific UMAP plots. Moreover, to illustrate how we assigned malignant cell states to patient-specific malignant clusters, we have shown the average scaled expression for each Neftel state in malignant cells, which was used to assign Neftel states.

Fig. 1B-C



Text (Lines 92-97)

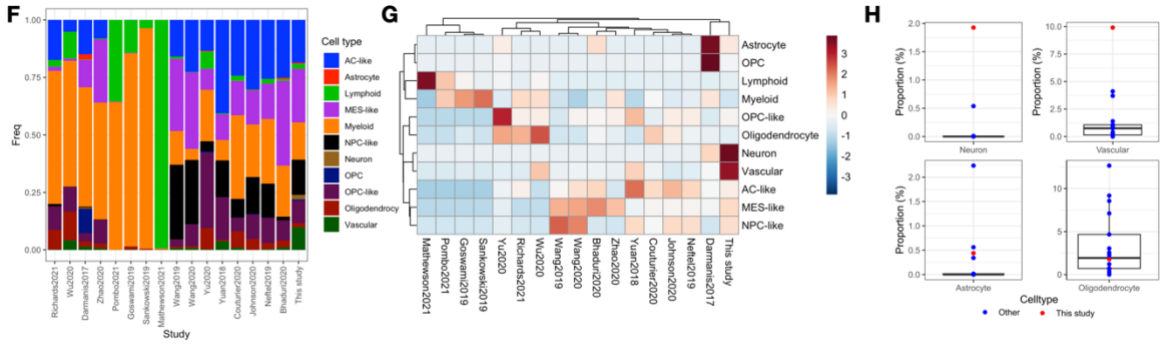
"Cell clusters were manually annotated as malignant or normal, using a combination of patient-specific clustering patterns and gene markers from publicly available databases from previous studies (9, 31, 32) (Fig. 1B). The patient-specific clustering of malignant cells supported cell annotation, as normal cells clustered regardless of their sample of

origin. Malignant cell states were then assigned using modified gene sets from previous studies (9) (Fig. 1C)."

2. Figure 2A, the authors should compare their cell composition with published dataset to backup their claim they captured increased proportions of normal cells.

Thank you for the suggestion of quantifying these differences. We have compared the proportion of cells within our dataset to those annotated in a recently published glioblastoma atlas. We find that while our dataset is similar to a high number of single-cell RNA sequencing datasets, we identify a very high number of vascular cells and neurons within our dataset.

Below are the relevant figures and rewritten text.
Fig. 1F-H



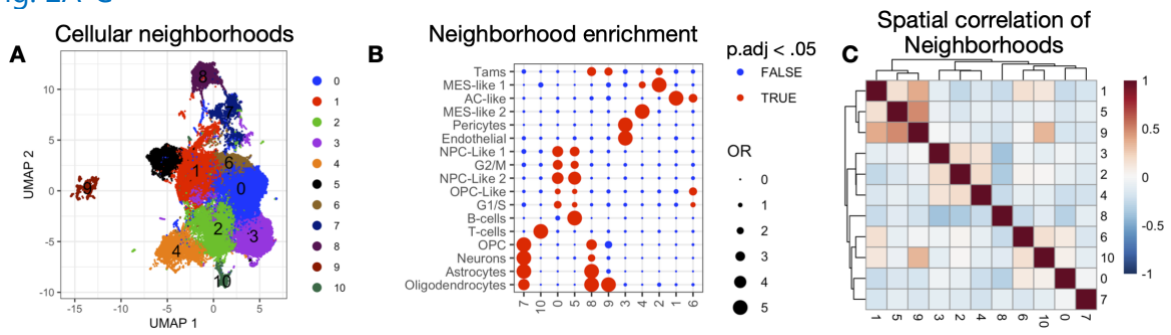
Text (Lines 103-113)
"Spatial transcriptomics, in contrast to single-cell sequencing, does not require tissue dissociation and has been used to quantify cellular proportions of tumors (34). We investigated how the composition of our dataset compares with other single-cell RNA sequencing studies annotated in a recently curated GBM atlas (GBMap) (32) (Fig. 1F-H). Some of these studies sorted cells prior to single-cell RNA sequencing, making their cellular proportions not representative of the tumor (e.g. Darmanis et al). However, the cellular composition of most studies were randomly sampled. Most malignant cell states and myeloid cells were of similar proportions in our data compared to the other studies. We found that astrocytes were found in high proportions in our study (with the highest being from Darmanis et al (11), where astrocytes were selected during sample collection), while oligodendrocytes were of average proportions. Lastly, our study showed remarkably high proportions of vascular cells and neurons compared to all other datasets."

3. Figure 2F, it is not clear how the authors define the boundary between CT, MVP and Necrosis.

Thank you for pointing this out. We have gone into more detail of how we have performed the spatial neighborhood analysis. Briefly, we identified the 50 nearest neighbors to each cell and clustered cells by counting their neighbors cell labels. Neighborhoods were compared to adjacent H&E stained sections and enriched for cell types for annotation.

Below are the relevant figures and rewritten text.

Fig. 2A-C



Text (Lines 122-126)

"To gain a global overview of different cellular neighborhoods, defined by areas with different cellular compositions, we counted cell annotations of the 50 nearest neighbors to each cell, creating an annotation by cell count matrix. Dimension reduction and clustering were performed to identify cellular neighborhoods (Fig. 2A) and results were visually assessed using local transcriptional neighborhoods (Fig. S2A-C) (see methods) and adjacent sections stained with H&E."

4. The authors used two distinct spatial RNA-sequencing platforms in their study. However, the data generated from these separate methodologies are not clearly labeled or specified in the figures.

We thank the reviewer for pointing this out to us. We have now clearly indicated which platforms are being used in both the figure legends and main text.

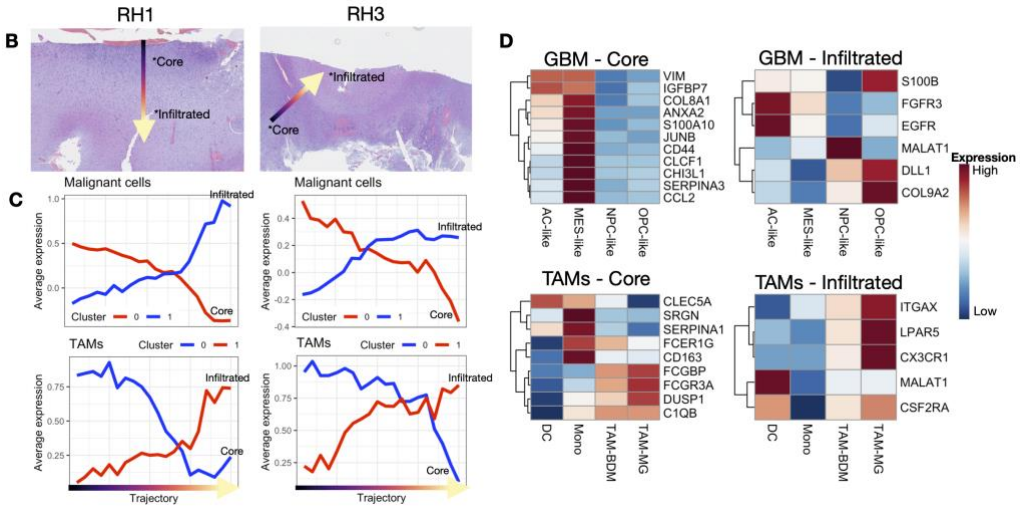
5. The authors should be consistent in their presentation of differential expression results, particularly regarding the use of p-values versus adjusted p-values. In Figure 1H, they used raw p-values, whereas in other sections they used adjusted p-values.

Thank you for this comment. We agree and have adjusted the analyses and manuscript accordingly.

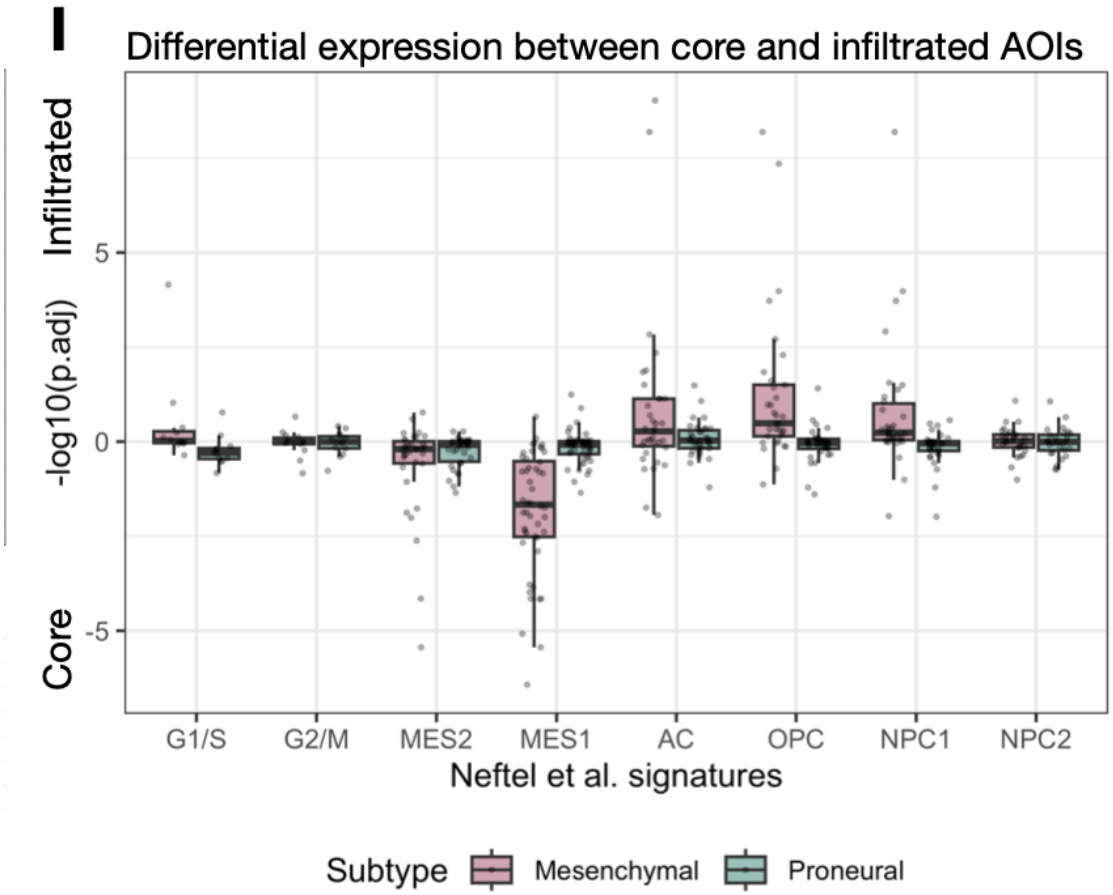
6. In Figure 3 and 5, the authors have identified differentially expressed genes between the core and infiltrative regions of the tumors. They should integrate these data sets to reinforce their conclusions.

We thank and agree with the reviewer regarding the integration of the GeoMx and CosMx analysis. Due to the large differences in the gene coverage between technologies, the overlap within genes associated with the infiltrating cells is scarce. The main findings from the initial CosMx analysis provided an idea of which malignant cell states are associated with the infiltrated brain tissue. Here we show the majority of genes upregulated in the tumor core to be most expressed in the MES-like state, while genes with highest expression in the infiltrated tissue were expressed most by NPC-, OPC- and AC-like states (Fig. 3B-D). In GeoMx data we see the same pattern (Fig. 4I).

Below are the relevant figures and rewritten text.



