

Multidimensional profiling of tumor heterogeneity in supratentorial ependymomas

Corresponding Author: Dr Mariella Filbin

Any redactions in this file are there to maintain patient confidentiality, the confidentiality of unpublished data, or to remove third-party material.

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

Version 0:

Reviewer comments:

Referee #1

(Remarks to the Author)

In this manuscript, Jeong, Danielli, and colleagues examine the developmental roots of supratentorial ependymomas and identify that these tumors harbor two distinct progenitor-like cell populations that differ in their extent of neuron vs. ependymal cell differentiation. The authors use cutting-edge single-cell and spatial transcriptomic techniques, as well as live-cell imaging to define the spatial architecture of these tumors and uncover a role for brain-resident cells in contributing to the heterogeneity of these tumors. Overall, the studies are nicely designed, the data are beautifully presented, and this work addresses an important gap in the field regarding childhood brain cancers. I do have a number of significant critiques that are mostly related to three points: (1) oftentimes the authors have overstated some of their data (including for plots where there is no statistical significance), (2) the in vitro & in vivo work is well-inspired but raises several concerns, outlined below, and (3) there is little discussion on how ST-EPN subtypes correlate to clinical prognoses, which greatly minimizes the ability to assess the impact of the work. Overall, my enthusiasm is mixed—the strength of the manuscript lies in the descriptive figures, and the more mechanistic findings, although interesting, are far less developed.

Major concerns:

1. From Fig. 1F, it appears as though early radial glia do not closely relate to any of the ST-EPN subtypes, but radial glia correlate to ZFTA-Cluster 2, as the authors also state. Apart from developmental time, what are the transcriptomic and functional differences between “early” radial glia vs. radial glia?
2. From Fig. 2D, it appears as though ZFTA-Cluster 2 tumors are mainly composed of neuroepithelial / MES / neuron-like cells, but not radial glia, which is contradictory to the findings presented in Fig. 1F and lines 136-138, wherein the authors show and state that ZFTA-Cluster 2 tumors correlate with radial glia (in addition to neuroepithelial cell types) from the reference atlases. Similarly, Fig. 2D suggests that ZFTA-Cluster 3 and 4 tumors are high in embryonic neuron-like populations, but data in Fig. 1F (and lines 138-139) suggests that these tumors are mainly mesenchymal. How do the authors explain these differences?
3. It is interesting that despite harboring the same ZFTA-RELA fusions, the ZFTA-Cluster 3 and ZFTA-RELA tumors have distinct cell populations. Could this be a result of 7 out of the 8 ZFTA-Cluster 3 tumors being recurrent, whereas the distribution of primary vs. recurrent tumors was more or less equal in the ZFTA-RELA tumors? Alternatively, could the differences be microenvironment-mediated?
4. This study raises several important questions about the main finding that there are two main progenitor-like cell states in ST-EPNs. For the two progenitor-like cell states (neuroepithelial vs. embryonic-like), can the authors elaborate on what is known about their role, function, and presence during normal fetal development? What are the differentiation trajectories of these cells? How do they differ from each other at the transcriptomic and functional levels? Is it known whether (and how) they give rise to the two mature lineages identified in this work (neuronal vs. ependymal)? Some of these would be important to examine to make the earlier half of the manuscript less “descriptive”.

5. The relevance of data presented in Ext. Fig. 3 and Table S6 to this manuscript is unclear. Furthermore, the authors state in lines 173-175 that therapeutic targets “in the context of metaprograms correlated with lower survival” were uncovered. This raises the question of what metaprograms, of the 8 that were identified, were the ones correlated with lower survival? Although this latter statement is important from a clinical perspective, it was not explained in detail or substantiated with experimental work.

6. Please provide correlational plots between the spatial coherence score and frequency of each of the other malignant cell states/types (and other non-malignant cell states, too), as was done for the MES/hypoxia cell states and vascular cells in Fig. 3C.

7. For these correlational plots (including Fig. 3C and Ext. Fig. 6E), please also display the “r” value to show the strength and directionality of the correlation, which is not fully captured by the “r²” value.

8. Related to the previous critiques – although the authors nicely show that a higher spatial coherence score (i.e., tendency to be a “structured” tumor) is correlated with a higher frequency of MES/hypoxia cell states, it is important to note that MES/hypoxia cell states are not the dominant cell states in any of the structured tumors. In fact, MES/hypoxic cell states comprise only ~6% of the tumor with the second highest spatial coherence score (STEPN-01). Reporting the “r” value as noted in #7, above, would help clarify whether or not the MES/hypoxic cells are strongly correlated with a tendency to be a more structured tumor, although it appears that this link may be overstated and likely driven by a single sample (STEPN-19). However, there may be other cell states that are more likely to be correlated with being a structured tumor, in which case the authors should show the plots, r-values, and p-values correlating these other cell states and the coherence score, as noted in #6 and #7, above.

9. The correlational plot between vascular cells and the spatial coherence score in Fig. 3C may show an inverse directionality, but the p-value is highly non-significant, so the authors should not over-interpret these findings – particularly in the absence of r-values to depict the strength of the link. Please temper the language in the Results section accordingly.

10. Ext. Fig. 10E: How are neuroepithelial-like 1 different from neuroepithelial-like 2, at the transcriptomic and functional levels? It appears that there is a sharp decline in neuroepithelial-like 1 cells in coculture, but a concomitant increase in neuroepithelial-like 2 (in addition to increases in EPN-glia and neuronal-like cells).

11. Although the coculture data are quite interesting, the authors performed cocultures of patient-derived tumor cells with rat neurons/astrocytes, thereby (potentially) skipping species-specific differences in microenvironmental cues. Ideally these experiments could be performed with human iPSC-based neurons/astrocytes.

12. In Ext. Fig. 10F-G, can the authors include a fourth group—a coculture group—using the same cell lines that were included as adherent cultures, sphere cultures, or PDX models? This would help understand which of the cell state transitory effects are attributed to growing cells as 3D spheres vs. 2D (as adherent monocultures vs. cocultures) vs. in PDX models, as each of these four groups are very different in the culture conditions and microenvironmental cues.

13. Did the authors confirm that the neuron-like enrichment of cells in coculture are true “malignant” cells, as FACS may give a false td-tomato-positive population? A CNV plot, as was shown for the earlier figures, could also be included for the coculture scRNA-seq analyses.

14. The authors state in line 342-343 that normal brain-resident cells induce a “shift” of neuroepithelial-like progenitor cells toward neuron-like lineages, but instead of this being a plasticity effect, could it be attributed to cell stress/death of progenitor cells, or perhaps, a differentiation effect, wherein the progenitor cells simply differentiate to neuron-like cells?

15. In lines 348-349, the authors state that the “overall invasion speed of ZFTA-RELA cells significantly increased due to a higher fraction of neuronal-like cell states in the coculture model”, but this may be an overstatement, given that the experiments and observations are merely correlative and not causal: 1) Neuroepithelial-like 2 cells are also increased on coculture, but it is unclear whether this subtype of neuroepithelial cells also show higher invasion speed. After all, data in Fig. 5C suggest that there may be a bimodal distribution of cells in the neuroepithelial group as there are some cells that are equally mobile as neuron-like cells, 2) some of the higher somatokinesis noted in Fig. 5k could be attributed to a sharp decline in neuroepithelial-like 1 cells, which may be the slower moving cells from Fig. 5C. One way to answer some of these caveats would be to repeat the quantification of Fig. 5K as was done in Fig. 5C and contrast the invasion speed of neuroepithelial vs. EPN-glia vs. neuron-like cells in monoculture vs. coculture (analyzed by a two-way ANOVA to control for multiple variables).

Minor concerns:

1. Lines 73-75: The authors nicely state that it is unknown whether ST-EPN subtypes have distinct cells-of-origin and malignant cell states and whether these differences contribute to varying outcomes and therapeutic resistance. However, the authors should include some discussion of what those differences in outcomes are/can be. For example, do some ST-EPN subtypes show better clinical prognoses? Do some ST-EPN subtypes show lesser therapeutic defiance?

2. It is unclear why the UMAPs in Ext. Fig. 1C vs. main Fig. 1C look different, given that these are the same samples and cells across both figures? I'd recommend keeping the same UMAP, as a comparison with Fig. 2C would also help clarify which ST-EPN subtype correlates with which clusters.

3. I recommend putting appropriate citations for the reference atlases within the relevant figures (i.e., for example, name reference #29 and #30 next to Fig. 1D and Ext. Fig. 1D), as the authors projected their data onto these atlases.

4. Line 148 should be "extent" and not "extend"

5. Can the authors find a different way to represent the data in Fig. 1D, because some small populations (e.g.: choroid) are hard to see and track over time? One way to do this is to modify main Fig. 1F, such that they could add another panel that reflects how non-malignant expression programs (reference) change over developmental time as a dotplot.

6. The authors could assess if data in Ext. Fig. 6F-G are non-normally distributed, and if so, use a Mann-Whitney test instead of the t-test. Furthermore, although the authors state the high variability in these datasets, the Results section that corresponds to the data on primary vs. recurrent tumors may be overstated, given the lack of statistical significance (or even a statistical "trend"). The authors should acknowledge this caveat in the Results/Discussion section more clearly. In addition, without additional data mining and experimentation, it is unclear if the small increase in MES/hypoxic cells in recurrent tumors has any biological significance, given that the proportion of this cell state is incredibly small (<10% in primary tumors and ~15% in recurrent tumors).

7. Ext. Fig. 7C and main Fig. 4C could use some more details in the legends. What do the rows vs. columns indicate?

8. In Ext. Fig. 8A, all the panels are at differing magnifications, making it difficult to fully appreciate the staining patterns. This may add to some discrepancies that are evident between the staining patterns shown in Ext. Fig. 8A and the mean intensity heat map in main Fig. 5B: From Ext. Fig. 8A, it seems that vimentin is expressed by all 3 cell types, but from Fig. 5B, it seems as though only neuroepithelial cells express this marker.

9. Related to Ext. Fig. 8D: Can the authors show a supplemental table listing the "top 20 marker genes of neuroepithelial-like cells (n = 16 cells, left panel), and neuronal-like cells (n = 16 cells, right panel)" from the co-culture groups?

(Remarks on code availability)

Code provided is acceptable for reproducing main findings from the paper. README file is quite sparse and I did not take time to install and run the code.

Referee #2

(Remarks to the Author)

In this study, Jeong and colleagues study the transcriptional signatures of supratentorial ependymomas. Using integrated single cell- and spatial transcriptomics they refine our knowledge of the malignant cellular subpopulations and uncover that the spatial organization of the disease is driven – similar to gliomas – by regions of hypoxia inversely correlated with vascularity and positively correlated with recurrence/aggressiveness. They provide functional validation of individual cell states that were identified by transcriptional profiling, demonstrating distinct morphological characteristics of the malignant subcellular compartments. Finally, they confirm that the neural microenvironment supports the cellular heterogeneity to faithfully recapitulate the disease and demonstrate that microenvironmental neural cells promote a more mature and motile tumor cell state.

The comprehensive evaluation of the cellular composition of ST-EPN in this manuscript is a valuable contribution to the field and contributes to the emerging field of cancer neuroscience in important ways. Critically, the authors highlight two distinct embryonic progenitor populations present in STN-EPN tumors, opening new avenues for future investigation. The study is further strengthened by the functional characterization of the malignant cell states, bringing in the importance of the neural microenvironment to tumor progression.

I have the following recommendations for improvement:

1. The manuscript appears to be written as two separate stories and the authors should aim to address this in the storytelling. Whilst neural elements of the tumor microenvironment are known to play a key role in brain tumor progression, this study has limited investigation into the contribution of the neuronal and glial non-malignant cells in the EPN tumor microenvironment prior to the functional data presented in Figure 5. The storytelling transition from Figure 1-4 to Figure 5 should be improved. For example, the glial-like tumor population does not appear as a meta program (or part of a meta program) and is therefore absent from analyses in Figure 2-4, yet this population is unexpectedly investigated in Fig. 5. To potentially address this, the authors could perform some spatial imaging confocal microscopy using neuronal (and axonal) and/or astrocyte markers to describe their contribution to the tumor composition.

2. The authors state that they have uncovered potential therapeutic targets of relevance in the context of metaprograms correlated with lower survival (line 174, 175). The Gojo et al. study demonstrated that the high ST-Ependymal-like metaprogram have a favorable prognosis, yet it did not present the Neuronal-precursor-like group (as performed in PF-A EPN). Here, the authors should analyze the prognosis of EPN according to the metaprograms they have identified, in

addition they should present the survival of each molecular subgroup in this patient cohort.

3. In Fig. 1f, it appears that the ZFTA-RELA group lacks any projection onto the Choroid cell type, whereas in Fig. 2d ZFTA-RELA expresses a significant proportion of Ependymal-like cells. Whereas the ZFTA-CI.1 group appears to identify this choroid projection and Ependymal-like cell population in both types of analyses. The authors should provide a possible explanation for this result.

4. In Extended Fig. Data 4c, it appears that ZFTA-RELA samples from the external patient cohort also appear to map onto both the neuronal and ependymal-like signature. This should be reflected in the text (line 201). In addition, in Extended Fig. Data 4c, it would be helpful to add a plot of ZFTA-RELA only with the neuronal- and ependymal-like gene expression module scores.

5. It would be informative to analyze the proliferative capacity of each metaprogram in the patient data.

6. In Figure 5a, the grouped classification of ependymal- and glial-like is confusing. The morphology and markers of the EPN-glial-like group suggest a more astrocytic-like cell, but this morphology is not convincing of an ependymal-like cell. A clearer rationale for this grouping should be provided. The 10X Xenium analysis and pearson correlation of the EPN-glial-like group should be included in Extended Data Fig. 8d, to support the groupings. Furthermore, are the processes in the neuronal-like cells TMs, or more axon-like neurites/processes? They are certainly thinner than the more TM-like processes in the other cell states....

7. For data shown in Figure 5b, please include the methodology for quantifying the staining intensity in the methods.

8. In Extended data Fig 8a it would be clearer to label the cells of interest in each image. The strong positive vimentin staining in the EPN-glial-like and Neuronal-like groups make it hard to identify the cells of interest. In Fig 5b, the Neuronal-like group is depicted as having no vimentin positivity, yet the representative image in Ext Fig 8a does not clearly reflect this analysis, the authors should replace with better and clearer representative images. Representative images for the Beta-III-tubulin and GFAP quantification should also be included.

9. Harvesting cortical neurons produces a mixed population of glutamatergic and GABAergic neurons in culture; the authors should quantify the different neuronal types (glutamatergic vs GABAergic) neurons present in the cocultures, and the possible presence of astrocytes.

10. In support of the data shown in Figure 5b, the expression level of the markers (Vimentin, Nestin, S100B, GFAP, B-tubulin, DCX) in each metaprogram should be included.

11. The experimental data presented in this manuscript beautifully depict the motility and migratory characteristics of the tumor cells, however they do not address the invasive capacity of these cells. The invasiveness of a cancer cell is measured by the cells' ability to restructure, breakdown and move through a 3D barrier/extracellular matrix. The authors should either employ 3D in vitro invasion assays in their study on the three cellular states depicted in Figure 5, or replace any "invasiveness/invasive state/invasion pattern" text with terms such as "migration", "motility" and "movement" (including the label on Figure 5g) when discussing in vitro 2D data.

12. Given the differences in cell state between in vitro and in vivo, measuring and comparing the rate/distance of protrusion, retraction and branching in the neuronal-like cells in in vivo (Figure 5f) and in vitro (Extended Data Figure 9) could further strengthen the observations that the neural microenvironment supports the neuronal-like population.

13. In line 331, the authors state their data strengthen the concept of cooperativity of cell states for tumor progression. Whilst cooperativity is an established concept in cancer biology, the data here does not confirm this in EPN. In order to assess the importance of cooperativity for tumorigenicity in EPN, the authors should sort the individual cellular compartments and inject orthotopically to compare bulk tumor to individual subcellular population growth. Alternatively, the authors can remove the concluding remark in the results.

14. The authors identify that the EPN-glial-like and neuronal-like cells in coculture are less proliferative than the neuroepithelial-like cells, however they also demonstrate there are far less of the EPN-glial-like and Neuronal-like cells in coculture compared to PDX. The authors should confirm the proliferation rate of each cell state in vivo, given the PDX provides a more even distribution of each cellular subpopulation.

15. Transition toward neuronal-like lineages: The authors suggest that normal brain cells induce the transition from neuroepithelial-like progenitor cells toward neuronal-like lineages (lines 342-343). Why is this mainly observed in ZFTA-RELA tumors and less so in other subtypes (e.g., clusters 1-4, main figure 2e), where the neuronal-like state is less apparent? The authors should test whether this transition also occurs in other subtypes using a similar experimental setup. Finally, is this cell state transition neuronal activity-regulated, such as blocked with TTX in co-culture??

16. Computational comments:

A. Assignment of malignant cells: After reading the authors' analytic framework (in lines 119-123 and in the methods section), it is unclear to me whether the classification of malignant cells is entirely accurate. In the methods section, they state that it was based on the "presence of copy-number alterations," but this process is complex, and they do not provide

sufficient detail or show the inferred chromosomal aberrations in the so-called malignant group versus the reference group. Extended Data Figure 1a shows many false positives and false negatives (i.e., cells classified as true by inferCNV in the TME clusters and vice versa). Please address these concerns.

B. Deriving cell states (Figure 2): I have several comments:

B1. The authors use NMF, a well-established method, to derive cell states, but they assume a rank of 6 in this process. Why was this specific rank chosen? Newer implementations of this method do not assume a fixed rank per sample a priori, but rather iterate across multiple ranks and retain the NMF programs that are consistent and recurrent.

B2. In reading the methods, I did not understand how the clustering was performed. They select the top 50 genes per NMF program (which would result in different lists of genes across programs and samples), and then cluster based on correlations? This approach doesn't make sense, as the genes being correlated are different. Please clarify.

B3. How were the metaprograms defined based on the clustering of the NMF programs? I could not find a reference to this. Was some consensus used between programs in each cluster? There are a few clusters in the heatmap that were not marked - did the authors investigate these?

B4. In most tumors, including EPN (as previously published), cell cycle programs clearly separate into G1S and G2M phases. I am surprised not to see this separation here.

c. Spatial transcriptomics analysis (Figure 3): Please explain the decision to exclude the EPN non-canonical subgroups from the spatial analysis. This is surprising, especially in light of the limitations of previous studies, which were highlighted in the introduction and necessitated "a comprehensive study examining tumors across all subgroups of ST-EPNs" (line 80).

Minor comments:

Extended Data Figure 4a (line 193): The authors performed differential expression analysis using single cells, which results in inflated p-values. From a statistical standpoint, single cells from the same sample are not independent observations, and this analysis should be avoided. The authors should consider repeating this analysis using pseudo-bulk profiles (per sample). Also, please arrange the legend group by malignant and non-malignant and re-label Mesenchymal in the legend as MES/Hypoxia for consistency.

Extended Data Figure 9e, include representative images of the manually counted cell division.

Include additional, newer references for the role of neuronal activity in driving tumor progression where applicable. For example, the authors should reference and discuss the recent findings on the role of neuronal signaling in EPN (Chen et al., Nature 2024).

Lines 74-75: Even if different subgroups harbor distinct malignant cell states, this does not necessarily imply that these states are driving different outcomes.

Lines 78-80: This sentence suggests that cell states are sample-specific due to the cohort chosen, but this is more related to the methodology used to define cell states. Specifically, the study cited as reference 14 defined cell states within individual samples and then aggregated them to form metaprograms.

Figure 1a: Basic information, such as age and sex (rather than gender), should be included

Figure 1b-c: The authors mention 4 samples from the ZFTA cluster 1, but Figure 1b shows 5 such samples plotted. In Figure 1c, to what extent is the clustering driven by the sample of origin? The statement "we find that the molecular subgroup classification drives clustering of all malignant cells" is quite strong - typically, in most cancers, the sample of origin is what primarily drives clustering of malignant cells.

Figure 1f and line 133: The authors generated pseudo-bulk profiles using cells from different samples of the same subtype, which is not standard practice and is prone to batch effects. Why didn't they perform this analysis on a per-sample basis and demonstrate consistency across samples within the same subtype? They should score each sample against the different programs and provide statistical analysis. Additionally, in the caption (line 452), the term "tumor cell type gene expression" is not clearly defined—did the authors derive subtype-specific signatures? What genes are driving the observed similarity?

Line 135: The authors only show expression scores, not correlations.

Line 148: typo ('extent').

Line 170 (Extended Data Figure 2c): The metaprograms are composed of lists of genes, so it's unclear what exactly the authors are correlating. Please clarify or consider using a different similarity metric.

Line 171: Figures 1f and 2a are addressing different aspects. Figure 2a derives patterns of variability within tumors, whereas Figure 1f claims that different subtypes globally resemble developmental processes. Therefore, the metaprograms do not corroborate the results shown in Figure 1f.

Figures 2c-d: These panels assume that each cell must be assigned to a metaprogram, which is generally not accurate - a cell's score should exceed a certain threshold for it to be assigned. This is evident in Figure 2e, where cells located in the middle of the cloud should not be assigned to any of the presented programs. Moreover, the varying proportions of metaprogram abundance in different tumor subtypes should be supported by more quantitative statistics. Finally, I do not understand (and could not find reference in the methods) whether the scoring was conducted per sample (i.e., with prior centering per sample) or after aggregating cells from all samples. The latter approach seems better suited for this analysis.

The methods section (lines 659-670): The description relating to scoring and the cell state plot lacks appropriate references. In their important 2020 paper (Gojo et al.), the authors also sequenced ST-RELA samples and demonstrated that these lack the ependymal signature (main Figure 4 in that paper). However, here the ependymal-like program is strongly seen in these tumors (line 177). Additionally, in the 2020 study, the authors analyzed one ST-YAP1 sample, which they indicated is dominated by a program they term ST-YAP1 and lacks an ependymal signature (Figure 4 in that paper). This finding appears to conflict with their current results (line 198) - can the authors please clarify?

Line 329: There is a missing word "and" in the phrase: "are highly invasive AND neuroepithelial-like cell states are the most proliferative." Additionally, in Figure 2a, the neuroepithelial-like metaprogram and the cycling metaprogram are anti-

correlated, while the cycling metaprogram is slightly positively correlated with the neuronal cell metaprogram, which the authors state in line 326 is the least proliferative. Can the authors please clarify this discrepancy?

(Remarks on code availability)

Referee #3

(Remarks to the Author)

A. Summary of the key results

I want to thank the authors for presenting an interesting manuscript with a compelling story, I appreciate the effort that went into addressing this important topic. The manuscript is overall well written, and data clearly presented.

The authors make use of a multimodal framework of cutting-edge technologies to address questions regarding the tumor biology and origin of supratentorial ependymomas (ST-EPN). They employ single cell/single nuclei and spatial transcriptomics, and live cell imaging addressing several important questions. I've summarized the main findings below:

- ST-EPN subgroups align with temporally restricted cells across the human cortical development and have distinct developmental signatures.
- ST-EPN modes of organization where high order architecture ranging from disorganized to structured correlate with mesenchymal and hypoxia cell state signatures.
- The neural tumor microenvironment promotes plasticity towards a neural developmental trajectory, which affects the morphology and behavior of the ST-EPN cells.

I tentatively recommend accepting the manuscript with major revisions.

B. Originality and significance: if not novel, please include reference

The main novel findings are:

- ST-EPN subgroup characterization using normal developmental cortex to align the tumors data for temporally distinct subgroup signatures. Previous suggested cell type of origin has been radial glia cells, here the authors map ST-EPN tumors to two even earlier progenitor populations (caveat, needs revision)
- ST-EPN spatial investigation of the canonical ZFTA-RELA subgroup, finding higher order organization correlating with mesenchymal and hypoxia cell state signatures (also found in study authors reference 24, but here the authors examine a different type of tumor). (caveat, revision, that more cell states need to be analyzed).
- Using cell morphology, marker genes and live cell imaging elucidating the effect of the neural tumor microenvironment in promoting invasive behavior and developmental trajectory of ST-EPN cells. (caveat, needs revision)

Caveats: depending on revision results assessments may change.

C. Data & methodology: validity of approach, quality of data, quality of presentation

Dear Authors, I have tried to structure my comments using citations from the manuscript, specific line numbers and figure names. I have tried to explain as clearly as possible my reasoning and feedback. It was an interesting read, and I see the future potential of a beautiful work. However, there are many things I deem is needed to be revised. If anything is unclear, please contact the editor and I will do my best to explain.

Abstract

The authors state "we integrated single cell- and spatial transcriptomics, as well as in vitro and in vivo live-cell imaging". However, reading the manuscript it does not include in vivo live-cell imaging, only in vitro live cell imaging, please rephrase.

The authors state "We find that individual tumor subgroups harbor two distinct progenitor-like cell states reminiscent of early human brain development and diverge in the extent of neuronal or ependymal differentiation.". However, in Figure 2e, four different cell states are displayed, and subgroups are mapped across these. Please clarify. E.g., was this specifically referring to the neuroepithelial-like and embryonic-like states?

The authors state "We further uncover several modes of spatial organization of these tumors". However, only MES/Hypoxia is statistically significant and supported by the data analysis, and many of the state/cell types were not tested for, at least not

in the provided material. See revision points.

Introduction

Clearly written, describes the current knowledge gap, brings up the main findings, and frames the authors scientific work in a proper context.

Lines 81-83: The authors use a reference to a review paper from 2012 stating that “it has recently become clear”. In my opinion ‘recently’ would be within the last ~ 5 years, I suggest the use of another term or removal of ‘recently’ in this context.

Lines 95-96: The authors claim to “find that malignant cells reflect two distinct temporally restricted progenitors [which are? Please name them] from early human embryonic cortical development”, please clarify here already by adding in parenthesis which these two progenitors are.

Results

The different sections of the results are overall clearly structured and easy to follow. Presenting the data and the authors investigation in a chronological order that makes sense and create an interesting story. However, there are multiple points that needs to be revised.

Line 110: The authors state “we compiled a patient cohort of 44 patient tumors”, but it is not immediately clear that several patients have been sampled multiple times. Furthermore, one patient has 2 samples in subgroup ZFTA Clusters 3 and 1 sample in subgroup ZFTA Clusters 4, please add an explanation somewhere in the paper of why the patient’s samples are spread across different subgroups. Although the authors illustrate this in Figure 1a, I suggest they further add in their main text when presenting the dataset that the 44 sample comes from 37 patients. If possible, it would also be of interest to know if the multiple samples from the same patients have been collected at different time-points or from different spatial areas of the tumors. Because sampling from different areas of a tumor could capture global heterogeneity and contain different genomic and transcriptomics profiles.

Lines 110-111: The author’s state: “we compiled a patient cohort of 44 patient tumors encompassing ZFTA-RELA (n = 20), ZFTA Clusters 1-4 (n = 20) and ST-YAP1 (n = 4) subgroups.” However, in figure 1a the number for each subgroup is clearly stated. The authors also make a point of their study including more diverse subgroups then previous studies. I suggest that instead of grouping only ZFTA Clusters 1-4 here, that they be presented individually also in the text. E.g., ZFTA Clusters 1 (n = 4), ZFTA Clusters 2 (n = 4), ZFTA Clusters 3 (n = 8), and ZFTA Clusters 4 (n = 4).

Line 121: The authors state they have retain 6933 high-quality malignant cells, however in Extended Data Figure 1, subpanel ‘Final cell type classification’ the number is instead 6680 malignant cells. Please clarify where the 253 additional cells come from.

Line 127: The authors state that “subgroup classification drives clustering of all malignant cells (Figure 1b-c)”, however in the c panel there are distinct differences between DNA methylation and sc/sn RNA-seq data. I suggest the authors test coloring the sc/sn RNA-seq data UMAP also based on patient ID. Many of the subgroup clusters looks to be separate by the patients, not their subgroup classification. I suggest a different analysis for this. One can treat each patient sample as a pseudo-bulk sample, then select e.g., top 3000 variable genes across all pseudo-bulk samples, then do a Pearson correlation between all samples and display them in a heat map with subgroups as color legends. Look to see if subgroups are clustering together using such an approach instead.

Line 139: The authors state that “ZFTA-Cluster 3 and 4 tumors mapped to mesenchymal cells found in the early first trimester”, which the figure supports, but for ZFTA-Cluster 3, in particular, it appears that cells map to OPC to an equal extent when considering cell types. Can the authors please explain why mesenchymal and not OPC is mentioned in their results interpretation?

Lines: 155-169: The authors state that annotation of the metaprograms are based on the expression of known cell type-specific marker genes in Table S4. They also highlight specific examples in the text for all the 8 metaprograms. However, in Table S4, when checking if all the mentioned program specific genes are within this table, I can only find 7 out of 16. From the text the reader assumes that the annotation of the metaprograms is based on the marker genes in Table S4. This is confusing to me. Could the authors please explain why 9 of the mentioned markers are not in Table S4 among the metadata program specific top 200 genes? Also, if these 9 genes are below the top 200 genes for each metaprogram, please specify and explain the rational for a cut-off for which genes are deemed to be metaprogram specific enough to merit their use as annotation markers for the different metaprograms. Furthermore, this should also be taken into consideration for GO enrichment analysis, that a cutoff and number of top genes are consistent and that the marker genes used for annotation is included in this top set of genes used for GO enrichment analysis (Extended Data Figure 2b). Currently this does not appear to be the case, I suggest setting motivated cutoffs and re-running the GO analysis based on these. See further comments in section F.

Lines 173-175: The authors include GO results for 6 of 8 metaprograms in Extended Data Figure 3b. Please add all 8. If no significant results were obtained for the other 2 metaprogram, please state so.

Lines 177-178: The authors state “While ZFTA-RELA tumors were composed predominantly of neuronal-like and

ependymal-like cells”, however looking at the subpanels for figure 3 d-e, I suggest they also mention neuroepithelial-like cells.

Line 178: The authors state “neuronal-like cells were absent in ZFTA-Cluster 1 and ST-YAP1 tumors”, however there are some proportions of neuronal-like cells in ZFTA-Cluster 1. I suggest adding “cells were few or absent”.

Lines 217-225: This is great! The authors present high-level results from the spatial experiment and explain their rationale for only going forward to analyze four non-malignant cell types and five malignant cell states.

Lines 227-228: The authors state “was highly consistent across consecutive sections from individual patients, confirming high reproducibility of this technology (Extended Data Figure 6b).” Although the cell proportions are visualized in Extended Data Figure 6b the authors do not show any calculations supporting the specific claim that proportions were highly consistent across consecutive sections from individual patients. To make a claim for the consistency of proportions across consecutive sections, please add calculations supporting this claim.

Lines 231-234: The authors state “After annotating spatially distributed cell states, we observed that tumors segregated into two global organization patterns – (1) high compartmentalization of cell states, which we term structured tumors, and (2) cell states/types scattered throughout the section, which we term disorganized tumors. (Extended Data Figure 5a).” However, there is no indication for case (1) and (2) in Extended Data Figure 5a. It would be valuable to have something similar to what is shown in Figure 3b. I suggest that the authors also reference Figure 3b here and add spatial coherence score visualizations to Extended Data Figure 6b.

Lines 242-245: The authors state “we performed linear regression analyses between the spatial coherence score of each section and the frequency of cells observed for each state/type.” By formulating this explorative hypothesis, one expected the authors to investigate all 9 cell states/types (4 non-malignant, 5 malignant). This should then result in 9 linear regression calculations and 9 p-values which then multiple correction is applied for, e.g., preferably using Bonferroni correction given the small number of p-values to correct for. However, the authors only show linear regression calculations and p-values for MES/Hypoxia and vascular cells (no mention of multiple correction). This needs to be addressed. Given 9 test, the adjusted p-value for MES/Hypoxia would be $p_{\text{adjusted}} = 0.018$ (still significant). The authors draw the conclusion that the MES/Hypoxia state is especially important for spatial organization without investigate and/or show results from the other cell types and state. Please address this, which could have implications for how the results are interpreted and lead to major revision of this section of the manuscript. See further comments in section F.

Lines 246-249: The authors hypothesize “that vascular perfusion might decrease hypoxia, and therefore revert this structural organization”. In my opinion, this is too early a hypothesis to form from this data. First the previous comment needs to be addressed, there could be more states than MES/Hypoxia that correlate with structural organization. In addition, in Figure 3c, the p-value (appearing to be non-adjusted for multiple testing) for linear regression between structural organization and vascular cells is not significant, i.e., the data and analysis does not currently support this line of reasoning and the authors conclusions. My suggestion is to remove this from the manuscript.

Lines 251-254: The authors state “This indicates that higher spatial organization of ZFTA-RELA tumors correlates with more aggressive tumors.” However, there is no mention of how this is calculated in the material and methods. The statistical test (mentioned in the figure legend) in Extended Data Figure 6f-g requires a normal distribution, which has not been assessed by the authors, please re-evaluate and used appropriate motivated statistics. The p-values are not significant, thus the claims and reasoning by the authors for this part is not supported by the data and data analysis. For this data, care needs to be taken here, because some tumors have been sampled from multiple regions (perhaps use an average measurement from each tumor?). Also, it is unclear if patients with primary tumors have had recurrent disease at a later time-point. My assessment is that this part is too speculative and the sample size in this cohort is too small to draw any conclusions here. I suggest removing this from the manuscript.

Lines 260-268: The authors describe the spatial Xenium data by dividing different regions into different spatial niches. However, to me as a reader it gets confusing. There are first niches 1-4 for each section, but results are only shown for one subregion from two sections. I’m not certain why the authors selected to only look for 4 niches across all tumors (I would like to see an extended data figure showing spatial niches for all samples), which also have different degrees of spatial organization, which could affect how many niches there could be. Please motivate why 4 niches were chosen. E.g., why not pick a higher number of niches, e.g., 9 that was investigated before (4 non-malignant, 5 malignant)? Especially since the authors do make observations that indicate that transcriptional state and cell types are directly correlated with morphology (Figure 4a). Please provide a clear motivation for this choice, and additional supplementary figures showing the results from the analysis.

Lines 269-274: Here the authors define recurrent spatial niches (methods lines 769-773) based on the frequency of cell types, reading the manuscript and looking at the corresponding subpanels and pie chart plots (Figure 4), I interpret this as also including the different cell states. The authors then conclude that “Most recurrent niches were predominantly composed of one malignant cell state (Figure 4c), which indicates that malignant cell states preferentially co-localize with themselves.” But according to the current definition of these recurrent cell types, this argument becomes circular (since recurrent spatial niches were also defined by frequency of cell types). Thus, in my opinion, the conclusion that “we uncovered local spatial patterns of cell states across ZFTA-RELA tumors, with a common observation of malignant cell states preferentially co-localizing with themselves.”, is not supported by this data analysis. See section F for suggestion of re-analysis.

Lines 310-311: The authors states “we characterized subcellular and cellular tumor cell behaviour with longitudinal live-cell imaging”, to me the use of ‘longitudinal’ is more used when studies span months and years, I suggest the use of ‘temporal’ instead.

The authors uses the word ‘somatokinesis’ to describe what I believe is cell movement. I can only find one mention of this word in ncbi. Is there some other word that is more commonly and broadly used? Please define this term clearly in the text for the reader.

The authors use the definitions ‘continuous invasion mode’ and ‘saltatory invasion patterns’. Please clearly describe and define to the reader what these entails. E.g., in the main text or more detailed in material and methods.

The authors use different terms to describe invasive cell behavior, such as: protrusion, retraction, branching, branching migration, locomotion and translocation. I suggest adding clear and descriptive definitions by the authors in the methods section of what these are and entail.

Line 320: The authors mention ‘EPN-glial-like’ cell for the first time in the manuscript. It is not clear to the reader which meta program or cell state these are related to. Please clarify. Is it Radial glial-like? If yes, then why in Extended Data Figure 10d does it appear to be more of this in monoculture but on line 341 is says that it is detected less?

Lines 338-340: The authors state that “in coculture versus monoculture and found significant enrichment of neuronal-like GO terms and gene signatures only when tumor cells are grown in coculture, as well as a decrease in neuroepithelial-like signature”, however only GO analysis is shown for coculture. Please include how this analysis was done in the methods section, also include how many genes were used for the GO analysis and do the same for the monoculture. Display these top GO enrichment terms in plots side by side in Extended Data Figure 10, to show that/if there is a difference among GO terms between monoculture and coculture.

Line 345-346: The authors state that “only ZFTA-RELA cells grown in vivo and cocultures are enriched in neuronal-like and ependymal-like states,” Extended Data Figure 10f. However, the used of the word ‘enriched’ is not supported by any calculations or statistical test. From the plot in Extended Data Figure 10f it does not look like ZFTA-RELA cells grown in vivo are present only in neuronal-like and ependymal-like states, it looks like they are spread out a bit everywhere. Please include a statistical test, if the p-value is not significant, then remove this statement and conclusion from the manuscript, since it would then not be supported by the data. E.g., one could divide the data into 2 categories i) Neuronal-like + Ependymal-like and ii) Neuroepithelial-like and Embryonic-like, based on the plot. Then perform a Chi-squared test.