Text (Lines 176-188)
"Two patient samples had intact trajectories between the core and infiltrated brain tissue (Fig. 3B, S3A-B). To enable the comparison of trajectories between samples, we identified spatially variant gene clusters with a low clustering resolution for the purpose of between-patient comparison (Fig. 3C). In malignant cells and myeloid cell populations, we identified genes with spatial variance that were unique and common between samples (Fig. 3D, S3C). To assess which cells and malignant states expressed these genes, we aggregated gene expression for each cell type or state in GBMap (32) (Fig. 3D). Genes commonly expressed in the tumor core were highly expressed in MES-like GBM states. Amongst these genes was CHI3L1, a gene recently associated with connected and non-invasive malignant cells (37). The few genes upregulated in invasive cells were most associated with the AC-like state and growth factor receptors (EGFR, FGFR3), or NPC-like/OPC-like states, indicating their increased presence in malignant cell composition outside the tumor. TAM genes upregulated in the core were expressed by both blood-derived macrophages and microglia, while those in the infiltrated area were, as expected, most expressed in microglia."



Text (Lines 262-266)
"To compare how Neftel et al. malignant state signatures (9) varied in expression across the tumor core and infiltrated brain tissue for each tumor subtype, we tested for differential expression between these areas in p53 AOIs (Fig. 4I). Mesenchymal tumors showed large differences, where MES-like 1 marker genes within the tumor core were replaced with marker genes of the AC-, OPC- and NPC-like 1 states in the infiltrated brain tissue. We observed minimal differences for proneural tumors across regions."

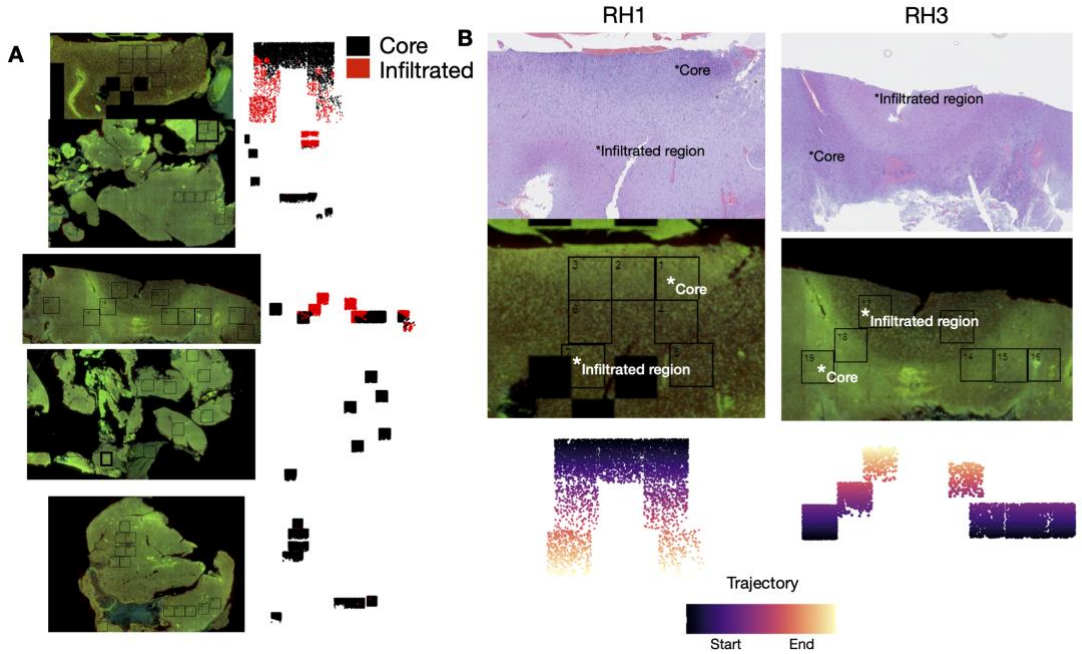
7. In Figure 3F, the authors chose to illustrate their point by showing the expression of a single gene, PTGSD, as an example. What about the other genes? Do they show similar cell-type specificity?

We thank the reviewer for pointing out the ambiguity of this choice. Due to all reviewers highlighting the importance of validation in external datasets, we have removed this analysis to make room for these findings. Briefly, we highlight the complications of performing differential expression in infiltrated brain regions using Visium data with multicellular resolution, and use non-negative matrix factorization factors to identify genes and cell states associated with the infiltrated brain tissue. These analyses are elaborated in comment #8.

8. The authors should provide the rationale why they only performed trajectory analysis on one sample out of 5.

We thank the reviewer for commenting on this. We have extended this analysis to another sample which also included adjacent field of views leading from the tumor core to infiltrated brain tissue (Fig. S2A-B). One other sample (RH2) had infiltrated brain tissue although this patient did not have adjacent core field of views and could not be used for our trajectory analysis (Fig. S2A). We changed our analysis for the comparison of two patient samples, and identified two main gene clusters which were spatially associated with either the tumor core or infiltrated brain tissue (Fig. 2C). For genes that were identified in both trajectories, we assessed which cells had highest expression of these markers for relevant cell types and states (Fig. 2D), and further characterized these genes in a Visium dataset with multicellular resolution (Fig. 2E-K). Our results across all analyses found MES-like gene markers to be associated with the tumor core, with genes associated with the NPC-, OPC, and AC-like cell states to be always or in part associated with the infiltrated brain tissue.

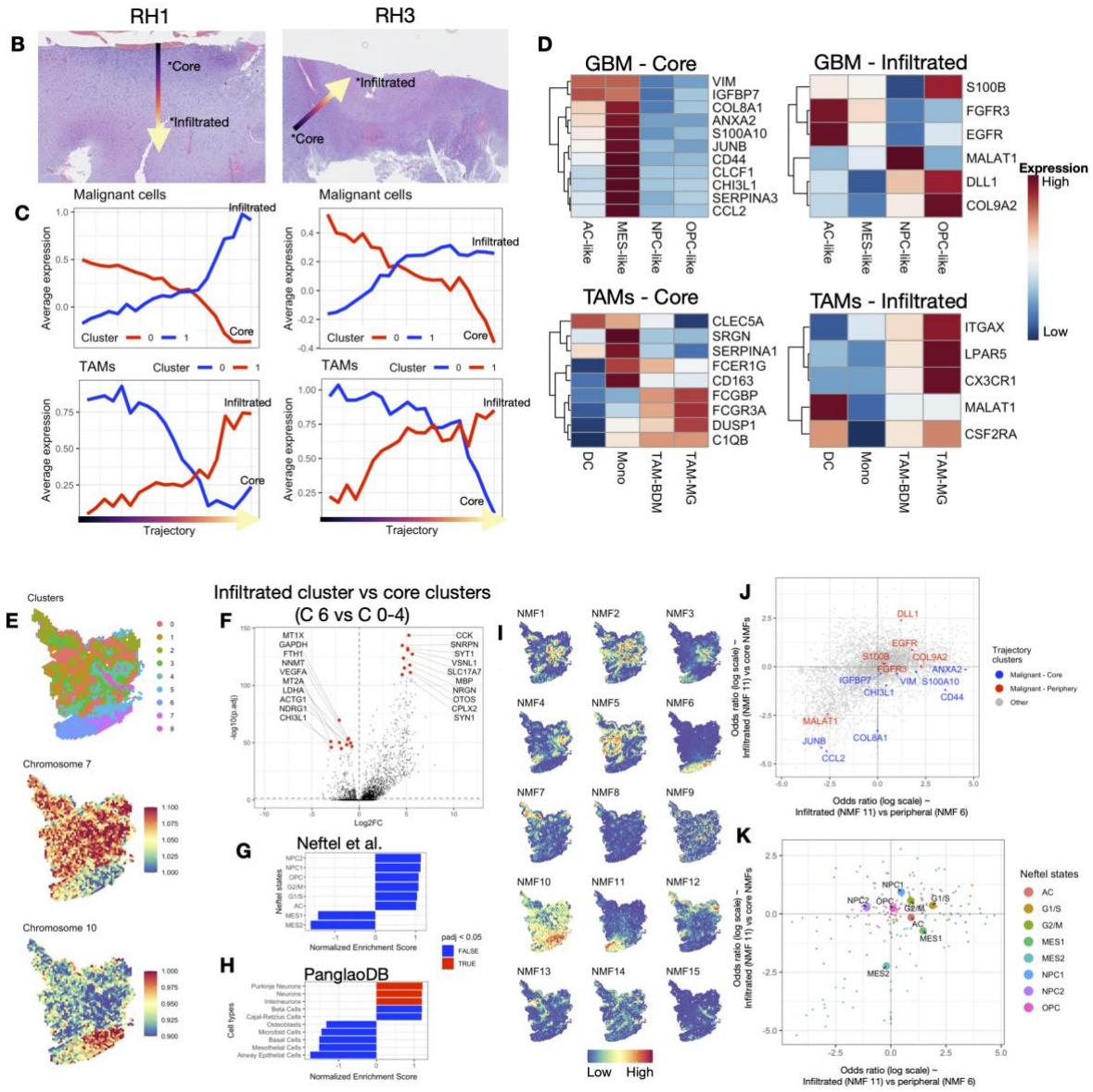
Fig. S2A-B



Text (Lines 176-178)

"Two patient samples had intact trajectories between the core and infiltrated brain tissue (Fig. 3B, S3A-B). To enable the comparison of trajectories between samples, we identified spatially variant gene clusters with a low clustering resolution for the purpose of between-patient comparison (Fig. 3C).

Fig. 2B-K



Text (Lines 178 – 221)

"In malignant cells and myeloid cell populations, we identified genes with spatial variance that were unique and common between samples (Fig. 3D, S3C). To assess which cells and malignant states expressed these genes, we aggregated gene expression for each cell type or state in GBMap (32) (Fig. 3D). Genes commonly expressed in the tumor core were highly expressed in MES-like GBM states. Amongst these genes was CHI3L1, a gene recently associated with connected and non-invasive malignant cells (37). The few genes upregulated in invasive cells were most associated with the AC-like state and growth factor receptors (EGFR, FGFR3), or NPC-like/OPC-like states, indicating their increased presence in malignant cell composition outside the tumor. TAM genes upregulated in the core were expressed by both blood-derived macrophages and microglia, while those in the infiltrated area were, as expected, most expressed in microglia.

Our findings indicated that MES-like signatures decreased in invasive GBM cells, while the remaining GBM states were associated with the infiltrated brain. However, these findings are limited to two patient samples, and a panel covering 960 genes. To further investigate infiltrated areas containing tumor cells, we identified infiltrative sections from a whole-transcriptome spatial transcriptomics dataset with multicellular resolution (14). We identified one patient with 3 distinct regions (core = high CNV alterations, invasive = intermediate CNV alterations, periphery = low CNV alterations) (Fig. 3E). When testing between the infiltrated and tumor core areas, we identified a large number of differentially expressed genes, although no malignant state signatures were significantly enriched in either area (Fig. 3F-G). Conversely, neuronal cell signatures were upregulated in infiltrated tissue (Fig. 3H). However, as GBM cells share transcriptional programs with resident brain cells, it is difficult to distinguish between which differentially expressed genes are derived from malignant and non-malignant populations.

To circumvent the limitations arising with multicellular resolution, we used non-negative matrix factorization (NMF) to deconvolute the count matrix to 15 individual factors (Fig. S3D). This identified NMF factors that were primarily expressed within the tumor core, the infiltrated area, and the tumor periphery (Fig. 3I). To identify genes that were more expressed in the infiltrating malignant cells, but not the cells of the healthy brain, we calculated the odds ratio of a gene being expressed in the factor associated with the infiltrated area compared with it being expressed in the peripheral factor (normal brain). We repeated this comparison with the factors found within the tumor core and were able to identify which genes were highly expressed in the infiltrated brain tissue compared to the tumor periphery (Fig. 3J, right vs left), and the infiltrated brain tissue compared to the tumor core (Fig. 3J, top vs bottom). This showed similar patterns to our previous findings with Delta Like Canonical Notch Ligand 1 (DLL1) being the gene most upregulated in the infiltrated area. DLL1 is also used as a NPC-like 1 marker gene, and to further investigate how genes associated with malignant states were associated with this invasive signature, we averaged the odds ratio of each malignant state from Neftel et al., and found the NPC-like 1 to be most associated with infiltrated brain tissue, with mesenchymal states being most present in the tumor core (Fig. 3K).

Our results show how transcription varies in malignant and immune cells spatially located along a trajectory from the tumor core to infiltrated brain tissue. Malignant cells within the tumor core have increased expression of genes associated with MES-like states, while cells in the infiltrated tissue increase expression of AC-, OPC-, and NPC-like states. These genes consist of growth factor receptors and notch signaling ligands, indicating their potential role in brain invasion. However, as these analyses were performed in only three patient samples, any results and interpretations may not be representative across GBM tumors."

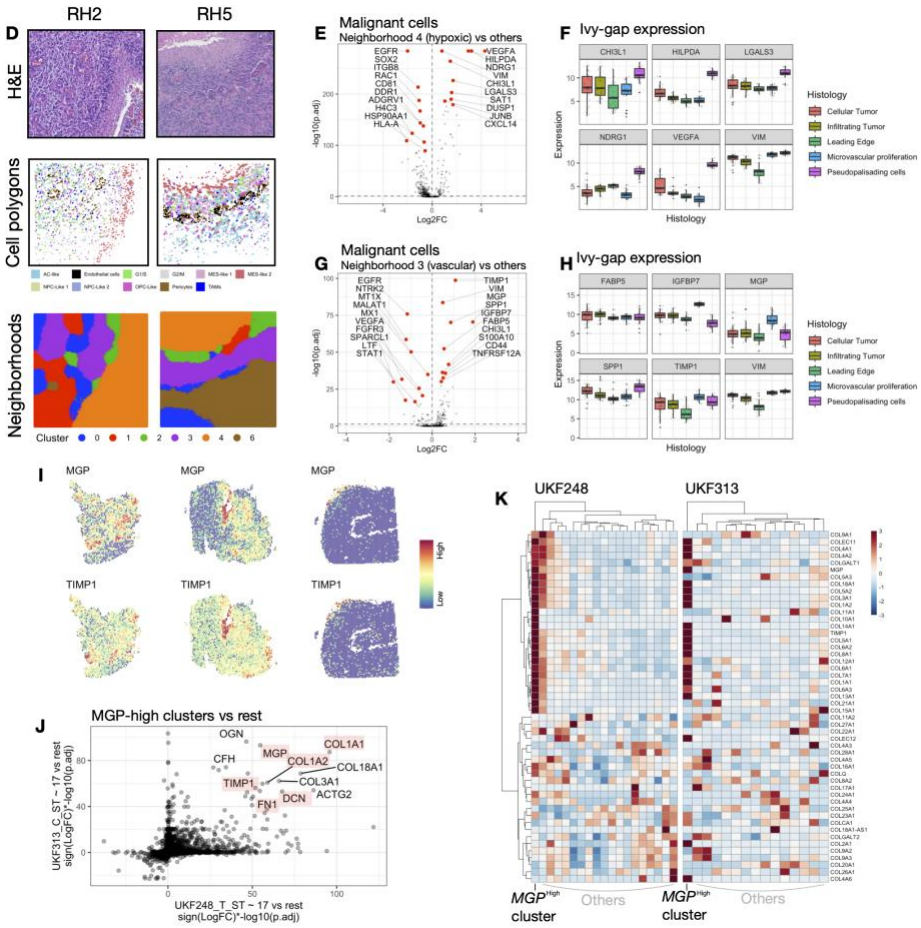
9. The authors should try to integrate and compare their datasets with recently published sequencing-based spatial transcriptomics on GBM to confirm their major conclusions.

Thank you for this comment. While very few spatial datasets have sufficient resolution to investigate malignant cells in the infiltrated brain tissue, we have integrated our analyses with other datasets whenever possible.

We identify multiple genes being overexpressed in malignant cells that are spatially located in areas with high amounts of necrosis or microvascular proliferation (Fig. 2D,E,G). We show that these genes are also upregulated in similar areas in the Ivy-gap dataset, which does not have single-cell resolution (Fig. 2F,H). Moreover, some of these genes that are more specifically expressed in areas of MVP were further found to be associated in areas with high collagen deposition in a Visium glioblastoma dataset (Fig. 2I-K).

In Fig. 3, we also use this Visium dataset to identify infiltrated brain tissue. Due to the multicellular resolution of this method, our results obtained by conventional analysis of differential expression are highly influenced by the high abundance of resident brain cells (Fig. 3E-H). Therefore, we utilize a novel approach, that identifies transcriptional factors by non-negative matrix factorization. This enables the identification of an "infiltrated brain signature", that is highly enriched for genes of the NPC-like 1 malignant state, and similar to our findings using GeoMx data in Fig. 4I.

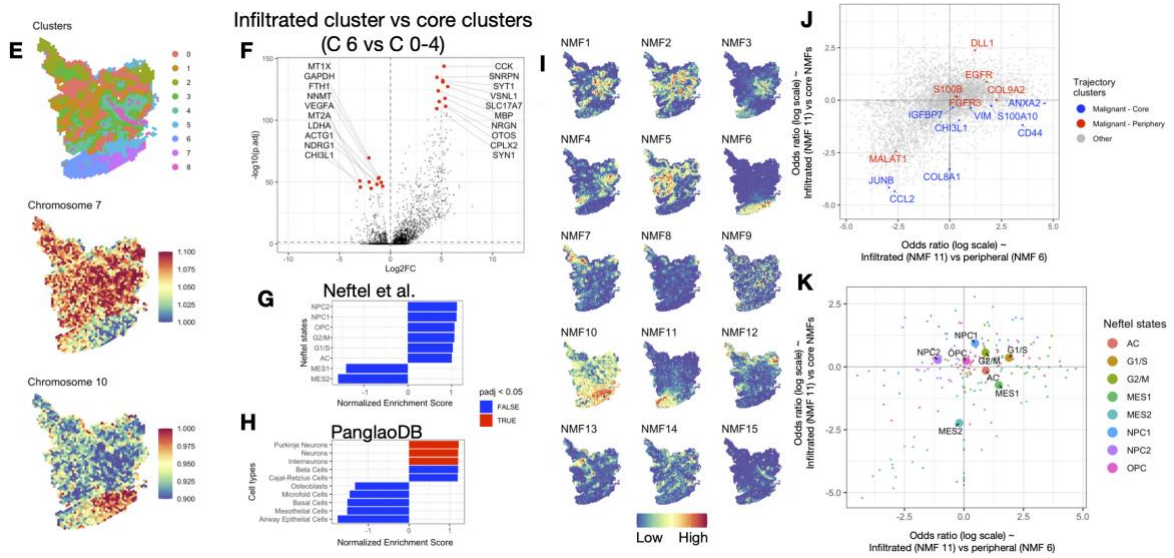
Fig 2E-K



Text (Lines 150-157)

"To further investigate the roles of these genes, we investigated the expression of MGP in a spatial transcriptomics dataset with multicellular resolution (14), and found its expression to be highest in 3 patient samples, with spatial colocalization with TIMP1 in two samples (Fig. 2I). MGPhigh and TIMP1high clusters (Fig. S2E) were tested for differential expression(Fig. 2J), and showed increased expression of MGP and TIMP1, pericyte/fibroblast markers (DCN, FN1) and collagens (e.g. COL1A1, COL1A2). Further analysis showed both patient MGPhigh clusters had higher expression of genes encoding collagen proteins, compared to the other transcriptional clusters (Fig. 2K)."

Fig 3E-H



Text (Lines 189 – 214)

"Our findings indicated that MES-like signatures decreased in invasive GBM cells, while the remaining GBM states were associated with the infiltrated brain. However, these findings are limited to two patient samples, and a panel covering 960 genes. To further investigate infiltrated areas containing tumor cells, we identified infiltrative sections from a whole-transcriptome spatial transcriptomics dataset with multicellular resolution (14). We identified one patient with 3 distinct regions (core = high CNV alterations, invasive = intermediate CNV alterations, periphery = low CNV alterations) (Fig. 3E). When testing between the infiltrated and tumor core areas, we identified a large number of differentially expressed genes, although no malignant state signatures were significantly enriched in either area (Fig. 3F-G). Conversely, neuronal cell signatures were upregulated in infiltrated tissue (Fig. 3H). However, as GBM cells share transcriptional programs with resident brain cells, it is difficult to distinguish between which differentially expressed genes are derived from malignant and non-malignant populations.

To circumvent the limitations arising with multicellular resolution, we used non-negative matrix factorization (NMF) to deconvolute the count matrix to 15 individual factors (Fig. S3D). This identified NMF factors that were primarily expressed within the tumor core, the infiltrated area, and the tumor periphery (Fig. 3I). To identify genes that were more expressed in the infiltrating malignant cells, but not the cells of the healthy brain, we calculated the odds ratio of a gene being expressed in the factor associated with the infiltrated area compared with it being expressed in the peripheral factor (normal brain). We repeated this comparison with the factors found within the tumor core and were able to identify which genes were highly expressed in the infiltrated brain tissue compared to the tumor periphery (Fig. 3J, right vs left), and the infiltrated brain tissue compared to the tumor core (Fig. 3J, top vs bottom). This showed similar patterns to our previous findings with Delta Like Canonical Notch Ligand 1 (DLL1) being the gene most upregulated in the infiltrated area. DLL1 is also used as a NPC-like 1 marker gene, and to further investigate how genes associated with malignant states were associated with this invasive signature, we averaged the odds ratio of each malignant state from

Neftel et al., and found the NPC-like 1 to be most associated with infiltrated brain tissue, with mesenchymal states being most present in the tumor core (Fig. 3K)."

10. In Figure 4, the authors used mutant p53 protein staining to label malignant cells. However, it is well documented that even in p53-mutant tumors not all tumor cells express high level of mutant p53. Some of p53-negative/low cells may be malignant cells as well. Also, GFAP could be expressed by both astrocytes and tumor cells. Thus, the cell identities should be verified by their expression of malignant gene signatures.

Thank you for pointing out these concerns. When screening for TP53 mutated tumors, we also included a criterion that there should be high expression of the p53 protein. We therefore only included tumors where the vast majority of tumor cells expressed the p53 protein. We have rewritten the text to make it more clear to the reader:

Text (Lines 225 -229)
"Using our cohort of TP53 mutated tumors with overexpression of Tumor Protein P53 (p53) in the majority of cells within the tumor core, we were able to identify regions of diffuse brain invasion and selected regions of interest (ROIs) from the tumor core, infiltrated brain tissue, periphery (few malignant cells), and brain tissue void of malignant cells (normal)."

Furthermore, we performed copy-number analysis, which identified a high correlation between gene amplifications in genomic data and the predicted copy-number of the transcriptional data (Fig S1B, 4D). Finally, principle component analysis showed a good separation between p53 and Iba1 areas of interest (Fig. 4C), and our p53 segments showed high expression of the Neftel malignant states (Fig. 4F).

We agree with the reviewer that GFAP could be astrocytes outside the tumor core. We did not use this data for any analysis in the study for this matter and have removed it in the manuscript to avoid any confusion.