Line 350-351: The authors state “Overall, these results highlight how the neural microenvironment promotes plasticity towards a neuronal trajectory, affecting the morphology and behavior of ZFTA-RELA cells.”. At this point I would say that the results partly do. E.g., Figure 5j does. But, as mentioned in the comment for lines 345-346, calculations and a statistical test in needed for this to also be a supported statement for Extended Data Figure 10f.

Figure 5j-k. Check for normal distribution and then decide on appropriate test to use, explain in material and methods.

Figures

Figure 1a: For methylation data legend, it seems like there is also a grey color, e.g., for 3 of the ZFTA-RELA samples, cell lines and PDXs. If this mean Not Applicable (NA), please add a grey box denoting NA at the bottom for the Methylation score legend.

Figure 1e: Please write out the abbreviations in the figure legend for CP, OSVZ and VZ. Also indicate with a text label that the green cells on-top are neurons I believe. Currently only the newborn and maturing neurons are text labeled.

Figure 1e: Would it be possible to also color code layers within the rectangle of the cortex? The cell type illustration has some background colors that seems to indicate different layers.

Figure 1f and Extended data figure 2a: I found the different expression score labels to be a bit confusing. Although it is explained in the figure legend, I would find it helpful with an addition to the labels. E.g., instead of writing (tumor cells), write (normal profile in tumor cells), to indicate that normal cell profiles have been projected in tumor cell data. For (normal cells), it would be (tumor profile in normal cells). I’ve seen this somewhere also in the Extended Data, would be great to have it updated there to.

Figure 2e: Just wanted to say that this representation looks great and makes it a lot easier to interpret the cell proportions from Figure 2d!

Extended Data Figure 1d: In general, it is difficult to see the placement of all the cell types in these UMAPs. Perhaps using different colors and/or transparency, or dividing the visualization into more plots? E.g., The third developmental trimester is not at all visible.

Extended Data Figure 4a: I find the labeling a bit confusing. Which is the case and control? From the y-axis labeling I would assume that ZFTA-RELA is the case and ZFTA-Cluster 3 is the control. But in the plot the upregulated genes are shown for ZFTA-Cluster 3, I would expect the plot to show upregulated genes for the case. I suggest doing something similar to what

was done by the authors in Extended Data Figure 10a instead.

Figure 4 and Extended Data Figure 7: Please use a different color scale for the heatmaps, it is difficult to interpret when the scale is from 0-60%, but the color range go from blue-white-red. Use e.g., white-yellow-red, something that steadily increases. The blue-white-red is appropriate when looking at number ranges e.g., between -1 to 1. Furthermore, for these heatmaps, please include all cells and states in the color legend. I suggest having two color legends, keep the one with niches A-E but add one including all cells and cell types, e.g., more annotations are shown in the pie charts.

Extended Data Figure 10a: This looks really nice, something similar could be done for easier interpretation of Extended Data Figure 4a.

Figure 5h: The illustration showcases monocultures and cocultures as if they are performed in the same Petri dish with some form of separation wall. However, from reading the manuscript this does not appear to be the case. Please revise the image to reflect the experimental design.

Methods

I'm unable to find anything in the methods related to the methylation analysis in Figure 1b. Please add experimental protocol procedures and data analysis.

I'm unable to find any information regarding how the samples fusion gene statuses were determined, shown in Figure 1a. If this was provided from the hospital as metadata, I suggest stating which type of metadata was provided by the hospital in the material and methods section regarding the samples.

Line 533: The authors have included ethical approval information for the human subjects, but not for the animal studies. This information appears to be included in the Reporting Summary document, please also include this information in the main manuscript ethical approval section.

Line 637: The authors mention the use of the InferCNV R package to identify malignant and non-malignant cells. However, they do not state how InferCNV was run or if a reference was used. This is important to state in the methods, e.g., if no reference is used then amplification and deletions that are broadly present in the dataset will be missed.

Lines 684-688: The authors describe 'Correlation to reference datasets', which in the main text (line 133-135) is described and referred to as a way to project the gene expression from the different subgroups onto reference cell atlases. It was not immediately clear to me as a reader that this part of the methods corresponds to this part of the results. I suggest some clarification in the methods description to help the reader navigate. Furthermore, I want to see a more detailed description of how this was done. Currently, it is explained that marker genes are found for each cell types in the reference datasets, and "within ST-EPN molecular subgroups or cellular states", however in Figure 2d only cellular states seems to be presented. How was the ST-EPN molecular subgroups score used? Then the authors state "We then scored all ST-EPN malignant cells with marker genes identified across the developing human cortex, and other way around", however, there is no mention of what score cut-off values or ranking were considered for the annotation itself, please add this to the methods section.

Please add a methods part describing the correlation between Gojo et al 2020. It is mentioned in the text (line 170-171), but I can't find a description in the methods part. E.g., from what values were these correlations calculated? Were the top 200 genes from the metaprograms used to calculate a score (how was that score calculated) and then how many top genes for the different gene programs from Gojo et al 2020 was used to calculate scores that were then used in the correlation analysis?

Please add differential gene expression analysis and GO enrichment analysis shown in Extended Data Figure 4a to the methods section.

In general, for all GO enrichment analysis, please add information regarding version control, these databases update regularly and using the same list of genes for a search in a few months might yield slightly different results. E.g., if using g:profiler for GO one can specify which Ensembl version and Ensembl Genomes version the database was built on (<https://biit.cs.ut.ee/gprofiler/page/archives>). Does clusterProfiler used here have something similar? Please also indicate which settings the enrichGO function was run under, e.g., if there were some non-default settings. Also, regardless mention the p-value cut-off and multiple testing method selected even if default was used. Also mentioned if ranked or non-ranked gene lists were used or a possible choice with the enrichGO function search.

Line 740 Metaprogram cell assignment. Which figure and data is this paired with? As a reader I'm confused, is this for Figure 4 a subpanel 'Cell types' from the Xenium spatial data? Please clarify.

Line 764 Spatial niche. After reading the paper and this methods part I'm still unsure as a reader how the different niches are defined. I think it is because the authors both refer to 'spatial niches' which are mentioned as niche 1, 2, 3, and 4. Then the authors define recurrent niches A, B, C, D and E.

Discussion

Lines 354-356: The authors claim that "Our work uncovers different developmental signatures as well as malignant cell

states across ST-EPN subgroups including previously uncharacterized non-canonical ST-ZFTA tumors, and demonstrates that these cell states have distinct spatial, morphological, and behavioral characteristics.” However, this framing is misleading, because Xenium spatial data analysis and cell experiments were only performed in canonical ST-ZFTA tumors. Please reframe and clarify.

Lines 382-383: The authors state “In addition, the strong ependymal signatures in ST-YAP1 corroborates previous work demonstrating ciliogenic marker expression in these tumors^{14,55}”, referencing previous work. But the dataset presented in this manuscript is new, please check for high ciliogenic marker expression in this ST-YAP1 cohort data. E.g., add a plot in the Extended Data.

Lines 396-397: Part about relapses was not supported by the data, please remove.

Please include a section in the discussion bringing up limitations of the study. Some suggestions for this are:

- tumor global heterogeneity and sample site bias.
- Too few primary/relapse data points to establish if relapses generally are more structured than primary tumors.
- That spatial data was only generated on the canonical subtype and not for the non-canonical.
- That the ST-SE subgroup was not included in the study (authors ref. 13).
- That mesenchymal and hypoxia cell state signatures are not direct measurements of e.g., hypoxia and acts as a proxy.

D. Appropriate use of statistics and treatment of uncertainties

General comment on statistics. Please include a section of this in the material and methods part. All distributions tested needs to first be assessed if they are normally distributed (e.g., using a Shapiro-Wilk test). Then select appropriate parametric or non-parametric tests. Make clear how multiple testing has been considered and performed. For test with few p-values, it is most robust to use Bonferroni correction.

Please add the statistical test used in Extended Data Figure 4c to the methods part (line 695). Also, test if the data is normal distributed, which is an assumption required for the use of the student's t-test. If not normally distributed, please re-calculate using appropriate statistical test, e.g., Wilcoxon rank sum test. Also, because 3 tests are made, use multiple testing correction of the 3 p-values, describe in the methods section (and figure legend) which multiple correction method was used.

Figure 3c: Linear regression between structural organization and vascular cells is not significant, but results are interpreted as if they were. This needs to be removed from the manuscript in the places it occurs.

Extended Data Figure 6f-g: The tests between primary and relapses are not significant, but results are interpreted as if they were. This needs to be removed from the manuscript in the places it occurs.

Please add for Extended Data Figure 6a how R² and the p-value was calculated in the methods section. Also explain the rational for why “For the smart-seq2 dataset, neuronal-like and embryonic-neuronal like metaprograms were combined (“neuronal-like”), as well as neuroepithelial-like and embryonic-like (“neuroepithelial-like”).” i.e., motivate/explain more clearly why combine and not keep them separate. This division and grouping of metaprograms does not appear to make sense, especially since the authors in the main text state, lines 184-187, “To summarize the distribution of metaprograms in each ST-EPN subgroup, we scored all single cells for the two most immature (neuroepithelial-like and embryonic-like) and the two more mature (neuronal-like and ependymal-like) lineage-related expression programs and visualized the results as a cell state plot (Figure 2e).”

General comment, please add a section in methods describing in detail how and using what programs the linear regressions were done in the paper, currently there is only some description in the figure legends.

Extended Data Figure 10c: The statistical test required normal distribution, please test for normal distribution and use appropriate test and describe rational in methods section.

Please include description of the statistical test with rational for selecting the test in Figure 5c, and Extended Data Figure 9e, in the methods section.

Extended Data Figure 9e: Please add statistical test and rational to material and methods.

E. Conclusions: robustness, validity, reliability

The authors eloquently argue for their conclusions, basing them on their data and previous work by others. However, due to some statistical test being not significant, but interpreted as if they were, and what appears to me to be a circular argument in on section, some major revision is required here.

The authors claim, “that individual tumor subgroups harbor two distinct progenitor-like cell states reminiscent of early human brain development and diverge in the extent of neuronal or ependymal differentiation”, looks to be supported by the data, but there is some uncertainty regarding the annotations of the different metaprograms based on the authors selection of genes

used for annotation that need to be addressed, see comment for lines 155-169.

The authors claim that high order architecture is influenced by MES/Hypoxia is supported by the data and analysis. However, this analysis is incomplete until test also have been done for all 9 cell types and states. There could be more cell types and states contributing to this higher order architecture. See comment about Lines 246-249.

The authors claim that “This indicates that higher spatial organization of ZFTA-RELA tumors correlates with more aggressive tumors.”, is not supported by the data, see comments for lines 251-254.

The section of the manuscript ‘ZFTA-RELA tumors are locally organized into recurrent tumor niches’ needs to be revised due to what appears to be circular arguments.

F. Suggested improvements: experiments, data for possible revision

I have tried to collect the main suggested improvements here, but some are also under other sections.

Please include an additional inferCNV analysis using normal cells as reference to see if there are any potentially different CNV patterns between and within subgroups. This could be relevant supporting information when investigating subgroup and inter-tumoral differences of the cohort.

Line 127: Try a different approach too comparing the transcriptomes of the different subgroups. Now they do not cluster based on subgroups but look to cluster based on patient ID. See comments in section C for more details.

Line 157-158: In addition to marker genes. The authors also use GO terms to annotate the metaprograms (Extended Data Figure 2b). From the methods section (line 673) it states that top 30 genes from each metaprogram is used. E.g., in Table S4 genes MKI67 and TOP2A (used to annotate the cycling metaprogram) are not within top 30. Neither is DCX for neuronal-like. I would suggest to re-run the GO analysis using all the top 200 genes for each metaprogram. Although, in the Extended Data file at the bottom there is a list of supplementary materials, stating that Table S5 “Gene ontology terms enriched in NMF metaprograms identified in ST-EPN patients (top 200 marker genes/metaprogram selected; GO terms with P-value < 0.01)”. This is inconsistent with the main material and methods part. Please clarify and update to correct and use consistent information.

Lines 242-245: Given that the probe panel was design to capture cells from all eight metaprograms and multiple cell types, and that the authors have stated that their intention was to investigate 9 of these after data collection. I suggest additional analysis to be made (with multiple testing corrections), including all 9 cell types and metaprograms in separate linear regression analysis. E.g., included in an additional Extended Data Figure. There might be additional interesting biology not explored here that could have an impact on the higher order architecture of the tumor organization. This could substantially change the story of the paper in this section. Especially since the current data and analysis does not support the authors investigation into vascular cells as an explanation for disorganization (Figure 3 c and d), (p-value is not significant, could be study is underpowered, but can’t draw these conclusions with this data, please discuss in a revised discussion).

For the section of the manuscript ‘ZFTA-RELA tumors are locally organized into recurrent tumor niches’. Consider characterizing recurrent niches using an unsupervised approach to avoid a circular argument. One alternative could also be to set out to investigate the recurrence of the 5 different malignant states captured by the Xenium data. Do the same analysis using BuildNicheAssay’s function from Seurat (v. 5), setting k=5. Then annotate the 5 different spatial niches for each section according to the same cell types/states as in Figure 4c.

Line 345-346: Perform calculations and statistical test to support statement of enrichment. See section C for more details.

G. References: appropriate credit to previous work?

Lines 67-68: the authors state “differences across species underscore the challenge in extrapolating findings from these studies in mice.”, please provide/mark a reference for this statement.

Lines 160-169: the authors mention genes associated with different metaprograms, please add literature references where missing.

Line 230: In this section the author could use their ref 24 (Greenwald, A. C. et al. Integrative spatial analysis reveals a multi-layered organization of glioblastoma. Cell 187, 2485-2501.e26 (2024), to high-light previous work done by others looking at spatial organization and hypoxia.

H. Clarity and context: lucidity of abstract/summary, appropriateness of abstract, introduction and conclusions

This is described under C.

I. Comments on the Data repository

Please add a separate file about the metadata on GEO for the spatial data. The current files are large and having a separate small metadata file makes the data more accessible and more in line with FAIR data principles. E.g., a version of the authors supplementary table 1 could be used here.

In the manuscript the authors state that they are analyzing 23 tissue sections using the spatial transcriptomics method Xenium. However, in the GEO database only 22 samples are provided. Please update and add all samples to the GEO database, or if it only was 22 sections, please update and correct this number in the manuscript.

For the single cell and single nuclei data in the GEO, please add more metadata information to facilitate easier re-use of data. I suggest an additional supplementary file that helps the reader navigate the 42 sample GEO upload. E.g., 3 columns: 1) GSM number (GSM8115764), 2) Sample group (1), 3) Description (scRNAseq? Xenograft ZFTA-RELA?). From the GEO database entry, I'm currently having difficulties to easily understand what these samples are, I think that adding such a Description would greatly increase the data's adherence to FAIR principles. Due to this I'm unable to ascertain if all data has been uploaded in accordance with the authors figure 1a. However, I do note that in Figure 1a it looks like there should be 47 samples for sc/snRNAseq, but only 42 are added to the GEO database, please double check that all data has been deposited.

(Remarks on code availability)

The authors provide extensive and descriptive annotation of the code for their work. I have not tested the code.

Version 1:

Reviewer comments:

Referee #1

(Remarks to the Author)

The authors have truly added a great amount of content to the manuscript and have largely addressed many of the initial concerns (not only by this referee, but the others as well). I would like to commend them for this impressive work, and do feel the manuscript is suitable for publication at this point.

(Remarks on code availability)

Referee #2

(Remarks to the Author)

The authors have addressed all my concerns in this comprehensive revision. It is an important contribution to the literature.

(Remarks on code availability)

Referee #3

(Remarks to the Author)

Dear Authors, thank you for the revised version of the manuscript. I would like to acknowledge the effort, care, and attention to detail that has gone into improving the work, as well as the clear and extensive communication about the revisions. After reading the updated manuscript, I still have a few comments and questions, outlined below.

Main

Lines 92–94: The authors state: “However, these studies were limited by patient number and did not include the non-canonical ST-ZFTA subgroups, which led to many cell states being sample- or subgroup-specific.” In this new cohort—2–3 times larger—the authors still observe (lines 151–152) that “the malignant cells cluster both by sample and molecular subgroup classification.” This suggests that similar patterns emerge despite the increased sample size.

I would recommend adjusting the wording around lines 92–94 to something like: “However, these studies were limited by patient number and did not include the non-canonical ST-ZFTA subgroups, which could have led to many cell states appearing sample- or subgroup-specific.”

If this change is made, I also suggest including a comment elsewhere in the manuscript to note that, despite the increased sample size, global similarities remain evident at both the sample and subgroup level. Alternatively, you might consider avoiding patient number as a differentiating factor altogether, e.g.: “However, these studies did not include the non-canonical ST-ZFTA subgroups.” The current wording conveys certainty about sample size as the cause of observed differences. Do

we actually know that an increased cohort would change the sample- or subgroup-specific clustering, or could this reflect true underlying heterogeneity?

Line 118: The authors write "... with neuronal-like cancer cells being the most migratory...". Would it be possible, based on the available metadata, to explore whether patients with a higher proportion of neuronal-like metaprograms show a greater tendency toward relapse, compared with other patients—assuming the number of relevant cases permits such an analysis?

Results

Line 189: The authors provide example genes for 7 out of 8 metaprograms among the top 30 genes used to define these metaprograms. Could the authors clarify why this was not done for (2) the neuroepithelial-like metaprogram? For consistency, I suggest either including representative gene examples for this metaprogram in Table S4 or briefly explaining why this was not possible.

Lines 196–199: The authors state that neuroepithelial-like and embryonic-like cells are the most proliferative, based on Extended Data Figures 2e–f. This is consistent with Extended Data Figure 2f; however, in Extended Data Figure 2e the pie charts suggest a much larger difference between neuroepithelial-like and embryonic-like cells. Could the authors clarify this discrepancy?

One possible explanation is that Extended Data Figure 2e includes all cells, so that the cell-cycle signal is measured not only in neuroepithelial-like cells but also in others that fall within the neuroepithelial-like square of the plot. If this is the case, I suggest reconsidering whether the pie charts in Extended Data Figure 2e should be included, as they can be confusing when compared to Extended Data Figure 2f. For example, in 2e the difference between neuroepithelial-like and embryonic-like appears large, while in 2f it is small; conversely, in 2f the difference between embryonic-like and neuronal-like is clear, but in 2e they appear identical.

Lines 229–231: The authors state: "While there was a continuum of gene expression to both neuronal and ependymal-like cells in ZFTA-RELA tumors, the other ST-EPN subgroups only harbored one of the mature cell states." However, for ZFTA-Cluster 4 this does not appear to be the case. It seems to derive from a neuroepithelial-like continuum but remain "stuck" between the two mature metaprograms. In my reading, it looks to have roughly equal numbers of cells in the neuronal-like, ependymal-like, and embryonal-like regions.

Could the authors clarify whether this interpretation is correct? If so, I suggest revising the sentence to reflect this distinction. For example: "While there was a continuum of gene expression to both neuronal and ependymal-like cells in ZFTA-RELA tumors, the other ST-EPN subgroups showed unique cell state signatures, with ZFTA-Clusters 1–3 harboring only one mature cell state, and ZFTA-Cluster 4 appearing to remain in a premature state."

Figure 2a: It looks like the heatmap color bars for the cycling metaprogram is dark-purple on the y-axis (to the middle-bottom-right) but dark-grey in all other places in the manuscript. Color-typo? Please update for consistency.

Spatial transcriptomics section: I suggest including an analysis of the cycling metaprogram in the spatial samples. Since this metaprogram was identified in the sc/snRNA-seq data and there are at least four probes related to cycling in the Xenium panel, exploring its spatial distribution could provide valuable insights into tumor growth dynamics.

For example, it would be interesting to see whether the cycling metaprogram is anti-correlated with the MES/Hypoxia metaprogram in a gradient-like manner. This could help understand whether dividing cells are spatially excluded from hypoxic regions, and whether slow-cycling or quiescent cells near hypoxic regions might contribute to recurrent disease.

If the number of matched primary and relapse samples is too small for detailed analysis, a preliminary profiling of the cycling metaprogram could still be informative. At minimum, the authors could provide a rationale if this analysis is not technically feasible given the panel design. Including this would strengthen the spatial characterization of the tumor landscape and complement the MES/Hypoxia findings.

Figure 3c–d: The authors provide a clear overview of the data, including helpful visual examples. For consistency, I suggest ensuring that the same samples shown visually in 3c are also highlighted in 3e. Additionally, in 3d the y-axis is labeled "percentage %," whereas in Extended Data Figure 4e it is labeled "Frequency %." I recommend standardizing this terminology for consistency across figures.

Figure 3d and Extended Data Figure 4e: The authors' initial approach of correlating each cell type proportion with the global spatial coherence score and applying a Bonferroni correction is acceptable for exploratory analysis, and the plots are informative for visualizing these relationships. However, because the follow-up test—removing MES/Hypoxia cells and recalculating the spatial coherence score—is conditional on MES/Hypoxia cells being significant, the resulting p-value is not fully independent and should be interpreted as a post hoc test rather than a confirmatory test.

A more rigorous approach would be to perform a multiple linear regression including all cell types as predictors of the spatial coherence score. This would provide a direct estimate of each cell type's unique contribution while adjusting for correlations

among them. The authors could retain their original analysis and plots for descriptive purposes, while using the multiple regression to confirm whether MES/Hypoxia cells remain significantly associated with spatial coherence after accounting for other cell types. This approach would give a statistically more robust assessment of each cell type's independent effect and reduce the risk of inflated type I error from sequential testing.

Lines 309–310: The authors state: “We overall identified six recurrent niches, each enriched in distinct malignant cell states or non-malignant cells (Extended Data Figure 6b).” The term “enriched” implies that a statistical test has been applied, which is not the case here. I suggest using “enhanced” instead when describing the meta-niches to avoid implying statistical significance.

Lines 321–322: Similarly, please update the wording to describe the clusters as “enhanced” rather than “enriched” if no statistical test has been performed. This recommendation applies to all downstream analyses of these clusters.

Lines 328–330: The authors state: “The distinct MES/Hypoxia-enriched spatial cluster underscores the distinct spatial organization of hypoxic regions found by calculating spatial coherence across cell states/types (Figure 3e).” However, comparing Figures 3e and 4e, similar patterns could also be observed for neuroepithelial-like, endothelial, and neuronal-like cells. As written, the current phrasing may overemphasize the MES/Hypoxia cluster and underrepresent other data patterns. I suggest reframing the sentence to acknowledge these additional spatial patterns while still highlighting the MES/Hypoxia finding as appropriate.

Extended Data Figure 6f: The plots and analyses appear to be strongly influenced by single outlier points. I suggest reconsidering their inclusion, as they may not reliably represent broader trends.

Figure 4d and Extended Data Figure 6g: This is a very interesting analysis! However, in Figure 4d, it appears that spatial clusters 22, 25, and 0 are not sample-specific, whereas cluster 6 may be sample-specific. I suggest revisiting the selection of these clusters, which may require updating the statistical test in Extended Data Figure 6g. Additionally, for consistency, please label the two groups on the x-axis as “enhanced” rather than “enriched.”

#3–19: The authors test whether there is a difference in spatial coherence scores between primary and recurrent tumors. Although the result remains non-significant with the increased number of data points, negative results are still valuable to report. I suggest including a brief statement in the manuscript that this comparison was investigated and that no significant difference was found. This would help provide a complete view of spatial coherence patterns, which is central to one of the main findings of the study.

#3–28: Upon reviewing the revised manuscript and Figure R46 (Extended Data Figure 7e), it appears that each condition includes only three biological replicates, which is too few to calculate p-values reliably. Treating each single cell as an independent observation would inflate statistical significance.

I suggest instead calculating the proportion of Progenitor-like versus Differentiated cells per replicate for each condition (adherent, spheres, coculture, PDX, and patient). The plot could then be presented as a bar plot showing the mean proportion across replicates, with error bars representing variability, and individual replicate points overlaid (n=3). Use descriptive language such as “higher proportion of...” or “predominantly...” rather than statistical terminology. Given the small sample size, avoid using terms like “enrichment” for this analysis.

#3–29: Given the small number of biological replicates (see #3–28), I suggest softening the conclusion from “Overall, these results highlight how...” to something like “Overall, these results indicate that...”. Additionally, use “enhanced” instead of “enriched” to avoid implying statistical significance.

Methods

For reproducibility, please add package version for the different analysis packages used. E.g., Seurat (v.x.x.x). It looks like it is missing for some packages in the methods section.

Lines 886–888: The authors state: “ST-EPN cells/nuclei were annotated based on the maximum expression score for the respective MP signatures after excluding the one related to cell cycle (‘cycling’).” However, in Figure 2, cycling cells appear in the UMAP in 2c but not in 2d. Could the authors clarify: are these cells excluded from all downstream analyses and only shown in 2c, or are they only excluded in 2d?

If they are only excluded in 2d, I suggest including them for consistency across metaprograms. If there is a reason for their exclusion from the main figure, an additional supplementary figure including the cycling cells would be helpful. Also, please clarify whether cycling cells are included or excluded in Extended Data Figures 2e–f.

Line 1049: I think this could be a typo for the name of the R-package. Should it not be ‘stats’ instead of ‘stast’? I think it is

present at more places in the manuscript.

Lines 1158–1161: The statistical strategy should be revised. From the methods, it appears that the Kruskal–Wallis test may not have been applied correctly (e.g., Figure R55). The Kruskal–Wallis test is a global test that only indicates whether at least one group differs; it does not specify which groups are different. To identify specific group differences, a significant Kruskal–Wallis result should be followed by post hoc pairwise Wilcoxon rank-sum tests with appropriate multiple-comparison correction.

A recommended framing would be: “If one or more groups did not follow a normal distribution (Shapiro–Wilk p-value < 0.05), we used the non-parametric Kruskal–Wallis test for comparisons across more than two groups. If the Kruskal–Wallis test was significant, we performed post hoc pairwise comparisons using the Wilcoxon rank-sum test with appropriate (e.g., Bonferroni) multiple testing correction.”

Extended Data Figure 1d (Figure R26): Please define the abbreviations CNA and NM in the figure legend. If CNA and CNV refer to the same concept, consider using only one term for consistency and easier interpretation.

I was unable to access the deposited datasets. For publication, just a reminder to please make sure to include the annotation metadata of all the metaprograms and meta-niches the spatial and sc/snRNAseq data was annotated with in the final manuscript version.

(Remarks on code availability)

Version 2:

Reviewer comments:

Referee #3

(Remarks to the Author)

Dear Authors,

Thank you for the revised version of the manuscript, and the response to my feedback. For me, only one small concern remains before the manuscript is suitable for publication.

Line 260. The authors added the sentence “We did not observe a statistically significant correlation between global structural organization and sampling of the tissue (data not shown).”. However, as it reads now, it is not clear what this text associates to, and in the worst case could be interpreted as evidence against some of the authors main conclusions. It should be clarified that this refers to differences in spatial coherence scores between primary and recurrent tumors, (which is an important negative result to report).

(Remarks on code availability)

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

RE: Nature manuscript 2024-08-16431

“Single-cell multidimensional profiling of tumor cell heterogeneity in supratentorial ependymomas”

Point-by-point response to reviewer comments

We are grateful to the reviewers for their thorough and constructive comments. We appreciate the overall enthusiasm from all three reviewers, and the recognition that this study represents an impactful contribution to the fields of ependymoma biology and pediatric cancer.

The reviewers’ suggestions have led us to perform additional analyses and modify the manuscript. We believe that these changes have strengthened our original manuscript.

We trust that all comments have been adequately addressed through additional experiments or with textual edits. Please see below for a point-by-point response to each reviewer comment.

List of Figures included in this rebuttal letter:

- Figure R1.** Projections of ST-EPN subgroups onto developmental time points in human cortex
- Figure R2.** Comparison of cell type proportions in ZFTA-RELA and ZFTA-Cluster 3 tumors
- Figure R3.** Comparison of neuroepithelial-like and embryonic-like cell states
- Figure R4.** Kaplan-Meier survival plots of top 30 genes for each metaprogram
- Figure R5.** Correlation between the spatial coherence score and the percentage of individual cell states/types in ZFTA-RELA tumor samples profiled by 10X Xenium
- Figure R6.** Comparison of neuroepithelial-like-1 and neuroepithelial-like-2 cell states
- Figure R7.** Live-cell imaging of ZFTA-RELA cells cocultured with human iPSCs-derived cells
- Figure R8.** scRNA-seq of ZFTA-RELA cells cocultured with human iPSCs-derived cells
- Figure R9.** Immunofluorescence and 10X Xenium of immature ependymal-like cells
- Figure R10.** scRNA-seq of 3 ZFTA-RELA cell lines
- Figure R11.** RNA velocity of ZFTA-RELA cells in coculture and PDX
- Figure R12.** Cell migration speed of cell states in monoculture vs coculture
- Figure R13.** ZFTA-RELA and non-ZFTA-RELA survival curves
- Figure R14.** UMAP of ST-EPN cells/nuclei
- Figure R15.** Amended reference annotations in Figure 1f and Extended Data Figure 1f
- Figure R16.** Dotplot of cell type proportions in reference developmental dataset
- Figure R17.** Spatial coherence score and cell state composition in primary vs recurrent samples
- Figure R18.** Immunofluorescence panel of cell states in coculture with rat E19 cortical cells
- Figure R19.** Gene expression module scores in ZFTA-RELA tumors of external cohort
- Figure R20.** Cycling status of cell states in ST-EPN tumors
- Figure R21.** Quantification and imaging of neuronal-like TM thickness
- Figure R22.** Comparison of TM patterns *in vitro* and *in vivo*
- Figure R23.** Cycling status of cell states in patients and patient-derived models
- Figure R24.** scRNA-seq of ZFTA-Cluster 4 patient-derived cells
- Figure R25.** scRNA-seq of TTX treated ZFTA-RELA cells
- Figure R26.** CNV profiles of ST-EPN frozen and fresh patient tumors
- Figure R27.** Comparison of different NMF ranks
- Figure R28.** Pearson correlation of NMF programs in ST-EPN tumors
- Figure R29.** 10X Xenium of non-ZFTA-RELA tumors
- Figure R30.** Pseudobulk differentially expressed genes of ZFTA-RELA versus ZFTA-Cluster 3 tumors
- Figure R31.** Representative images of ZFTA-RELA cells undergoing cell division
- Figure R32.** Revised oncoplot of profiled ST-EPN samples
- Figure R33.** ST-EPN clustering by sample and subgroup
- Figure R34.** Projection of each ST-EPN tumor sample to developmental timepoints in reference dataset
- Figure R35.** Statistical comparison of metaprogram proportion in each ST-EPN subgroup
- Figure R36.** Correlation to Gojo et al. 2020 study
- Figure R37.** PCA and Pearson Correlation of 500-3000 HVGs across samples
- Figure R38.** Top 30 genes and GO analysis of metaprograms in ST-EPN tumors
- Figure R39.** Correlation of consecutive tumor sections in 10X Xenium
- Figure R40.** Proportion of cell states/types in ZFTA-RELA tumors profiled by 10X Xenium
- Figure R41.** Pearson correlation of niches in ZFTA-RELA tumors by different niches.k values

- Figure R42.** Spatial maps of ZFTA-RELA tumor sections by different niches.k values
- Figure R43.** Overlay of cell types and spatial niches of different niches.k values
- Figure R44.** CellCharter analysis of ZFTA-RELA tumors
- Figure R45.** Differentially expressed genes and GO analysis of ZFTA-RELA cells in monoculture vs coculture
- Figure R46.** Chi-square test of progenitor-like and differentiated-like cells in patient-derived models
- Figure R47.** Shapiro-wilk test of metaprogram module score in ZFTA-RELA cells grown in monoculture versus coculture
- Figure R48.** Schematic of cell types present in human cortical development
- Figure R49.** Projection of ST-EPN subgroups and metaprograms to the reference dataset
- Figure R50.** Graphical summary of ZFTA-RELA cells profiled by sc/snRNA-seq
- Figure R51.** Violin plots of ciliogenic gene expression across ST-EPN subgroups
- Figure R52.** Shapiro-wilk test of gene expression module scores in ZFTA-RELA and ST-YAP1 tumors in external cohort
- Figure R53.** Correlation of percentage of cells per cell state/type in sc/snRNA-seq and percentage of cells per cell state/type in 10X Xenium
- Figure R54.** Shapiro-wilk test of gene expression in ZFTA-RELA cells grown in monoculture versus coculture
- Figure R55.** Shapiro-wilk normality test of cell migration speed values of different cell states

Referee #1:

In this manuscript, Jeong, Danielli, and colleagues examine the developmental roots of supratentorial ependymomas and identify that these tumors harbor two distinct progenitor-like cell populations that differ in their extent of neuron vs. ependymal cell differentiation. The authors use cutting-edge single-cell and spatial transcriptomic techniques, as well as live-cell imaging to define the spatial architecture of these tumors and uncover a role for brain-resident cells in contributing to the heterogeneity of these tumors. Overall, the studies are nicely designed, the data are beautifully presented, and this work addresses an important gap in the field regarding childhood brain cancers. I do have a number of significant critiques that are mostly related to three points: (1) oftentimes the authors have overstated some of their data (including for plots where there is no statistical significance), (2) the *in vitro* & *in vivo* work is well-inspired but raises several concerns, outlined below, and (3) there is little discussion on how ST-EPN subtypes correlate to clinical prognoses, which greatly minimizes the ability to assess the impact of the work. Overall, my enthusiasm is mixed—the strength of the manuscript lies in the descriptive figures, and the more mechanistic findings, although interesting, are far less developed.

We thank the reviewer for their kind feedback and providing excellent suggestions. We aimed to address these concerns, by: (1) performing additional statistical analyses to amend or remove claims that are overstated, (2) strengthening our *in vitro* and *in vivo* work using additional ZFTA-RELA models, human iPSCs-derived neurons and novel coculture systems, and (3) including more discussion regarding the limitations of the study regarding the clinical prognoses of ST-EPN subgroups. We thank the reviewer for helping us improve our manuscript, which unveil for the first time the tumor heterogeneity across ST-EPN subgroups, their spatial localization, and the overlay of transcriptomic and phenotypic characteristics of these different cell states, that the field can use as a resource to build upon.

Major concerns:

#1-1. From Fig. 1F, it appears as though early radial glia do not closely relate to any of the ST-EPN subtypes, but radial glia correlate to ZFTA-Cluster 2, as the authors also state. Apart from developmental time, what are the transcriptomic and functional differences between “early” radial glia vs. radial glia?

We thank the reviewer for this thoughtful comment and question. The reviewer is correct that there are several key differences between early radial glia and radial glia apart from developmental time. Firstly, gene expression patterns are distinct between these two cell types, where early radial glia have upregulated proneural transcription factors upregulated (e.g. *NEUROG2*, *NEURO6D6*, *NEUROD1*)¹ compared to radial glia. Secondly, signaling pathways differ, where early radial glia cells have a conserved gene network similar to Wnt signaling^{1,2}, whereas later-stage radial glia are more enriched in genes involved in gliogenesis (e.g. *GPX3*, *ETV5*). Finally, in contrast to early radial glia, later-stage radial glia cells are composed of distinct subtypes, including ventricular radial glia (vRG), outer radial glia (oRG), and truncated radial glia (tRG), which have distinct spatial localization and signaling pathway activation throughout the developing cortex³. While we have merged the gene expression profile of these three subtypes of radial glia into a single later-stage “radial glia” cluster for our analyses, the diversification of later-stage radial glia highlights the dynamic differences that occur as radial glia mature through development. The functional differences across radial glia subtypes are actively ongoing studies in the neurodevelopment field.

#1-2. From Fig. 2D, it appears as though ZFTA-Cluster 2 tumors are mainly composed of neuroepithelial / MES / neuron-like cells, but not radial glia, which is contradictory to the findings presented in Fig. 1F and lines 136-138, wherein the authors show and state that ZFTA-Cluster 2 tumors correlate with radial glia (in addition to neuroepithelial cell types) from the reference atlases. Similarly, Fig. 2D suggests that ZFTA-Cluster 3 and 4 tumors are high in embryonic neuron-like populations, but data in Fig. 1F (and lines 138-139) suggests that these tumors are mainly mesenchymal. How do the authors explain these differences?