Fig S1B

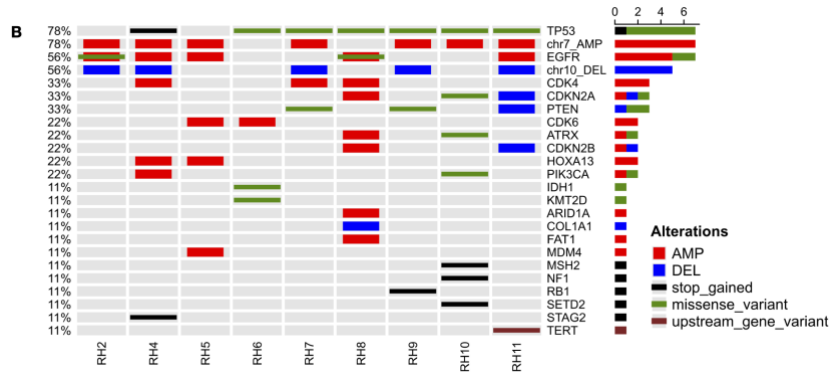
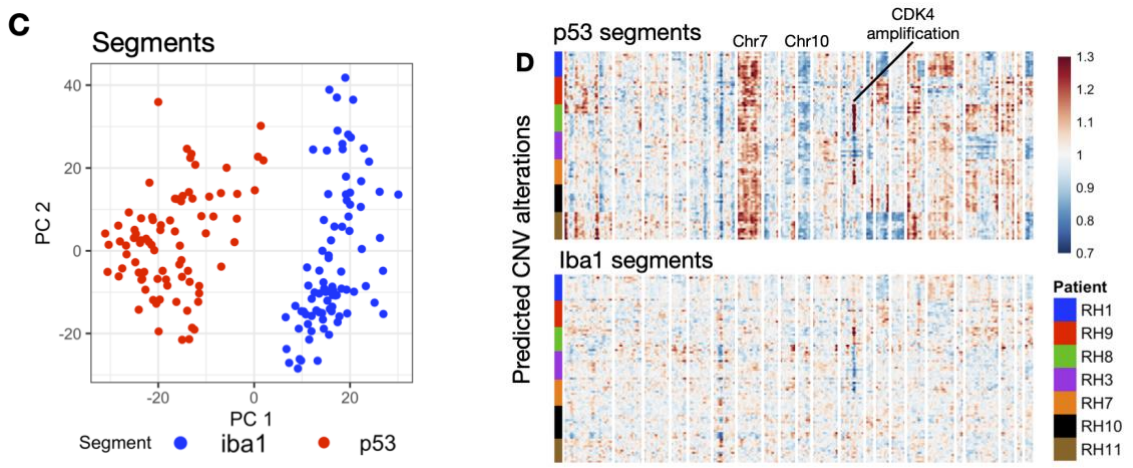


Fig 4C-D

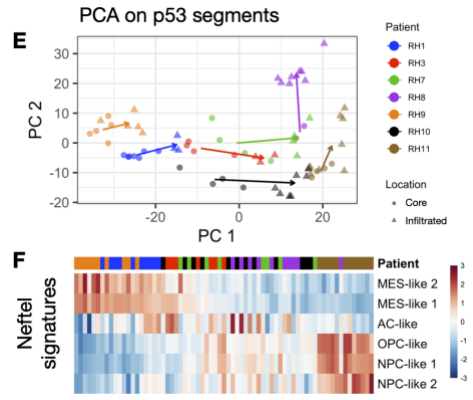


Text (Lines 232 – 235)
"Principal component analysis showed a clear difference between p53 and Iba1 segments (Fig. 4C). To ensure that the p53 segmentation captured the transcriptomic profile of malignant cells, we performed copy-number variation (CNV) analysis to approximate the malignant cell purity of each segment (Fig. 4D)."

11. Figure 4C, what about the expression of classical and neural subtype signature genes?

Thank you for this concern. We agree that our analysis did not take each subtype or malignant cell state into account. Due to recent papers identifying improved normalization methods, we revisited these analyses and used PCA analysis to identify which malignant segments were either mesenchymal or proneural (Fig. 4E). As tumor subtypes also include genes associated with the microenvironment, we quantified all Neftel et al. signatures in each segment instead (Fig. 4F).

Fig 4E-F



Text (Lines 241-247)
"While the GeoMx technology can identify single populations of cells, it is not at single-cell resolution, making each area of interest (AOI) a mix of multiple malignant states. Using principal component analysis on p53 AOIs, we found the first component to identify a mesenchymal to proneural axis (Fig. S4C-D), a pattern commonly observed in

single-cell RNA sequencing experiments (16, 38). We found most tumors (n = 5) transitioned towards a proneural composition in infiltrated brain tissue (Fig. 4E-F), while those that were most proneural (n = 2) stayed proneural but with an increase in genes associated with neuronal markers, neurotransmitter transportation, and presynaptic and axonal genes (Fig. S4D)."

Reviewer #2 (Remarks to the Author): Expert in cancer genomics, tumour microenvironment, spatial and single-cell omics

The manuscript by Harwood et al. investigates gene expression and cellular composition of glioblastoma in the tumour core and periphery using spatial transcriptomics on 11 samples. The study reveals that the core and periphery exhibit different gene expression patterns and cellular features. The core exhibits mesenchymal, proneural, and classical subtypes, while the periphery cells show gene expressions related to neurodevelopment and synaptic functions. This study improves our understanding of glioblastoma's invasive behaviour and provides information on pathways crucial for tumour growth and its interaction with the microenvironment. However, I have severe reservations about the claims made in the manuscript due to using a small number of samples without functional validation.

We thank the reviewer for spending their time to help us improve our study. We also thank the reviewer for mentioning how our study improves our understanding of the invasive cells. While one of the reasons for conducting these studies were the lack of spatial transcriptomics datasets that are able to characterize the invasive glioblastoma cells, we agree that our study does not have a sufficient number of patients to cover the full spatial heterogeneity of glioblastoma tumor cores and infiltrated areas. However, all of our samples were chosen based on the neuropathological expertise of our group and covers a large proportion of the tissue morphology that we see in patients. To ensure reproducibility of our findings,, we have now performed an extensive analysis of external datasets that have similar tissue patterns. Finally, we have reorganized our results and reanalyzed our data to provide a more biological context of our findings, and have identified a gene module that is prognostic of poor patient survival.

1. The study used cutting-edge spatial transcriptomics CosMx but only on five primary GBM tumours (Fig1). Due to interpatient heterogeneity, it is unclear how universal these observations are, especially since they are not validated in other single-cell or bulk RNA-seq datasets. For example, do these samples have very similar genetic drivers of glioblastoma?

We thank the reviewer for these important concerns. We have included genomic information for all possible samples. Our samples include a high proportion of TP53 mutated tumors, due to the criterions of the cohort in our GeoMx analysis. However, this pathway is mutated in a large number of glioblastoma tumors, and has not been shown to influence the tumor subtype like CDK4, NF1, and EGFR. Apart from this gene, we found no unusual mutations in the patients that were able to be genomically

characterized. We have now added a supplementary figure with the genomic characterization of patient tumors (Fig. S1B). Moreover, we have compared the proportion of cells within our dataset to those annotated in a recently published glioblastoma tumor atlas (Fig. 1F-H). Our dataset is similar to a high number of single-cell RNA sequencing datasets that did not select for cells prior to single-cell RNA sequencing, and we identify a very high number of vascular cells and neurons within our dataset. These differences are interpreted to be a strength of the spatial method and not attributed that our patients are not representative of the general glioblastoma population.

We also thank the reviewer for suggesting the validation of our results in external datasets. For our most interesting results, we have now investigated our findings in external datasets as written below:

In Fig. 2 we identify multiple genes being overexpressed in malignant cells that are spatially located in areas with high amounts of necrosis or microvascular proliferation (Fig. 2D,E,G). We show that these genes are also upregulated in similar areas in the Ivy-gap dataset, which does not have single-cell resolution (Fig. 2F,H). Moreover, some of these genes that are more specifically expressed in areas of MVP were further found to be associated in areas with high collagen deposition in a Visium glioblastoma dataset (Fig. 2I-K).

In Fig. 3, we use the same Visium dataset to identify samples with infiltrated brain tissue. Due to the multicellular resolution of this method, our results obtained by conventional analysis of differential expression are highly influenced by the high abundance of resident brain cells (Fig. 3E-H). Therefore, we utilize a novel approach, that identifies transcriptional factors by non-negative matrix factorization. This enables the identification of an "infiltrated brain signature", that is highly enriched for genes of the NPC-like 1 malignant state, and similar to our findings using GeoMx data in Fig. 4I.

In Fig. 5, we identify gene modules that are differentially expressed between the tumor core and infiltrated brain tissue (Fig. 5A-B). The significantly regulated gene modules in the infiltrated brain tissue show a high correlation in TCGA patients (n ~ 120 patients). Furthermore, we show in data from Darmanis et al, 2017, that the cells (n ~50 cells) that are found in the tumor periphery of two patients follow almost a perfect tendency with the gene modules we found (Fig. 5D).

Below are the relevant figures and rewritten text.

Fig. S1B

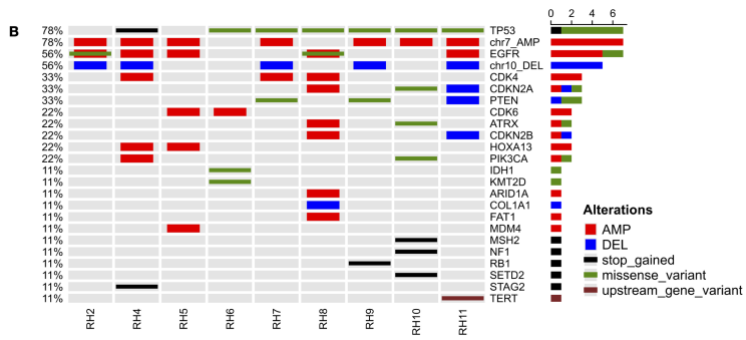
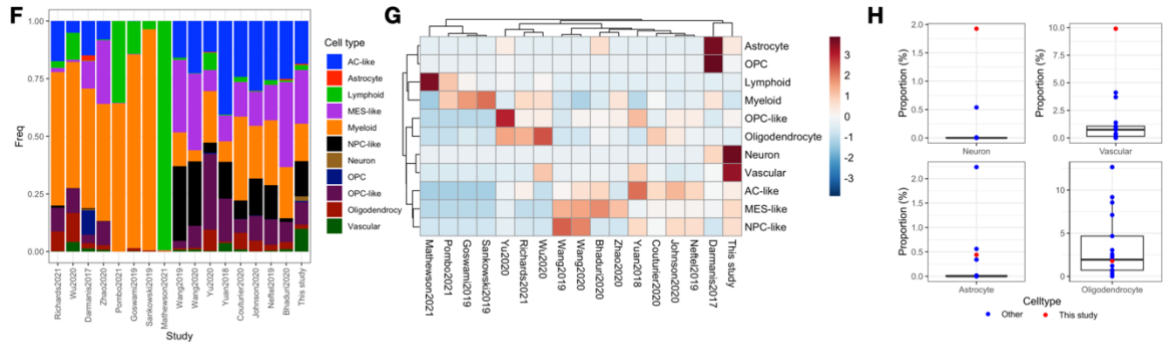


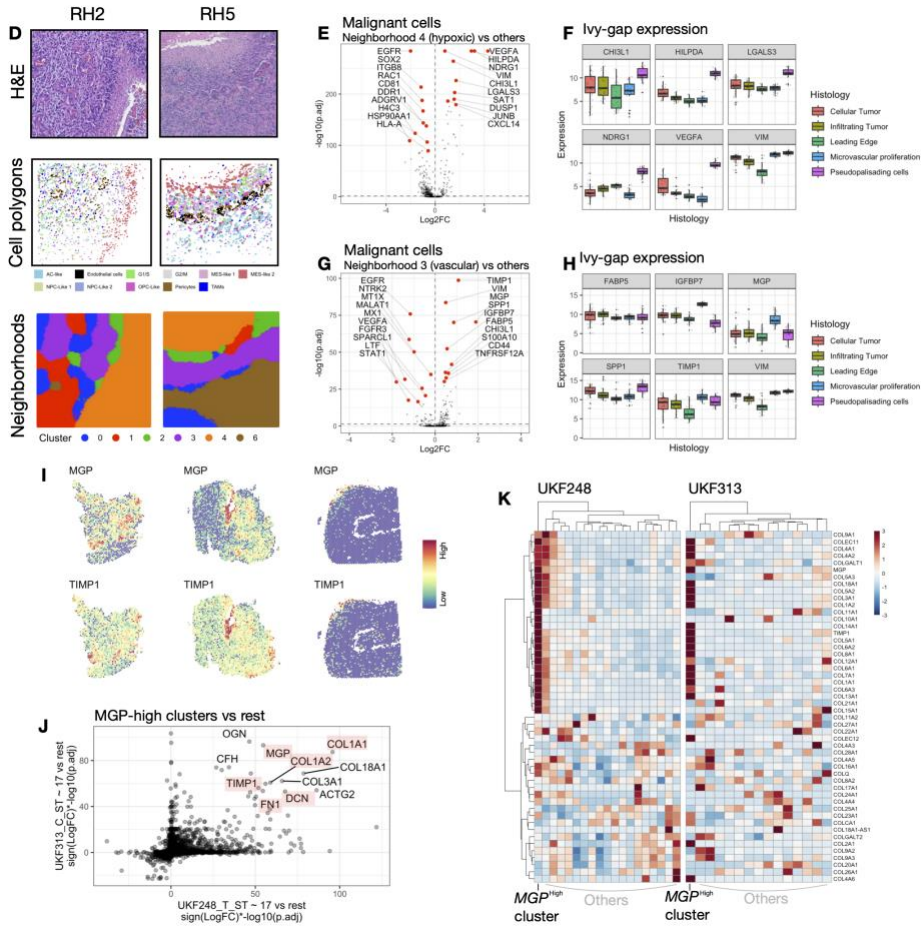
Fig. 1F-H



Text (Lines 103-113)