We thank the reviewer for pointing out this apparent inconsistency between findings of Fig.1F and Fig. 2D. This comes from fundamental differences in analysis methods, as outlined below:

1. Fig. 1F depicts correlation based on aggregate/ “pseudobulk”. First, single cell counts of each tumor subgroup were aggregated (“pseudobulk”), which were then scored by the top 30 marker genes of cell types in the reference dataset (Nowakowski et al., 2017 + Eze et al., 2024)^{1,4}. This is labeled “expression score (tumor cells)”. Similarly, single cell counts of each cell type in the reference dataset were pseudobulked and scored by the top 30 differentially expressed genes of tumor subgroups in the tumor sc/snRNAseq data. This is labeled “expression score (normal cells)”. The dotplot in Fig. 1F therefore displays correlations of samples aggregated and projected to the developmental dataset, representing **an aggregate** confounded by the heterogeneity of cell states of each patient per subgroup.
2. For Fig. 2D, we performed NMF analysis of tumor sc/snRNAseq data and identified recurrent gene expression programs across samples (“metaprograms”). While each development-like metaprogram was annotated by the comparison of gene signatures to the developmental reference scRNA-seq dataset as well as gene ontology (GO) analysis, these are *de-novo* identified malignant programs in the tumor. As such, the plot in Fig. 2D more accurately reflects the true malignant cell composition across tumor samples.

Due to the confusing and sometimes conflicting findings of these two figures, we have **removed** the pseudobulk projection of tumor subgroups to cell types in the reference dataset and only retained the projection to time (**Figure R1**). We have also respectively revised the manuscript results section to explain these discrepancies.

Figure R1

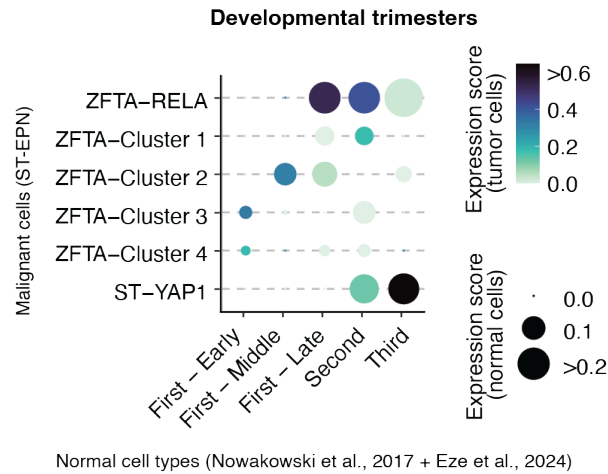


Figure R1. Projections of ST-EPN subgroups (y-axis) onto developmental time points found across the developing human cortex (x-axis). First-Early: post-conception week (PCW) 4, First-Middle: PCW 5-6, First-Late: PCW 7-12. PCW, post-conceptional week.

Results: Next, we projected gene expression profiles of non-malignant cells from the reference atlas^{27,28} onto individual ST-EPN cells (**Figure 1f, Extended Data Figure 2a**). While ZFTA-Cluster 3 and ZFTA-Cluster 4 tumors projected to the earliest time window of early and middle first trimesters, ZFTA-Cluster 2 and ZFTA-Cluster 1 tumors mapped onto the later windows of middle first and second trimesters. Finally, canonical ZFTA-RELA tumors projected broadly from the late first to second, and third trimester, while ST-YAP1 aligned with the latest timepoints. (**Figure 1f, Extended Data Figure 2a**). Overall, the correlation of ST-EPN subgroups to different developmental stages suggest divergent developmental patterns across subgroups.

Changes made: Revised Figure 1e, Results

#1-3. It is interesting that despite harboring the same ZFTA-RELA fusions, the ZFTA-Cluster 3 and ZFTA-RELA tumors have distinct cell populations. Could this be a result of 7 out of the 8 ZFTA-Cluster 3 tumors being recurrent, whereas the distribution of primary vs. recurrent tumors was more or less equal in the ZFTA-RELA tumors? Alternatively, could the differences be microenvironment-mediated?

We thank the reviewer for the interesting suggestion. According to the reviewer's hypothesis, we would expect recurrent ZFTA-RELA tumors to have a similar cell state composition to ZFTA-Cluster 3 tumors (as the samples that we profiled from this group are mostly derived from recurrent tumors). Specifically, we would expect to see an increase in embryonic-like and embryonic-neuronal-like programs that are highly enriched in ZFTA-Cluster 3 tumors. Therefore, we performed statistical comparison of primary and recurrent ZFTA-RELA tumors (**Figure R2a**).

We used the propeller function in R package *Speckle* (v1.8.0)⁵, a method for performing statistical tests on compositional data, to determine statistical significance in cell proportions of primary and recurrent patient samples for ZFTA-RELA (**Figure R2a**). Briefly, this function fits a linear model for each cell type, using the group (in our case, primary versus recurrent) as the explanatory variable. It then performs Bayes shrinkage of the variances and applies multiple testing correction to obtain false discovery rates (FDR). Applying this method to our dataset, we find no statistical difference in the cell type proportions of primary versus recurrent samples in ZFTA-RELA patient samples, and specifically do not see an increased frequency of embryonic-like and embryonic-neuronal-like programs in the recurrent samples. This suggests that ZFTA-RELA and ZFTA-Cluster 3 tumors have distinct cell populations that do not seem to be associated with the sampling status of these tumors.

Furthermore, to examine whether these cell type compositions are driven by microenvironmental differences, we used the same statistical method to compare the non-malignant cell type proportions in sc/snRNA-seq data of ZFTA-RELA and ZFTA-Cluster 3 tumors (**Figure R2b**). Again, we do not find any statistically significant difference in proportion of non-malignant cells between the tumor subgroups in our sc/RNAseq data (**Figure R2b**). This analysis suggests that there are no significant microenvironmental differences between ZFTA-RELA and ZFTA-Cluster 3 that could drive the distinct cancer cell populations. As sc/snRNA-seq is limited in the number of cells captured (ZFTA-RELA: $n = 301$ cells, ZFTA-Cluster 3: $n = 263$ cells), we additionally performed preliminary 10X Xenium on two ZFTA-Cluster 3 samples (STEPN-33, STEPN-34) to capture the non-malignant population in both tumors (ZFTA-RELA: $n = 990,893$ cells, ZFTA-Cluster 3: $n = 61,437$ cells) (**Figure R2c**). Using the same statistical method, we did not detect significant changes in the proportion of tumor microenvironment-associated non-malignant cell types in ZFTA-RELA versus ZFTA-Cluster 3 tumors.

Overall, our analyses suggest that ZFTA-RELA and ZFTA-Cluster 3 tumors have distinct tumor subpopulations that are likely are not driven by sampling status or significant differences in cell types present in the tumor microenvironment. However, as our analyses are still limited by the number of samples available for ZFTA-Cluster 3 and only examine cell type proportions in the tumor microenvironment, it would be very beneficial for the field to further explore this question with more samples and functional validations (such as introduction of fusion genes *in utero* or explore non-genetic microenvironmental factors) in future studies.

Figure R2

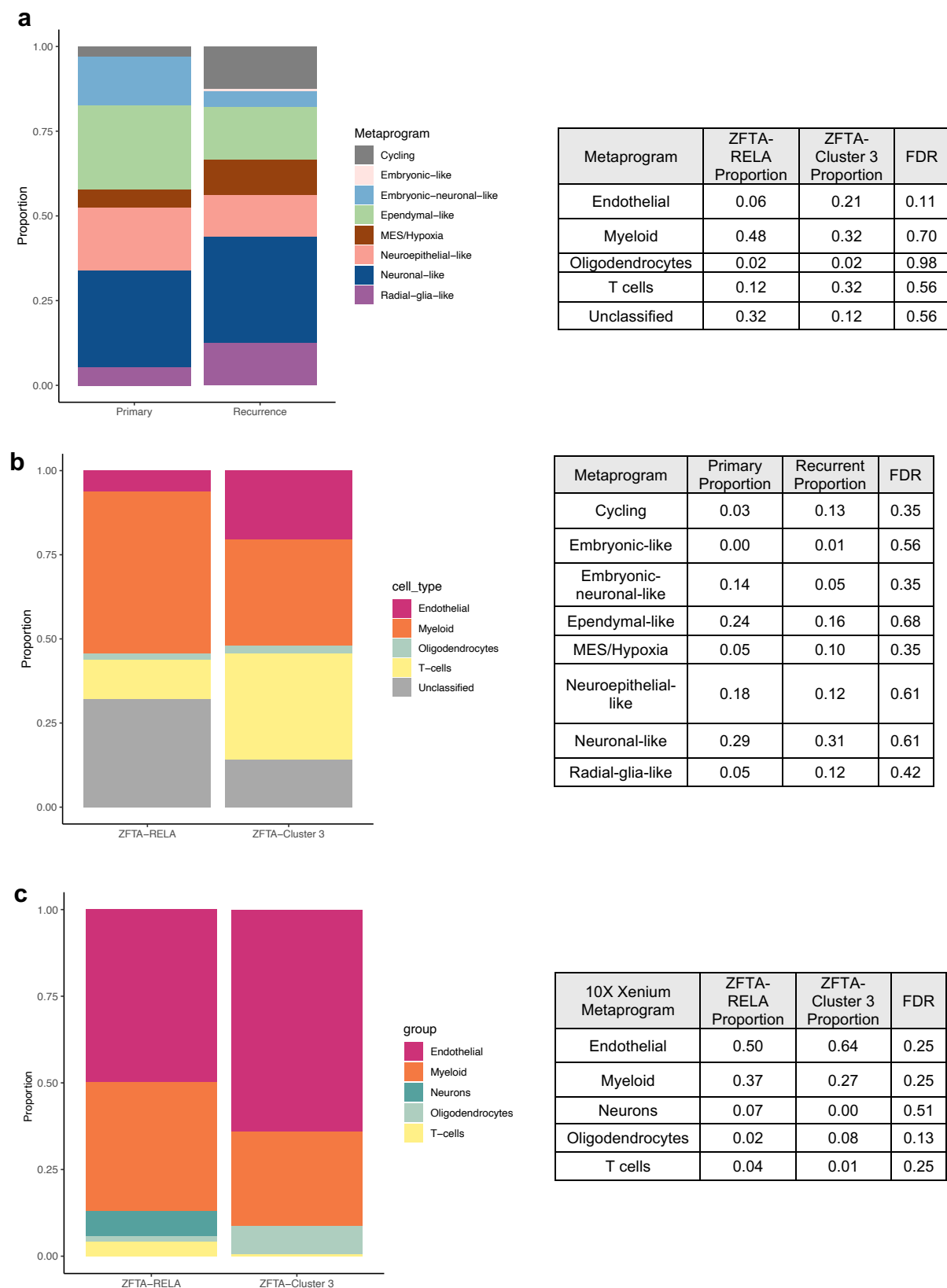


Figure R2. (a) Proportion of metaprograms (y-axis) in primary and recurrent patient samples (x-axis) in ZFTA-RELA subgroup (left). Statistical significance of cell proportions by primary and recurrent tumors in ZFTA-RELA using the propeller function from the R package *speckle* (right). **(b)** Proportion of non-malignant cell types (y-axis) in ZFTA-RELA ($n = X$ samples, $n = 301$ cells) and ZFTA-Cluster 3 ($n = X$ samples, $n = 263$ cells) patient tumors analyzed by sc/snRNAseq (x-axis)

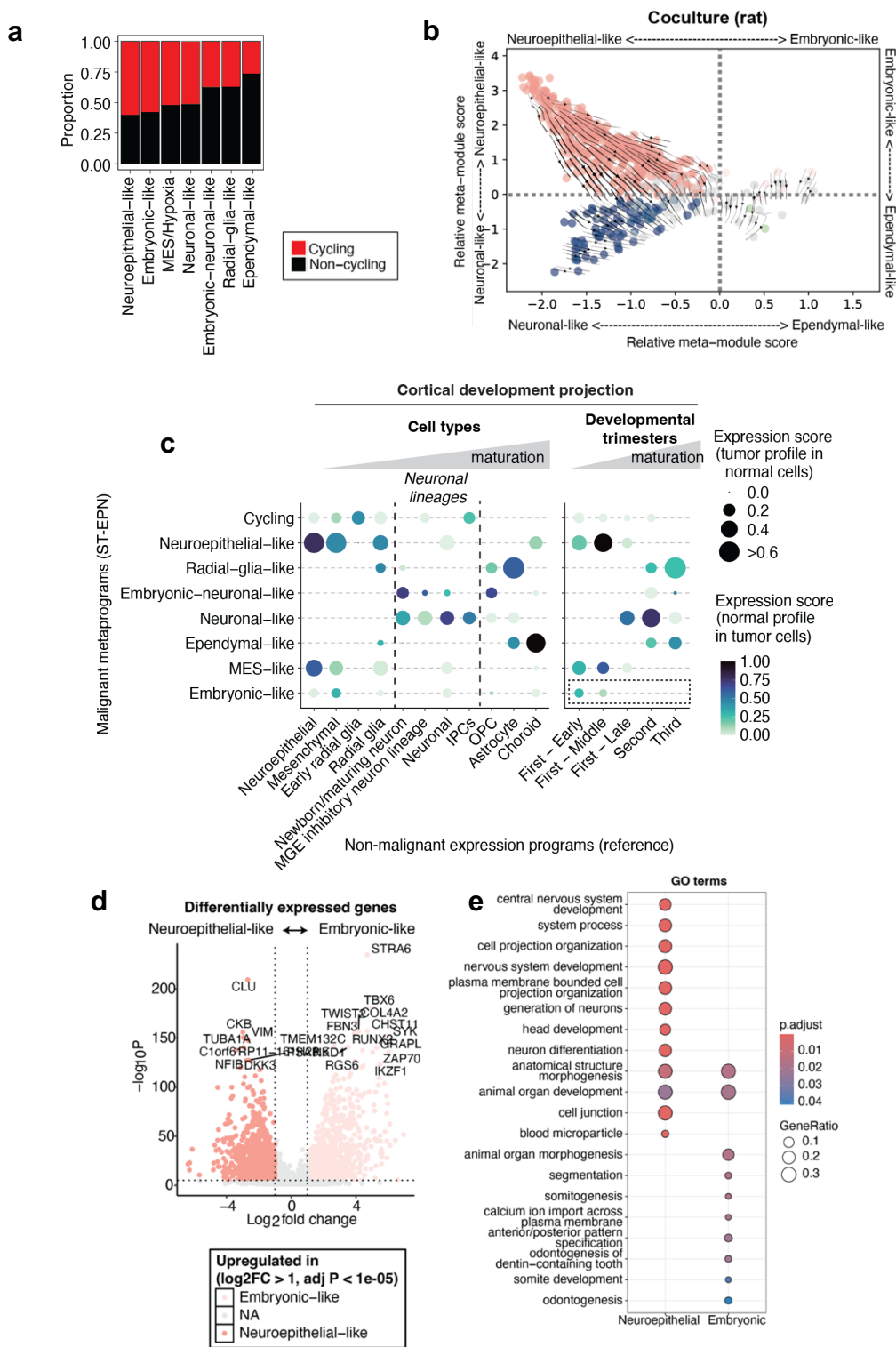
(left). Statistical significance of non-malignant cell type proportions in ZFTA-RELA and ZFTA-Cluster 3 using the propeller function from the R package *speckle* (right) (c) Proportion of non-malignant cell types (y-axis) in ZFTA-RELA ($n = 990,893$ cells) and ZFTA-Cluster 3 ($n = 61437$ cells) patient tumors analyzed by 10X Xenium (x-axis) (left). Statistical significance of non-malignant cell type proportions in ZFTA-RELA and ZFTA-Cluster 3 using the propeller function from the R package *speckle* (right).

#1-4. This study raises several important questions about the main finding that there are two main progenitor-like cell states in ST-EPNs. For the two progenitor-like cell states (neuroepithelial vs. embryonic-like), can the authors elaborate on what is known about their role, function, and presence during normal fetal development? What are the differentiation trajectories of these cells? How do they differ from each other at the transcriptomic and functional levels? Is it known whether (and how) they give rise to the two mature lineages identified in this work (neuronal vs. ependymal)? Some of these would be important to examine to make the earlier half of the manuscript less “descriptive”.

We thank the reviewer for these important questions. In normal development, neuroepithelial cells are foundational stem cells in the developing central nervous system during normal fetal development^{3,6}. These cells serve as primary neural stem cells during early embryogenesis, constituting the neural tube from weeks 4-6 in humans⁴. They are highly proliferative, undergoing symmetric divisions that allow for self-renewal and consequently expansion of the neural tube⁷. In addition, they undergo extensive asymmetric divisions that generate radial glia, neurons and various types of glial cells, including ependymal cells⁶⁻⁸. Firstly, we find that the malignant ST-EPN neuroepithelial-like cells mirror these phenotypic characteristics in firstly their proliferative nature (**Figure R3a**). Secondly, we find that neuroepithelial-like cells can transition to other cell states like their developmental counterpart. This is evidenced by ZFTA-RELA patient-derived tumor spheroid models which are mostly neuroepithelial-like in monoculture but acquire more mature cell states when cocultured with rat cells. We also performed RNA-velocity⁹ on single cell RNA-seq data of 3 fresh patient-derived ZFTA-RELA cell lines cocultured with rat E19 cortical cells, which revealed that in coculture, neuroepithelial-like cells progressed along a trajectory towards more mature cell states (**Figure R3b**). While the complex interplay of transcription factors and signaling pathways-such as Notch, hedgehog, and Olig2-drives neuroepithelial cell differentiation during normal development^{4,10,11}, how tumor neuroepithelial-like cells give rise to other mature lineages is unknown. It would be valuable for future studies to investigate the activation of these pathways and additional mechanisms in the neuroepithelial-like cells of ST-EPN tumors, in conjunction with methods such as lineage tracing.

In contrast to the distinct developmental presence of neuroepithelial-like cells, the embryonic-like cell state does not as clearly correspond to a distinct cell type in normal cortical development. We assigned this nomenclature to the tumor subpopulation based on: (1) their projection to early developmental timepoints in the reference dataset (**Figure R3c**), (2) their high proliferative capacity (**Figure R3a**), and (3) their distinct gene expression profile relative to neuroepithelial-like cells, enriched for **embryonic developmental processes**, such as somite development and organ morphogenesis (e.g. *TBX6*, *DLL1*, *NKDI*, *RUNX2*) (**Figure R3d-f**). Based on these findings, we attributed these cells to be “embryonic-like”, reflecting their acquisition of stem-like progenitor characteristics in the embryonic tissue, with the capacity to both self-renew and generate other cell states/types. Given the absence, to our knowledge, of any non-ZFTA-RELA patient-derived cell lines as of yet, it would be valuable for the field for future work to establish robust non-ZFTA-RELA patient-derived models to further validate the plasticity of the embryonic-like cells and to elucidate the mechanisms driving their cell state transitions. We have added these new analyses to **Extended data Figure 2b,f,g-i**, have amended **Results**, and mentioned these limitations in the **limitations of the study**.

Figure R3



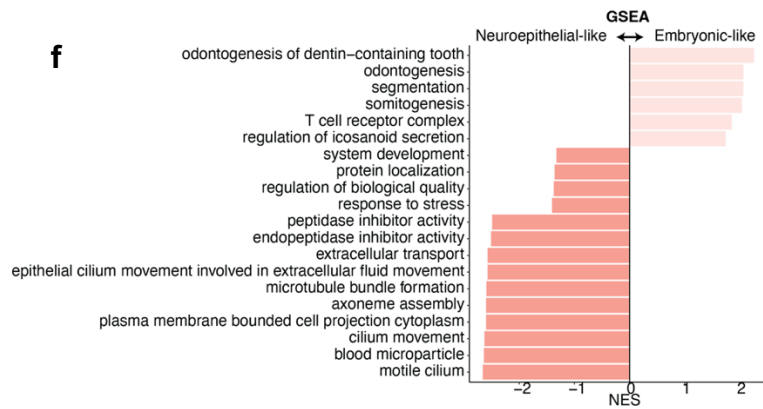


Figure R3. (a) Barplot showing the proportion of cycling and non-cycling cells across ST-EPN cell states, ordered from the most (left) to the least (right) proliferative (b) RNA velocity analysis^{9,12} of ZFTA-RELA cells in monoculture, coculture with rat E19 cortical cells, or PDX models. Cells represented in cell state plot scored by neuroepithelial-like, embryonic-like, neuronal-like and ependymal-like gene signatures. Cells colored by metaprogram assignment. (c) Projection of ST-EPN cell states (y-axis) onto cell types present during the first, second and third trimester of human cortical development^{1,4} (d) Volcano plot showing genes differentially expressed between neuroepithelial-like and embryonic-like ST-EPN cells based on Seurat's *FindAllMarkers* function. Genes with $|\log_2FC| > 1$ and adjusted P-value < 0.00001 are highlighted in color. (e and f) (e) Top 10 gene ontology (GO) terms (f) and pathways identified by GSEA analysis.

Results: Identification of differentially expressed genes between the two progenitor-like populations revealed that while neuroepithelial-like cells were enriched in genes associated with GO terms related to central nervous system development and neural differentiation (Extended Data Figure 2g-i), embryonic-like cells were enriched in genes associated with early embryonic developmental processes, including somite development (Extended Data Figure 2g-i)⁴⁶.

Methods: Differential gene expression analysis. We performed pseudo-bulk differential expression analysis to identify transcriptional programs that distinguish ZFTA-RELA from ZFTA-Cluster 3 tumors. We first extracted raw count matrices from each tumor sample, and calculated pseudo-bulked gene expression counts using the SingleCellExperiment's *aggregateAcrossCells* function. We then conducted differential gene expression analysis using edgeR. We filtered out genes with low expression counts using the *filterByExpr* function, based on our grouping variable (tumor subtype). We then normalized the counts using the *calcNormFactors* function. We constructed a design matrix to include tumor subtype, estimated negative binomial dispersions using the *estimateDisp* function, and fitted a generalized linear model using *glmFit*, followed by hypothesis testing. **Gene Ontology (GO) and Gene Set Enrichment (GSEA) analysis.** To identify biological terms that are over-represented in a specific gene set, we performed GO and GSEA analysis using *enrichGO*, respectively *gseGO* functions from the *clusterProfiler*¹³ package. As a database, we used the R package *org.Hs.eg.db* (v. 3.18.0), which is an organism annotation package for *Homo sapiens*, relying on the Ensembl version released on 2023-May10. Unless otherwise states, we used the following parameters: *ont* = "ALL", *universe* = all genes retained for differential gene expression analysis, *minGSSize* = 3, *maxGSSize* = 800, *pvalueCutoff* = 0.05, *OrgDb* = *org.Hs.eg.db*, *pAdjustMethod* = "BH".

Changes made: Extended Data Figure 4d-f, Results, Methods

#1-5. The relevance of data presented in Ext. Fig. 3 and Table S6 to this manuscript is unclear. Furthermore, the authors state in lines 173-175 that therapeutic targets “in the context of metaprograms correlated with lower survival” were uncovered. This raises the question of what metaprograms, of the 8 that were identified, were the ones correlated with lower survival? Although this latter statement is important from a clinical perspective, it was not explained in detail or substantiated with experimental work.

To address the correlation of metaprograms with survival, we evaluated the top 30 genes of 8 metaprograms in a bulk RNA-seq dataset of 46 ZFTA-RELA patients of an external cohort (Open Pediatric Brain Tumor Atlas). Using a method previously described¹⁴, we estimated the survival using univariate models calculated by the Kaplan-Meier Method and tested for significance by the log-rank test. Expression of metaprogram gene signatures were split into low and high expression based on the bottom and top quartiles, respectively. Due to the small patient cohort size ($n = 46$), most of the survival correlations were statistically

not significant except for the Cycling metaprogram (**Figure R4**). As we do not have clear correlation with metaprograms and lower patient survival, we have removed this statement in the manuscript.

In addition, we agree with the reviewer that the relevance of previous Ext. Fig 3 and Table S6 is not clear as these results are not further explored or connected with the rest of the manuscript. As a result, we have removed Table S6.

Figure R4

[FIGURE REDACTED]

Figure R4. [TEXT REDACTED]

#1-6. Please provide correlational plots between the spatial coherence score and frequency of each of the other malignant cell states/types (and other non-malignant cell states, too), as was done for the MES/hypoxia cell states and vascular cells in Fig. 3C.

& #1-7. For these correlational plots (including Fig. 3C and Ext. Fig. 6E), please also display the “r” value to show the strength and directionality of the correlation, which is not fully captured by the “r²” value.

& #1-8. Related to the previous critiques – although the authors nicely show that a higher spatial coherence score (i.e., tendency to be a “structured” tumor) is correlated with a higher frequency of MES/hypoxia cell states, it is important to note that MES/hypoxia cell states are not the dominant cell states in any of the structured tumors. In fact, MES/hypoxic cell states comprise only ~6% of the tumor with the second highest spatial coherence score (STEPN-01). Reporting the “r” value as noted in #7, above, would help clarify whether or not the MES/hypoxic cells are strongly correlated with a tendency to be a more structured tumor, although it appears that this link may be overstated and likely driven by a single sample (STEPN-19). However, there may be other cell states that are more likely to be correlated with being a structured tumor, in which case the authors should show the plots, r-values, and p-values correlating these other cell states and the coherence score, as noted in #6 and #7, above.

We thank the reviewer for these constructive suggestions. To address the points made in #1-6~8 and increase rigor of the study, we expanded our spatial cohort by adding **33** new ZFTA-RELA 10X Xenium samples of **13** new patients (from $n=23$ to $n=56$ sections; $n=10$ to $n=23$ patients). We also amended the cell annotation method to include label projection of previously defined tumor cell states and/or non-malignant cell types from patient tumor sn/scRNA-seq dataset for more unbiased and time-efficient analysis across all samples (**Methods**).

With this increased spatial cohort and amended annotation methods, we confirm our previous findings that MES/Hypoxia is still significantly correlated with the global spatial coherence score across all samples ($R=0.68$, adjusted p.value = 0.0047 by Bonferroni correction) (**Figure R5a**), while no other metaprogram or non-malignant cell type correlates significantly with the global spatial coherence score (Figure R5a). We also corrected the linear regression analysis by: (1) addition of “R” value to show the strength and directionality of the correlation, (2) corrected P-values for multiple comparisons using Bonferroni correction, (3) calculated the mean value of samples that have multiple sections.

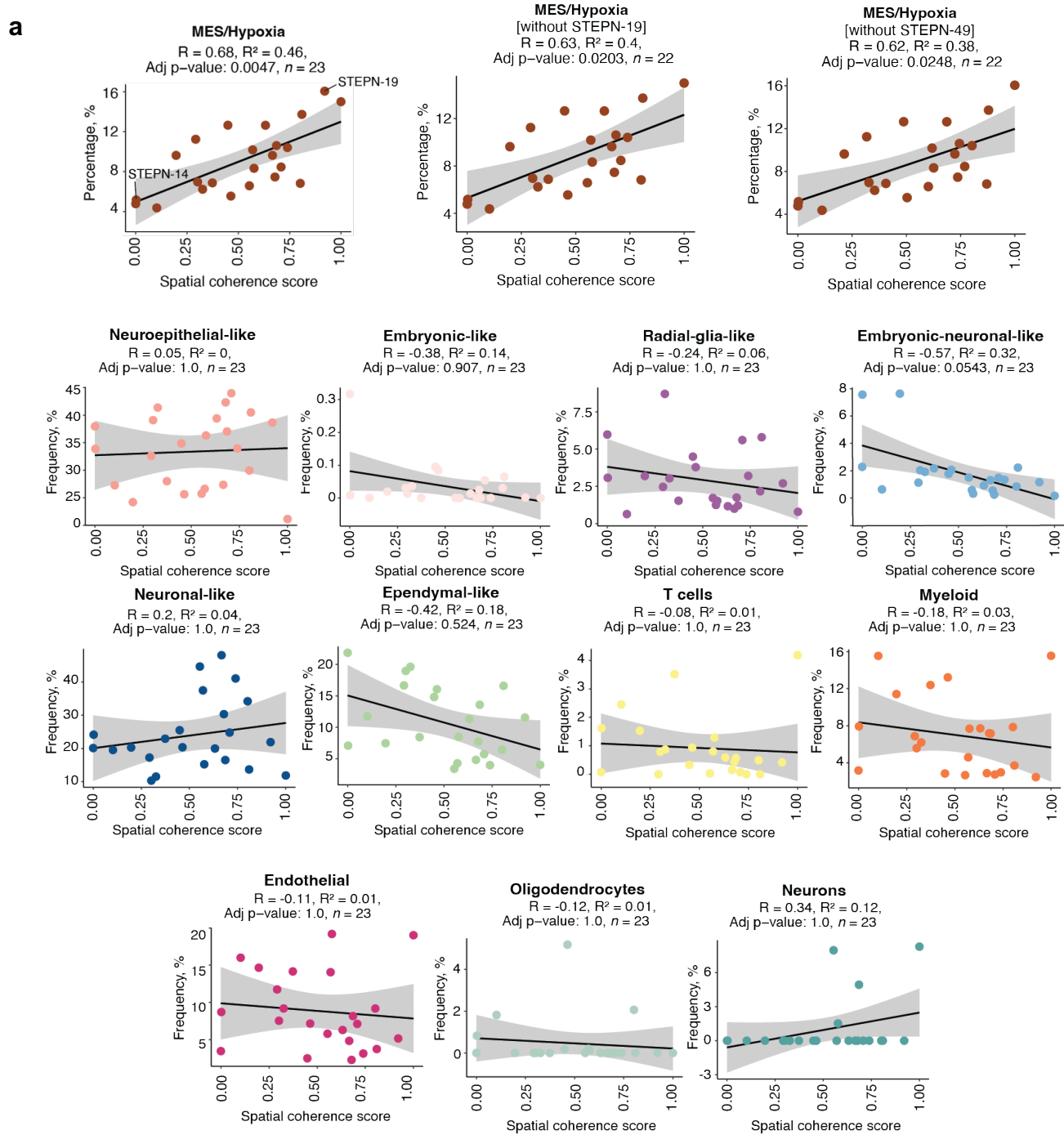
Furthermore, to address the reviewer’s comment that these results may be driven by a single sample, we removed STEP-19 and found that the correlation between MES/Hypoxia and spatial coherence correlation was still significant (adjusted p value = 0.0203). We also repeated the analysis after removing STEP-49, the most coherent sample out of our new spatial cohort and found similar results (adjusted p value = 0.0248). (**Figure R5a**).

As the reviewer astutely pointed out, while MES/Hypoxia is not the most dominant cell states in the tumors (**Figure R5b**), our revised analyses still indicate that the proportion of MES/Hypoxia and spatial coherence is still significantly correlated. Furthermore, when the spatial coherence is calculated for each cell state/type (**Figure R5c**), MES/Hypoxia scores highly. This shows that while MES/Hypoxia is not the most dominant program, cells of this transcriptional state both tend to cluster with themselves as well as correlate with the overall structure of the tumor.

Finally, to further solidify the association between MES/Hypoxia and global spatial coherence, we removed MES/Hypoxia cells from the calculation of mean spatial coherence of the samples and found that the samples still significantly correlated with a higher spatial coherence with increasing MES/Hypoxia (**Figure R5d**). This suggests that the increased spatial organization associated with MES/Hypoxia is likely beyond the cancer metaprogram itself, as also found in adult gliomas¹⁵.

In conclusion, we examined the correlation between each cell state/type and spatial coherence across a larger cohort of samples and continued to find that MES/Hypoxia is significantly correlated with global structure. In addition, we found that the MES/Hypoxia cell state is likely influencing the structure of the tumor beyond its own self. We have amended our findings to **Figure 3d-e**, **Extended Data Figure 4e-f**, **Results and Methods**.

Figure R5



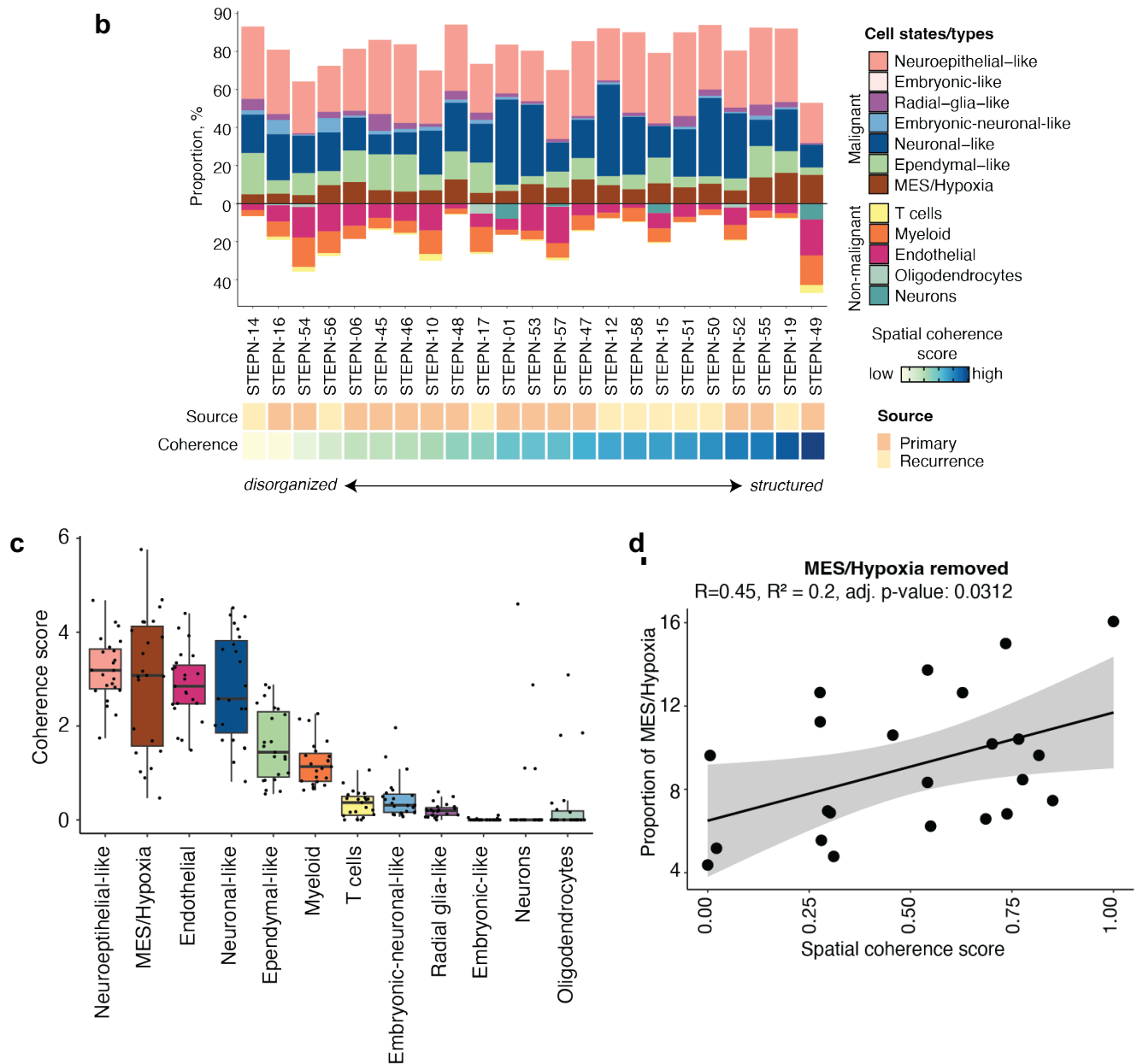


Figure R5. (a) Correlation between the spatial coherence score and the percentage of individual cell states/types in ZFTA-RELA tumor samples profiled by 10X Xenium. Data points (representing each tumor sample) are interpolated with a simple linear regression. For tumor samples with multiple sections, the average values are shown. The goodness of fit (R , R^2), Bonferroni-corrected P-values and total number of tumor sections (n) are displayed at the top. All samples included (left), STEP-N-19 is removed (middle), or STEP-N-49 is removed (right). **(b)** Proportion of malignant metaprograms and non-malignant cell types (y-axis) in the ZFTA-RELA patient tumors profiled by 10X Xenium (x-axis). For patients with multiple tumor sections analyzed, the average proportion is shown. Samples are ordered by increasing spatial coherence score. **(c)** Spatial coherence score of each metaprogram and non-malignant cell type in ZFTA-RELA tumors. **(d)** Correlation between the proportion of MES/Hypoxia and global spatial coherence score of ZFTA-RELA tumor samples when MES/Hypoxia cells are removed from the spatial coherence score calculation. For tumor samples with multiple sections, the average values are shown. The goodness of fit (R , R^2), Bonferroni-corrected P-values and total number of tumor sections (n) are displayed at the top.

Results: To determine whether the presence of a particular tumor state/type correlates with spatial organization, we performed linear regression analyses between the spatial coherence score of each sample and the percentage of cells observed for each malignant cell state or non-malignant cell type (**Extended Data Figure 4e**). This analysis revealed that higher spatial coherence scores were positively correlated with the percentage of MES/Hypoxia cells ($R = 0.67$, Bonferroni-corrected P-value = 0.0047, $n = 23$, **Figure**

3d, Table S7) out of all the cell states/types (**Extended Data Figure 4e**), suggesting that the presence of this malignant cell state is specifically associated with *structured* tumors. In addition, to distinguish whether global spatial coherence was driven simply by the organization of MES/Hypoxia cells themselves, we removed MES/Hypoxia cells from the calculation of spatial coherence and correlated with the proportion of MES/Hypoxia in the tissue. We found that samples enriched with MES/Hypoxia cells significantly correlate with higher spatial coherence even after removing MES/Hypoxia cells from the spatial coherence calculation ($R = 0.45$, $p = 0.03$, **Extended Data Figure 4f**). These findings suggest that hypoxia might exert an external effect on the overall organization of the tumor itself, as also observed in adult gliomas²⁵.

Methods: Xenium cell state/type cell assignment. To assign cell identities in Xenium data, we first performed unsupervised clustering using *Seurat* (v5.1.0), performing *RunUMAP* on the least number of PCs capturing 80% or more variance, shared-nearest-neighbor graph using *FindNeighbors*, and Louvain clustering (*FindClusters*) run at resolutions 0.1 to 1.0, in 0.1 increments. We then scored only for the normal cell types, by transferring labels from the single cell/nuc reference object from corresponding tumor subgroup using *FindTransferAnchors* and *TransferData* (dims = 20), yielding a label-transfer score for every cell-to-cell type pair. Independent module scores were calculated using *AddModuleScore_UCell* using marker genes present in the Xenium panel. For each cluster, average label-transfer score and module score were manually inspected and visualized through UMAP, and the clusters corresponding to normal cell types were identified in each sample. Cells belonging to the normal cell type annotation was removed, and malignant subset of each sample was annotated in the same way as normal cell types. To mitigate systemic score inflation, scores for each malignant cell type were mean centered across all cells: $centered_score_{ij} = score_{ij} - \overline{score}_{.j}$, where i corresponds to cells and j corresponds to cell types. Finally, each malignant cell was assigned with the cell type with highest mean centered score, and malignant and normal cells were combined to generate the final, fully annotated cells.

Changes made: Figure 3, Extended Data Figure 4, Results and Methods

#1-9. The correlational plot between vascular cells and the spatial coherence score in Fig. 3C may show an inverse directionality, but the p-value is highly non-significant, so the authors should not over-interpret these findings – particularly in the absence of r-values to depict the strength of the link. Please temper the language in the Results section accordingly.

Indeed, with the addition of new samples, the correlation between vascular (endothelial) cells and spatial coherence remains non-significant. We have removed this claim in **Results section: “ZFTA-RELA tumors segregate into globally structured and disorganized patterns”**

#1-10. Ext. Fig. 10E: How are neuroepithelial-like 1 different from neuroepithelial-like 2, at the transcriptomic and functional levels? It appears that there is a sharp decline in neuroepithelial-like 1 cells in coculture, but a concomitant increase in neuroepithelial-like 2 (in addition to increases in EPN-glia and neuronal-like cells).

We thank the reviewer for this interesting question. To address this point, we compared the gene expression profiles and phenotypic characteristics of the two cell states.

Firstly, we applied 10X Xenium to simultaneously assess cell morphology and targeted gene expression (**Figure R6a**). Using the same Xenium panel ($n = 358$ genes) previously used for patient tumor analysis, we confirmed that both neuroepithelial-like-1 and neuroepithelial-like-2 cells express cell cycle-related genes (neuroepithelial-like-1: *CENPF*, *TOP2A*; neuroepithelial-like-2: *TUBA1A*, *PRDX1*) (**Figure R6b**). Notably, neuroepithelial-like-2 cells additionally expressed development-associated genes including *SOX4*, *NES*, and early neuronal-like marker *DCX*. The acquisition of these progenitor-like and neuronal-like genes indicates that neuroepithelial-like-2 cells are transcriptionally more mature than neuroepithelial-like-1. This interpretation is further supported by the increased abundance of neuroepithelial-like-2 cells-alongside other mature cell states-in coculture, where interactions with normal brain-resident cells promote tumor cell plasticity towards more mature cell states.

Furthermore, we quantified somatokinesis, or cell migration speed, and the percentage of dividing cells for each morphological subtype (**Figure R6c-d**). In coculture with rat E19 cortical cells, we found that neuroepithelial-like-2 cells are both significantly more migratory and more proliferative (~2.5 fold) than neuroepithelial-like-1 cells (**Figure R6c-d**). The acquisition of migratory behavior further suggests that

neuroepithelial-like-2 cells represent a transitory cell state between the more immature, proliferative neuroepithelial-like-1 cells and the more mature, migratory neuronal-like cells. Notably, their simultaneously highly proliferative nature underscores their phenotypic distinction from all other cell states, as they exhibit both strong proliferative and migratory capacities. These complementary phenotypes could enable neuroepithelial-like-2 cells to have critical functions in tumor progression, driving both tumor expansion and facilitating invasion into other brain regions.

Overall, these new analyses suggest that these cell states may coexist along a continuum of transcriptional and phenotypic characteristics. We thank the reviewer for their insightful observation that helped us to uncover new findings. We have updated the manuscript accordingly to incorporate these results in **Figure 5, Extended Data Figure 9, Table S8**.

Figure R6

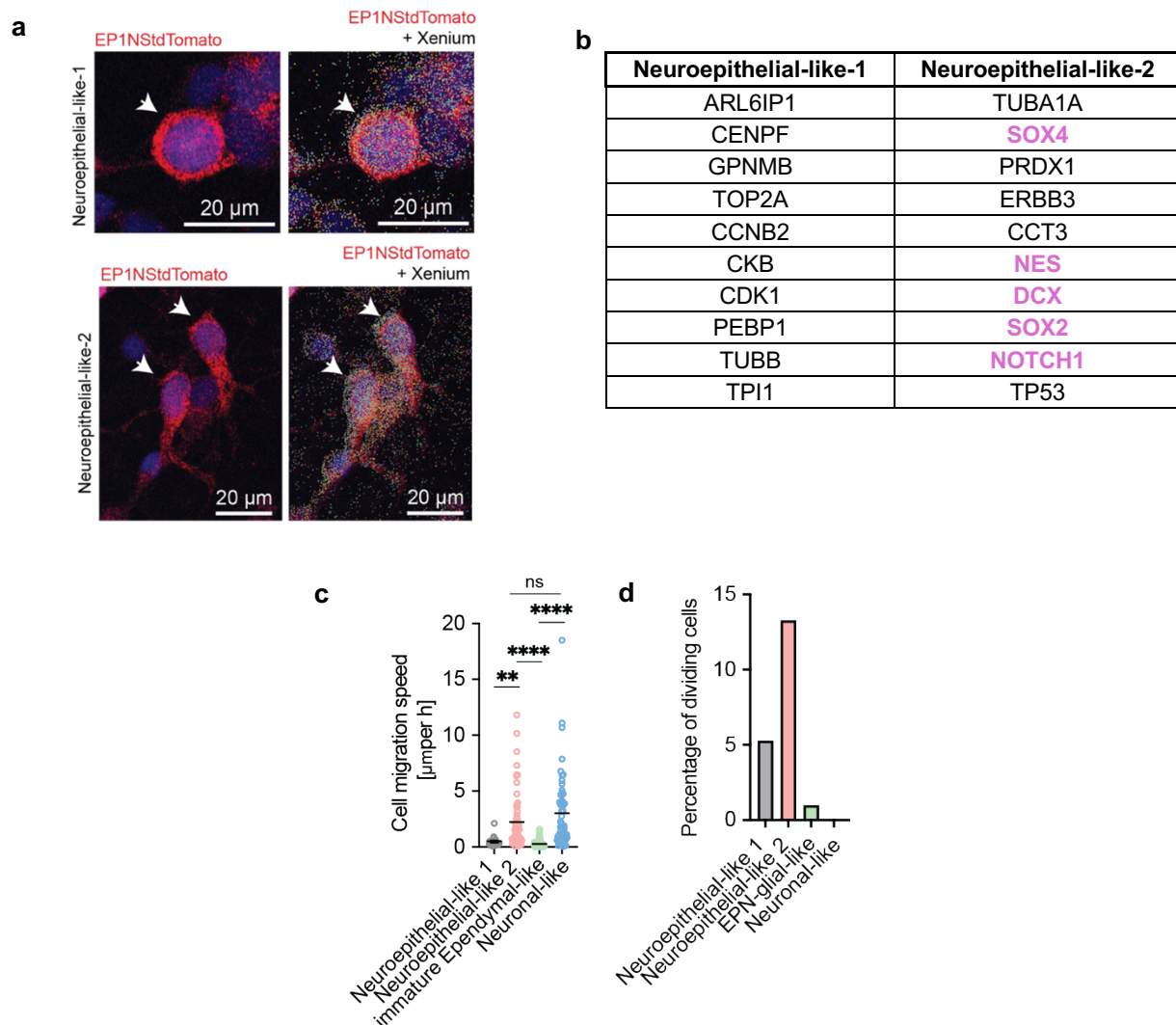


Figure R6. (a) Representative 10X Xenium analysis images of neuroepithelial-like-1 and neuroepithelial-like-2 EP1NS-tdTomato cells in coculture with rat E19 cortical cells. Left panels depicts immunofluorescence images of tdTomato, and right panels represent Xenium transcripts overlayed with immunofluorescence image, where each dot represents detected transcript from the custom ($n = 103$) and base panel ($n = 255$) probes used for patient tumor sections. **(b)** Table showing the unique and shared genes detected on 10X Xenium of each cell state in coculture. **(c)** Cell migration speed ($\mu\text{m}/\text{hour}$) of ZFTA-RELA cells in coculture with rat E19 cortical cells determined using timelapse imaging *in vitro*. Cells are categorized by morphological subtype ($n = 235$ cells acquired from $n = 3$ independent biological replicates). Individual data points with mean $\pm$ SE are shown. P values calculated by Kruskal-Wallis test. **(d)** Percentage of dividing cells quantified for each morphological subtype ($n = 235$ cells acquired from $n = 3$ independent biological replicates)

Results: Using this approach, we subclassified neuroepithelial-like cells into two groups: cells lacking TMs (neuroepithelial-like-1) and cells with 1-3 short TMs (neuroepithelial-like-2). While both groups expressed cell cycling markers (neuroepithelial-like-1: *ARL6IP1*, *CENPF*; neuroepithelial-like-2: *TUBA1A*, *PRDX1*) (**Figure 5c**, **Table S8**), neuroepithelial-like-2 cells expressed additional development-associated markers such as *NES* and *DCX* (**Figure 5c**, **Extended Data Figure 9a**).

Discussion: We further identify a distinct population-neuroepithelial-like-2-that possesses both migratory and proliferative capacity. These distinct behaviors of cellular states highlight different forces promoting tumor malignancy, where neuronal-like cells promote tumor migration while neuroepithelial-like cells maintain proliferation, and neuroepithelial-like-2 cells drive both functions.

Changes made: Figure 5, Extended Data Figure 9, Table S8, Results, Discussion

#1-11. Although the coculture data are quite interesting, the authors performed cocultures of patient-derived tumor cells with rat neurons/astrocytes, thereby (potentially) skipping species-specific differences in microenvironmental cues. Ideally these experiments could be performed with human iPSC-based neurons/astrocytes.

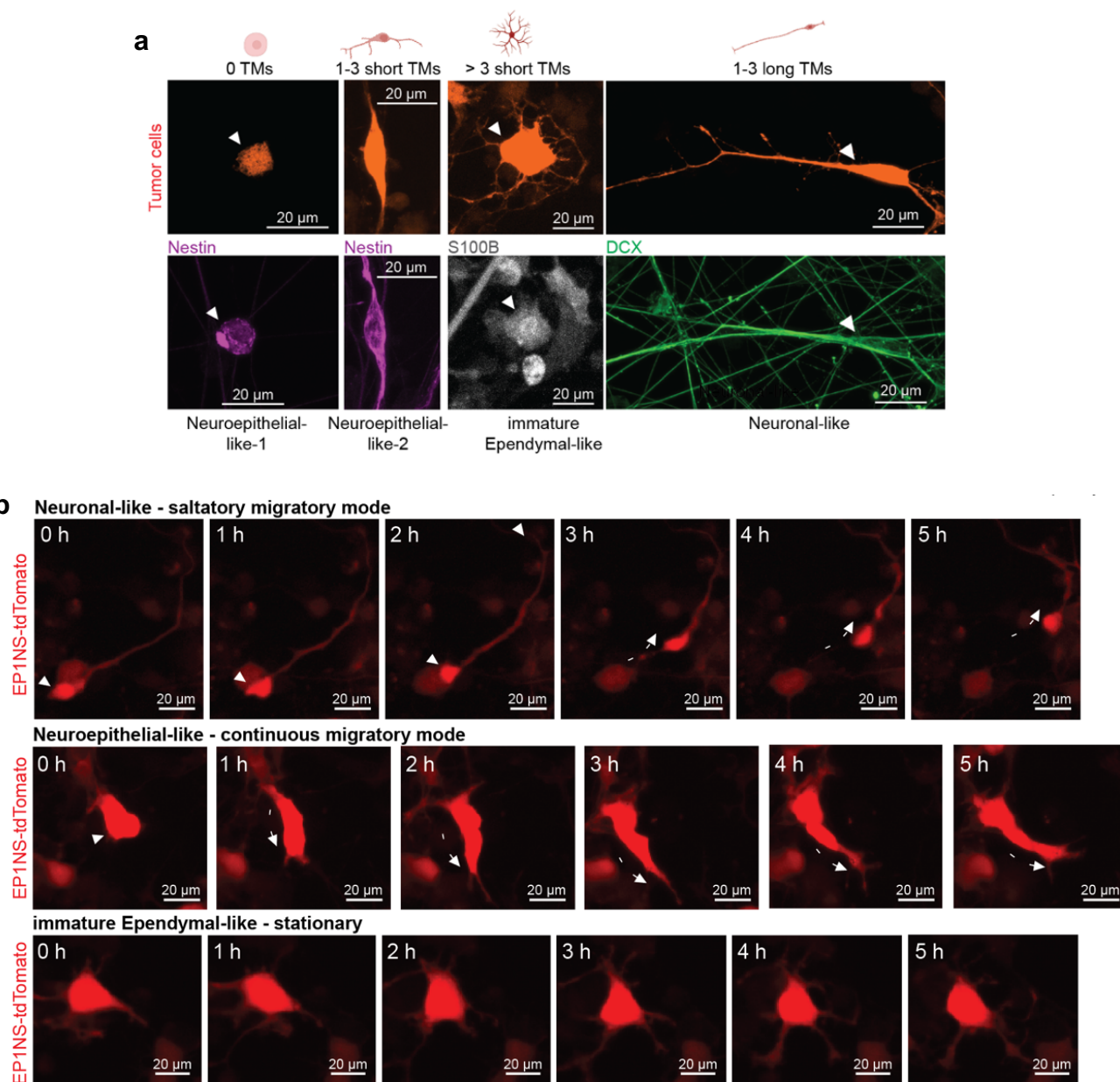
We thank the reviewer for this thoughtful suggestion. To address this point, we repeated the coculture experiments of patient-derived ZFTA-RELA tumor cells with human iPSC-derived neurons and astrocytes (**Figure R7**, **Extended Data Figure 10**). Briefly, human iPSCs were differentiated into glutamatergic neurons for 18 days, and cocultured with human iPSC-derived astrocytes at 10-15% of the total neuronal cell density. We then cocultured human iPSC-derived neurons and astrocytes with tdTomato labeled tumor cells for 7 days and performed both 10X scRNAseq and live-cell imaging to capture the molecular and phenotypic changes of tumor cells in coculture with human normal cells (**Figure R7-R8**).

Firstly, we found that the different morphologies of cell states and their migratory behavior are reproducible in the human cocultures (**Figure R7**). As seen in ZFTA-RELA tumor cells cocultured with rat E19 cortical cells, we found four different morphological subtypes of tumor cells in human cocultures (tumor cells with 0 tumor microtubes, or TMs; 1-3 TMs, >3 short TMs, and 1-3 long TMs) (**Figure R7a**). We performed immunofluorescence to confirm the identity of these tumor cells as neuroepithelial-like, *immature ependymal-like* (previously named “EPN-glial-like”, *please refer to the end of this rebuttal for detailed explanation clarifying the change of nomenclature, pages 21-24*) and neuronal-like, respectively (**Figure R7a**). In addition, we identified different migratory behaviors (cell migration speed) of cell states, including saltatory, continuous, and stationary patterns (**Figure R7b**). Neuronal-like cells specifically exhibited additional immature neuron-like cell behavior, including branching migration, locomotion, and translocation (**Figure R7c**). As also seen in the rat coculture system, neuronal-like cells were the most migratory, while neuroepithelial-like-1 and immature ependymal-like tumor cells were the most stationary (**Figure R7d**). Interestingly, in contrast to rat cocultures, neuronal-like cells were significantly more migratory than neuroepithelial-like-2 cells in the human coculture system (**Figure R7d**). In addition, a higher proportion of neuronal-like cells exhibited saltatory migratory mode in human cocultures compared to rat cocultures (**Figure R7e**). These differences between the rat and human cocultures may be attributed to the different cell type compositions of the two model systems (rat: E19 cortical cells, **Figure R7f**; human: **Figure R8**), as well as potential species-specific differences.

Furthermore, we performed 10X scRNAseq to capture transcriptional changes in concordance with these morphological changes in human cocultures (**Figure R8**). We captured 3427 cells high quality cells and distinguished between tumor and normal cells by both tdTomato expression and copy number variation (**Figure R8a-b**). We then annotated the normal cells by marker gene expression and found that the human cocultures were composed of excitatory neurons and astrocytes, as expected from the human iPSCs-derived neurons differentiation protocols¹⁶. In addition, we observed that the tumor cells in human cocultures were enriched in cells with intermediate gene expression signatures between neuroepithelial-like, neuronal-like and ependymal-like cell states (**Figure R8d**), while tumor cells in monoculture had high enrichment of only neuroepithelial-like gene expression.

Overall, we performed coculture experiments with human iPSC-derived neurons and astrocytes and found that the results were consistent with findings from rat coculture experiments. We believe this strengthens our finding that ZFTA-RELA tumor cells when grown in coculture with normal cells across different model systems and species, exhibit distinct morphologies and behaviors. We thank the reviewer for this suggestion that enriches our study and have added these new findings into **Extended Data Figure 10** and amended **Results, Methods**.

Figure R7



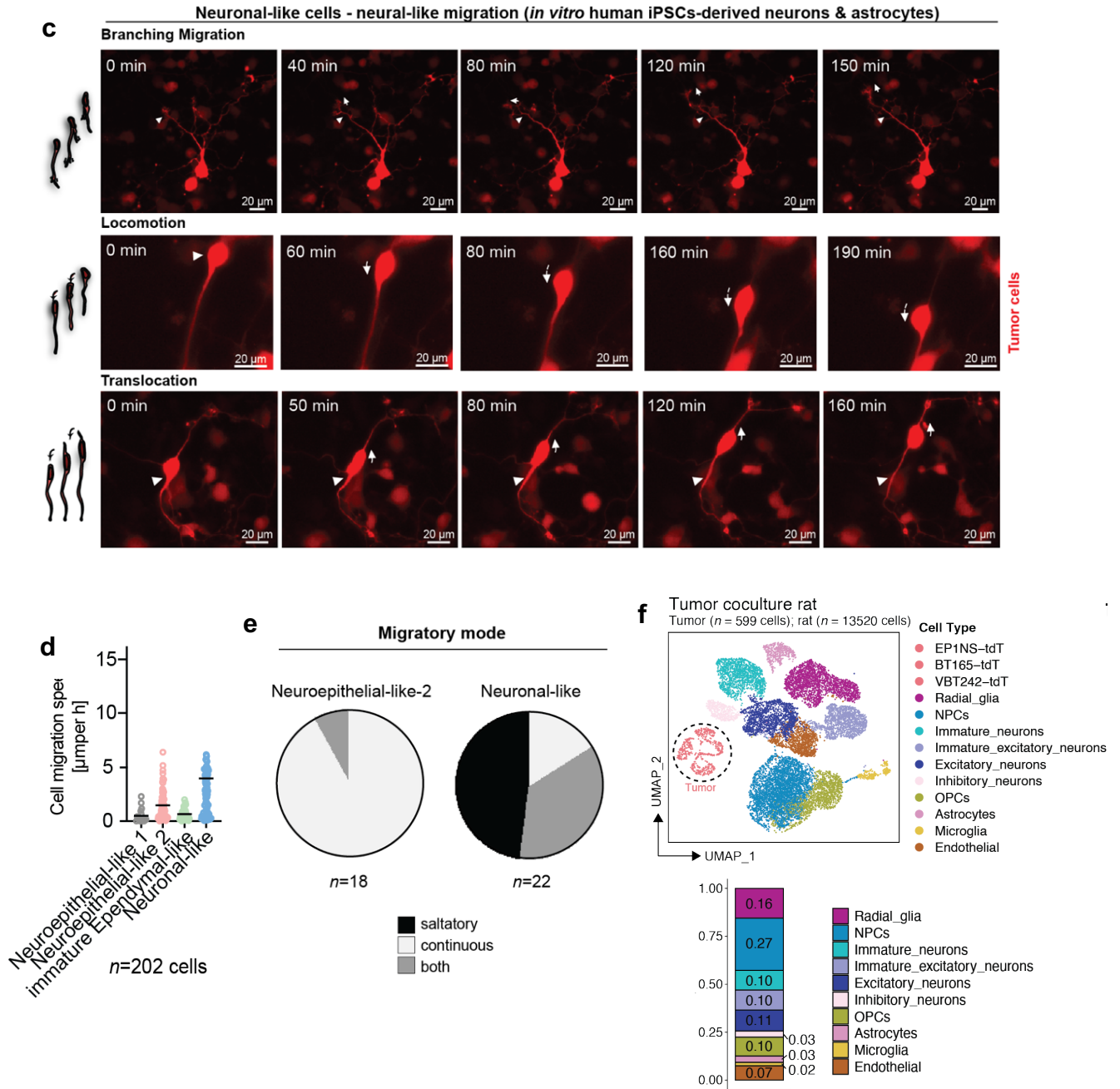


Figure R7. *In vitro* imaging of tumor microtubule dynamics in ZFTA-RELA cells expressing tdTomato cocultured with human iPSCs-derived neurons and astrocytes. **(a)** Representative immunofluorescence images of ZFTA-RELA cells expressing tdTomato, categorized by distinct morphologies of ZFTA-RELA cells (EP1NS), classified as: round cells without TMs (neuroepithelial-like-1, left), cells with 1-3 short TMs (neuroepithelial-like-2), cells with more than 3 short TMs (immature ependymal-like, middle right), and cells with 1-3 long TMs (neuronal-like, right). **(b)** Distinct migratory patterns depicted for each morphological subtype. Arrowheads point to the soma. **(c)** Neuronal-like migratory patterns of an *in vitro* model of ZFTA-RELA cells expressing tdTomato shown in a schematic and live-cell timelapse imaging examples. Arrows indicate direction of TM movement. Arrowhead pointing to branching point (branching migration), cell soma (locomotion) or stable TM tip (translocation). **(d)** Cell migration speed of ZFTA-RELA cells in coculture with human iPSC-derived cells in μm per hour, determined using timelapse imaging *in vitro*. Cells are categorized by morphological subtype ($n = 202$ total, acquired from $n = 3$ independent biological replicates). Individual data points with mean $\pm$ SE are shown. P values are calculated by Kruskal-Wallis test. **(e)** Quantification of migration modes across ZFTA-RELA cells classified as neuroepithelial-like-2 ($n = 18$), or neuronal-like cells ($n = 22$). **(f) (left)** UMAP of ZFTA-RELA tumor cells (EP1NS-tdTomato, BT165-tdTomato, VBT242-tdTomato) cocultured with rat E19 cortical cells. Plots are colored by the cell type, and tumor cells are highlighted by black dotted circle. **(right)** Proportion (y-axis) of cell types in the right E19 cortical cells (x-axis).

Figure R8

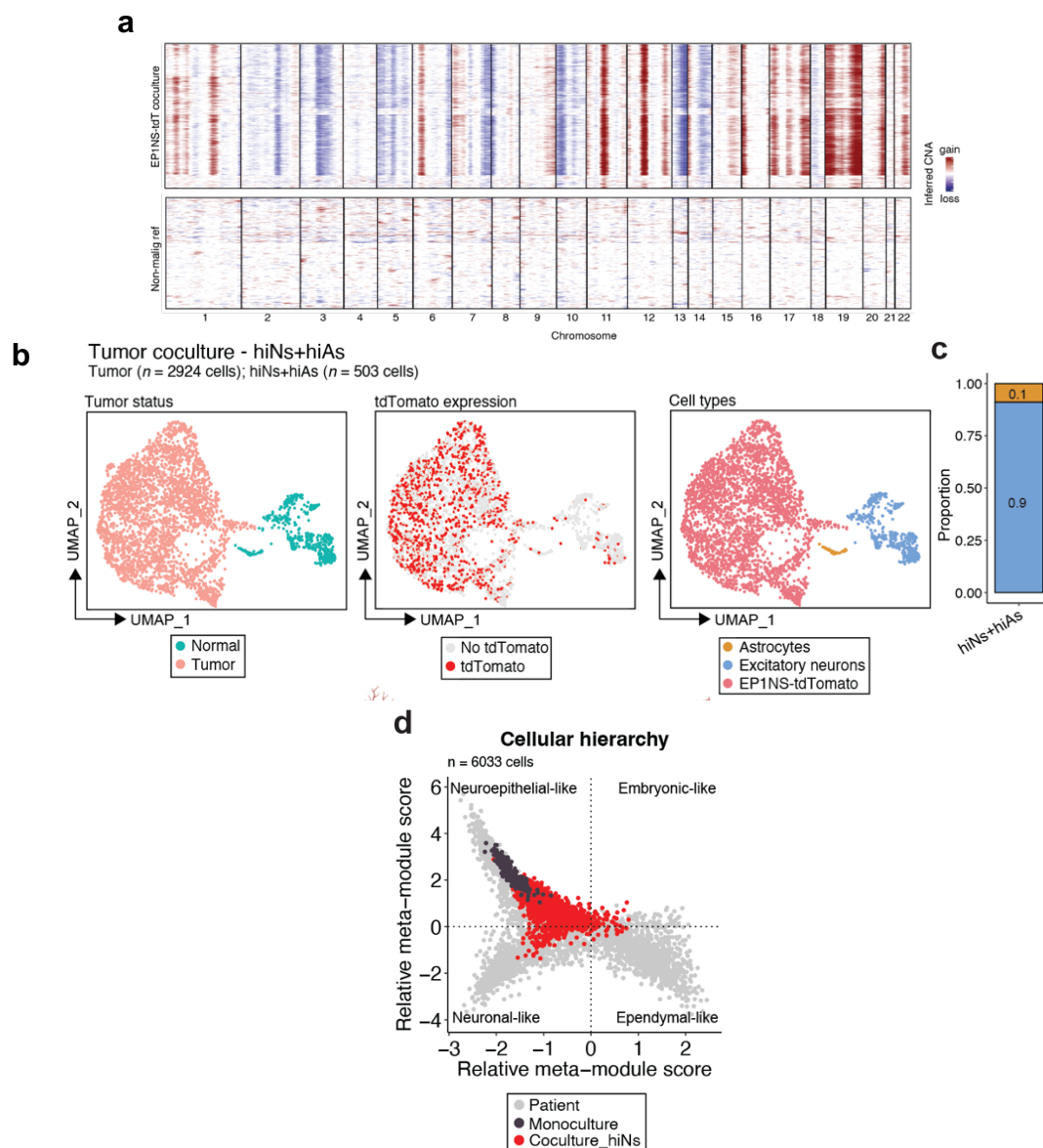


Figure R8. 10X scRNAseq of ZFTA-RELA tumor cells cocultured with human iPSC-derived neurons (hiNs) and astrocytes (hiAs). **(a)** CNV profiles of each cell (y-axis) across chromosomal location (x-axis) inferred from scRNA-seq data of EP1NS-tdT coculture with human iPSCs-derived neurons and astrocytes. **(b)** UMAP plots of all tumor ($n = 2924$) and human normal cells ($n = 503$), plots colored by tumor status (left), tdTomato expression (middle), and cell annotation (right). **(c)** Proportion barplot of normal cell types present in EP1NS-tdT coculture with human iPSCs-derived neurons and astrocytes **(d)** Cell state plot showing ZFTA-RELA tumor cells profiled by scRNAseq, colored by culture condition. ZFTA-RELA patients profiled by sc/snRNA-seq (grey) were combined as reference for cell state hierarchy.

Results: Finally, to distinguish species-specific as well as cell type differences in microenvironmental cues we cocultured ZFTA-RELA cells with human iPSCs-derived excitatory neurons and astrocytes and performed 10X scRNA-seq as well as live-cell imaging (**Extended Data Figure 10a-i**). We found a similar increase of neuronal-like signatures in coculture compared to cells grown in monoculture, with a very small increase in ependymal-like cells (**Extended Data Figure 10d**). These results are consistent with our findings from coculture with rat E19 cortical cells, suggesting that cell state transitions induced by brain-resident cells are found across species. Similarly, ZFTA-RELA cells cocultured with human iPSCs-derived neurons and astrocytes displayed the same morphological subtypes and distinct migratory modes (**Extended Data Figure 10e-i**), highlighting that these features are conserved across rat and human microenvironmental contexts.

Methods: Generation of human iPSC-derived neurons. Induced pluripotent stem cell (iPSC) lines were obtained from the Boston Children's Hospital Human Neuron Core, with all appropriate permissions for use. Induced neurons (iNs) were generated from the GON0515-03 iPSC line following established protocols¹⁶. Briefly, iPSCs were plated at a density of 95,000 cells/cm² on Geltrex-coated plates in mTeSR1 medium (StemCell Technologies, Cat #85850) for viral transduction. The following lentiviral plasmids were obtained from Addgene: FUDeltaGW-rtTA (Addgene #19780; RRID: Addgene #19780), a gift from Konrad Hochedlinger; pTet-O-Ngn2-puro (Addgene #52047; RRID: Addgene #52047), and Tet-O-FUW-eGFP (Addgene #30130; RRID: Addgene #30130), both gifts from Marius Wernig. High-titer lentiviruses ($\sim 1 \times 10^9$ TU/mL) were generated by Alstem and used at the following concentrations: 0.13 μ L per 50,000 cells for each virus. Following transduction, cells were dissociated with Accutase (StemCell Technologies, Cat#AT104-500) and replated at a density of 200,000 cells/cm² on Geltrex-coated plates in mTeSR1 medium supplemented with ROCK inhibitor (10 μ M, StemCell Technologies, Cat#72304) (day 0). On day 1, medium was replaced with N2 medium supplemented with doxycycline (2 μ g/mL, Clontech, Cat#NC0424034), brain-derived neurotrophic factor (BDNF; 10 ng/ μ L, Peprotech), neurotrophin-3 (NT3; 10 ng/ μ L, Peprotech, Cat#450-03), and laminin (0.2 μ g/ μ L, Life Technologies, Cat#23017015). Doxycycline was maintained throughout the differentiation protocol. On day 2, medium was switched to N2B medium containing BDNF, NT3, laminin (same concentrations), Ara-C (2 μ M, Sigma, Cat#1768-100MG), and puromycin (1 μ g/mL, Gibco, Cat#A11138-03). From day 3 onward, cells were cultured in N2B medium supplemented with 1 $\times$ B27 (Life Technologies, Cat#17504044), doxycycline (2 μ g/mL), puromycin (5 μ g/mL), BDNF, NT3, and laminin, with half-media changes performed every other day. At day 6, puromycin was withdrawn and cells were enzymatically dissociated using the Papain Dissociation System (Worthington Biochemical, Cat#LK003150) for downstream experimental setups. Induced neurons were plated on poly-D-lysine and laminin-coated 6-well plates or 15 mm glass coverslips in co-culture with iPSC-derived human astrocytes (Ncardia, Cat#M0605), which were added at 10–15% of the total neuron cell density. **Culture of Human iPSC-derived Astrocytes.** Human iPSC-derived Ncyte® Astrocytes (Ncardia, Cat# M0605) were cultured in T75 flasks (Thermo Fisher Scientific) pre-coated with poly-D-lysine (R&D Systems, Cat#3439-200-01). Cells were maintained in Astrocyte Medium (ScienCell, Cat#1801), consisting of 500 mL of basal medium (Cat# 1801), 10 mL of fetal bovine serum (FBS; Cat# 0010), 5 mL of astrocyte growth supplement (AGS; Cat# 1852), and 5 mL of penicillin/streptomycin solution (Cat# 0503). Half-media changes were performed every other day. Prior to co-culture, astrocytes were enzymatically dissociated using Accutase (StemCell Technologies) and subsequently plated together with day 6 iPSC-derived neurons. Astrocytes were added at a final density corresponding to 10–15% of the total neuronal cell count and maintained in neuronal medium. **In vitro coculture of ZFTA-RELA cells with human iPSC-derived neurons.** On day 18 of human iPSC-derived neurons differentiation, ZFTA-RELA tumor cells were added to neuronal cultures at a ratio of 1:3 (tumor:neuron). Cocultures were monitored for 7 days and maintained with half-media changes every other day with B27 medium, doxycycline (2 μ g/mL), puromycin (5 μ g/mL), BDNF (10 ng/ μ L), NT3 (10 ng/ μ L), and laminin (0.2 ng/ μ L). For all single-cell RNA sequencing (scRNA-seq) experiments, cocultures were dissociated using the Papain Dissociation System after the 7-day coculture period. To assess the role of neuronal activity on tumor cell behavior, cocultures were treated with either tetrodotoxin (TTX; 1 μ M, Tocris, Cat#1069) or vehicle control.

Changes made: Extended Data Figure 7f, Figure 10, Results, Methods

Nomenclature changes from EPN-glial-like to immature ependymal-like cells

To further characterize the identity of “EPN-glial-like” cells (cells with more than 3 short tumor microtubules (TMs), we performed additional immunofluorescence (IF) and 10X Xenium on all tumor morphological subtypes. We found that the tumor cells with more than 3 short TMs are **negative** for Vimentin and GFAP, and **positive** for S100B as well as **CCDC40 (Figure R9a)**, a protein specifically expressed in ependymal motile cilia of the brain ventricles^{17,18}. We confirmed that **CCDC40** is specifically expressed in ependymal-like cells of patient tumor samples (**Figure R9b**). Furthermore, by 10X Xenium of tumor cells in coculture with rat E19 cortical cells, we overlaid cell morphology to a panel of 366 transcripts. This allowed us to expand the number of markers examined for these cells. From these analyses, we found that tumor cells with more than 3 short TMs are positive for **MAP3K19** and **DNER (Figure R9c)**, which are also highly expressed by ependymal-like cells in patient tumor samples (**Figure R9b**). These cells also express other progenitor-like markers, including **SOX10 (Figure R9c)**, which is not expressed in the ependymal-like cells of patient tumors (**Figure R9b**).

To further clarify the difference between tumor cells in patient tumors and tumor cells in coculture, we analyzed the 10X scRNAseq data across patients and patient-derived models (monoculture, coculture, PDX) (**Figure R10**). Tumor cells in adherent and spheroid monoculture conditions were predominantly neuroepithelial-like, while subpopulations in patient and PDX were distinctly neuronal-like (bottom left quadrant in **Figure R10c-d**) and ependymal-like (bottom right quadrant) cells. Finally, tumor cells in coculture with rat E19 cortical cells were somewhere in between, with transitional cells midway progenitor-like and mature-like gene signatures. These findings suggest that tumor cells in coculture are less differentiated than

cells in patient tumors. In light of these analyses and that tumor cells with more than 3 short TMs are positive for both ependymal-like and progenitor-like markers, we reasoned that these cells are likely more immature ependymal-like cells. Accordingly, we amended the nomenclature of these cells to “**immature ependymal-like**” cells in manuscript, referring to cells that are simultaneously enriched with immature progenitor-like as well as tumor ependymal-like metaprogram gene signatures.