"Spatial transcriptomics, in contrast to single-cell sequencing, does not require tissue dissociation and has been used to quantify cellular proportions of tumors (34). We investigated how the composition of our dataset compares with other single-cell RNA sequencing studies annotated in a recently curated GBM atlas (GBMap) (32) (Fig. 1F-H). Some of these studies sorted cells prior to single-cell RNA sequencing, making their cellular proportions not representative of the tumor (e.g. Darmanis et al). However, the cellular composition of most studies were randomly sampled. Most malignant cell states and myeloid cells were of similar proportions in our data compared to the other studies. We found that astrocytes were found in high proportions in our study (with the highest being from Darmanis et al (11), where astrocytes were selected during sample collection), while oligodendrocytes were of average proportions. Lastly, our study showed remarkably high proportions of vascular cells and neurons compared to all other datasets."

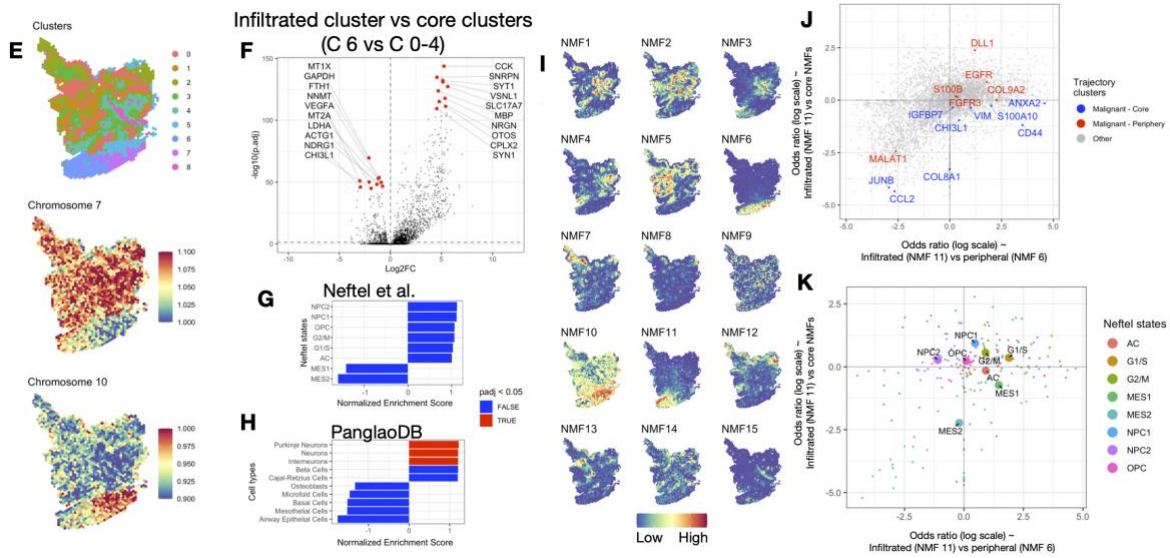
Fig 2D-K



Text (Lines 150-157)

"To further investigate the roles of these genes, we investigated the expression of MGP in a spatial transcriptomics dataset with multicellular resolution (14), and found its expression to be highest in 3 patient samples, with spatial colocalization with TIMP1 in two samples (Fig. 2I). MGPhigh and TIMP1high clusters (Fig. S2E) were tested for differential expression(Fig. 2J), and showed increased expression of MGP and TIMP1, pericyte/fibroblast markers (DCN, FN1) and collagens (e.g. COL1A1, COL1A2). Further analysis showed both patient MGPhigh clusters had higher expression of genes encoding collagen proteins, compared to the other transcriptional clusters (Fig. 2K)."

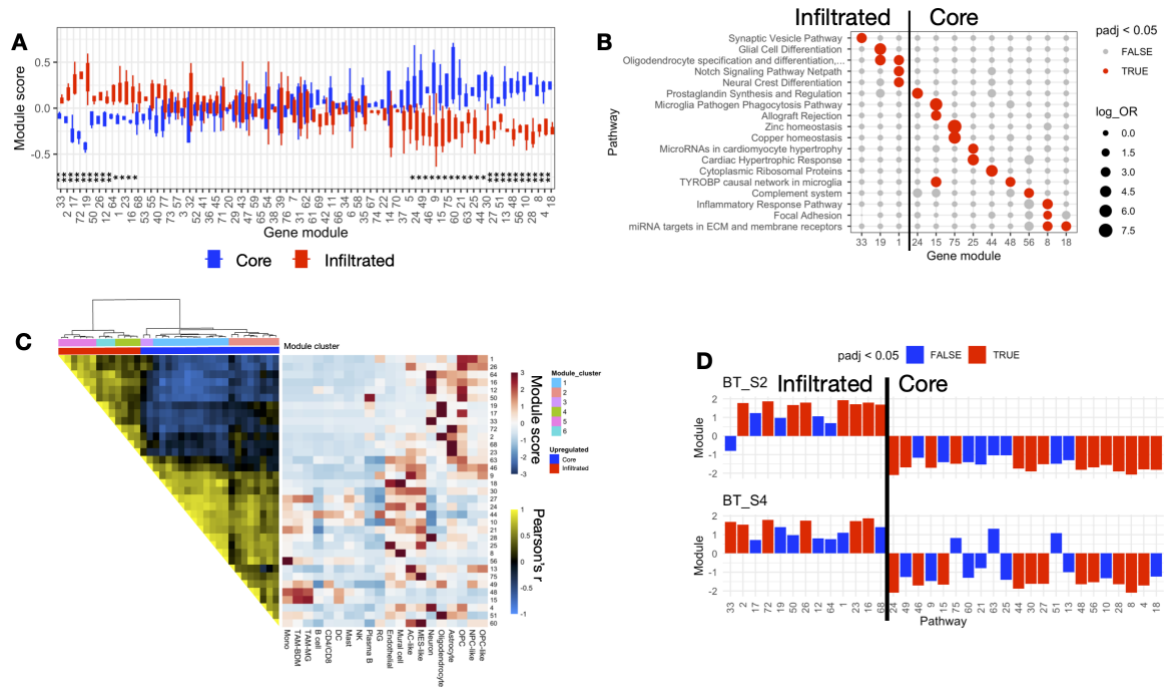
Fig 3E-H



Text (Lines 189-199)

"Our findings indicated that MES-like signatures decreased in invasive GBM cells, while the remaining GBM states were associated with the infiltrated brain. However, these findings are limited to two patient samples, and a panel covering 960 genes. To further investigate infiltrated areas containing tumor cells, we identified infiltrative sections from a whole-transcriptome spatial transcriptomics dataset with multicellular resolution (14). We identified one patient with 3 distinct regions (core = high CNV alterations, invasive = intermediate CNV alterations, periphery = low CNV alterations) (Fig. 3E). When testing between the infiltrated and tumor core areas, we identified a large number of differentially expressed genes, although no malignant state signatures were significantly enriched in either area (Fig. 3F-G). Conversely, neuronal cell signatures were upregulated in infiltrated tissue (Fig. 3H). However, as GBM cells share transcriptional programs with resident brain cells, it is difficult to distinguish between which differentially expressed genes are derived from malignant and non-malignant populations."

Fig. 5A-D



Text (Lines 274 - 305)

"Principal component analysis showed two main trajectories in how p53 segments changed from the tumor core to infiltrated brain tissue. Most tumors (n = 5) followed a mesenchymal to proneural transition, while the two most proneural tumors became more neural-like (Fig. S4D). These results illustrate that GBM invasion is, again, heterogeneous across tumors and likely is influenced by genomics and the microenvironments inside and outside of the tumor. As five of the tumors followed a similar pattern, we continued to analyze these tumors to test which gene modules become upregulated in this infiltrative pattern. As we sought to identify regulatory gene modules associated with brain invasion, we identified co-expressed gene modules within our dataset and tested whether they were differentially expressed between p53 core and brain infiltrated AOIs (Fig. 5A, S5A). To annotate these gene modules based on previous literature, we performed gene set enrichment analysis to test which ontologies were overrepresented in each gene module (Fig. 5B). While most gene modules upregulated in the infiltrated brain tissue were not significantly enriched with a known ontology, three modules were enriched with ontologies associated with synaptic vesicles, glial cell differentiation, oligodendrocyte specification and differentiation, notch signaling and neural crest differentiation.

To address how our gene modules are expressed and co-expressed across many patient samples, we correlated each module's expression score in TCGA data (Fig. 5C). Hierarchical clustering identified 6 module clusters which consisted exclusively of either gene modules that were upregulated in either the tumor core or infiltrated brain tissue. To inspect which cell types expressed these modules, we aggregated the expression of GBMap across all cell types or cell states. Tumor core modules were mostly expressed by the mesenchymal malignant states, endothelial cells and pericytes (annotated mural cells

in GBMap), and myeloid cells (monocytes, dendritic cells, TAMs). Conversely, modules significantly expressed in the infiltrated tissue regions were most expressed in NPC- and OPC-like malignant states, and resident brain cells (astrocytes, neurons, oligodendrocytes, OPCs). However, the GBMap atlas is generated primarily from resected patient tumors with little tissue derived from infiltrated brain tissue. To compare our findings with those from a previous study using single-cell RNA sequencing on biopsies from the core and infiltrated brain tissue (11), we downloaded and processed their data using standard processing pipelines (40). Using copy-number analysis to identify malignant cells, we found two patient samples containing malignant cells from both the core and infiltrated brain tissue, and differential expression analysis was performed for each patient for malignant cells residing in either the tumor or infiltrated brain tissue (Fig. 5D, S5B). The direction of regulation of each significant gene module in our dataset was almost perfectly in alignment with those found in the single-cell dataset, supporting that the upregulated genes in the infiltrated brain tissue were not only expressed by non-malignant resident brain cells."

2. Furthermore, the authors performed some excellent bioinformatics analysis of spatial transcriptomics data but did not provide a clear biological context for these observations. For instance, one of the essential findings is that the Notch signalling pathway is more active in the tumour periphery. However, it is not clear what is the biological relevance of this observation. Does high-notch activity in the periphery influence the immune landscape? If yes, How? Does the high-notch activity make malignant cells in the periphery more EMT-like? Can the role/function of the peripheral notch be tested by blocking notch activity?

We thank the reviewer for the kind words regarding our bioinformatics analysis. We agree that our findings could be put more into biological context. We have attempted to emphasize the biological context of major findings. We discuss how MES-like cells may both initiate angiogenesis in hypoxic tumors and further facilitate vascular development and homeostasis around areas of microvascular proliferation (Fig. 2). In Figure 4, we identify how glioblastoma subtypes differ in expression within tumors, while the major transcriptional differences in the infiltrated brain are those associated with genomic alterations and not from microenvironmentally induced genes (Fig 4). Moreover, we identify OS-9, a protein that regulates the cellular response to hypoxia and is encoded immediately adjacent to CDK4 on chromosome 12. As the major differences in gene expression between proneural and mesenchymal subtypes are associated with hypoxia, which may lead to immune infiltration and microvascular proliferation, some of the transcriptional differences may lie in OS-9 being a passenger on CDK4 amplification.