Figure R9

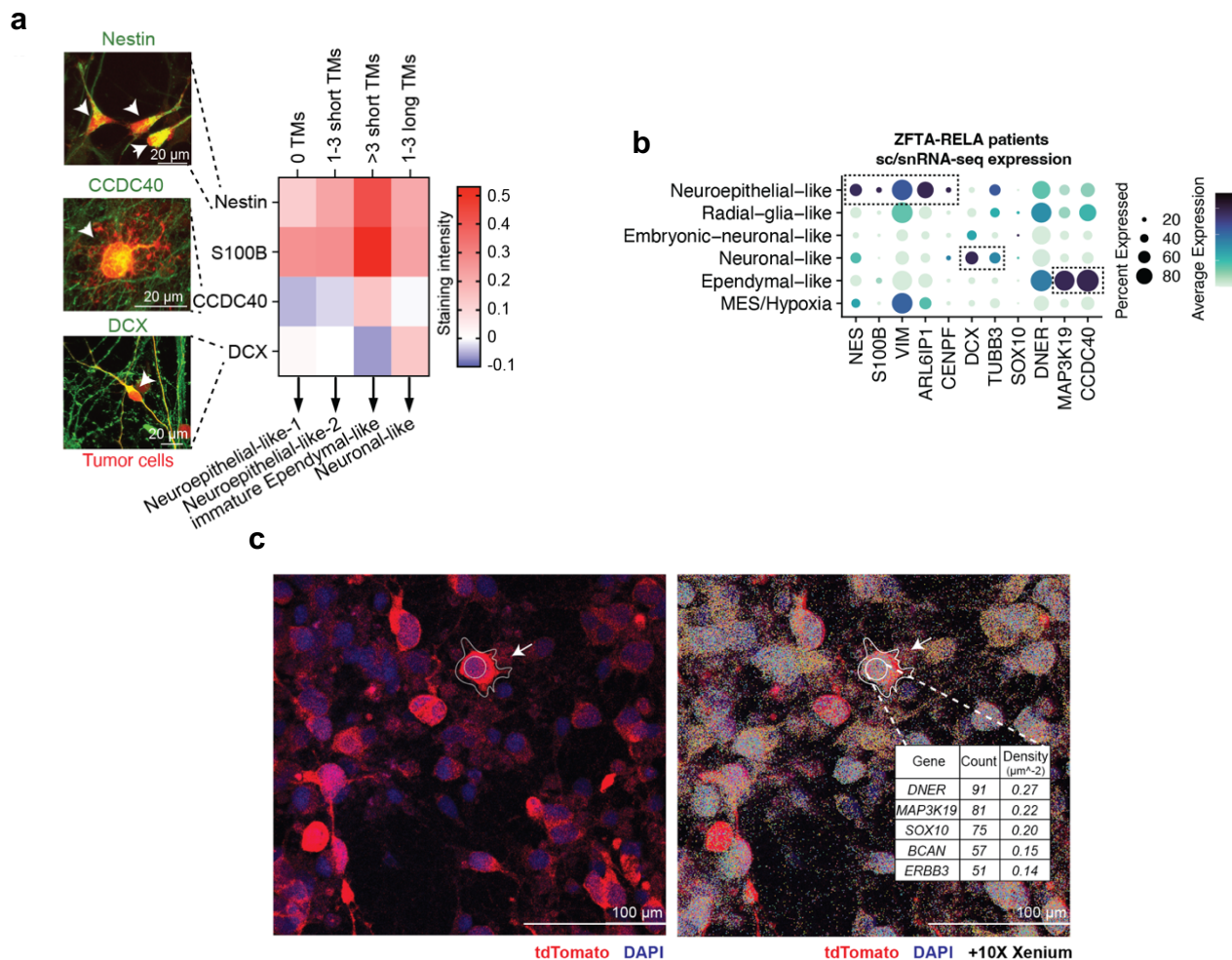
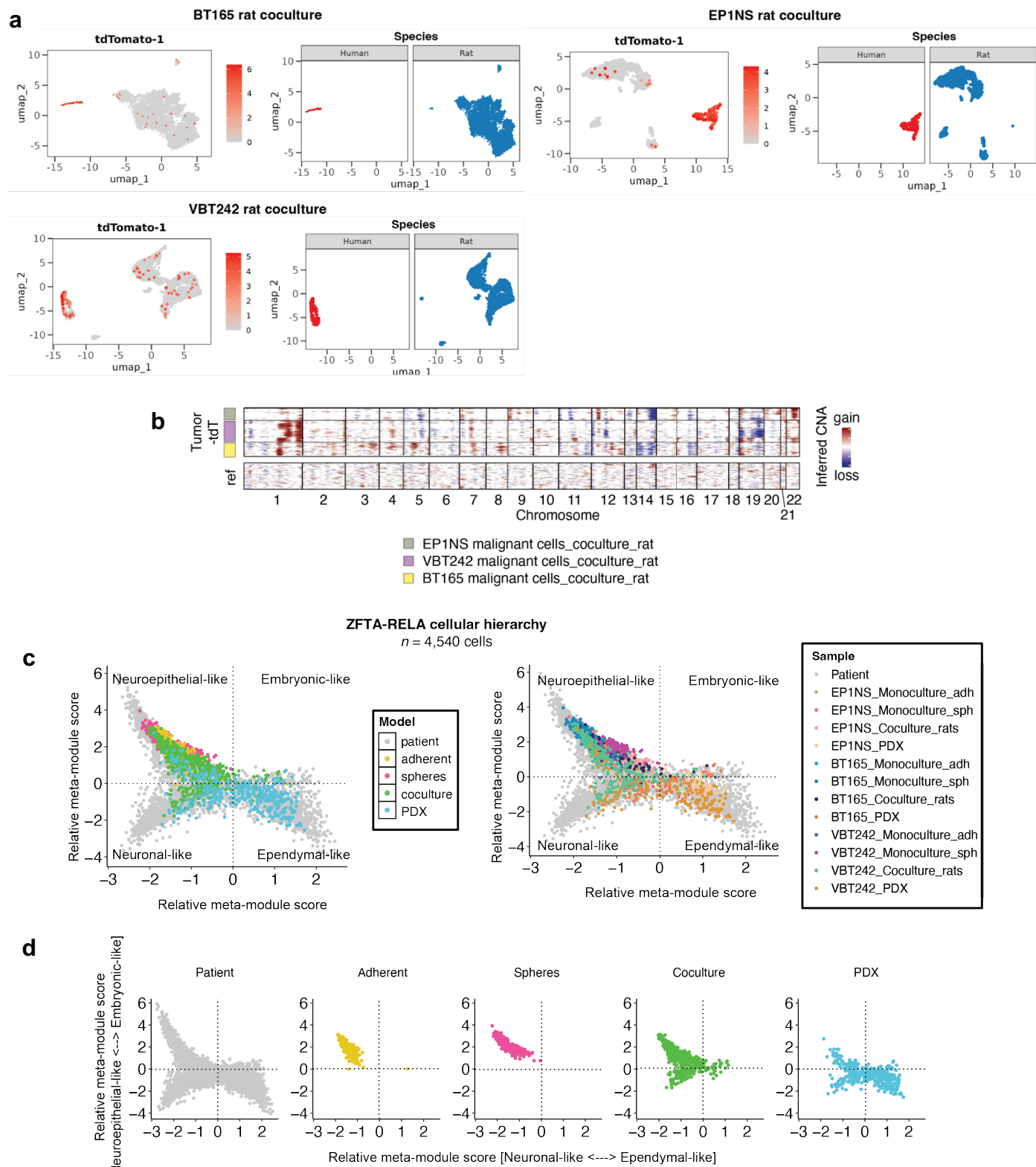


Figure R9. (a) Representative immunofluorescence images and mean intensity of signal of an *in vitro* co-culture model of ZFTA-RELA cells expressing tdTomato, categorized by morphological cell subtypes. Data represent the mean intensity value of $n = 155$ cells acquired from $n = 3$ independent biological replicates per marker. **(b)** Dotplot showing scaled mRNA expression of marker genes for each metaprogram in ZFTA-RELA patient tumors profiled by sc/snRNAseq. Dot size represents percent expressed, while color scale represents the average expression. **(c) (left)** Immunofluorescence image of ZFTA-RELA (EP1NS cells expressing tdTomato) cocultured with rat E19 cortical cells, stained for anti-tdTomato and DAPI. **(right)** Immunofluorescence image overlaid with 10X Xenium transcripts. Each dot represents detected transcript from the custom ($n = 103$) and base panel ($n = 255$) probes used for patient tumor sections. Table indicates top 5 enriched markers in demarcated region (white).

Figure R10



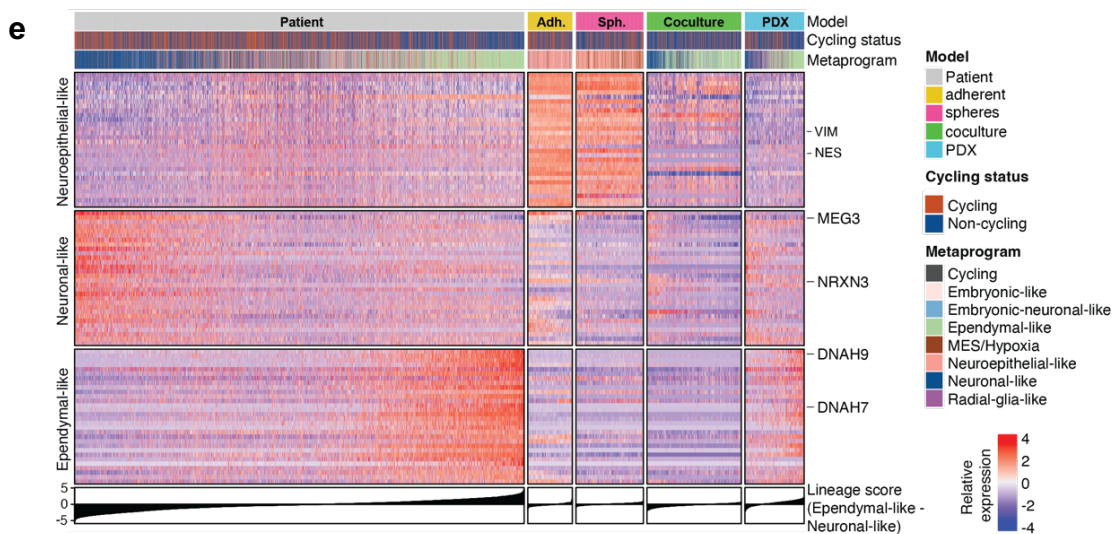


Figure R10. (a) UMAP plots of each ZFTA-RELA tumor cell line (EP1NS, BT165, VBT242) cocultured with rat E19 cortical cells, colored by detection of tdTomato (left) or species (right). **(b)** Copy number alterations (CNA) profiles inferred from scRNAseq data of ZFTA-RELA and ZFTA-Cluster 4 patient-derived tumor cell lines cocultured with rat E19 cortical cells. Cells are ordered by cell line as rows (EP1NS, BT165, VBT242, BT2214), and clustered by their pattern of CNAs across chromosomal locations (columns). Representative fresh nonmalignant cells lacking CNAs are shown on top. **(c)** Cell state plots of ZFTA-RELA patient tumor cells profiled by snRNAseq overlaid with single cells derived from PDX ($n=3$), adherent ($n=3$), neurosphere ($n=3$) and coculture with rat E19 cortical cells ($n=3$) models of EP1NS, BT165, and VBT242. Plots are colored by model (left) and by sample (right). **(d)** Cell state plots, separated by ZFTA-RELA patient tumors profiled by snRNAseq, adherent monoculture, spheroid monoculture, coculture with rat E19 cortical cells, and PDX profiled by sc/snRNAseq. **(e)** Heatmap showing relative expression of the top 50 genes identified by NMF analysis specific to neuroepithelial-like, neuronal-like, and ependymal-like signatures. Cells are sorted according to the overall expression of ependymal-like versus neuronal-like gene programs. Model (patient, adherent, spheroid, coculture, or PDX), cycling status, and metaprogram assignment are indicated on the top bars. Select genes of interest are indicated on the right.

#1-12. In Ext. Fig. 10F-G, can the authors include a fourth group—a coculture group—using the same cell lines that were included as adherent cultures, sphere cultures, or PDX models? This would help understand which of the cell state transitory effects are attributed to growing cells as 3D spheres vs. 2D (as adherent monocultures vs. cocultures) vs. in PDX models, as each of these four groups are very different in the culture conditions and microenvironmental cues.

We thank the reviewer for this thoughtful suggestion. To address this, we included a fourth coculture group for all 3 cell lines, including EP1NS, BT165, and VBT242 cocultured with rat E19 cortical cells (**Figure R10**). We cocultured each of the cell lines with rat E19 cortical cells for 7 days, after which we dissociated both tumor and rat cells to process for 10X scRNAseq. We distinguished between tumor and rat cells by the enrichment of human or rat genes and expression of tdTomato, and additionally confirmed the malignancy of the human tumor cells by *inferCNV* (**Figure R10a-b**). Interestingly, we did detect low tdTomato expression in some non-human normal cells, potentially due to transfer of fluorescence protein encoding genes as reported between malignant cells and non-malignant cells in GBM organoids¹⁹. By comparing culture conditions across all three cell lines (**Figure R10c-e**), we see that cell transitions more towards more mature cell states is more apparent in cells cocultured with rat E19 cortical cells and in PDX models compared to adherent or neurosphere monocultures. We have included these new results into **Extended Data Figure 7b-e**.

Changes made: Extended Data Figure 7b-e

#1-13. Did the authors confirm that the neuron-like enrichment of cells in coculture are true “malignant” cells, as FACS may give a false td-tomato-positive population? A CNV plot, as was shown for the earlier figures, could also be included for the coculture scRNA-seq analyses.

As we now also show in **Reviewer comment #1-12** and **Figure R10b**, we have examined the copy number alterations in the tdTomato-positive and human genes enriched population using R package *inferCNV*. We detected clear CNVs in the malignant population of the rat coculture samples across all 3 cell lines (EP1NS, BT165, and VBT242).

#1-14. The authors state in line 342-343 that normal brain-resident cells induce a “shift” of neuroepithelial-like progenitor cells toward neuron-like lineages, but instead of this being a plasticity effect, could it be attributed to cell stress/death of progenitor cells, or perhaps, a differentiation effect, wherein the progenitor cells simply differentiate to neuron-like cells?

This is an excellent point. To exclude that the differences in lineages between monoculture and coculture are attributed to increased cell stress/death of progenitor cells, we measured the percentage of genes associated with mitochondrial or ribosomal genes in monoculture versus coculture conditions. A high percentage of mitochondrial or ribosomal genes in a particular cell is associated with cell stress/death^{20–22}. Therefore, we expected an increase in the percentage of mitochondrial/ribosomal genes if cells in coculture were undergoing increased stress/death. As shown in **Figure R11a**, cells in coculture display decreased mitochondrial and ribosomal signatures, suggesting that they are not undergoing increased stress/death compared to cells in monoculture. We therefore conclude that the shift towards neuron-like lineages is not due to cell stress and death.

Furthermore, to examine whether there truly is a ‘shift’ of neuroepithelial-like cells towards more mature cell states, we performed RNA velocity⁹ to infer direction of gene expression changes in cells from scRNA-seq data. We analyzed scRNAseq data from three cell lines (EP1NS, BT165, VBT242) across the different models: monoculture (adherent and spheroid), coculture with rat E19 cortical cells, and PDX (**Figure R11b**). The predicted trajectory started from neuroepithelial-like cells in all three models, with direction towards neuronal-like cells in rat coculture and towards both neuronal-like and ependymal-like cells in PDX. Overall, these findings suggest that the neuroepithelial-like cells transition towards more mature cell states in the presence of normal brain-resident cells.

The question raised by the reviewer regarding the difference between “plasticity” and “differentiation” is very interesting. Our understanding is that “differentiation” implicates a unidirectional and irreversible cell state transition¹⁶, while “plasticity” more broadly includes all cell state transitions including ones that are bidirectional and/or reversible. While we do observe that the progenitor-like neuroepithelial-like cells move towards more mature-like neuronal and ependymal-like cells when cocultured with normal brain-resident cells, we do not have evidence to support that this is unidirectional and irreversible. As this is the case, we attribute this phenomenon to be more broadly “cell state shift/transition” or “plasticity”. In addition, it would be very interesting in future studies to experimentally examine these findings such as lineage tracing in *in vitro/in vivo* model systems.

Figure R11

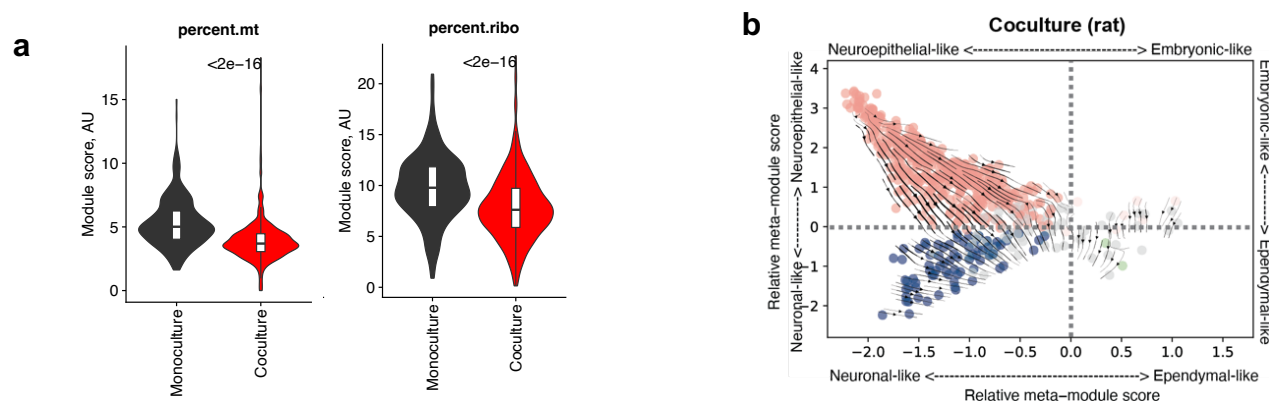


Figure R11. (a) Violin plots showing module score of mitochondrial genes (left) and ribosomal genes (right) in ZFTA-RELA cells cultured in monoculture (black) or coculture with rat E19 cortical cells (red). Boxplots denote Tukey’s whiskers (25-75 percentile represented by minima-maxima; statistical median as center). *P* values calculated by Wilcoxon-rank sum test. **(b)** RNA velocity [ref] analysis of ZFTA-RELA cells in monoculture, coculture with rat E19 cortical cells, or PDX models. Cells represented in cell state plot scored by neuroepithelial-like, embryonic-like, neuronal-like and ependymal-like gene signatures. Cells colored by metaprogram assignment.

#1-15. In lines 348-349, the authors state that the “overall invasion speed of ZFTA-RELA cells significantly increased due to a higher fraction of neuronal-like cell states in the coculture model”, but this may be an overstatement, given that the experiments and observations are merely correlative and not causal: 1) Neuroepithelial-like 2 cells are also increased on coculture, but it is unclear whether this subtype of neuroepithelial cells also show higher invasion speed. After all, data in Fig. 5C suggest that there may be a bimodal distribution of cells in the neuroepithelial group as there are some cells that are equally mobile as neuron-like cells, 2) some of the higher somatokinesis noted in Fig. 5k could be attributed to a sharp decline in neuroepithelial-like 1 cells, which may be the slower moving cells from Fig. 5C. One way to answer some of these caveats would be to repeat the quantification of Fig. 5K as was done in Fig. 5C and contrast the invasion speed of neuroepithelial vs. EPN-glia vs. neuron-like cells in monoculture vs. coculture (analyzed by a two-way ANOVA to control for multiple variables).

We thank the reviewer for this important remark. We have removed these claims from the manuscript accordingly. As we do not see any neuronal-like cells in monoculture, we could not contrast invasion (or cell migratory) speed by two-way ANOVA as suggested. Instead, we compared monoculture versus coculture conditions for each cell state with appropriate statistical tests after testing for normality in **Figure R12**. These quantifications suggest that neuroepithelial-like-1, neuroepithelial-like-2, and immature ependymal-like cells have no significant difference in migratory speed when grown in monoculture versus coculture. Neuronal-like cells were highly migratory in coculture, while no neuronal-like cells were detected in monoculture conditions. Overall, as we do not have enough evidence to support a causal relationship between enrichment of cell states versus the overall cell migration speed in coculture models, we have now removed this statement in **Results section: “Superimposing molecular and morphological cell state characteristics in ZFTA-RELA tumors”**

Figure R12

a

Shapiro-Wilk test	neuroepithelial-like-1	neuroepithelial-like-1	neuroepithelial-like-2	neuroepithelial-like-2	neuronal-like	Immature Ependymal-like	Immature Ependymal-like
	monoculture	coculture	monoculture	coculture	coculture	monoculture	coculture
W	0.6375	0.6896	0.8565	0.7506	0.7623	0.5402	0.7476
P value	<0.0001	<0.0001	0.0069	<0.0001	<0.0001	<0.0001	<0.0001
passed normality test?	No	No	No	No	No	No	No
Test used	Mann-Whitney	Mann-Whitney	Mann-Whitney	Mann-Whitney	Mann-Whitney	Mann-Whitney	Mann-Whitney

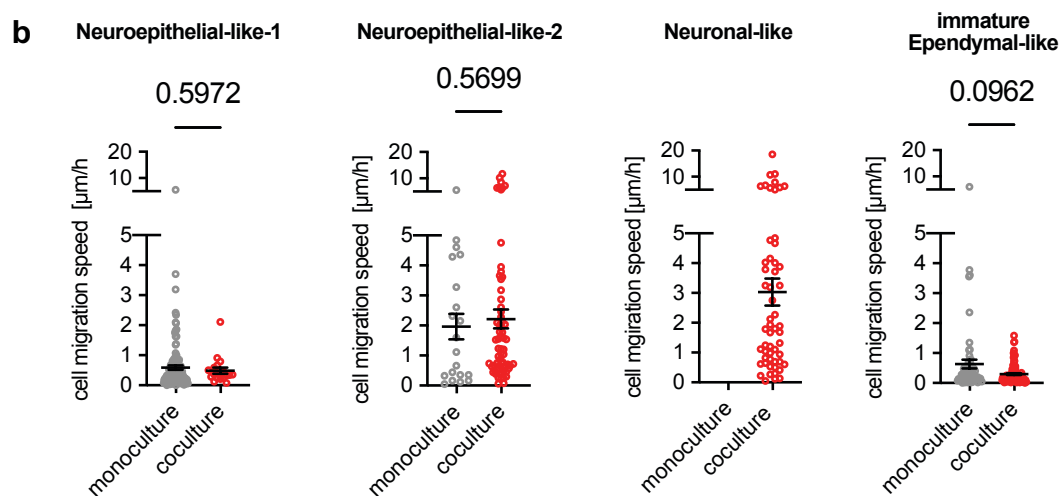


Figure R12. (a) Representative time-lapse images of cell states in coculture with rat E19 cortical cells. (b) Table of results from Shapiro-Wilk test determining whether somatokinesis data of each morphological subtype are normally distributed. (c) Quantification of somatokinesis speed of different cell states in monoculture compared to coculture, Mann-Whitney test ($n = 125$ neuroepithelial-like-1 monoculture, 19 neuroepithelial-like-1 coculture, 20 neuroepithelial-like-2 monoculture, 64 neuroepithelial-like-2 coculture, 55 neuronal-like coculture, 59 EPN-glia monoculture, 97 EPN-glia coculture cells).

Changes made: Results section: “Superimposing molecular and morphological cell state characteristics in ZFTA-RELA tumors”

Minor concerns:

#1-m1. Lines 73-75: The authors nicely state that it is unknown whether ST-EPN subtypes have distinct cells-of-origin and malignant cell states and whether these differences contribute to varying outcomes and therapeutic resistance. However, the authors should include some discussion of what those differences in outcomes are/can be. For example, do some ST-EPN subtypes show better clinical prognoses? Do some ST-EPN subtypes show lesser therapeutic defiance?

This is indeed a critical point that is challenging to address due to the currently still sparse clinical information regarding the non-canonical ST-EPN subgroups in patients. To our knowledge, there is currently no comprehensive study of clinical prognoses and survival outcomes stratified by these recently defined DNA methylation subgroups, highlighting the importance of follow-up studies. Several of our collaborators are currently working on accumulating all available patient data to finally arrive at a statistically solid answer to this question.

In short of this, several smaller studies have alluded to clinical outcome differences between ST-subgroups with ZFTA-RELA fusion negative compared to ZFTA-RELA fusion positive tumors. Firstly, very clear differences between ZFTA-RELA and ST-YAP1 in clinical outcome have been documented^{23,24}, where ST-YAP1 tumors have significantly more favorable clinical outcomes compared to ZFTA-RELA tumors.

Secondly, in the study of Tauziède-Espariat et al., 2021²⁴, clinical outcomes of ZFTA-Cluster 2 and ZFTA-Cluster 4 tumors harboring non-ZFTA-RELA fusions including ZFTA-NCOA1, ZFTA-NCOA2, and ZFTA-MAML2, were compared with ZFTA-RELA tumors. PFS for ZFTA-Cluster 2 and ZFTA-Cluster 4 (Figure R13) was significantly shorter than ZFTA-RELA with $p < 0.001$, while OS for ZFTA-Cluster 2 and ZFTA-Cluster 4 was comparable to ZFTA-RELA. These findings are also supported by a recent clinical report²⁵, where ZFTA-Cluster 2 patients showed shorter PFS and significantly younger age of onset. In addition, ZFTA-Cluster 2 and ZFTA-Cluster 4 tumors exhibited distinct histology of sarcoma-like poorly differentiated features that correlated with this shorter PFS.

Finally, in Zheng et al. 2021²⁶ the authors introduce multiple different ZFTA fusions (that are only present in non-canonical ST-EPN subgroup tumors) by *in utero* electroporation in mice and examined survival.

ZFTA-MAML2 and ZFTA-NCOA2 overexpressed mouse tumors led to worse survival outcomes than mice growing tumors driven by ZFTA-RELA fusion overexpression. These results suggest that the likely clinical prognosis of non-canonical ST-EPN subgroups in patients might also be low.

Another important part is that differences in clinical outcome between ZFTA-RELA and non-canonical ST-subgroups harboring ZFTA-RELA fusions (ZFTA-Cluster 1 and ZFTA-Cluster 3) is not known because previous studies grouped these subgroups together based on the common ZFTA-RELA fusion. However, it is notable that ZFTA-Cluster 3 tumors have distinctive undifferentiated histological features²⁶ in addition to clustering differently by DNA methylation arrays, which also agrees with our findings that ZFTA-Cluster 3 are more enriched in embryonic-like transcriptional programs. Based on ZFTA-Cluster 2 and ZFTA-Cluster 4 tumors with sarcoma-like poorly differentiated features that correlated with this shorter PFS, we also predict that clinical outcomes for ZFTA-Cluster 3 tumors will be poor compared to ZFTA-RELA.

In conclusion, it is likely that the non-canonical ZFTA tumors have lower survival outcomes than the canonical ZFTA-RELA tumors, which still critically needs to be comprehensively examined in future studies. We have incorporated these comments into **Discussion**.

Figure R13

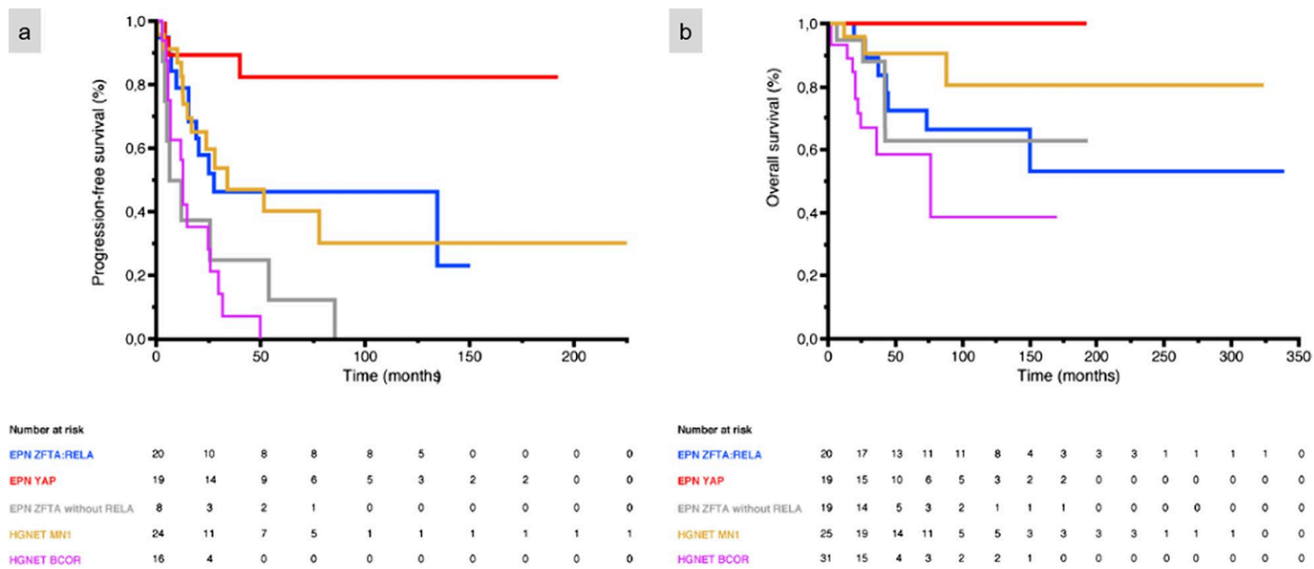


Figure R13. Non-ZFTA-RELA PFS and OS in Tauziède-Espariat et al., 2021²⁴.

Discussion: These distinct developmental signatures and perturbed lineage differentiation across subgroups lays the groundwork for developing much needed therapeutic strategies for these tumors. Importantly, while clear patient outcome differences are described comparing ZFTA-RELA to ST-YAP1^{23,24}, as well as some noted case studies comparing ZFTA-RELA, ZFTA-Cluster 2 and ZFTA-Cluster 4^{24,25}, there is a dire need for comprehensive clinical study comparing patient outcomes and therapeutic resistance across all most recently defined molecular subgroups in ST-EPN tumors to further apply findings from this study.

Changes made: Discussion

#1-m2. It is unclear why the UMAPs in Ext. Fig. 1C vs. main Fig. 1C look different, given that these are the same samples and cells across both figures? I'd recommend keeping the same UMAP, as a comparison with Fig. 2C would also help clarify which ST-EPN subtype correlates with which clusters.

We thank the reviewer for highlighting this lack of clarity in our previous representations and for the nice suggestion. The inconsistencies were due to differences in the type of data displayed (scRNA-seq, snRNA-seq, or both datasets), and/or the use of harmony²⁷ to regress out inter-sample (batch) differences. Taking advantage of the reviewer suggestion, we now revised all figures to display UMAP plots of all cells (scRNA-

seq and snRNA-seq) without any batch-correction applied (**Figure R14**). We have included these amended analyses in the manuscript.

Figure R14

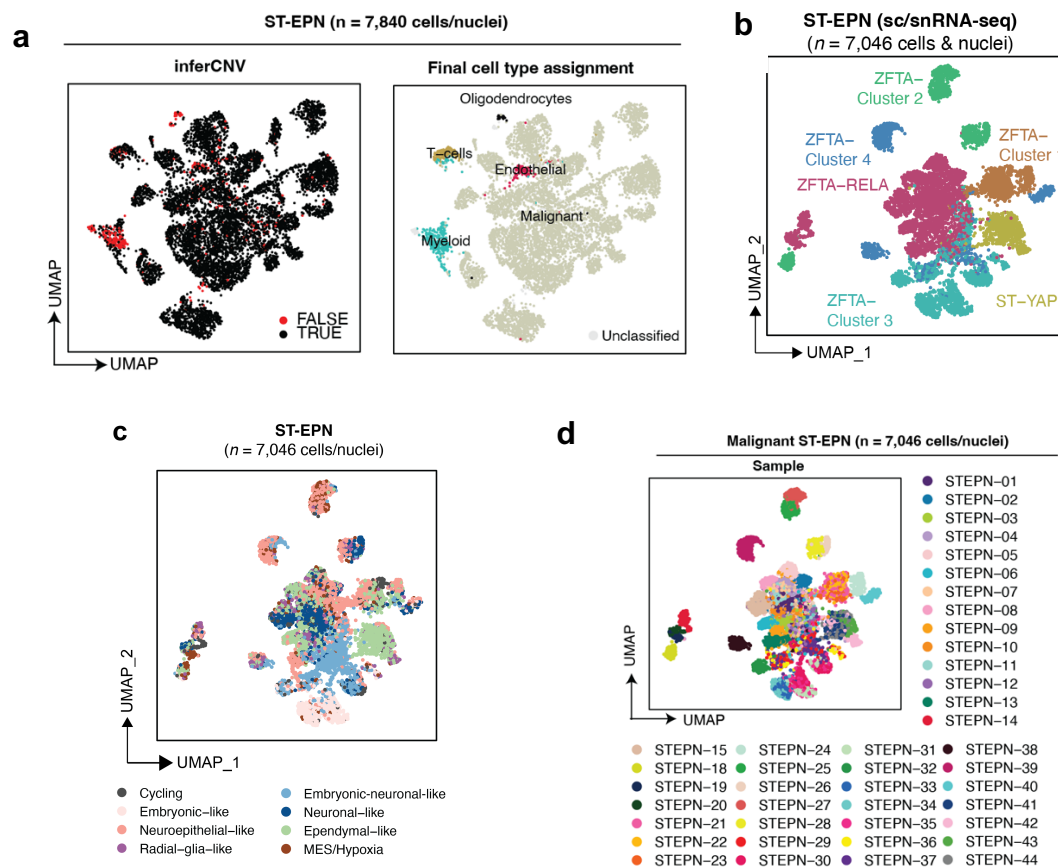


Figure R14. (a) UMAP plot of all ST-EPN patient cells/nuclei ($n = 7,840$) profiled by sc/snRNA-seq. Plots are colored by: inferred malignant status based on copy number variation (CNV, left panel) and final cell type classification (right panel). (b) UMAP plot of malignant ST-EPN cells profiled by sc/snRNA-seq and colored by molecular subgroup. (c) UMAP plot of malignant ST-EPN cells colored by identified cellular state. (d) UMAP plot of all ST-EPN patient cells/nuclei colored by sample of origin.

Changes made: Figure 1c, Figure 2c, Extended Data Figure 1a, Extended Data Figure 1e.

#1-m3. I recommend putting appropriate citations for the reference atlases within the relevant figures (i.e., for example, name reference #29 and #30 next to Fig. 1D and Ext. Fig. 1D), as the authors projected their data onto these atlases.

We have revised to include appropriate citations for the reference atlases (**Figure R15**), and have included these in **Figure 1f** and **Extended Data Figure 1f**.

Figure R15

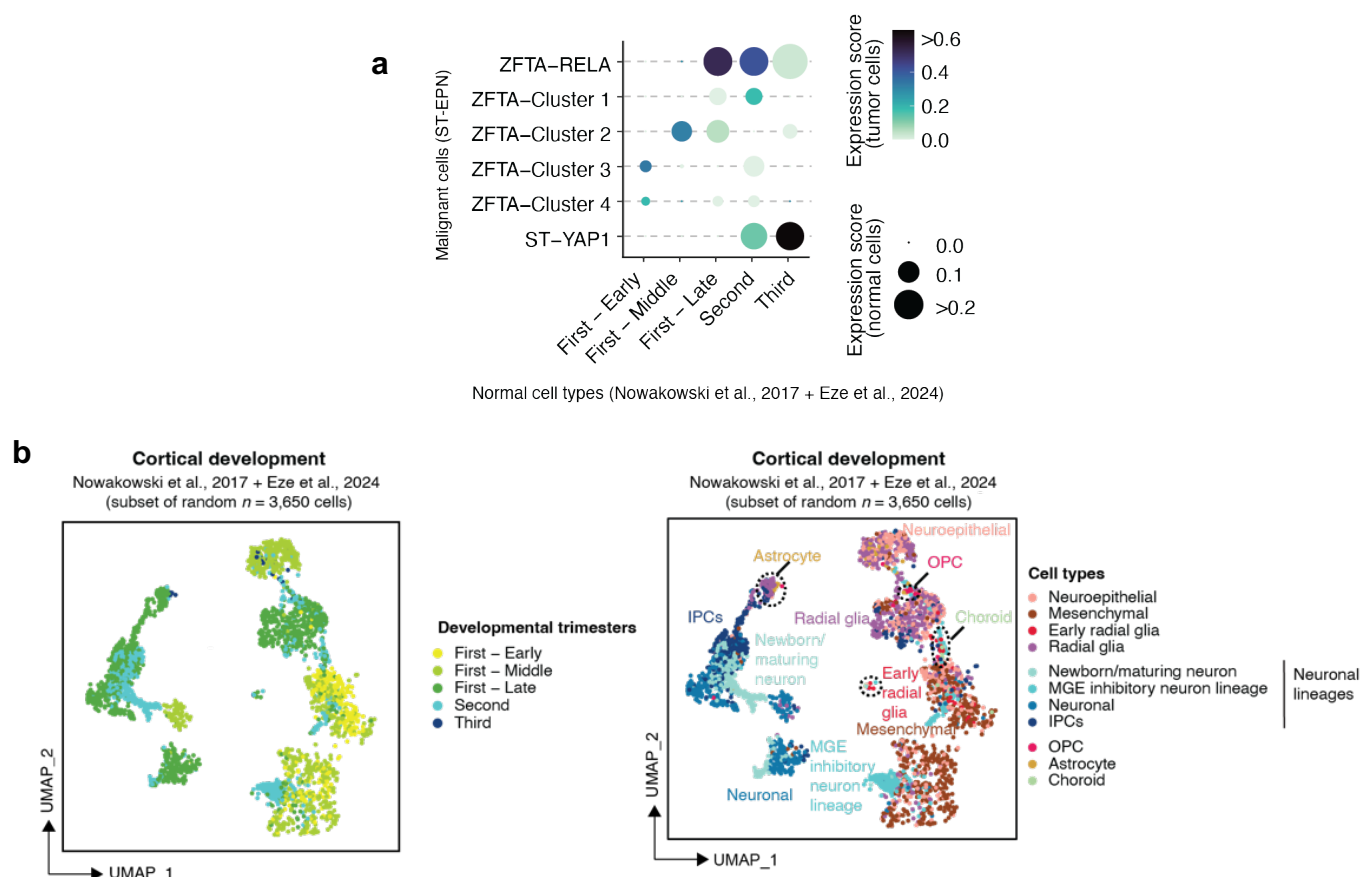


Figure R15. (a) Projection of ST-EPN subgroups (y-axis) onto cell types or developmental time points found across the developing human cortex (x-axis). Color scale depicts expression score of normal cell type gene expression in tumor cells, and dot size represents expression score of tumor cell type gene expression in normal cells. **(b)** UMAP plot of cell types present during the first, second and third trimester of human cortical development [refs], colored by developmental trimester (left) and cell types (right). Original publications listed for each plot.

Changes made: Figure 1f, Extended data figure 1f

#1-m4. Line 148 should be “extent” and not “extend”

We thank the reviewer for this observation. With additional edits, we have removed this sentence of the manuscript.

#1-m5. Can the authors find a different way to represent the data in Fig. 1D, because some small populations (e.g.: choroid) are hard to see and track over time? One way to do this is to modify main Fig. 1F, such that they could add another panel that reflects how non-malignant expression programs (reference) change over developmental time as a dotplot.

As the reviewer suggested, we now plotted the non-malignant cell types in the referenced dataset as dotplot over time (age in weeks) (**Figure R16**). We included this plot in **Extended Data Figure 1g**.

Figure R16

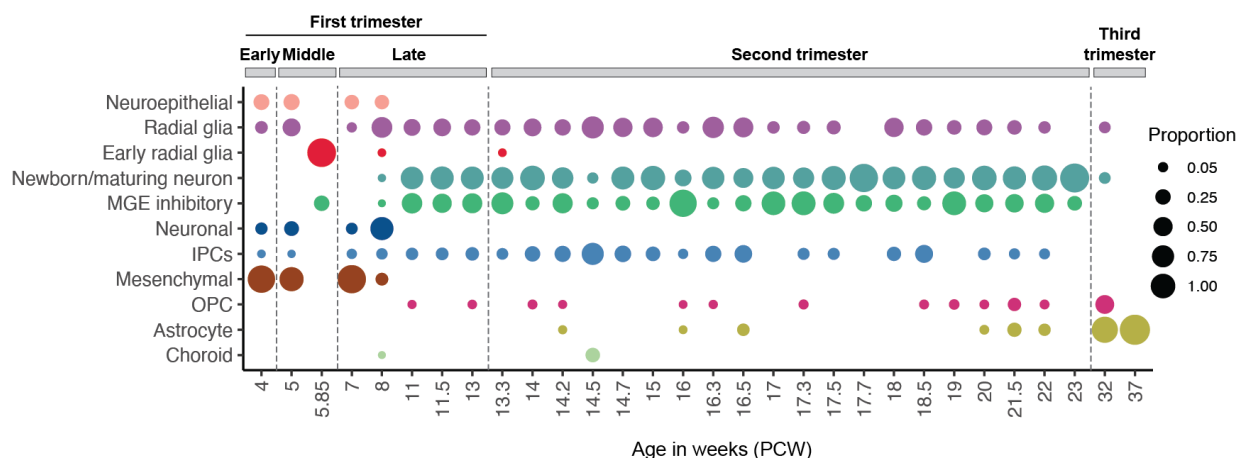


Figure R16. Dotplot depicting proportion of cell type in reference dataset^{1,4} (y-axis) to developmental time (age in weeks, x-axis).

Changes made: Extended Data Figure 1g.

#1-m6. The authors could assess if data in Ext. Fig. 6F-G are non-normally distributed, and if so, use a Mann-Whitney test instead of the t-test. Furthermore, although the authors state the high variability in these datasets, the Results section that corresponds to the data on primary vs. recurrent tumors may be overstated, given the lack of statistical significance (or even a statistical “trend”). The authors should acknowledge this caveat in the Results/Discussion section more clearly. In addition, without additional data mining and experimentation, it is unclear if the small increase in MES/hypoxic cells in recurrent tumors has any biological significance, given that the proportion of this cell state is incredibly small (<10% in primary tumors and ~15% in recurrent tumors).

We thank the reviewer for this important point. After extensive statistical analyses of both the spatial coherence score and cell state composition in sc/snRNA-seq and spatial data, we have **removed** analysis and conclusions from primary versus recurrence tumors in **Results**, and relevant places in **Discussion**. In addition, as described in **Reviewer comment #1-6-8**, we expanded our spatial cohort for ZFTA-RELA patient tumors from $n = 23$ to $n = 56$ sections and have improved cell annotation methods. Our analyses was as follows:

- **Spatial coherence:** We assessed whether the spatial coherence scores of primary versus recurrent tumors (**Figure R17a**) are normally distributed by Shapiro Wilk’s test. Both primary and recurrent tumor scores were normally distributed (Shapiro Wilk’s test p-value for Primary: 0.978, p-value for Recurrence: 0.116), and therefore we performed a student’s t-test. From this comparison, we can conclude that there is no statistical difference in spatial coherence between primary and recurrent tumors as observed in the previous manuscript.
- **Cell state composition of sc/snRNA-seq:** To determine statistical difference in cell state composition, we used the R package *speckle* (as mentioned also in **Reviewer comment #1-6-8**), a method tailed specifically for compositional data. From this comparison, we detected **no significant change in the proportion of MES/Hypoxia metaprogram nor any other metaprograms in primary versus recurrent tumors (Figure 17b)**.
- **Cell state composition of spatial ZFTA-RELA sections:** We further compared the composition of primary vs recurrent ZFTA-RELA tumor sections profiled by 10X Xenium. Again, we used the propeller function R package *speckle* to perform statistical comparisons of compositional data. From this comparison, we detected **no significant change in the proportion of malignant cell states between primary and recurrent 10X Xenium tumor data (Figure R17c)**.

Figure R17

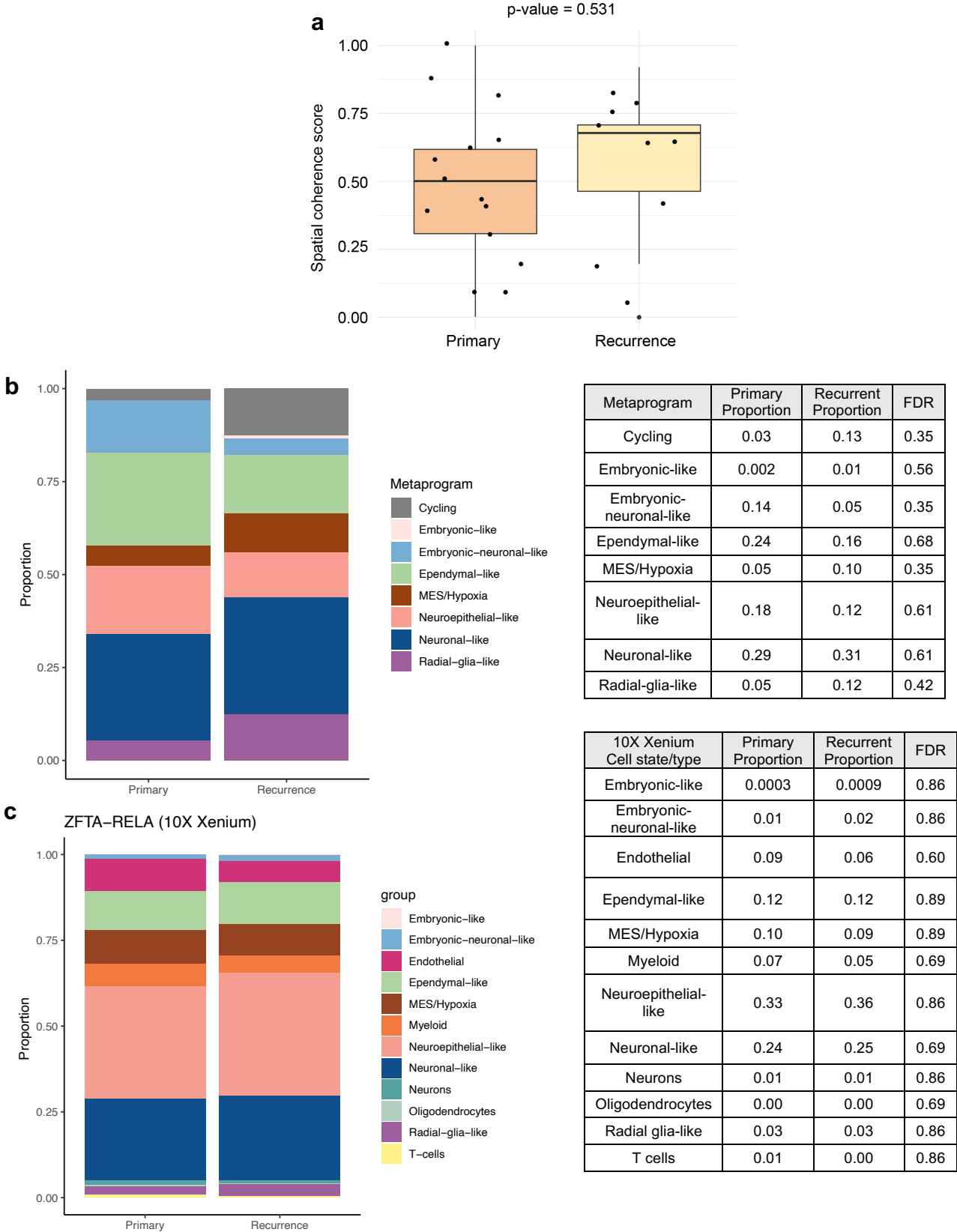


Figure R17. (a) Scaled spatial coherence score across ZFTA-RELA derived from primary or recurrent patients. Each dot depicts the average value of sections for each sample. P-value calculated by two-sided student t-test. **(b)** Proportion of metaprograms (y-axis) in primary and recurrent patient samples (x-axis) in ZFTA-RELA subgroup analyzed by sc/snRNAseq

(left). Statistical significance of cell proportions by primary and recurrent tumors in ZFTA-RELA using the propeller function from the R package *speckle* (right). **(c)** Proportion of metaprograms (y-axis) in primary and recurrent patient samples (x-axis) in ZFTA-RELA subgroup analyzed by 10X Xenium (left). Statistical significance of cell proportions by primary and recurrent tumors in ZFTA-RELA using the propeller function from the R package *speckle* (right).

#1-m7. Ext. Fig. 7C and main Fig. 4C could use some more details in the legends. What do the rows vs. columns indicate?

Each row and column represent one individual spatial niche identified in one tumor section (as we have 56 tumor sections and 6 spatial niches for each section, this makes a total of 336 rows/columns), summarized by its relative frequency of cell states/types. By correlating the composition of spatial niches identified across our cohort of 56 ZFTA-RELA samples, we identify spatial niches with similar cell state/type composition, which we term *metaniches*. We have added this clarification into the figure legends for **Extended Data Figure 6b**.

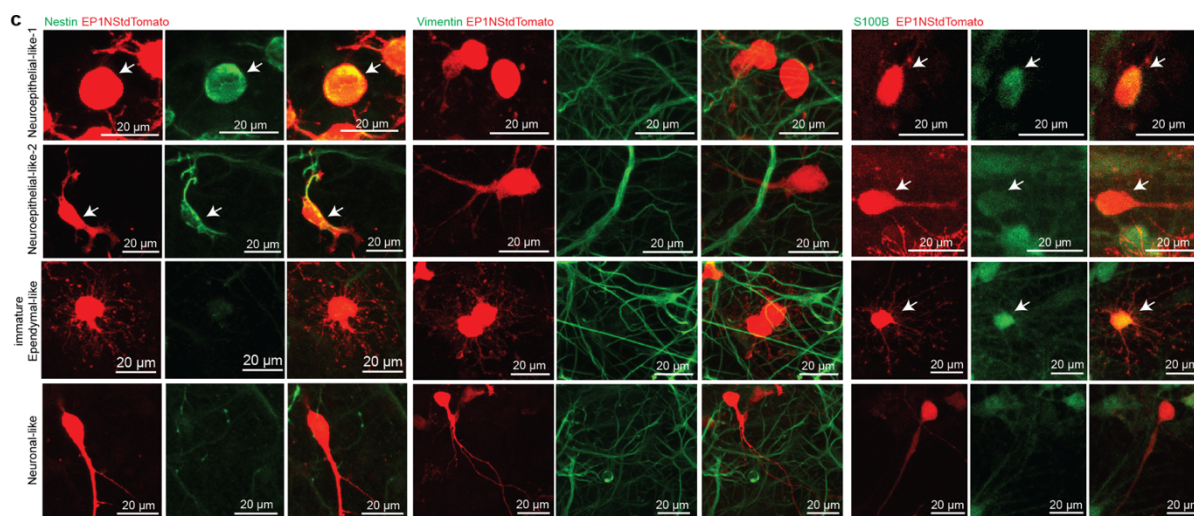
Extended Data Figure 6b legends: Pairwise Pearson correlation (color scale) representing similarity in cell type composition of the spatial niches identified in 56 ZFTA-RELA tumor sections analyzed by 10X Xenium. Each row and column represent one individual spatial niche identified in one tumor section, summarized by its relative frequency of cell states/types. Recurrent niches (metaniches) are highlighted by dashed line boxes. Niches are annotated based on the most frequent cell types (displayed on the right side in the pie charts).

Changes made: Extended Data Figure 6b legends

#1-m8. In Ext. Fig. 8A, all the panels are at differing magnifications, making it difficult to fully appreciate the staining patterns. This may add to some discrepancies that are evident between the staining patterns shown in Ext. Fig. 8A and the mean intensity heat map in main Fig. 5B: From Ext. Fig. 8A, it seems that vimentin is expressed by all 3 cell types, but from Fig. 5B, it seems as though only neuroepithelial cells express this marker.

As mentioned in **Reviewer comment #1-12**, we have re-done all the immunofluorescence staining and fluorescence quantification methods to capture the full staining patterns for each tumor cell morphological subtype (**Figure R8a,b,d, Figure R18**). We have also amended the panel (**Figure R18**) to include all images at the same magnification. We have included these new analyses to **Extended data Figure 8**.

Figure R18



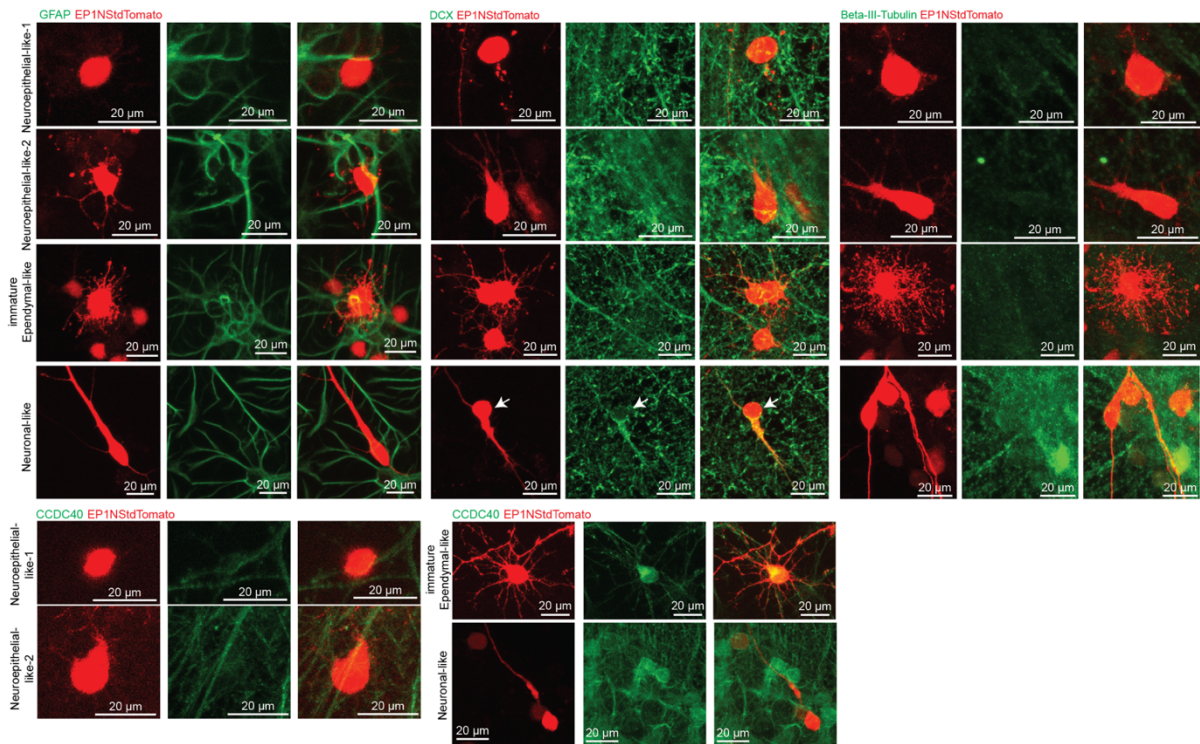


Figure R18. Representative immunofluorescence images of ZFTA-RELA (EP1NS cells expressing tdTomato) cells in coculture with rat E19 cortical cells ($n = 155$ cells acquired from $n = 3$ independent biological replicates per marker).

#1-m9. Related to Ext. Fig. 8D: Can the authors show a supplemental table listing the “top 20 marker genes of neuroepithelial-like cells ($n = 16$ cells, left panel), and neuronal-like cells ($n = 16$ cells, right panel)” from the co-culture groups?

We thank the reviewer for this helpful suggestion. We further narrowed to include the top **10** genes to include the most specific genes of each morphological subtype (neuroepithelial-like-1, $n = 20$ cells; neuroepithelial-like-2, $n = 20$ cells, immature ependymal-like, $n = 14$ cells, and neuronal-like, $n = 19$ cells) in **Table S8**.

Changes made: Table S8

Referee #1 (Remarks on code availability):

Code provided is acceptable for reproducing main findings from the paper. README file is quite sparse and I did not take time to install and run the code.

We have added additional information and codes in the README file for clarity.

Referee #2 (Remarks to the Author):

In this study, Jeong and colleagues study the transcriptional signatures of supratentorial ependymomas. Using integrated single cell- and spatial transcriptomics they refine our knowledge of the malignant cellular subpopulations and uncover that the spatial organization of the disease is driven – similar to gliomas – by regions of hypoxia inversely correlated with vascularity and positively correlated with recurrence/aggressiveness. They provide functional validation of individual cell states that were identified by transcriptional profiling, demonstrating distinct morphological characteristics of the malignant subcellular compartments. Finally, they confirm that the neural microenvironment supports the cellular heterogeneity to faithfully recapitulate the disease and demonstrate that microenvironmental neural cells promote a more mature and motile tumor cell state.

The comprehensive evaluation of the cellular composition of ST-EPN in this manuscript is a valuable contribution to the field and contributes to the emerging field of cancer neuroscience in important ways. Critically, the authors highlight two distinct embryonic progenitor populations present in ST-EPN tumors, opening new avenues for future investigation. The study is further strengthened by the functional characterization of the malignant cell states, bringing in the importance of the neural microenvironment to tumor progression.

I have the following recommendations for improvement:

#2-1. The manuscript appears to be written as two separate stories and the authors should aim to address this in the storytelling. Whilst neural elements of the tumor microenvironment are known to play a key role in brain tumor progression, this study has limited investigation into the contribution of the neuronal and glial non-malignant cells in the EPN tumor microenvironment prior to the functional data presented in Figure 5. The storytelling transition from Figure 1-4 to Figure 5 should be improved. For example, the glial-like tumor population does not appear as a meta program (or part of a meta program) and is therefore absent from analyses in Figure 2-4, yet this population is unexpectedly investigated in Fig. 5. To potentially address this, the authors could perform some spatial imaging confocal microscopy using neuronal (and axonal) and/or astrocyte markers to describe their contribution to the tumor composition.

We thank the reviewer for this very constructive feedback regarding the overall storytelling of the manuscript. After identifying distinct transcriptional cell states in ZFTA-RELA tumors in **Figure 1-2**, and their spatial localization in tumors in **Figure 3-4**, our goal in **Figure 5** was to examine whether these cell states also have distinct phenotypes, potentially leading to different functions. To answer this question, we aimed to identify patient-derived models that best re-capitulate the different cell states. We found that intratumor heterogeneity was better preserved either in PDX or in the coculture system and decided to use the latter to investigate the distinct morphologies and behaviors of the different ZFTA-RELA transcriptomic cell states. We agree with the reviewer that how the manuscript was written previously did not fully reflect our reasoning and have now re-written the **Results** section regarding **Figure 5** to more clearly explain the goal of exploring both transcriptional and phenotypic heterogeneity in tumor cells. Importantly, we have toned down the language regarding the effect of neuronal and glial non-malignant cells on the tumor cells, as this is not the main investigatory question that we aimed to answer.

Regarding the EPN-glial-like tumor cells, we also agree that this is a new term that is disconnected from the rest of the manuscript. After further investigation we have changed the nomenclature of these tumor cells to *immature ependymal-like* cells. The reasoning for this is described in detail in our rebuttal for **Reviewer comment #2-6**. As a result of using consistent nomenclature to describe tumor populations throughout the manuscript, our impression is that we now have a more cohesive story.

#2-2. The authors state that they have uncovered potential therapeutic targets of relevance in the context of

metaprograms correlated with lower survival (line 174, 175). The Gojo et al. study demonstrated that the high ST-Ependymal-like metaprogram have a favorable prognosis, yet it did not present the Neuronal-precursor-like group (as performed in PF-A EPN). Here, the authors should analyze the prognosis of EPN according to the metaprograms they have identified, in addition they should present the survival of each molecular subgroup in this patient cohort.

We thank the reviewer for these important points. To address the correlation of newly identified metaprograms with survival, we evaluated the top 30 genes of 8 metaprograms in a bulk RNA-seq dataset of 46 ZFTA-RELA patients of an external cohort (Open Pediatric Brain Tumor Atlas). Using a method previously described¹⁴, we estimated the survival using univariate models calculated by the Kaplan-Meier Method and tested for significance by the log-rank test. Expression of metaprogram gene signatures was split into low and high expression based on the top and bottom quartiles, respectively. Due to the small patient cohort size used for this analysis ($n = 46$), most of the survival correlations were statistically not significant except for the Cycling metaprogram (**Figure R4**). As we do not have a clear correlation between the newly identified metaprograms and patient survival, we have removed this statement in the manuscript.

In addition, we currently lack the information regarding the survival of every patient in each molecular subgroup in this patient cohort. This is indeed a critical point that is challenging to address due to the currently still sparse clinical information regarding the non-canonical ST-EPN subgroups in patients. To our knowledge, there is currently no comprehensive study of clinical prognoses and survival outcomes stratified by these recently defined DNA methylation subgroups, highlighting the importance of follow-up studies. Several of our collaborators are currently working on accumulating all available patient data to finally arrive at a statistically solid answer to this question. In short of this, several smaller studies have alluded to clinical outcome differences between ST-subgroups with ZFTA-RELA fusion negative compared to ZFTA-RELA fusion positive tumors, which we have included in **Discussion**.

Figure R4

[FIGURE REDACTED]

[FIGURE REDACTED]

Figure R4. [TEXT REDACTED]

Discussion: These distinct developmental signatures and perturbed lineage differentiation across subgroups lay the groundwork for developing much-needed therapeutic strategies for these tumors. Importantly, while clear patient outcome differences are described comparing ZFTA-RELA to ST-YAP1^{14,68}, as well as some noted case studies comparing ZFTA-RELA, ZFTA-Cluster 2 and ZFTA-Cluster 4^{68,69}, there is a dire need for a comprehensive clinical study comparing patient outcomes and therapeutic resistance across all most recently defined molecular subgroups in ST-EPN tumors to further apply findings from this study.

Changes made: Discussion

#2-3. In Fig. 1f, it appears that the ZFTA-RELA group lacks any projection onto the Choroid cell type, whereas in Fig. 2d ZFTA-RELA expresses a significant proportion of Ependymal-like cells. Whereas the ZFTA-Cl.1 group appears to identify this choroid projection and Ependymal-like cell population in both types of analyses. The authors should provide a possible explanation for this result.

We thank the reviewer for pointing out this inconsistency between findings of Fig.1F and Fig. 2D. This comes from fundamental differences in analysis methods, as outlined below:

1. Fig. 1F depicts correlation based on aggregate/"pseudobulk". First, single cell counts of each tumor subgroup were aggregated ("pseudobulk"), which were then scored by the top 30 marker genes of cell types in the reference dataset (Nowakowski et al., 2017 + Eze et al., 2024). This is labeled "expression score (tumor cells)". Similarly, single cell counts of each cell type in the reference dataset were pseudobulked and scored by the top 30 differentially expressed genes of tumor subgroups in the tumor sc/snRNAseq data. This is labeled "expression score (normal cells)". The dotplot in Fig. 1F displayed the correlations of samples aggregated and projected to the developmental dataset, representing **an aggregate** confounded by the heterogeneity of cell states of each patient per subgroup.
2. For Fig. 2D, we performed NMF analysis of tumor sc/snRNAseq data and identified recurrent gene expression programs across samples ("metaprograms"). While each development-like metaprogram was annotated by the comparison of gene signatures to the developmental reference

scRNA-seq dataset as well as gene ontology (GO) analysis, these are *de-novo* identified malignant programs in the tumor. As such, the plot in Fig. 2D more accurately reflects the true malignant cell composition across tumor samples.

Due to the confusing and sometimes conflicting findings of these two figures, we have **removed** the pseudobulk projection of tumor subgroups to cell types in the reference dataset and only retained the projection to time (**Figure R1**). We have also revised the manuscript to only include projection to developmental time.

Figure R1

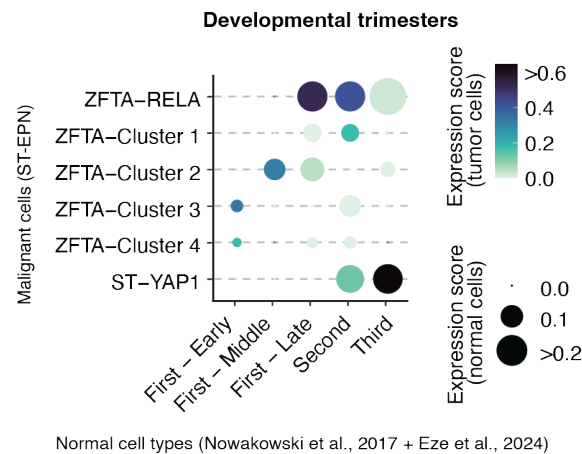


Figure R1. Projections of ST-EPN subgroups (y-axis) onto developmental time points found across the developing human cortex (x-axis). First-Early: post-conception week (PCW) 4, First-Middle: PCW 5-6, First-Late: PCW 7-12. PCW, post-conceptional week.

Results: Next, we projected gene expression profiles of non-malignant cells from the reference atlas^{27,28} onto individual ST-EPN cells (**Figure 1f, Extended Data Figure 2a**). While ZFTA-Cluster 3 and ZFTA-Cluster 4 tumors projected to the earliest time window of early and middle first trimesters, ZFTA-Cluster 2 and ZFTA-Cluster 1 tumors mapped onto the later windows of middle first and second trimesters. Finally, canonical ZFTA-RELA tumors projected broadly from the late first to second, and third trimester, while ST-YAP1 aligned with the latest timepoints. (**Figure 1f, Extended Data Figure 2a**).

Changes made: Figure 1f, Results

#2-4. In Extended Fig. Data 4c, it appears that ZFTA-RELA samples from the external patient cohort also appear to map onto both the neuronal and ependymal-like signature. This should be reflected in the text (line 201). In addition, in Extended Fig. Data 4c, it would be helpful to add a plot of ZFTA-RELA only with the neuronal- and ependymal-like gene expression module scores.

We thank the reviewer for this helpful feedback. We amended the statements in **Results** to address the neuronal and ependymal-like signatures. In addition, we plotted the neuronal-like and ependymal-like gene expression module scores for the ZFTA-RELA tumors only (**Figure R19**) as a violin plot and included this figure in **Extended Data Figure 3g**. We checked for normal-distribution of module scores for the two metaprograms using the Shapiro-Wilk test and found a p-value < 0.05 for Ependymal-like in ZFTA-RELA, indicating non-normal distribution. Hence, we used Wilcoxon rank sum test to check for significantly different distribution of module scores of the two metaprograms.

Results: We validated these findings using an external patient cohort¹⁴ and confirmed that ST-YAP1 patient tumors map onto ependymal-like signatures, whereas ZFTA-RELA patients onto both neuronal-like and ependymal-like signatures (**Extended Data Figure 3e-g**).

Figure R19

Score	Subtype	Shapiro W	Shapiro P	Comments
Neuronal-like	ZFTA-RELA	0.979	0.676	Normally distributed
Ependymal-like	ZFTA-RELA	0.935	0.0265	Not normally distributed

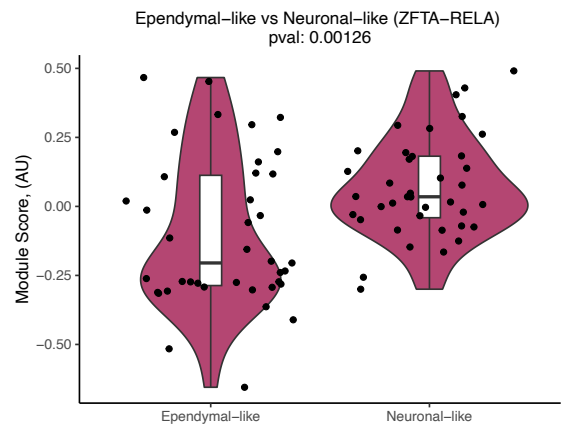


Figure R19. (top) Results of Shapiro-Wilk normality test of gene expression module scores in ZFTA-RELA tumors. **(bottom)** Boxplot of Ependymal-like and Neuronal-like module scores for all samples in ZFTA-RELA tumors only in the external cohort. Each dot represents a sample.

Changes made: Results, Extended Data Figure 3g

#2-5. It would be informative to analyze the proliferative capacity of each metaprogram in the patient data.

We appreciate the reviewer’s insight. We have included an additional graph showing the proliferative capacity of each metaprogram in the ST-EPN patient data (**Figure R20**). The neuroepithelial-like and embryonic-like programs appear as the most proliferative programs, consistent with their similarity to immature developmental cell types. The more differentiated ependymal-like cells are the least proliferative, consistent with the decrease in proliferation associated with increasing cell maturation. In addition, the neuronal-like cells are not the least proliferative and as such we removed the sentence from the manuscript. We have included these findings into **Extended Data Figure 2e-f** and **Results**.

Figure R20

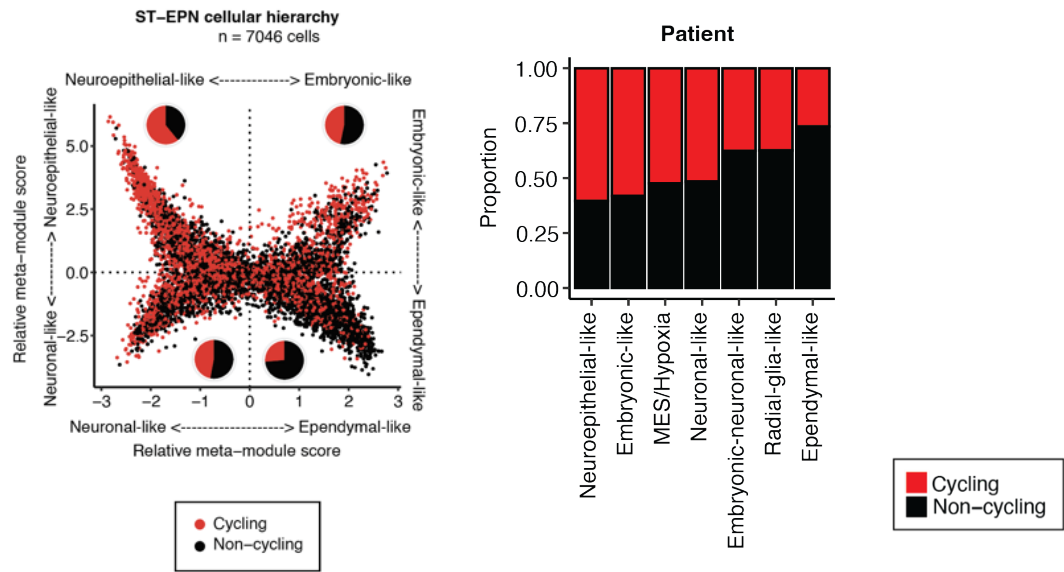


Figure R20. (left) Cell state plot colored by inferred cycling status. The frequency of cycling and non-cycling cell within each quadrant is displayed in the pie charts. **(right)** Barplot showing the proportion of cycling and non-cycling cells across ST-EPN cell states, ordered from the most (left) to the least (right) proliferative.

Results: We further examined the proliferative capacity of each metaprogram and found that the most progenitor-like populations, namely neuroepithelial-like and embryonic-like cells, were the most proliferative, while radial glia-like and ependymal-like were the least proliferative (**Extended Data Figure 2e-f**).

Changes made: Extended Data Figure 2e-f, Results

#2-6. In Figure 5a, the grouped classification of ependymal- and glial-like is confusing. The morphology and markers of the EPN-glial-like group suggest a more astrocytic-like cell, but this morphology is not convincing of an ependymal-like cell. A clearer rationale for this grouping should be provided. The 10X Xenium analysis and pearson correlation of the EPN-glial-like group should be included in Extended Data Fig. 8d, to support the groupings. Furthermore, are the processes in the neuronal-like cells TMs, or more axon-like neurites/processes? They are certainly thinner than the more TM-like processes in the other cell states....

We thank the reviewer for this important question. To address this point, we aimed to further characterize the EPN-glial-like (tumor cells with greater than 3 short TMs) by performing additional immunofluorescence (IF) and 10X Xenium on all tumor morphological subtypes. We first amended the methods for IF staining and increased the panel of markers to increase consistency and rigor of the study. For example, we improved the detection of tumor cells by amplifying the tdTomato signal using an anti-tdTomato antibody. This allowed us to more clearly distinguish the tumor cells and helped us identify the previous background signals that led to misleading results. In addition, we improved our method of quantifying immunofluorescence signal to a semi-automated system by performing segmentation of tdTomato signal of each cell to capture the morphology and area of the cell. Consequently, we found that the tumor cells with more than 3 short TMs are **negative** for Vimentin and GFAP, and **positive** for S100B as well as **CCDC40** (**Figure R9a**), a protein specifically expressed in ependymal motile cilia of the brain ventricles^{17,18}. We confirmed that *CCDC40* is specifically expressed in ependymal-like cells of patient tumor samples (**Figure R9b**). Furthermore, by 10X Xenium of tumor cells in coculture with rat E19 cortical cells, we overlaid cell morphology to a panel of 366 transcripts. This allowed us to expand the number of markers examined for these cells. From these analyses, we found that tumor cells with more than 3 short TMs are positive for *MAP3K19* and *DNER* (**Figure R9c**), which are also highly expressed by ependymal-like cells in patient tumor samples (**Figure R9b**). These cells also express other progenitor-like markers, including *SOX10* (**Figure R9c**), which is not expressed in the ependymal-like cells of patient tumors (**Figure R9b**). These findings reinforce that these cells display ependymal gene signatures, hence confirming their “ependymal-like” state despite their star-shaped morphology.

To further clarify the difference between tumor cells in patient tumors and tumor cells in coculture, we analyzed the 10X scRNAseq data across patients and patient-derived models (monoculture, coculture, PDX) (**Figure R10**). Tumor cells in adherent and spheroid monoculture conditions were predominantly neuroepithelial-like, while subpopulations in patient samples and PDX were distinctly neuronal-like (bottom left quadrant in **Figure R10c-d**) and ependymal-like (bottom right quadrant) cells. Finally, tumor cells in coculture with rat E19 cortical cells were somewhere in between, with transitional cells midway progenitor-like and mature-like gene signatures. These findings suggest that tumor cells in coculture are less differentiated than cells in patient tumors. In light of these analyses, and with our experimental findings showing that tumor cells with more than 3 short TMs are positive for both ependymal-like and progenitor-like markers, we reasoned that these cells are likely more immature ependymal-like cells. Accordingly, we amended the nomenclature of these cells to “**immature ependymal-like**” cells in manuscript, referring to cells that are simultaneously enriched with immature progenitor-like as well as tumor ependymal-like metaprogram gene signatures.