In Figure 5, we identify genes associated with both cell differentiation and notch signaling to be upregulated in the infiltrated brain tissue. Notch signaling is highly associated with the NPC-like malignant state, and these findings highlights the infiltrative capacity as a malignant state. We have observed that notch signaling receptors and ligands are expressed across multiple cell states (NOTCH1 is expressed in

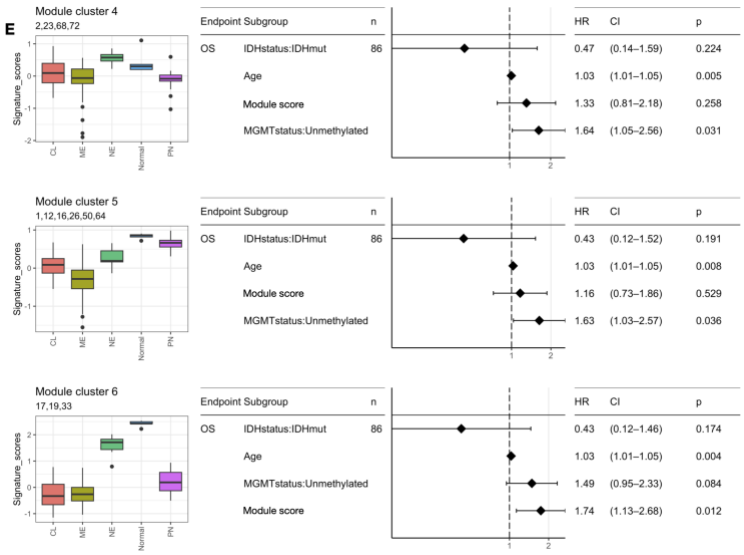
AC-like cells; NOTCH2 in OPC/NPC-like cells; DLL1 and DLL3 is expressed in NPC-like cells; DLL4 is expressed by vascular cells). Due to the complexity of this signaling pathway, we do not think that we can substantially assess it's role in brain infiltration and the effect on the tumor microenvironment – in this study. However, we fully agree that future studies investigating how notch signaling plays a role in brain infiltration and modulating immune responses is an exciting research avenue. Lastly, we do test which modules are associated with poor patient prognosis and find that a module associated with cell differentiation and not notch signaling was prognostic of poor patient survival.

3. Moreover, several published single-cell and bulk RNA-seq studies on glioblastoma exist. Still, the authors did not use their CosMx data effectively to correlate their findings with existing literature and the broader clinical relevance of their observations. This work is highly descriptive and has been obtained on very few samples, making conclusions unpredictable.

Thank you for this comment. Very few spatial datasets have sufficient resolution to investigate malignant cells in the infiltrated brain tissue, but we have now tried to use these datasets whenever possible. We have addressed the validation of our findings in external datasets in comment #1.

We thank the reviewer for suggesting assessing the clinical significance of our findings. We have therefore investigated whether the findings in Fig. 5 were prognostic of poor patient survival. We found that a module cluster, enriched for modules associated with cell differentiation, was prognostic of patient survival.

Fig 5E



Text (Lines 306 –312)
"Finally, to investigate whether our invasive gene module clusters were associated with patient outcomes, we used a Cox regression model to test whether they were significantly associated with survival. Using gender, age, MGMT- and IDH-status as

covariates (Fig. 5E), we found module cluster 6 was significantly associated with poor survival. Interestingly, this module cluster was most expressed in healthy brain tissue. Similar findings of a neural phenotype have recently been shown to be associated with poor patient outcomes (41), and our findings indicate that this gene module is found to a higher degree in infiltrating malignant cells."

Reviewer #3 (Remarks to the Author): Expert in glioma genomics, tumour microenvironment, and spatial omics

In this study, Harwood et al utilize spatial transcriptomic approaches to compare the expression profiles of malignant cells in tumor core regions to those in the tumor periphery. They find that malignant cells in the tumor cores of mesenchymal tumors are different than those of proneural tumors, but these differences do not expand to the tumor periphery. They also show that malignant cells in the core tend to be more mesenchymal in nature, while those from the periphery tend to be more stem-like or neuronal. The analyses presented in this manuscript are interesting and make good use of the spatial data generated. However, the results are descriptive in nature and their applicability is hindered by a limited sample size, a lack of statistical testing, and ambiguity in some of the methods.

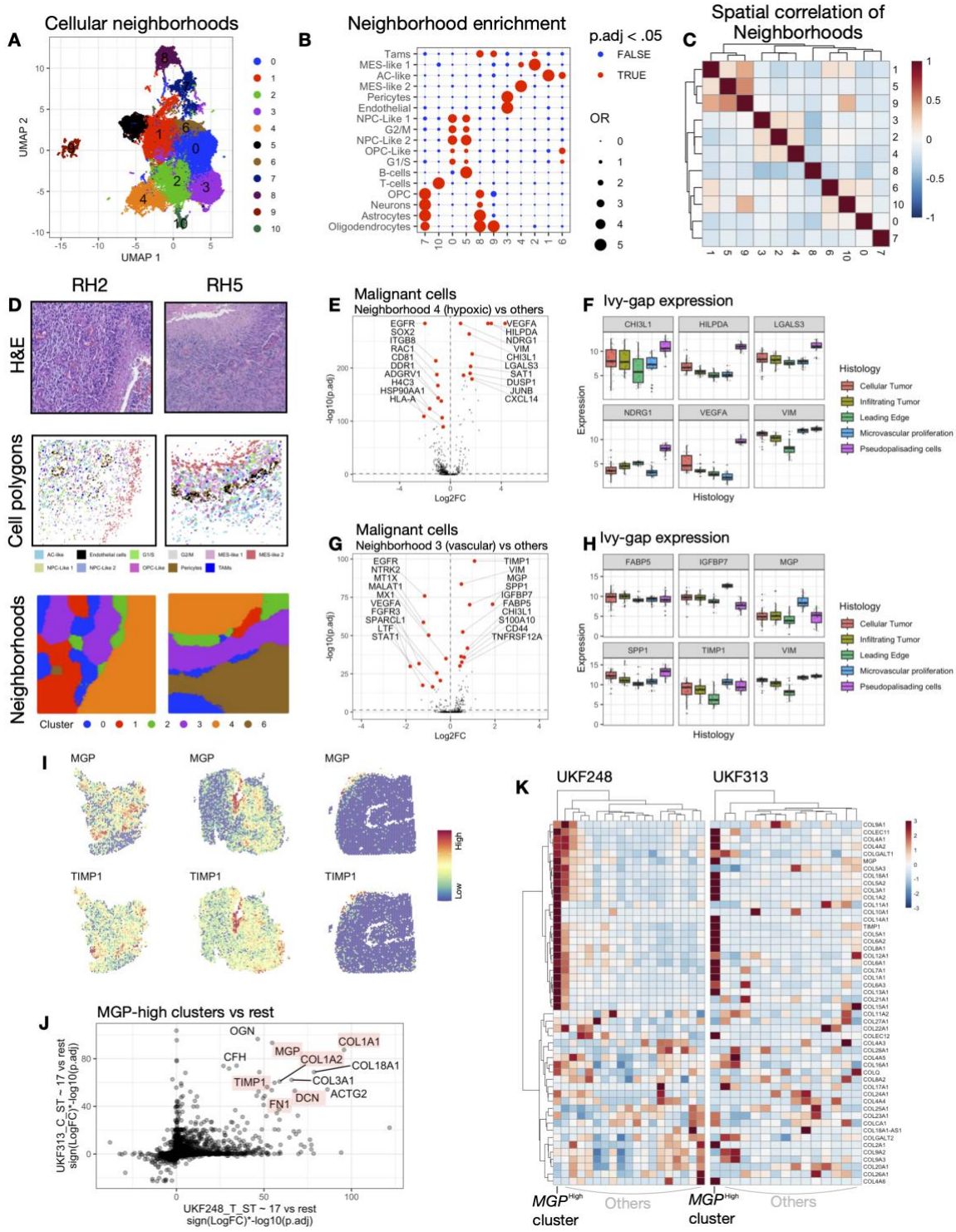
We thank the reviewer for highlighting that our results are interesting, and for the useful comments below. We agree with all of the comments and address them specifically below.

Major Comments
1. The analyses described in Figure 2 (lines 125-136) need statistics in order to be properly evaluated. While the increase in MES-like 2 near necrosis is visually striking in the presented image, proper metrics that compare the distribution of MES-like 2 cells to other cell types will allow the readers to determine whether this behavior is cell type specific. Similarly, the differential expression analyses presented in this section should be presented with p-values adjusted for multiple testing as opposed to the unadjusted p-values they are currently presented with.

Thank you for pointing this out. We have changed our analysis, so that we identify neighborhoods across all patient samples by identifying the 50 nearest neighbors to each cell and clustered cells by counting their neighbors cell labels. We could then cluster neighborhoods and use Fisher's exact test to assess which cell types and states were significantly enriched in a neighborhood. Moreover, we identify multiple genes being overexpressed in malignant cells that are spatially located in areas with high amounts of necrosis or microvascular proliferation (Fig. 2D,E,G). We show that these genes are also upregulated in similar areas in the Ivy-gap dataset, which does not have single-cell resolution (Fig. 2F,H). Moreover, some of these genes that are more specifically expressed in areas of MVP were further found to be associated in areas with high collagen deposition in a Visium GBM dataset (Fig. 2I-K).

Below are the relevant figures and rewritten text.

Fig. 2



Text (Lines 122-127)

"To gain a global overview of different cellular neighborhoods, defined by areas with different cellular compositions, we counted cell annotations of the 50 nearest neighbors

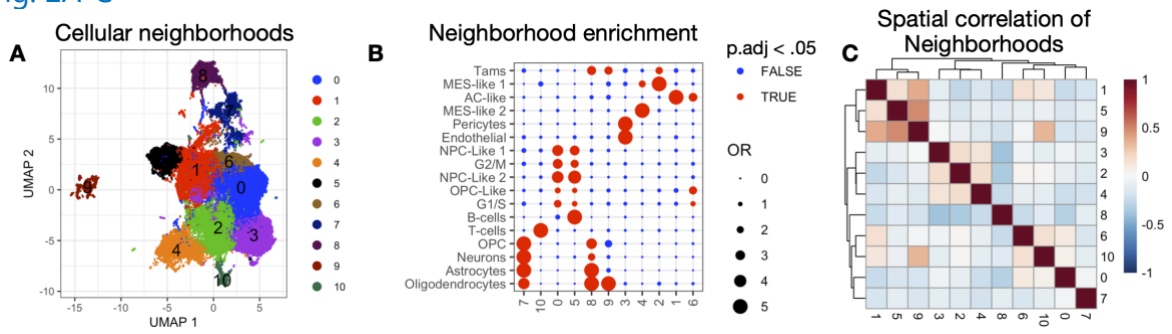
to each cell, creating an annotation by cell count matrix. Dimension reduction and clustering were performed to identify cellular neighborhoods (Fig. 2A) and results were visually assessed using local transcriptional neighborhoods (Fig. S2A-C) (see methods) and adjacent sections stained with H&E."

2. It's unclear how "histological neighborhoods" were defined. The methods describes cellular neighborhoods, which are defined as the 100 cells closest to a given cell type but does not describe how this approach was extended to histological entities. The histological neighborhoods depicted in Figure 2F do not seem to correspond to the H&Es. For instance, regions of MVP in the H&E slide seem to frequently be in neighborhoods of cells that are labelled as cellular tumor. Collectively, this makes it difficult to interpret the conclusions reached in Lines 137-153.