Figure R9

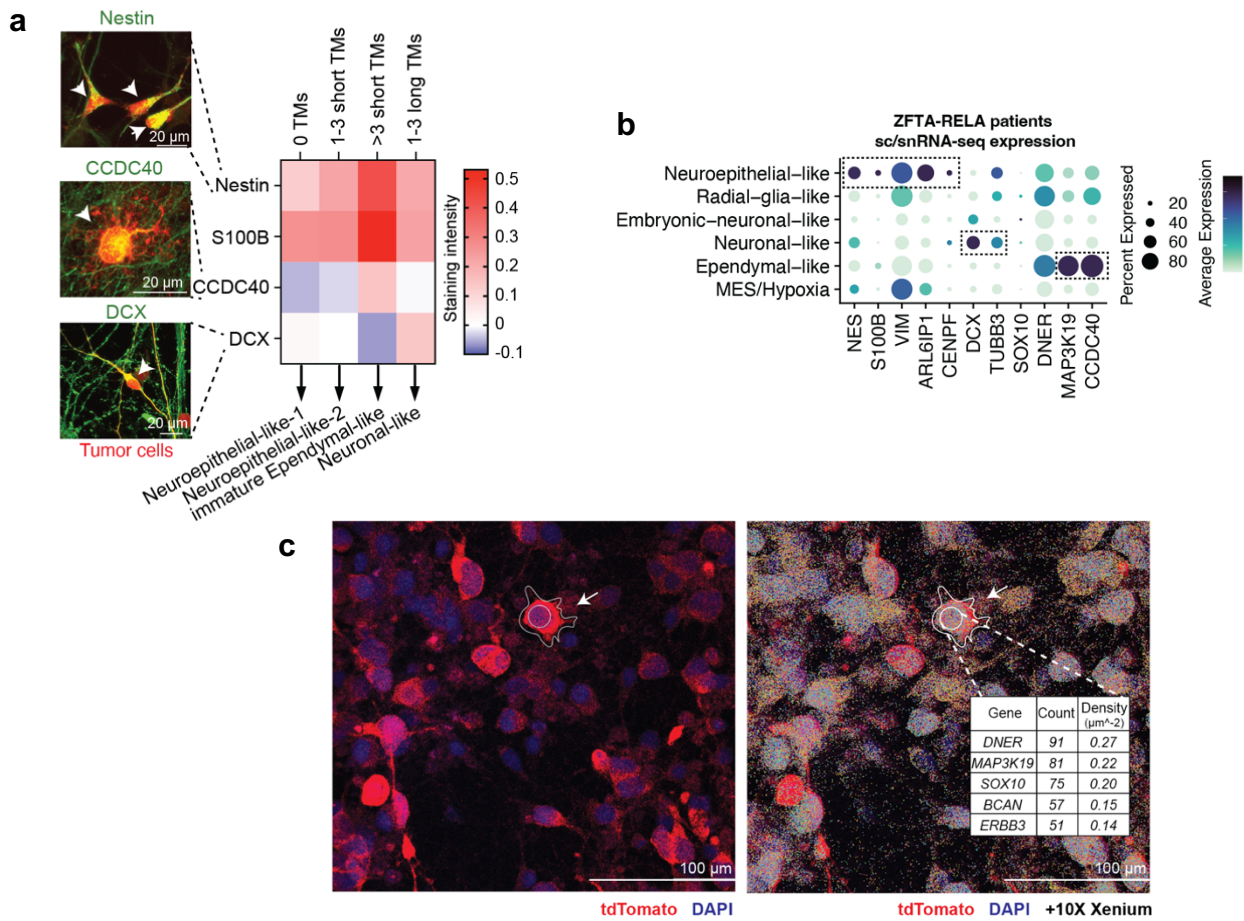
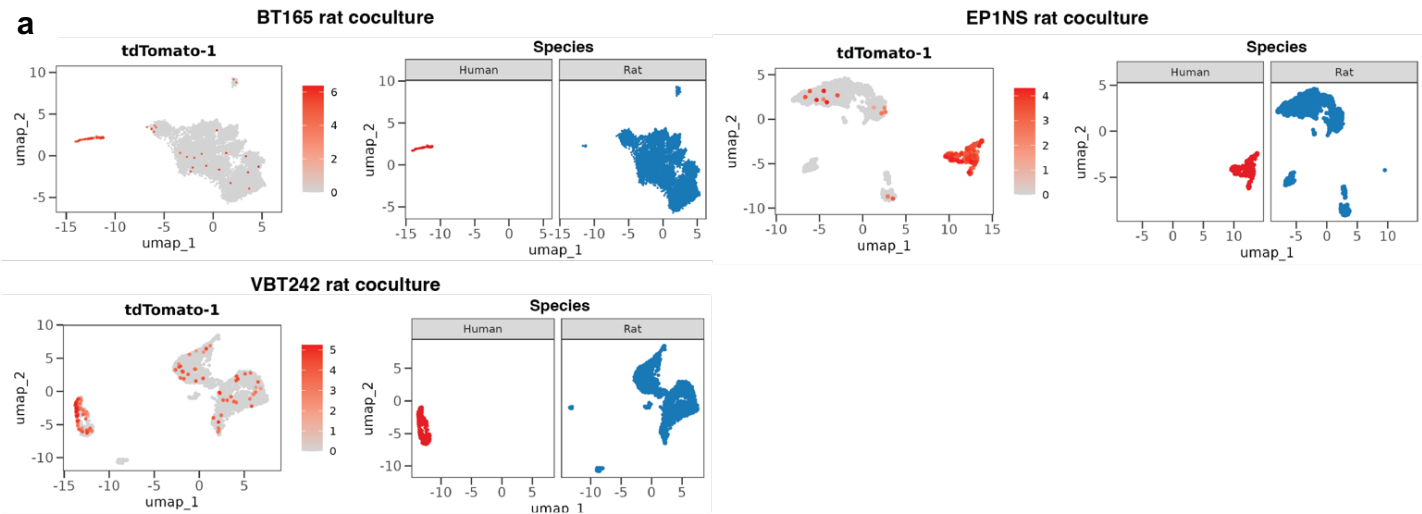


Figure R9. (a) Representative immunofluorescence images and mean intensity of signal of an *in vitro* co-culture model of ZFTA-RELA cells expressing tdTomato, categorized by morphological cell subtypes. Data represent the mean intensity value of $n = 155$ cells acquired from $n = 3$ independent biological replicates per marker. (b) Dotplot showing scaled mRNA expression of marker genes for each metaprogram in ZFTA-RELA patient tumors profiled by sc/snRNAseq. Dot size represents percent expressed, while color scale represents the average expression. (c) (left) Immunofluorescence image of ZFTA-RELA (EP1NS cells expressing tdTomato) cocultured with rat E19 cortical cells, stained for anti-tdTomato and DAPI. (right) Immunofluorescence image overlaid with 10X Xenium transcripts. Each dot represents detected transcript from the custom ($n = 103$) and base panel ($n = 255$) probes used for patient tumor sections. Table indicates top 5 enriched markers in demarcated region (white).

Figure R10



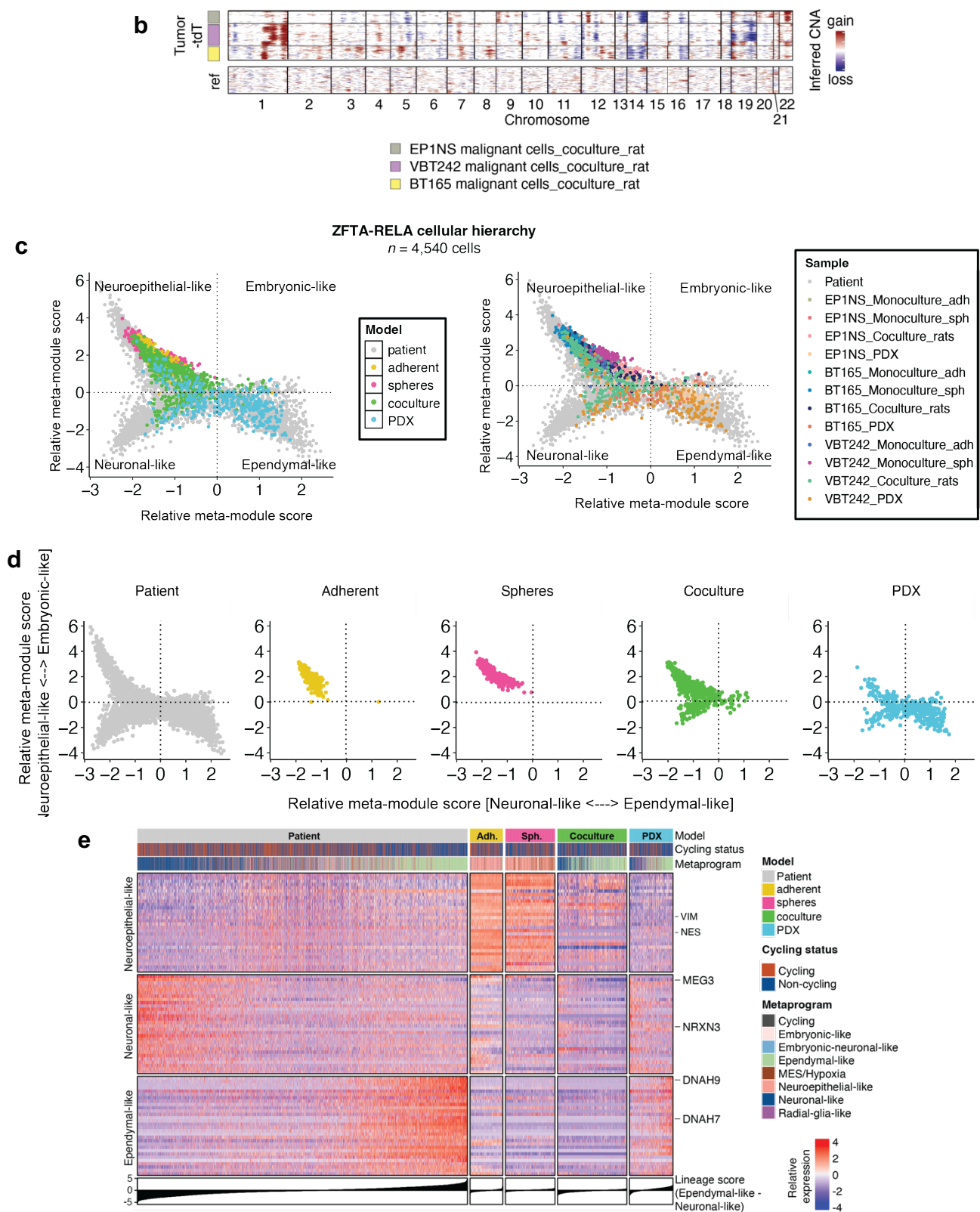


Figure R10. (a) UMAP plots of each ZFTA-RELA tumor cell line (EP1NS, BT165, VBT242) cocultured with rat E19 cortical cells, colored by detection of tdTomato (left) or species (right). **(b)** Copy number alterations (CNA) profiles inferred from scRNAseq data of ZFTA-RELA and ZFTA-Cluster 4 patient-derived tumor cell lines cocultured with rat E19 cortical cells. Cells are ordered by cell line as rows (EP1NS, BT165, VBT242, BT2214), and clustered by their pattern of CNAs across chromosomal locations (columns). Representative fresh nonmalignant cells lacking CNAs are shown on top. **(c)** Cell state plots of ZFTA-RELA patient tumor cells profiled by snRNAseq overlaid with single cells derived from PDX ($n=3$), adherent

($n=3$), neurosphere ($n=3$) and coculture with rat E19 cortical cells ($n=3$) models of EP1NS, BT165, and VBT242. Plots are colored by model (left) and by sample (right). **(d)** Cell state plots, separated by ZFTA-RELA patient tumors profiled by snRNAseq, adherent monoculture, spheroid monoculture, coculture with rat E19 cortical cells, and PDX profiled by sc/snRNAseq. **(e)** Heatmap showing relative expression of the top 50 genes identified by NMF analysis specific to neuroepithelial-like, neuronal-like, and ependymal-like signatures. Cells are sorted according to the overall expression of ependymal-like versus neuronal-like gene programs. Model (patient, adherent, spheroid, coculture, or PDX), cycling status, and metaprogram assignment are indicated on the top bars. Select genes of interest are indicated on the right.

Furthermore, the question regarding neuronal-like cell TMs is very interesting. To address this, we examined the length and thickness of tumor processes across cell states. Overall, neuronal-like tumor processes were significantly longer than other cell states, spanning multiple hundred microns (**Figure R21a**). This extensive length is reminiscent of axons in neurons that span from a couple hundred microns to millimeters^{28,29}. Previous studies have defined TMs as neurite-like membrane protrusions with a minimum length of 10 μm ^{30–33}. Thus, by length, neuronal-like cells satisfy the definition of both TMs and axon-like processes. In addition, overall tumor processes of neuronal-like cells are significantly thinner than neuroepithelial-like-2 cells, but not significantly thinner than the immature ependymal-like cells (**Figure R21a**). Axons of normal neurons have been shown to have a heterogeneous inner diameter ranging from 0.16 to 9 μm in the human cortical white matter³⁴. Overall, we cannot distinguish whether neuronal-like TMs are axon-like by the thickness of TM processes due to the lack of differences in TM thickness across cell states and the heterogenous range of axon diameters.

Interestingly, we observed different widths along the length of neuronal-like TMs, where the middle part was the thinnest, and the end part was the thickest (**Figure R21c**). This is a common structural feature of axons, where growth cones at the end part of axons are structurally thicker and appear as bulbous or expanded structure^{35,36}. Indeed, we observed that the end part of neuronal-like cells in both *in vivo* and *in vitro* models morphologically resemble axonal growth cones with a bulbous structure (**Figure R21b-c**). These observations need further functional validation in future studies such as immunofluorescence staining of growth cone-related proteins and/or other systematic studies of the dynamics of TM end parts over time.

In summary, while there is some evidence alluding to neuronal-like TMs as axon-like processes based on their distinctly long length and morphologies at the end of the processes, we do not yet have enough evidence to conclusively define these TMs as axon-like processes. We thank the reviewer for this interesting discussion and continue to describe the neuronal-like TMs as structures “reminiscent of developing neuronal axons” in the manuscript.

Figure R21

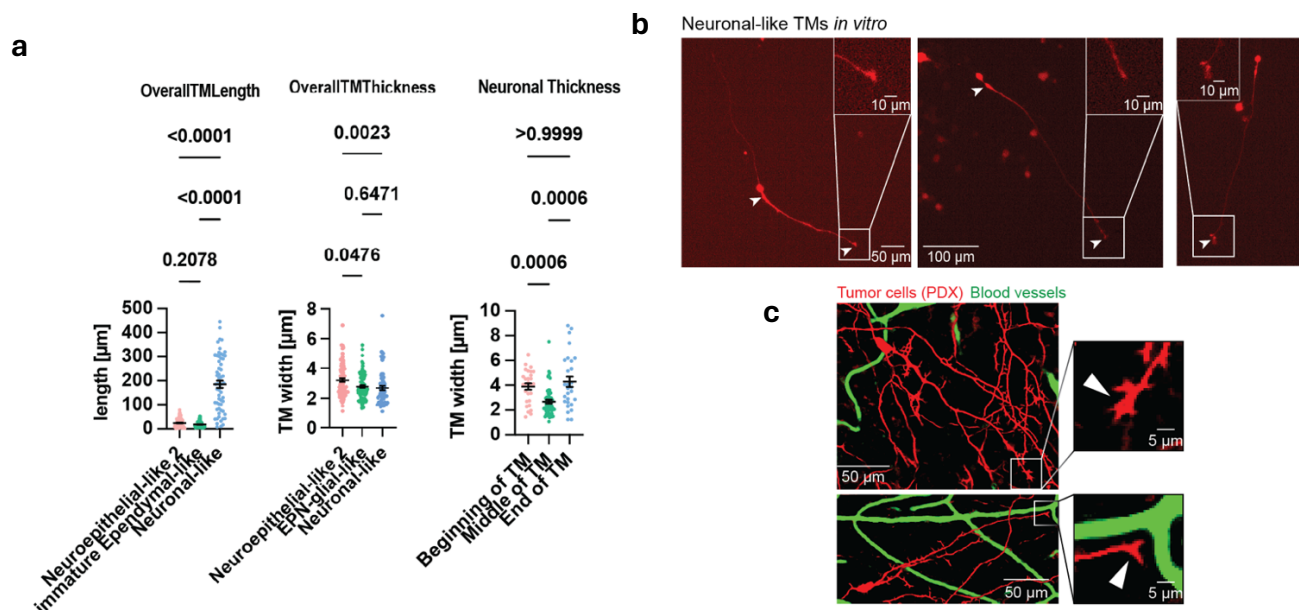


Figure R21. (a) Quantification of TM thickness *in vitro*, measured in the middle of each TM in μm ($n = 88$ neuroepithelial-like, 92 immature ependymal-like and 59 neuronal-like TMs, Kruskal-Wallis test). Middle: Thickness of TM of neuronal-like cells measured at the beginning of each TM ($n = 28$), the middle part ($n = 59$) and end tip ($n = 28$) of the TMs, Kruskal-Wallis test). **(b)** Representative images of neuronal-like ZFTA-RELA cells *in vitro* with long, tumor microtubule protrusions, depicting differences in TM width in different regions of the TM (arrowhead pointing to wider TM regions). Insets show TM tips. **(c)** Probability maps of representative images of neuronal-like ZFTA-RELA cells *in vivo* with long tumor microtubules. The ends of TMs resemble axon growth cones (inset, arrowhead).

Changes made: Result section “Superimposing molecular and morphological cell state characteristics in ZFTA-RELA tumors”

#2-7. For data shown in Figure 5b, please include the methodology for quantifying the staining intensity in the methods.

We thank the reviewer for this comment. As addressed in **Reviewer comment #2-1**, we amended the methodology to an automated fluorescence quantification method. We have modified **Methods** to include this new methodology.

Methods: Quantification of immunofluorescence of ZFTA-RELA coculture cells. tdTomato signal of each cell was segmented by setting a pre-set threshold to capture the morphology and area of the cell. Automated colocalization of fluorescence analysis was then performed using *Coloc 2* to derive Pearson’s Correlation coefficient for colocalization of tdTomato signal and other channels. All images were processed using ImageJ software (National Institutes of Health).

Changes made: Methods

#2-8. In Extended data Fig 8a it would be clearer to label the cells of interest in each image. The strong positive vimentin staining in the EPN-glial-like and Neuronal-like groups make it hard to identify the cells of interest. In Fig 5b, the Neuronal-like group is depicted as having no vimentin positivity, yet the representative image in Ext Fig 8a does not clearly reflect this analysis, the authors should replace with better and clearer representative images. Representative images for the Beta-III-tubulin and GFAP quantification should also be included.

We thank the reviewer for this important observation. As also addressed in Reviewer comment #2-1, we performed additional IF images, and quantification to capture the full staining patterns for each morphological subtype (**Figure R18**). We also included representative images for Nestin, Vimentin, S100B, GFAP, and Beta-III-tubulin (**Figure R18, Extended Data Figure 8c**). We see that with this new staining, Vimentin, GFAP, and Beta-III-tubulin expression is absent in all morphological subtypes. Nestin is expressed exclusively by neuroepithelial-like-1 and neuroepithelial-like-2, as expected, suggesting that this marker is specific for the progenitor cell type. In addition, DCX is specific to neuronal-like cells, and S100B is the most highly expressed in immature ependymal-like cells, while also expressed at lower levels in neuroepithelial-like-1 and neuroepithelial-like-2. Finally, we included IF stainings of CCDC40, which is specifically expressed in immature ependymal-like cells.

Figure R18

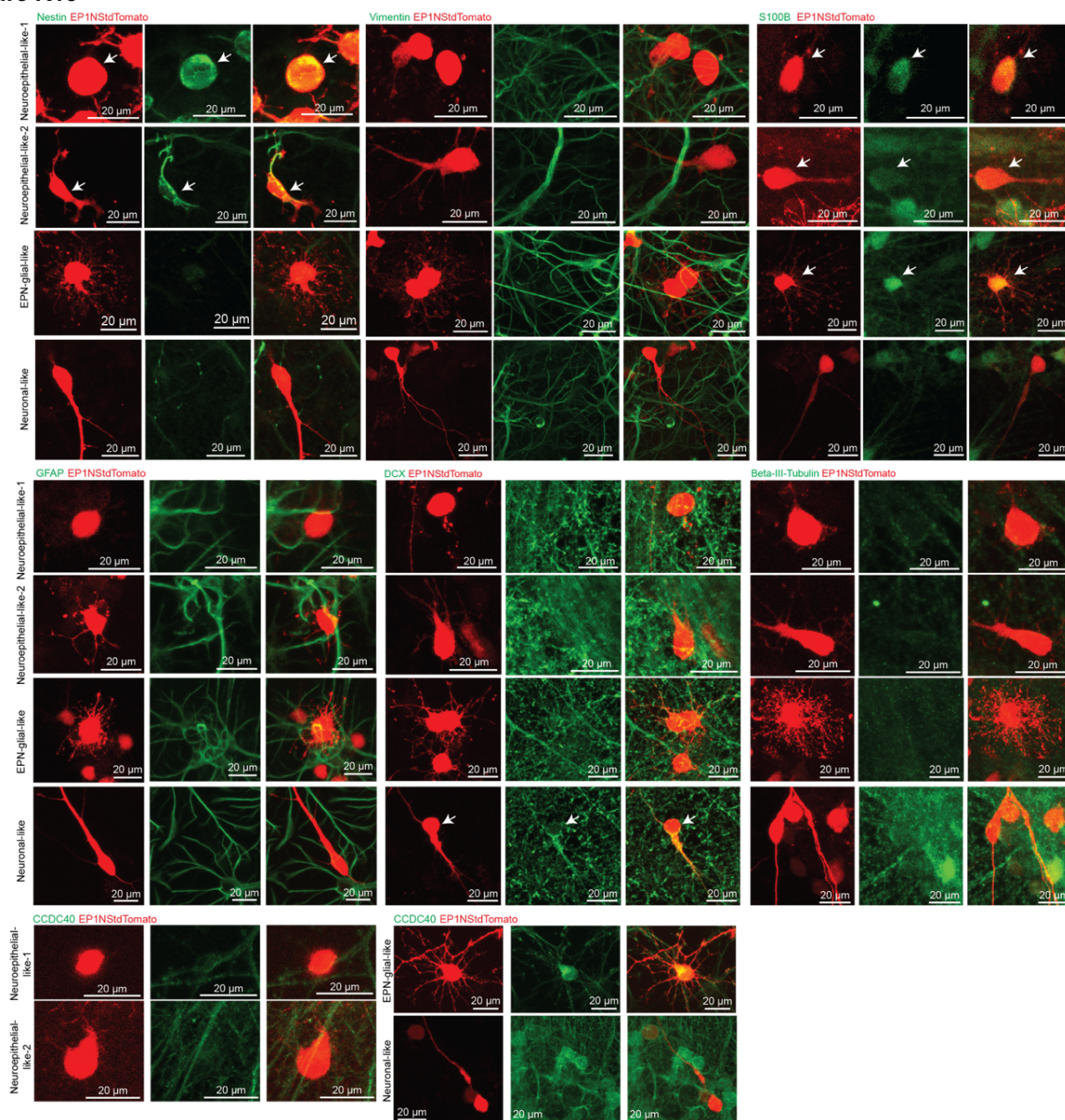


Figure R18. Representative immunofluorescence images of ZFTA-RELA (EP1NS cells expressing tdTomato) cells in coculture with rat E19 cortical cells.

Changes made: Extended Data Figure 8c

#2-9. Harvesting cortical neurons produces a mixed population of glutamatergic and GABAergic neurons in culture; the authors should quantify the different neuronal types (glutamatergic vs GABAergic) neurons present in the cocultures, and the possible presence of astrocytes.

As also addressed in **Reviewer comment #2-1**, we performed 10x scRNAseq of 3 different tumor cell lines (EP1NS, BT165, and VBT242) cocultured with rat E19 cortical cells. We separated the tumor and non-malignant populations by the enrichment of human versus rat genes (**Figure R10a-b**) and examined the tdTomato expression and copy number variation to validate the malignancy of the tumor cells in coculture. We

then isolated and characterized the rat cells by marker genes expression and found that there is indeed a mixture of excitatory neurons (21%), GABAergic neurons (3%), and astrocytes (3%) (**Figure R7f**). We additionally detected other cell types, including radial glia, neural progenitor cells (NPCs), immature neurons, oligodendrocyte precursor cells (OPCs), microglia and endothelial cells (**Figure R7f**). As these rat E19 cortical cells are composed of a mixture of neurons, glial cells, immune and vascular cells, we adjusted the nomenclature as “rat E19 cortical cells” in the manuscript and included these new analyses in **Extended Data Figure 7f**.

Figure R7

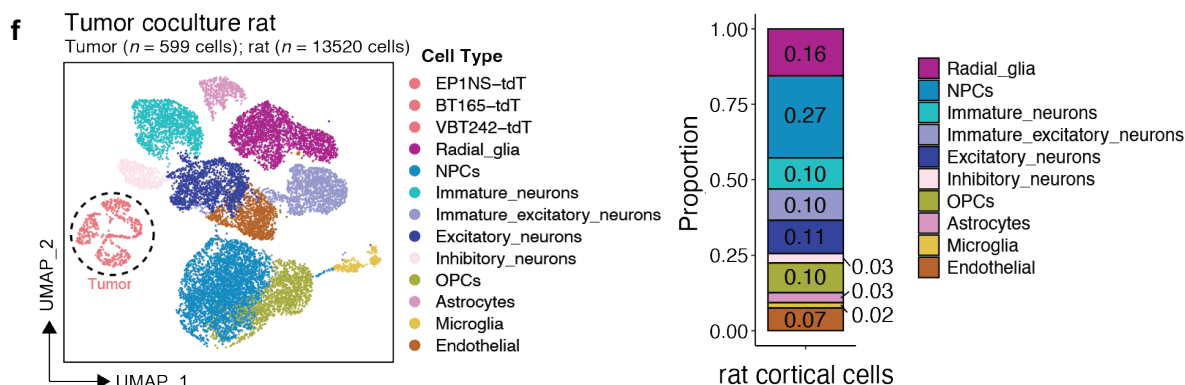


Figure R7. (f) (left) UMAP of ZFTA-RELA tumor cells (EP1NS-tdTomato, BT165-tdTomato, VBT242-tdTomato) cocultured with rat E19 cortical cells. Plots are colored by the cell type, and tumor cells are highlighted by black dotted circle. **(right)** Proportion (y-axis) of cell types in the right E19 cortical cells (x-axis).

Changes made: Results and Discussion [replaced “primary rat neurons and astrocytes” to “rat E19 cortical cells”], Extended Data Figure 7f

#2-10. In support of the data shown in Figure 5b, the expression level of the markers (Vimentin, Nestin, S100B, GFAP, B-tubulin, DCX) in each metaprogram should be included.

We appreciate the reviewer’s insight. To address this point, we included an additional dotplot showing the expression level of the markers in ZFTA-RELA patient tumors (**Figure R9e**). We in addition included *ARL6IP1*, *CENPF*, *SOX10*, *DNER*, *MAP3K19*, and *CCDC40*. Cells belonging to the Embryonic-like program were excluded from the plot representation due to their low frequency ($n = 9$ out of 2,854 total cells). We included this plot in **Extended Data Figure 8a**.

Figure R9

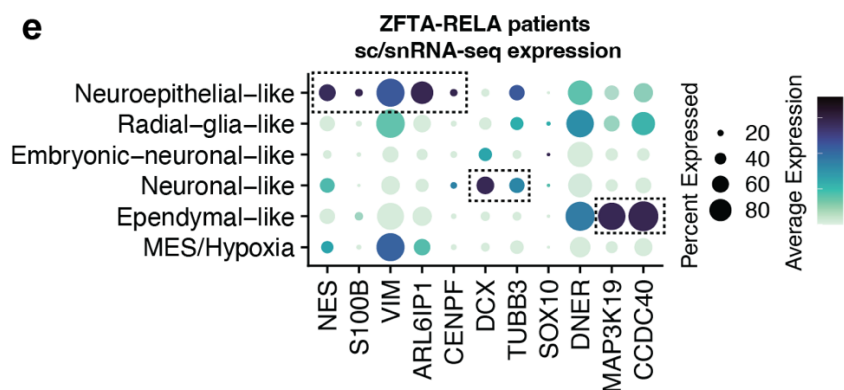


Figure R9. (e) Dotplot showing scaled mRNA expression of marker genes for each metaprogram in ZFTA-RELA patient tumors profiled by sc/snRNAseq. Dot size represents percent expressed, while color scale represents the average expression.

Changes made: Extended Data Figure 8a

#2-11. The experimental data presented in this manuscript beautifully depict the motility and migratory characteristics of the tumor cells, however they do not address the invasive capacity of these cells. The invasiveness of a cancer cell is measured by the cells' ability to restructure, breakdown and move through a 3D barrier/extracellular matrix. The authors should either employ 3D in vitro invasion assays in their study on the three cellular states depicted in Figure 5, or replace any "invasiveness/invasive state/invasion pattern" text with terms such as "migration", "motility" and "movement" (including the label on Figure 5g) when discussing in vitro 2D data.

We thank the reviewer for this clarification. We have changed the wording from "invasion" to "migration", "motility", or "movement" in the text and relevant figures and extended data figures.

Changes made: Results, Discussion, Figure and Extended Data Figure legends [replaced "invasion" with "migration"]

#2-12. Given the differences in cell state between in vitro and in vivo, measuring and comparing the rate/distance of protrusion, retraction and branching in the neuronal-like cells in in vivo (Figure 5f) and in vitro (Extended Data Figure 9) could further strengthen the observations that the neural microenvironment supports the neuronal-like population.

We thank the reviewer for this constructive suggestion. To address this point, we quantified the turnover of TMs that include protrusion and retraction, as well as branching patterns of the different cell states, both *in vitro* and *in vivo*. In both *in vitro* and *in vivo* conditions, neuronal-like cells have a significantly higher overall TM turnover than other cell states (**Figure R22a-b**). In addition, neuroepithelial-like-2 cells did not exhibit different TM turnover to immature ependymal-like cells in either condition (**Figure R22a**). Furthermore, we analyzed the ratio of branching TMs between the cell states and found no significant differences across cell states in coculture and PDX conditions (**Figure R22c-d**). Overall, we find strong similarities in the TM dynamics of cell states in coculture and PDX models, further strengthening the importance of the neural microenvironment in supporting the neuronal-like population. We have included these new analyses in **Extended Data Figure 9f** and **Results**.

Figure R22

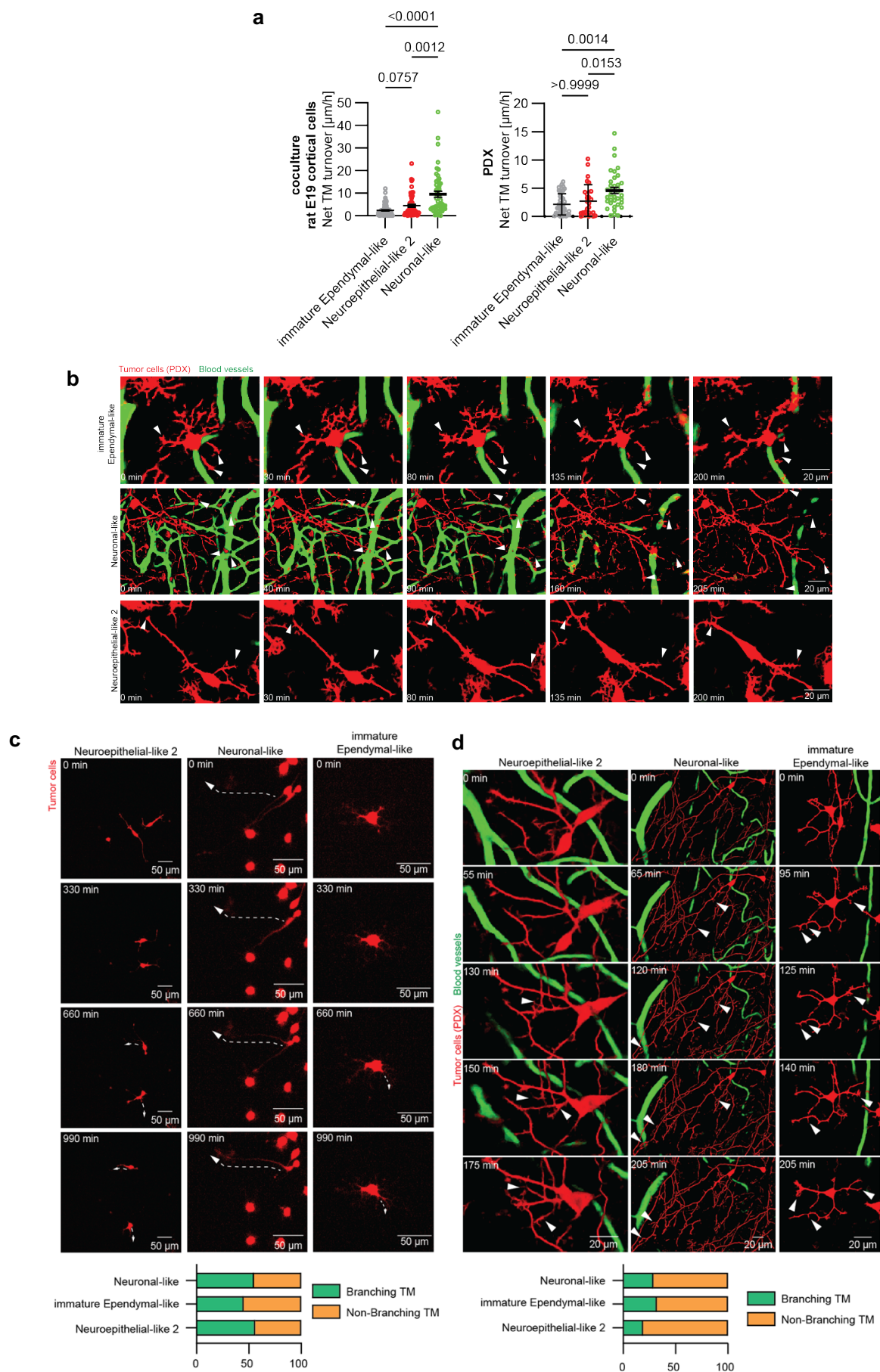


Figure R22. (a) (left) Net TM turnover over *in vitro* over 4 hours, including positive values for protrusion and retraction ($n = 35$ immature Ependymal-like, $n = 27$ neuroepithelial-like-2, $n = 37$ neuronal-like TMs, Kruskal-Wallis test). **(right)** Net TM turnover *in vivo* over 4 hours, including positive values for protrusion and retraction ($n = 96$ immature ependymal-like, $n = 57$ neuroepithelial-like-2, $n = 50$ neuronal-like TMs, Kruskal-Wallis test). **(b)** Probability maps of exemplary time-lapse images of TM dynamics in three cell states *in vivo*. Arrowheads point to TM tips. **(c) (left)** Exemplary time-lapse images of branching TMs in three different cell states in coculture. Dashed arrows show exemplary TM branches. **(right)** Fraction of branching and non-branching TMs in three different cell states in coculture ($n = 55.7\%$ branching TMs out of 88 neuroepithelial-like-2 TMs, 44.9% branching TMs out of 89 immature Ependymal-like TMs, and 55.17% branching TMs out of 57 neuronal-like TMs, Fischer's exact test 0.2477). **(d) (left)** Probability maps of exemplary time-lapse images of branching TMs in three different cell states *in vivo*. Arrowheads point to TM branches. **(right)** Fraction of branching and non-branching TMs in three different cell states *in vivo* ($n = 19\%$ branching TMs out of 27 neuroepithelial-like-2 TMs, 32% branching TMs out of 35 immature Ependymal-like TMs, and 29% branching TMs out of 41 neuronal-like TMs, Fischer's exact test 0.0939).

Results: Overall neuronal-like cells exhibited the most net TM turnover, which includes both protrusion and retraction, out of all cell states in both *in vitro* coculture with rat E19 cortical cells and *in vivo* models (Extended Data Figure 9f).