We thank the reviewer for these concerns which were also raised by reviewer #1.
We have gone into more detail of how we have performed the spatial neighborhood analysis. Briefly, we identified the 50 nearest neighbors to each cell and clustered cells by counting their neighbors cell labels. Neighborhoods were compared to adjacent H&E stained sections and enriched for cell types for annotation. Furthermore, we thank the reviewer for pointing out how we were not clear regarding the use of H&E sections. These sections were performed on adjacent tissue sections and are therefore not 100% overlapping with those found in the CosMx analysis. We have clarified this in the manuscript in line 124-126, where it states: "… Dimension reduction and clustering were performed to identify cellular neighborhoods (Fig. 2A) and results were visually assessed using local transcriptional neighborhoods (Fig. S2A-C) (see methods) and **adjacent** sections stained with H&E."

Below are the relevant figures and rewritten text.

Fig. 2A-C



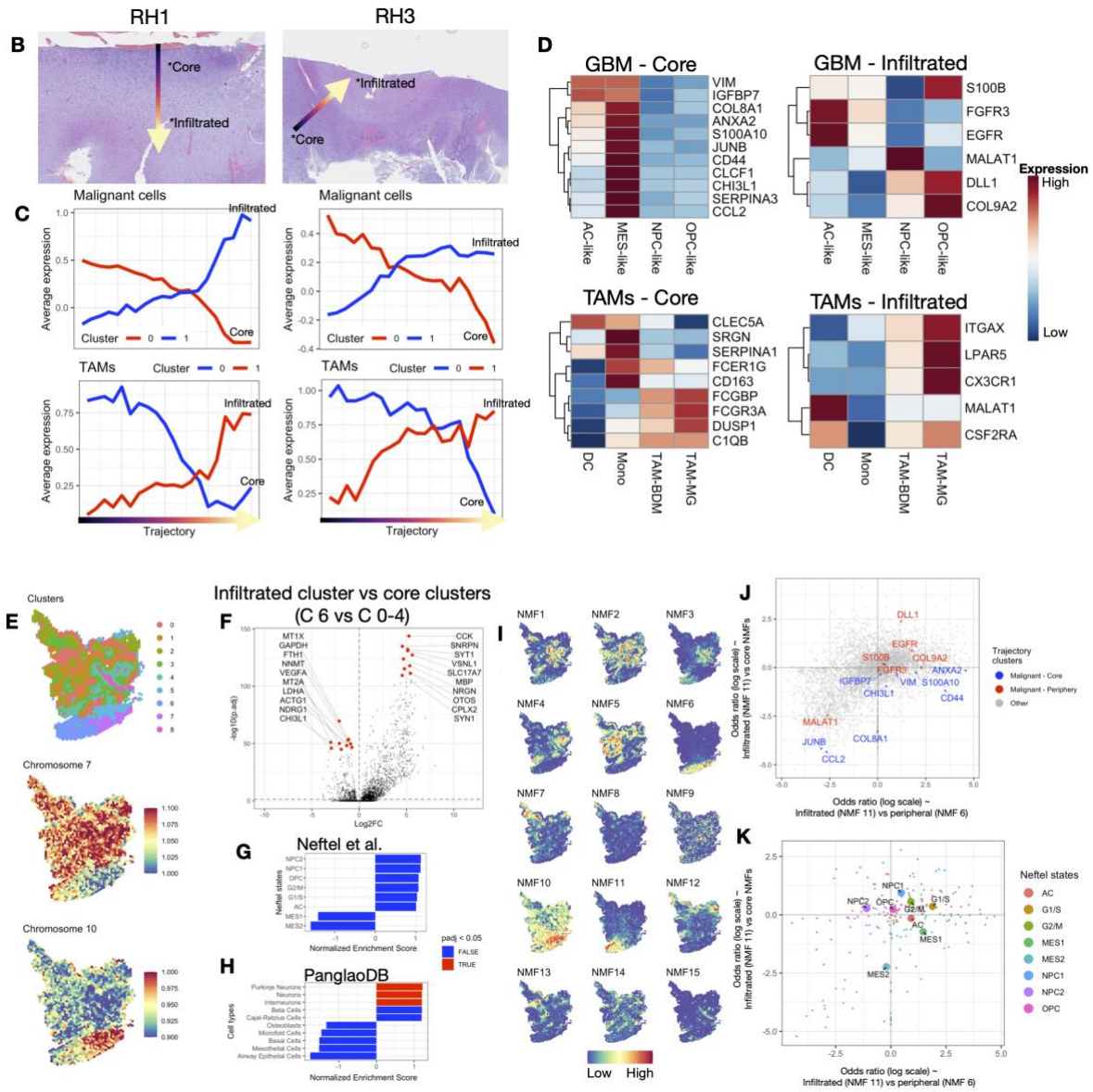
Text (Lines 122-126)
"To gain a global overview of different cellular neighborhoods, defined by areas with different cellular compositions, we counted cell annotations of the 50 nearest neighbors to each cell, creating an annotation by cell count matrix. Dimension reduction and clustering were performed to identify cellular neighborhoods (Fig. 2A) and results were

visually assessed using local transcriptional neighborhoods (Fig. S2A-C) (see methods) and adjacent sections stained with H&E."

3. Similar to comment 1, the analyses described in Figure 3 (lines 161-197) are interesting and produce intriguing qualitative results. However, without proper statistics, including p-values that compare cell abundance and gene expression across regions, it is difficult to evaluate how reproducible and meaningful these results are.

We thank and agree with the reviewer, that our statistics are more qualitative than quantitative. We have expanded the analysis to another patient in the CosMx data (Fig. 3B-D), and an alternative approach to quantify malignant states in infiltrated tissue from 10x Visium data (Fig. 3E-K). However, the limited number of samples and spread across technologies (CosMx = 2, Visium = 1) and heterogeneity of spatial trajectories in glioblastoma data does make it difficult to consider our results more than explorative. Therefore, we continued this analysis using the GeoMx platform, where the majority of these findings are replicated in Fig. 4I, where we perform statistics on 7 glioblastoma patients where we use FDR corrected p values. Briefly, all of the analyses indicate MES-like gene markers to be associated with the tumor core, and genes associated with the NPC-, OPC, and AC-like cell states to be always or in part associated with the infiltrated brain tissue.

Below are the relevant figures and rewritten text.
Fig 3B-K



Text (Lines 176 – 214)

"Two patient samples had intact trajectories between the core and infiltrated brain tissue (Fig. 3B, S3A-B). To enable the comparison of trajectories between samples, we identified spatially variant gene clusters with a low clustering resolution for the purpose of between-patient comparison (Fig. 3C). In malignant cells and myeloid cell populations, we identified genes with spatial variance that were unique and common between samples (Fig. 3D, S3C). To assess which cells and malignant states expressed these genes, we aggregated gene expression for each cell type or state in GBMap (32) (Fig. 3D). Genes commonly expressed in the tumor core were highly expressed in MES-like GBM states. Amongst these genes was CHI3L1, a gene recently associated with connected and non-invasive malignant cells (37). The few genes upregulated in invasive cells were most associated with the AC-like state and growth factor receptors (EGFR, FGFR3), or NPC-like/OPC-like states, indicating their increased presence in malignant cell composition outside the tumor. TAM genes upregulated in the core were expressed by both blood-

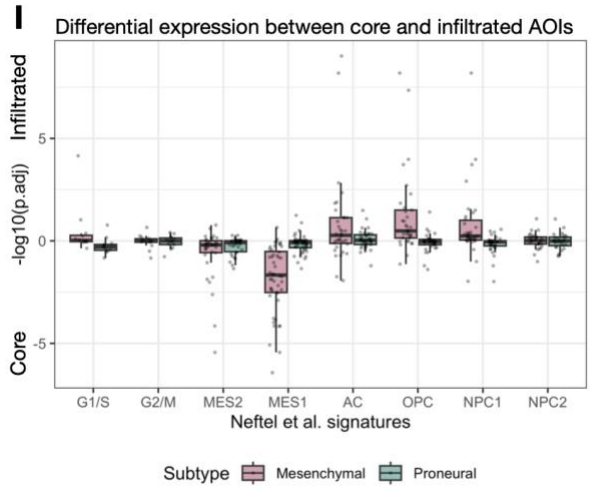
derived macrophages and microglia, while those in the infiltrated area were, as expected, most expressed in microglia.

Our findings indicated that MES-like signatures decreased in invasive GBM cells, while the remaining GBM states were associated with the infiltrated brain. However, these findings are limited to two patient samples, and a panel covering 960 genes. To further investigate infiltrated areas containing tumor cells, we identified infiltrative sections from a whole-transcriptome spatial transcriptomics dataset with multicellular resolution (14). We identified one patient with 3 distinct regions (core = high CNV alterations, invasive = intermediate CNV alterations, periphery = low CNV alterations) (Fig. 3E). When testing between the infiltrated and tumor core areas, we identified a large number of differentially expressed genes, although no malignant state signatures were significantly enriched in either area (Fig. 3F-G). Conversely, neuronal cell signatures were upregulated in infiltrated tissue (Fig. 3H). However, as GBM cells share transcriptional programs with resident brain cells, it is difficult to distinguish between which differentially expressed genes are derived from malignant and non-malignant populations.

To circumvent the limitations arising with multicellular resolution, we used non-negative matrix factorization (NMF) to deconvolute the count matrix to 15 individual factors (Fig. S3D). This identified NMF factors that were primarily expressed within the tumor core, the infiltrated area, and the tumor periphery (Fig. 3I). To identify genes that were more expressed in the infiltrating malignant cells, but not the cells of the healthy brain, we calculated the odds ratio of a gene being expressed in the factor associated with the infiltrated area compared with it being expressed in the peripheral factor (normal brain). We repeated this comparison with the factors found within the tumor core and were able to identify which genes were highly expressed in the infiltrated brain tissue compared to the tumor periphery (Fig. 3J, right vs left), and the infiltrated brain tissue compared to the tumor core (Fig. 3J, top vs bottom). This showed similar patterns to our previous findings with Delta Like Canonical Notch Ligand 1 (DLL1) being the gene most upregulated in the infiltrated area. DLL1 is also used as a NPC-like 1 marker gene, and to further investigate how genes associated with malignant states were associated with this invasive signature, we averaged the odds ratio of each malignant state from Neftel et al., and found the NPC-like 1 to be most associated with infiltrated brain tissue, with mesenchymal states being most present in the tumor core (Fig. 3K)."

4. The comparisons in Figure 4F need p-values in order to be interpreted properly.

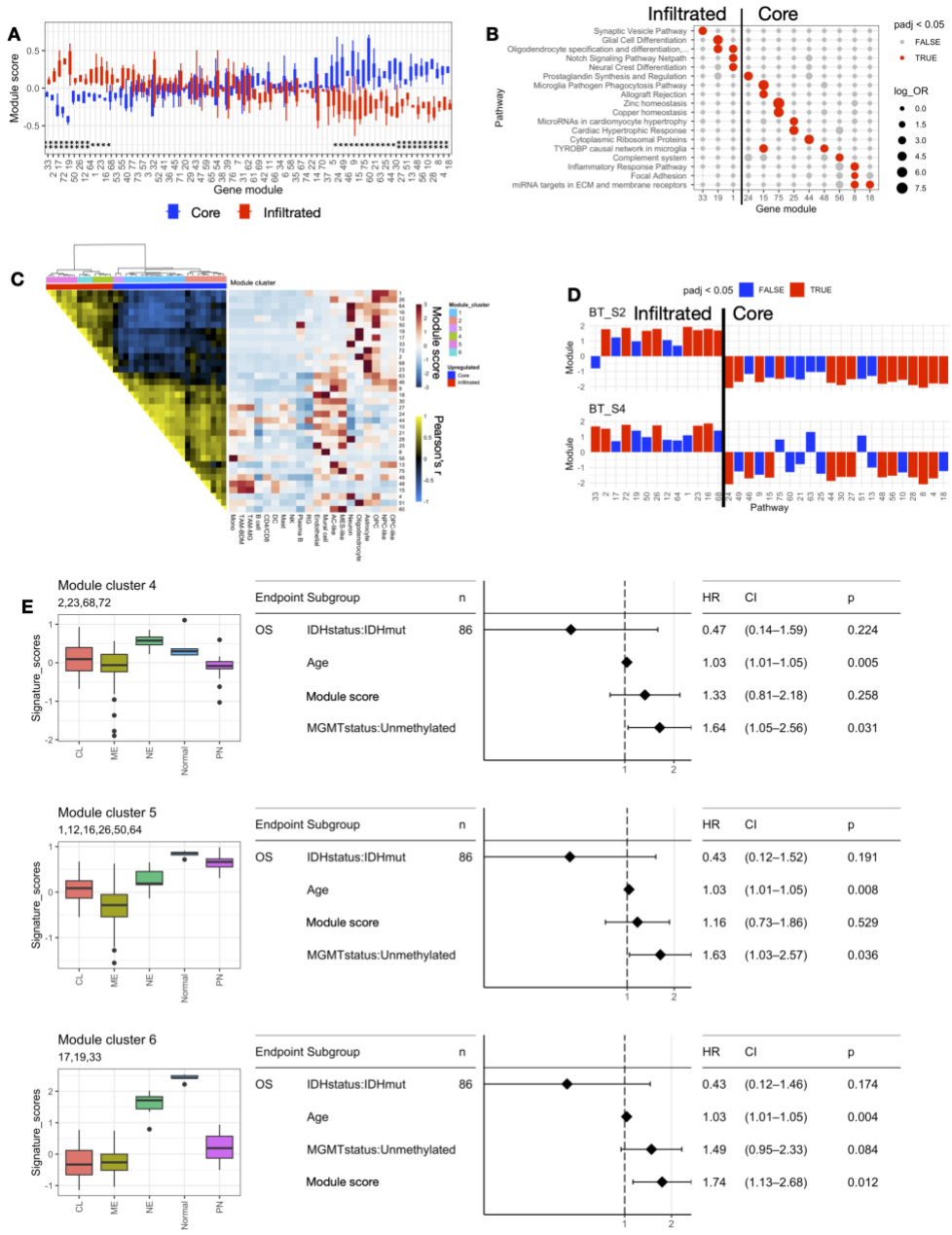
Thank you for pointing this out. We have now adjusted this analysis to include p.adj values instead of LogFC values.



5. It is difficult to understand the analyses described in Figure 5. Which single-nucleus RNA dataset did the authors use? The dataset provided in the methods is a single-cell RNAseq dataset of cultured glioma stem cells rather than a single-nucleus RNAseq dataset containing a mix of malignant and resident brain cells. The authors refer to this approach as deconvolution, but it is not clear what deconvolution is being done. Instead, the authors seem to have clustered genes based on single-cell expression and then selected the gene clusters that were enriched in tumor cells. Once this is done, they then use these clusters to interpret their results from a differential expression analysis of periphery vs core-associated malignant cells, though its not clear in the text what this annotation adds. The authors should provide additional details on what hypotheses they were testing here and how it relates to their previous spatial analyses.

We thank the reviewer for the critical concerns in figure 5. We have now improved the processing and analysis of GeoMx data (normalization methods and QC filters). Our hypothesis was to identify gene modules that were differentially expressed in the infiltrated brain compared to the tumor core, and to establish how these modules are related to patient outcomes. In the new approach, we identify differentially expressed gene modules in GeoMx data to identify phenotypes or pathways that are associated with brain infiltration (Fig. 5A). To further characterize these modules we perform gene set enrichment analysis (Fig. 5B). To assess how these modules were expressed in a larger cohort of patients, we correlated significant modules across TCGA bulk RNA-seq data and their expression in GBMap single-cell mRNA sequencing data (Fig. 5C). These modules were then assessed using a small single-cell RNA sequencing dataset from Darmanis et al, 2017 which included cells of the tumor core and periphery (n = ~50), where we see the same trend for all significant gene modules (Fig. 5D). Using the infiltrated brain tissue clustered modules, one was prognostic of poor patient survival independent of the most importantly known variables in glioblastoma (Age, IDH- and MGMT-status) (Fig. 5E). Together, our results indicate that gene modules associated with Notch signaling are not prognostic of patient survival, while those that are associated with a differentiated phenotype (neuronal) are.

Fig 5



Text (Lines 274 - 317)

"Principal component analysis showed two main trajectories in how p53 segments changed from the tumor core to infiltrated brain tissue. Most tumors (n = 5) followed a mesenchymal to proneural transition, while the two most proneural tumors became more neural-like (Fig. S4D). These results illustrate that GBM invasion is, again, heterogeneous across tumors and likely is influenced by genomics and the microenvironments inside and outside of the tumor. As five of the tumors followed a similar pattern, we continued to analyze these tumors to test which gene modules become upregulated in this infiltrative pattern. As we sought to identify regulatory gene modules associated with brain invasion, we identified co-expressed gene modules within

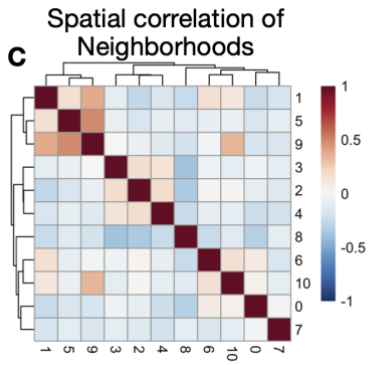
our dataset and tested whether they were differentially expressed between p53 core and brain infiltrated AOIs (Fig. 5A, S5A). To annotate these gene modules based on previous literature, we performed gene set enrichment analysis to test which ontologies were overrepresented in each gene module (Fig. 5B). While most gene modules upregulated in the infiltrated brain tissue were not significantly enriched with a known ontology, three modules were enriched with ontologies associated with synaptic vesicles, glial cell differentiation, oligodendrocyte specification and differentiation, notch signaling and neural crest differentiation. To address how our gene modules are expressed and co-expressed across many patient samples, we correlated each module's expression score in TCGA data (Fig. 5C). Hierarchical clustering identified 6 module clusters which consisted exclusively of either gene modules that were upregulated in either the tumor core or infiltrated brain tissue. To inspect which cell types expressed these modules, we aggregated the expression of GBMap across all cell types or cell states. Tumor core modules were mostly expressed by the mesenchymal malignant states, endothelial cells and pericytes (annotated mural cells in GBMap), and myeloid cells (monocytes, dendritic cells, TAMs). Conversely, modules significantly expressed in the infiltrated tissue regions were most expressed in NPC- and OPC-like malignant states, and resident brain cells (astrocytes, neurons, oligodendrocytes, OPCs). However, the GBMap atlas is generated primarily from resected patient tumors with little tissue derived from infiltrated brain tissue. To compare our findings with those from a previous study using single-cell RNA sequencing on biopsies from the core and infiltrated brain tissue (11), we downloaded and processed their data using standard processing pipelines (40). Using copy-number analysis to identify malignant cells, we found two patient samples containing malignant cells from both the core and infiltrated brain tissue, and differential expression analysis was performed for each patient for malignant cells residing in either the tumor or infiltrated brain tissue (Fig. 5D, S5B). The direction of regulation of each significant gene module in our dataset was almost perfectly in alignment with those found in the single-cell dataset, supporting that the upregulated genes in the infiltrated brain tissue were not only expressed by non-malignant resident brain cells. Finally, to investigate whether our invasive gene module clusters were associated with patient outcomes, we used a Cox regression model to test whether they were significantly associated with survival. Using gender, age, MGMT- and IDH-status as covariates (Fig. 5E), we found module cluster 6 was significantly associated with poor survival. Interestingly, this module cluster was most expressed in healthy brain tissue. Similar findings of a neural phenotype have recently been shown to be associated with poor patient outcomes (41), and our findings indicate that this gene module is found to a higher degree in infiltrating malignant cells. Our findings identify multiple gene modules associated with the tumor core and infiltrated brain tissue. Many modules were not enriched in known pathways, although few were significantly associated with cell fate, synapses, and notch signaling. Our gene modules were similarly correlated across TCGA tumors, and one gene module cluster had highest expression in the healthy brain and was also associated with poor patient survival."

Minor Comments

1. The color scheme in Figure 2B is confusing to read as it is centered at R ~ 0.3. The

authors should rescale this figure so that white, the center of the color scheme, corresponds to R = 0 and the min and max values are -1 and 1, respectively.

We thank the reviewer for these concerns and have corrected all heatmaps to be more centered as shown below.
Fig. 2C



2. Line 139: The method for defining histological neighborhoods is important to understanding this section and should briefly be outlined in the results section.

We thank the reviewer for pointing this out along the other reviewers, and have rewritten the section to be more clear, as expanded below:

Text (Lines 122-126)
"To gain a global overview of different cellular neighborhoods, defined by areas with different cellular compositions, we counted cell annotations of the 50 nearest neighbors to each cell, creating an annotation by cell count matrix. Dimension reduction and clustering were performed to identify cellular neighborhoods (Fig. 2A) and results were visually assessed using local transcriptional neighborhoods (Fig. S2A-C) (see methods) and adjacent sections stained with H&E."

3. The authors should provide more information in the results section on the GeoMX and CosMX methods, as it may not be clear to readers what these assays are measuring and how they differ.

We thank the reviewer for their comment. We have mentioned the technology in the beginning of every section and have written the technologies into the figure legend as well.

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

The authors have properly addressed my concerns.

Reviewer #2 (Remarks to the Author):

After reviewing the authors' revisions, it is evident that they have addressed many of the concerns raised in the initial review. However, despite these revisions, the study still needs validation of their observations on the Notch signalling pathway. The authors state that 'due to the complexity of this pathway, they cannot assess its role in brain infiltration and its effect on the tumour microenvironment'. Nevertheless, without validation, these complex biological pathways remain mere speculations. The authors must take steps to validate their observations to strengthen the study's findings.

Reviewer #3 (Remarks to the Author)

EDITORIAL NOTE: This Reviewer was not able to submit a report. However, Reviewer #1 assessed your response, and considers that most of Reviewer #3's concerns were addressed. Regarding previous point #4, the revised panel was not entirely clear; LogFC and Padj. values need to be shown.

REVIEWERS' COMMENTS
Reviewer #1 (Remarks to the Author):
The authors have properly addressed my concerns.

We thank the reviewer for helping us improve our manuscript.

Reviewer #2 (Remarks to the Author):
After reviewing the authors' revisions, it is evident that they have addressed many of the concerns raised in the initial review. However, despite these revisions, the study still needs validation of their observations on the Notch signalling pathway. The authors state that 'due to the complexity of this pathway, they cannot assess its role in brain infiltration and its effect on the tumour microenvironment'. Nevertheless, without validation, these complex biological pathways remain mere speculations. The authors must take steps to validate their observations to strengthen the study's findings.

We thank the reviewer for helping us improve our manuscript. We agree fully that experimental work is necessary to elucidate the role of notch signaling in glioblastoma. We have revised part of our discussion to emphasize this.
Lines 344-347
Two ligands of the notch signaling pathway (DLL1 and DLL3) are markers of NPC-like 1 cells (9) and are known to play a role in maintaining quiescence and regulating asymmetric mitosis in adult neural stem cells (43), as well as mediating neuronal migration. However, more research is needed to determine whether these ligands directly influence brain infiltration.

Reviewer #3 (Remarks to the Author)
EDITORIAL NOTE: This Reviewer was not able to submit a report. However, Reviewer #1 assessed your response, and considers that most of
Reviewer #3's concerns were addressed. Regarding previous point #4, the revised panel was not entirely clear; LogFC and Padj. values need to be shown.

We thank the reviewer for taking the extra time to help us complete our manuscript. We have clarified Fig 4I in the figure legend:
Lines 819 - 825
”(I) Gene markers found by Neftel et al. and the significance of a gene being expressed in the core or infiltrated regions of a tumor, split into mesenchymal (n = 4) and proneural (n = 3) tumors. -log10(P.adjusted) values are shown on the y-axis. To visualize whether a gene is differentially expressed across malignant programs, we assign a value equal to the sign of the LogFC (1 or -1) multiplied by -log10(p.adj). This approach allows us to distinguish between genes upregulated in the core or infiltrated regions. Genes upregulated in the core tissue appear on the bottom half of the plot, while genes upregulated in the infiltrated tissue appear on the top half.”