Changes made: Extended Data Figure 9f, Results

#2-13. In line 331, the authors state their data strengthen the concept of cooperativity of cell states for tumor progression. Whilst cooperativity is an established concept in cancer biology, the data here does not confirm this in EPN. In order to assess the importance of cooperativity for tumorigenicity in EPN, the authors should sort the individual cellular compartments and inject orthotopically to compare bulk tumor to individual subcellular population growth. Alternatively, the authors can remove the concluding remark in the results.

We thank the reviewer for this important clarification. We have removed the concept of cooperativity in the concluding remarks in Results.

Results: Taken together, these results suggest that ZFTA-RELA transcriptomic cellular states have distinct morphological and behavioral attributes in the context of their native neuroglial microenvironment, in which neuronal-like cell states are both highly migratory and adopt distinct migratory patterns reminiscent of immature neurons during development. We further confirm these patterns across species, underscoring the phenotypic heterogeneity of these cancer cell states influenced by the tumor microenvironment, which collectively contributes to tumor progression.

Changes made: Results

#2-14. The authors identify that the EPN-glial-like and neuronal-like cells in coculture are less proliferative than the neuroepithelial-like cells, however they also demonstrate there are far less of the EPN-glial-like and Neuronal-like cells in coculture compared to PDX. The authors should confirm the proliferation rate of each cell state *in vivo*, given the PDX provides a more even distribution of each cellular subpopulation.

We thank the reviewer for this insightful comment. To first validate our *in vitro* live-cell imaging results regarding the proliferation rate of each cell state, we examined the proportion of cycling cells of each cell state in the 10X sc/snRNA-seq data of PDX models and compared it to patients' tumor data as well as our coculture models with rat E19 cortical cells (Figure R23). We confirm that the neuroepithelial-like cells are the most proliferative cell state, and both neuronal-like and ependymal-like cells are less proliferative in PDX, patients, and coculture data. Interestingly, we also observed that neuronal-like cells in patient and PDX are more proliferative than in coculture models. This suggests that additional tumor microenvironmental factors in patients and PDX promotes the proliferation of neuronal-like cells, and it would be interesting in future studies to explore further what those factors specifically are. Overall, we also noted some discrepancy between calling cell cycle by using scRNA-seq data and our *in vitro* live-cell imaging data, which count actual morphological cell division: While in scRNA-seq, where a proportion of neuronal-like cells express cell cycle genes (~20%), while we did not detect dividing neuronal-like cells by imaging. Possible reasons could be that the cell doubling

time is different for each cell state, and the neuronal-like cells need to be observed for a longer period to capture cell division, or we did not capture the dividing neuronal-like cells in our coculture models *in vitro* using imaging as the majority of them (~80%) are not proliferative. Overall, our conclusion is that neuroepithelial-like cells are the most proliferative and ependymal-like cells are the least proliferative cells across models. We have amended our findings in **Results**, and included these analyses in **Extended Data Figure 2f, 9h**.

Figure R23

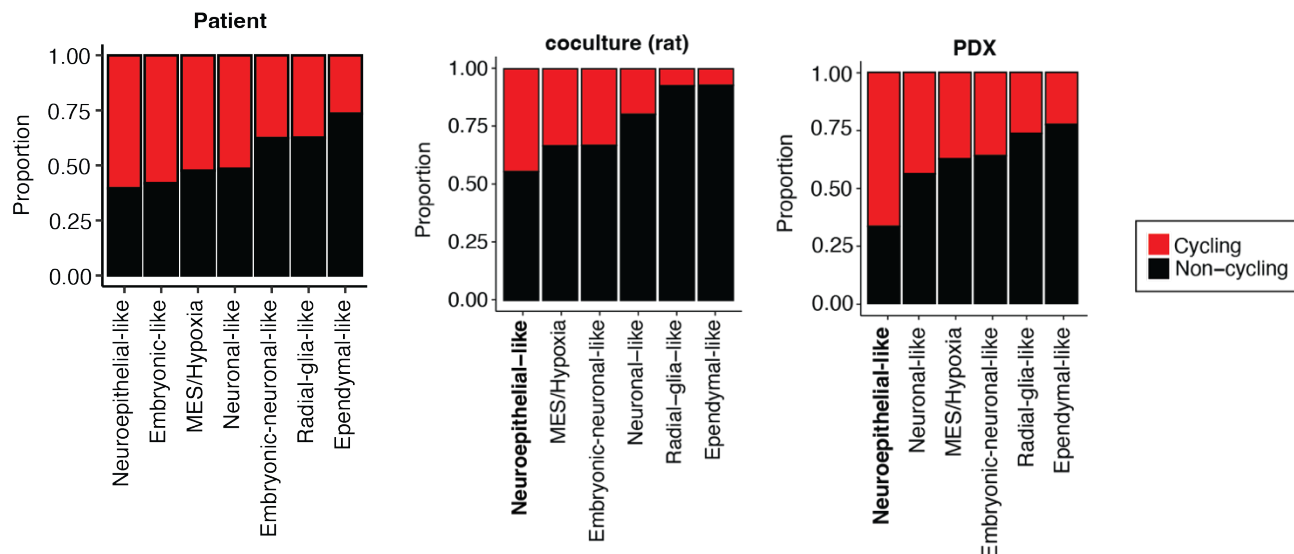


Figure R23. Proportion (y-axis) of cycling (red) and non-cycling (black) cells in each metaprogram in patients (left), coculture with rat E19 cortical cells (middle), and PDX (right) profiled by sc/snRNAseq.

Results: Furthermore, we measured whether different cellular states have different proliferation rates. Neuroepithelial-like-2 cells were the most proliferative compartment (**Figure 5i**, **Extended Data Figure 9g**), which was consistent with our sc/snRNA-seq results (**Extended Data Figure 9h**), and with the expected behavior of immature progenitor cells. These findings overall suggest that neuroepithelial-like-2 cells are both highly migratory (**Figure 5d**) and highly proliferative (**Figure 5i**).

Changes made: Results, Extended Data Figure 2f, 9h

#2-15. Transition toward neuronal-like lineages: The authors suggest that normal brain cells induce the transition from neuroepithelial-like progenitor cells toward neuronal-like lineages (lines 342-343). Why is this mainly observed in ZFTA-RELA tumors and less so in other subtypes (e.g., clusters 1-4, main figure 2e), where the neuronal-like state is less apparent? The authors should test whether this transition also occurs in other subtypes using a similar experimental setup. Finally, is this cell state transition neuronal activity-regulated, such as blocked with TTX in co-culture??

We thank the reviewer for this interesting comment. While cell lines derived from non-ZFTA-RELA subgroups are non-existent, we derived one cell line that is ZFTA-Cluster 4 subgroup from a patient tumor sample (the first to our knowledge). We then examined whether coculture with rat E19 cortical cells induces cell state transition as observed in ZFTA-RELA tumor cell lines. We cocultured ZFTA-Cluster 4 tumor cell line BT2214 in monoculture as well as with rat E19 cortical cells for 7 days and performed 10X scRNAseq (**Figure R24**). We saw a statistically significant decrease in the embryonic-like cells, while no difference in the neuroepithelial-like signature (**Figure R24b**). In addition, we saw a significant decrease in radial glia-like cells and an increase in embryonic-neuronal and neuronal-like signatures. We did not see an increase in the ependymal-like signatures. These findings suggest that there are specific progenitor populations (embryonic-like and radial glia-like) that transition in the presence of rat E19 cortical cells towards embryonic-neuronal and neuronal-like cells, while not to ependymal-like cells. This generally concurs with the metaprograms detected in ZFTA-Cluster 4 patient tumors, where there are embryonic-neuronal-like and neuronal-like cells,

with a low proportion of ependymal-like cells. Overall, these analyses suggest that both ZFTA-RELA and ZFTA-Cluster 4 tumor cells differentiate in the presence of rat cortical cells. It would be invaluable for the field going forward to acquire cell lines of other subgroups (ZFTA-Cluster 1-3, ST-YAP1) to comprehensively compare the exact intrinsic and extrinsic mechanisms that direct these transitions. We have addressed these aspects in **Limitations of the Study**.

Figure R24

[FIGURE REDACTED]

Figure R24. [TEXT REDACTED]

Furthermore, to determine whether the blocking activity of normal neurons could abrogate the cell state transition of ZFTA-RELA tumor cells, we conducted the proposed experiment by adding TTX (1 μ M) to tumor cells at the start of coculture with rat E19 cortical cells, and continuing coculture for 7 days before dissociating the cells for scRNAseq. We saw a statistically significant decrease of neuroepithelial-like cells, and an increase of neuronal-like cells (**Figure R25a-b**). In addition, we saw a very small decrease in the proportion of cycling cells (**Figure R25c**). These findings suggest that blocking overall neuronal activity may decrease the proliferation of cells and, concordantly, increase cell state transition. While neuronal activity promotes glioma proliferation³⁷, it would be interesting in future studies to examine further the effect of neuronal activity on cell state transitions.

Figure R25

[FIGURE REDACTED]

Figure R25. [TEXT REDACTED]

Limitations of the Study: **Limitations of the study.** In this study, we characterized tumor heterogeneity of patient samples diagnosed as supratentorial ependymomas. However, there may be a sample site bias in our data set due to random sampling of the whole tumor mass during surgery as well as dividing the taken tumor into pieces for sc/snRNAseq and spatial transcriptomics. Therefore, our findings are limited due to the sampling of tumors, and do not have a conclusive claim of the entire tumor mass present in patients. In addition, due to limited availability of non-canonical ZFTA tumors and corresponding patient-derived cell lines, we have not included spatial transcriptomics and other functional live-cell imaging data of these tumor subgroups. We also did not include ST-EPN-SE subgroup tumor samples in our analyses. Finally, the mesenchymal and hypoxia cell state signatures are not a direct measurement of hypoxia itself, and rather act as proxy for potential effects or consequences of local and global hypoxic microenvironments.

Changes made: Limitations of the Study

#2-16. Computational comments:

#2-A. Assignment of malignant cells: After reading the authors' analytic framework (in lines 119-123 and in the methods section), it is unclear to me whether the classification of malignant cells is entirely accurate. In the methods section, they state that it was based on the "presence of copy-number alterations," but this process is complex, and they do not provide sufficient detail or show the inferred chromosomal aberrations in the so-called malignant group versus the reference group. Extended Data Figure 1a shows many false positives and false negatives (i.e., cells classified as true by inferCNV in the TME clusters and vice versa). Please address these concerns.

We thank the reviewer for pointing out the lack of clarity in the description of our methods. To discriminate between malignant and non-malignant cells, we performed an initial annotation of cell lineages by unsupervised Louvain clustering and examined clustering patterns as well as expression of known marker genes of non-malignant cells, including immune, endothelial, and oligodendrocyte cells. We corroborated our initial cell assignment by the automated annotation using the package *SingleR*³⁸, which annotates single nuclei based on their similarity to a reference collection of bulk transcriptomes of healthy cells (Human Primary Cell Atlas data). We then inferred genome-wide evidence for somatic chromosomal copy number variation (CNVs) using the *inferCNV* R package. We spiked-in non-malignant immune and oligodendrocyte nuclei/cells from previous publications^{14,39,40} as reference normal karyotype for CNV inference, and used hierarchical

clustering of the single-cell copy-number profiles within each sample with the reference non-malignant copy-number profiles to classify the presence of CNVs. Due to the presence of false positive/negative rates of inferring CNVs in single-nucleus RNA-seq data from frozen samples, CNV results were used as secondary validations.

CNVs should be detected in cell clusters classified as malignant cells by transcriptional clustering. In contrast, CNVs should be absent from the cell cluster classified as a normal cell type. Reassuringly, we observed that nuclei lacking CNVs typically clustered together and were identified as macrophage, T cell, or endothelial cell nuclei based on their high expression of myeloid, T cell, or endothelial markers, respectively. All other clusters consistently exhibited CNVs and were classified as malignant tumor cells. False positive and false negative cells were classified by clustering, *SingleR* annotations, and marker gene expression. Only cells identified as malignant were selected for subsequent NMF analysis.

In response to the reviewer's comment, we have added additional information in **Methods** and the inferred chromosomal aberration plot of frozen (snRNA-seq) and fresh (scRNA-seq) cells identified as malignant and non-malignant in **Extended Data Figure 1d (Figure R26)**.

Figure R26

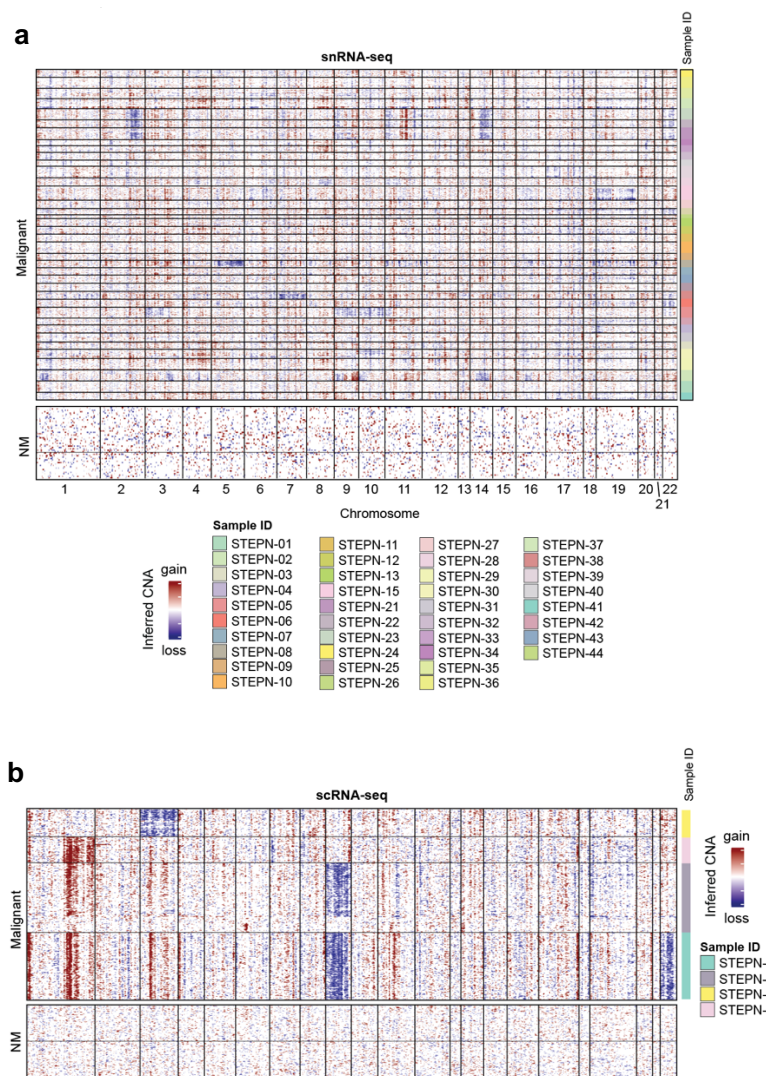


Figure R26. CNV profiles of each cell (y-axis) across chromosomal location (x-axis) inferred from sc/snRNA-seq data. Cells are colored by tumor subtype and ordered according to hierarchical clustering. Cells are split by inferred malignant/non-malignant origin. Representative fresh spike-in nonmalignant cells are shown at the bottom. **(a)** Frozen snRNA-seq and **(b)** Fresh scRNA-seq data.

Methods: To discriminate between malignant and non-malignant cells, we used: (1) Seurat clustering, (2) unbiased cell type annotation using the automated annotation package SingleR (v1.6.1) using the Human Primary Cell Atlas⁴¹ data as a reference, and (3) inference of copy-number variation using the InferCNV R package (v1.8.0). We began by performing an initial annotation of cell lineages through dimensionality reduction using UMAP and unsupervised Louvain clustering. To refine our classifications, we analyzed clustering patterns and the expression of established immune, endothelial, and oligodendrocyte gene markers. We validated these initial lineage assignments using the automated annotation package SingleR (v1.6.1), which assigns cell identities based on their similarity to a reference set of bulk transcriptomes from healthy cells (Human Primary Cell Atlas data⁴¹). We then inferred genome-wide evidence for somatic chromosomal copy number variation (CNVs) using the *inferCNV* R package as previously described^{14,39,42–44}. We spiked-in non-malignant immune and oligodendrocyte nuclei from previous publications^{14,39,44} as reference normal karyotype for CNV inference, and used hierarchical clustering of the single-cell copy-number profiles within each sample with the reference non-malignant copy-number profiles to classify the presence of CNVs. Due to presence of false positive/negative rates of inferring CNVs with nuc-seq data from frozen samples, CNV results were used as secondary validations. CNVs should be detected in cell cluster classified as malignant cells by transcriptional clustering. In contrast, CNVs should be absent from the cell cluster classified a normal cell type. Reassuringly, we observed that nuclei lacking CNVs typically clustered together and were identified as macrophage, T cell, or endothelial cell nuclei based on their high expression of myeloid, T cell, or endothelial markers, respectively. All other clusters consistently exhibited CNVs and were classified as malignant tumor cells. Only cells identified as malignant were selected for subsequent NMF analysis.

Changes made: Extended Data Figure 1d, Methods

#2-B. Deriving cell states (Figure 2): I have several comments:

B1. The authors use NMF, a well-established method, to derive cell states, but they assume a rank of 6 in this process. Why was this specific rank chosen? Newer implementations of this method do not assume a fixed rank per sample a priori, but rather iterate across multiple ranks and retain the NMF programs that are consistent and recurrent.

We thank the reviewer for this question. To address this point regarding the validity of rank $k=6$, we performed additional NMF analyses using other ranks, specifically variable ranks from 5 to 8 (**Figure R27a**). These results show that ranks 5-8 yield metaprograms that are comparable and consistent to the ones identified with a rank of 6.

In addition, to evaluate the robustness of our NMF decomposition and choice of factorization rank we performed a rank survey ranging from 3 to 14 with 100 runs per rank. We assessed multiple metrics in the rank survey, including cophenetic correlation, dispersion, explained variance (evar), residuals, residual sum of squares (RSS), silhouette score, and sparseness. The cophenetic correlation remained high and stable between ranks 3 and 6, indicating consistent clustering solutions (**Figure R27b**). Dispersion decreased and plateaued around rank 5-6, further supporting model stability within this range. The explained variance increased with higher ranks but showed marginal gains beyond rank 6, suggesting limited additional explanatory power. Similarly, the silhouette score peaked at rank 5 and remained relatively stable around rank 6, indicating that the cluster separation did not deteriorate significantly as the rank increased. Finally, RSS and residuals both show an inflection point around rank 6, further supporting that 6 is an optimal rank to use. Overall, these metrics demonstrate that the decomposition is stable across the tested ranks, with rank 6 offering a good balance between reconstruction accuracy and cluster consistency. Importantly, varying the rank within this range did not result in major changes in clustering structure, supporting the robustness of our choice of factorization rank.

Figure R27

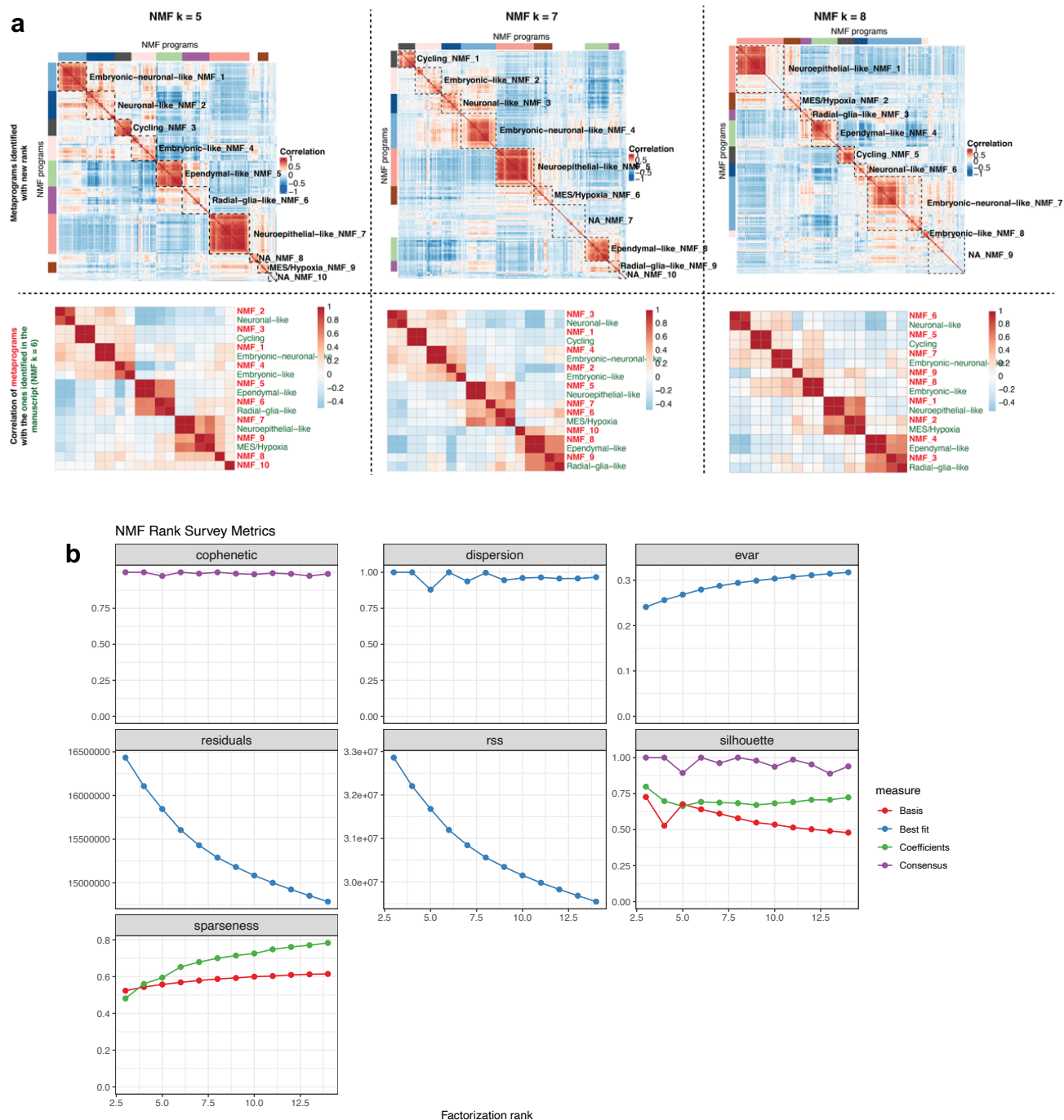


Figure R27. (a) (Top rows) Heatmaps showing Pearson correlation coefficient between individual NMF programs identified within each ST-EPN sample profiled by snRNA-seq using different NMF ranks (5, 7, and 8). NMF programs are grouped into eight recurrent metaprograms (MPs, dashed lines). **(Bottom rows)** Pairwise Pearson correlation between MPs identified with NMF ranks 5, 7 and 8 (red font) and MPs identified with NMF rank of 6 (green font), as used in the manuscript. **(b)** NMF rank survey of factorization ranks 3 to 14 with 100 runs per rank. Each point represents the results of each rank.

#2-B2. In reading the methods, I did not understand how the clustering was performed. They select the top 50 genes per NMF program (which would result in different lists of genes across programs and samples), and then cluster based on correlations? This approach doesn't make sense, as the genes being correlated are different. Please clarify.

We thank the reviewer for identifying the lack of clarity in the description of our approach. To identify metaprograms, we first selected the top 30 genes per NMF program and scored each cell in our sc/snRNA-seq data for those genes. We then correlated the **resulting scores** to identify the set of NMF programs that are highly correlated. We merged these correlated programs into metaprograms by selecting the top 30 genes with the highest average NMF weight within each correlated program set. We amended **Methods** to more clearly explain the clustering method.

Methods: Scoring gene sets. We first aggregated ST-EPN single cells/nuclei from all samples, and performed normalization and centering. Then, we scored these cells for NMF/developmental signatures as previously described^{14,39,42–44}. We did this by calculating for each cell i a score $SC_j(i)$, quantifying the average relative expression (Er) of the top 30 genes within each MP (G_j) and comparing the score to the average relative expression of a control gene (G_j^{cont}): $SC_j(i) = \text{average}[Er(G_j, i)] - \text{average}[Er(G_j^{cont}, i)]$. The control gene-set was defined by binning all genes into 30 bins of aggregate expression levels (Ea) and randomly selecting 100 genes from the same expression bin for each gene in the gene-set G_j . ST-EPN cells/nuclei were annotated based on the maximum expression score for the respective MP signatures after excluding the one related to cell cycle (“cycling”).

Changes made: Methods

#2-B3. How were the metaprograms defined based on the clustering of the NMF programs? I could not find a reference to this. Was some consensus used between programs in each cluster? There are a few clusters in the heatmap that were not marked - did the authors investigate these?

We thank the reviewer for this question. We first performed hierarchical clustering of the NMF programs using a correlation-based distance metric (distance metric: $1 - \text{Pearson correlation}$; linkage: Ward’s linkage). We then used the function *cutree* from the ‘*stast*’ R-package to automatically extract $n = 15-20$ highly correlated metaprograms (**Figure R28**). We finally curated the resulting clusters manually, by: (i) removing clusters that were associated with little sample number (and therefore considered not to reflect relevant general biology of the tumors), (ii) by removing clusters that were not associated with any significant gene enrichment ontology term, (iii) by grouping together clusters that reflected similar underlying biological processes. We have included the amended methods in **Methods**.

Figure R28

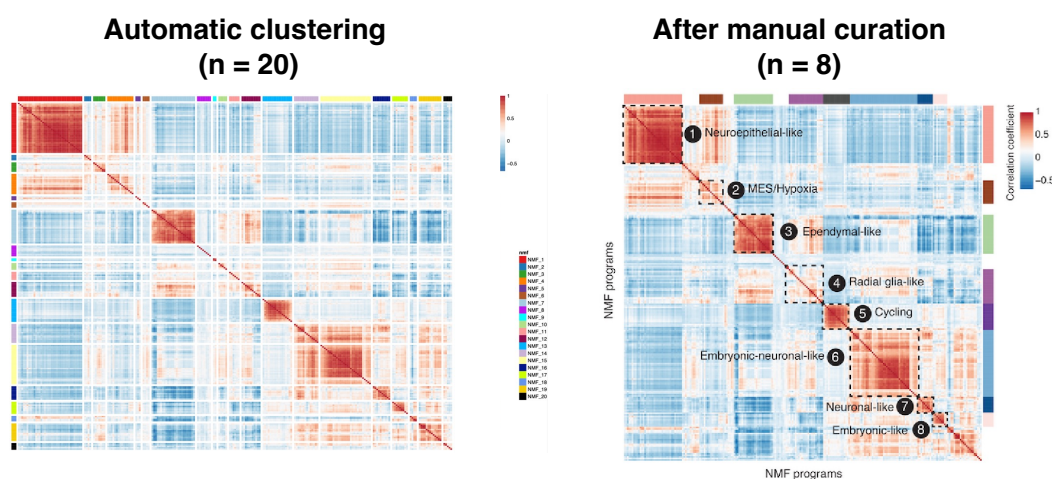


Figure R28. Heatmaps showing Pearson correlation of individual NMF programs identified within each ST-EPN sample profiled by snRNA-seq. Plots depict the automatic clustering results (left) and manual curation of final metaprograms (right).

Methods: Defining Nonnegative Matrix Factorization (NMF) programs and metaprograms. To identify heterogeneous transcriptional programs within each sample, we performed NMF analysis using frozen tumors profiled by snRNA-seq. As input for NMF, we selected the top 10,000 over-dispersed genes using PAGODA2 (v0.1.4) identified on the malignant cell compartment of each sample separately. We set the negative values to zero and ran NMF using rank = 6. To identify MPs, we first scored all malignant cells for the top 30 genes with the highest score from each NMF factor as explained in section: Scoring gene sets. We then clustered

the resulting scores across all NMF programs by hierarchical clustering as explained in section: Program-wise hierarchical clustering and used the function *cutree* from the ‘stast’ R-package to automatically extract $n = 15-20$ highly correlated MPs. We finally curated the resulting clusters manually, by: (i) removing clusters that were associated with limited samples (and therefore considered not to reflect relevant general biology of the tumors), (ii) removing clusters that were not associated with any significant gene enrichment ontology term, (iii) grouping together clusters that reflected similar underlying biological processes. This approach revealed eight highly correlated sets of programs or MPs within ST-EPN tumors.

Changes made: Methods

#2-B4. In most tumors, including EPN (as previously published), cell cycle programs clearly separate into G1S and G2M phases. I am surprised not to see this separation here.

We thank the reviewer for this interesting point. Tumors profiled in the previous EPN paper¹⁴ that have G1S and G2M phases clearly separated out are predominantly single **cell** (fresh) RNAseq samples. We have noticed that for single **nuclei** RNAseq samples, we rarely see the separation of these two phases in the cell cycle program due to stronger noise of the data. This has been observed in snRNAseq of other tumor entities³⁹. As our current cohort is predominantly single nuclei RNAseq (frozen $n=38$; fresh $n=3$), we expect to not see the separation of G1S and G2M phases.

#2-C. Spatial transcriptomics analysis (Figure 3): Please explain the decision to exclude the EPN non-canonical subgroups from the spatial analysis. This is surprising, especially in light of the limitations of previous studies, which were highlighted in the introduction and necessitated “a comprehensive study examining tumors across all subgroups of ST-EPNs” (line 80).

We thank the reviewer for this thoughtful question. To address this point, we have now performed 10X Xenium on 34 sections of 11 patient samples spanning the non-canonical ST-EPN subgroups: ZFTA-Cluster 1 ($n = 2$), ZFTA-Cluster 2 ($n = 2$), ZFTA-Cluster 3 ($n = 2$), ZFTA-Cluster 4 ($n = 3$), ST-YAP1 ($n = 2$) (**Figure R29**). We saw that the proportion of cell states detected in these tumors was comparable to the results of snRNAseq (**Figure R29a**). In addition, we derived the global spatial coherence score of each sample and found that ZFTA-Cluster 3 were the most coherent samples out of all the non-canonical subgroups. (**Figure R29b-c**). As further statistical tests and analyses were difficult due to the small number of samples available per subgroup, we did not include these analyses in the main manuscript. These results highlight the need for a comprehensive collection of these rare samples as well as the patient-derived cell lines, which we have included in **Limitations of the Study**. We have also amended the language in **Introduction** to more clearly state that spatial transcriptomics was only done on ZFTA-RELA samples.

Figure R29

[FIGURE REDACTED]

[FIGURE REDACTED]

Figure R29. [TEXT REDACTED]

Limitations of the Study: **Limitations of the study.** In this study, we characterized tumor heterogeneity of patient samples diagnosed as supratentorial ependymomas. However, there may be a sample site bias in our data set due to random sampling of the whole tumor mass during surgery as well as dividing the taken tumor into pieces for sc/snRNAseq and spatial transcriptomics. Therefore, our findings are limited due to the sampling of tumors, and do not have a conclusive claim of the entire tumor mass present in patients. In addition, due to limited availability of non-canonical ZFTA tumors and corresponding patient-derived cell lines, we have not included spatial transcriptomics and other functional live-cell imaging data of these tumor subgroups. We also did not include ST-EPN-SE subgroup tumor samples in our analyses. Finally, the mesenchymal and hypoxia cell state signatures are not a direct measurement of hypoxia itself, and rather act as proxy for potential effects or consequences of local and global hypoxic microenvironments.

Introduction: These limitations highlight the need for a comprehensive study examining the composition of malignant cell states of tumors across all subgroups of ST-EPNs.

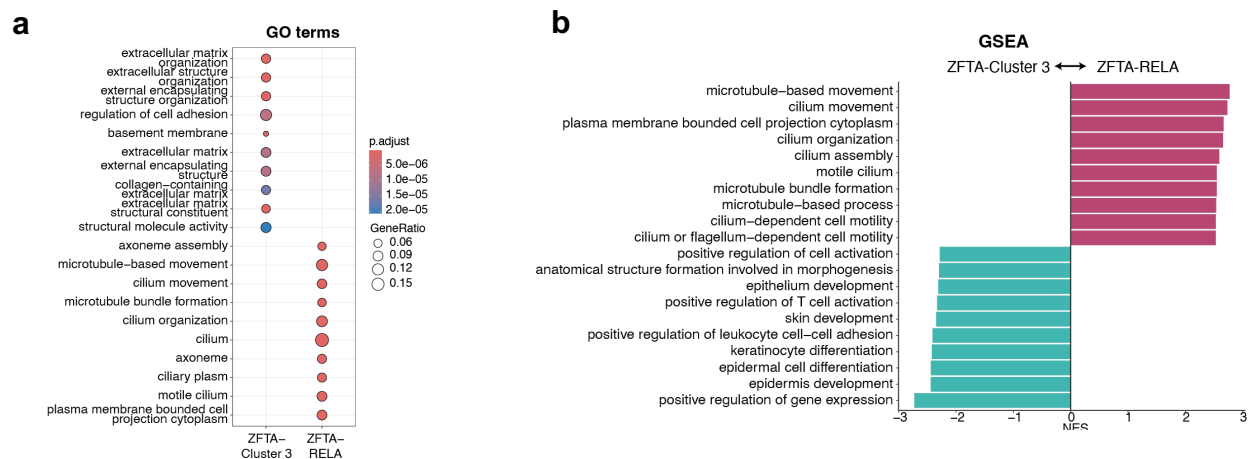
Changes made: Introduction, Limitations of the Study

Minor comments:

#2-m1. Extended Data Figure 4a (line 193): The authors performed differential expression analysis using single cells, which results in inflated p-values. From a statistical standpoint, single cells from the same sample are not independent observations, and this analysis should be avoided. The authors should consider repeating this analysis using pseudo-bulk profiles (per sample). Also, please arrange the legend group by malignant and non-malignant and re-label Mesenchymal in the legend as MES/Hypoxia for consistency.

We repeated this analysis using pseudo-bulk profiles per sample (**Figure R30**), included these new analyses in **Extended Data Figure 3b-d** and amended **Methods**.

Figure R30



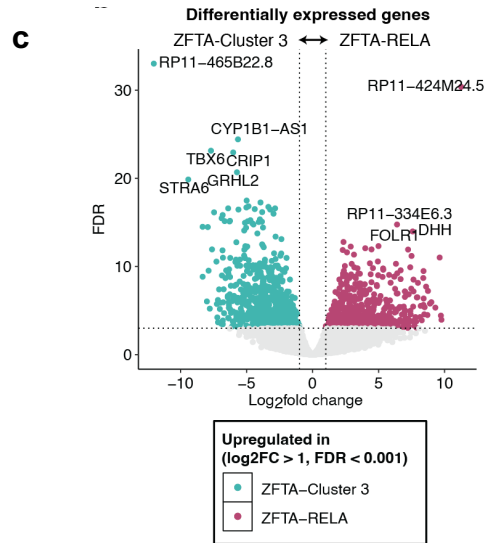


Figure R30. (a) Volcano plot showing genes differentially expressed between ZFTA-RELA and ZFTA-Cluster 3 tumors. Cells were pseudobulked by sample, and differential gene expression analysis was run using *edgeR*. Genes with $|\log_2FC| > 1$ and $FDR < 0.001$ are highlighted in color. **(b and c)** Top 10 gene ontology (GO) terms **(c)** and pathways identified by GSEA analysis **(c)** comparing genes differentially expressed (input for GO: $|\log_2FC| > 1$ and $FDR < 0.001$, input for GSEA: $FDR < 0.001$) between ZFTA-RELA and ZFTA-Cluster 3 tumors based on the genes identified in **(a)**.

Methods: Differential gene expression analysis. Pseudobulk. We performed pseudo-bulk differential expression analysis to identify transcriptional programs that distinguish ZFTA-RELA from ZFTA-Cluster 3 tumors. We first extracted raw count matrices from each tumor sample, and calculated pseudo-bulked gene expression counts using the SingleCellExperiment's *aggregateAcrossCells* function. We then conducted differential gene expression analysis using *edgeR*. We filtered out genes with low expression counts using the *filterByExpr* function, based on our grouping variable (tumor subtype). We then normalized the counts using the *calcNormFactors* function. We constructed a design matrix to include tumor subtype, estimated negative binomial dispersions using the *estimateDisp* function, and fitted a generalized linear model using *glmFit*, followed by hypothesis testing with *glmLRT* for differential expression. We considered significant genes based on a false discovery rate (FDR) threshold of <0.001 . For downstream gene ontology (GO) analysis we used the top 500 genes (ordered by $\log_2FC$) with $FDR < 0.001$ and $|\log_2FC| > 1$, for gene set enrichment analysis (GSEA), we used all genes with $FDR < 0.001$.

Changes made: Extended Data Figure 3b-d, Methods

#2-m2. Extended Data Figure 9e, include representative images of the manually counted cell division.

We thank the reviewer for this comment. We have included the representative images of the manually counted cell divisions in **Extended Data Figure 9g (Figure R31)**.

Figure R31

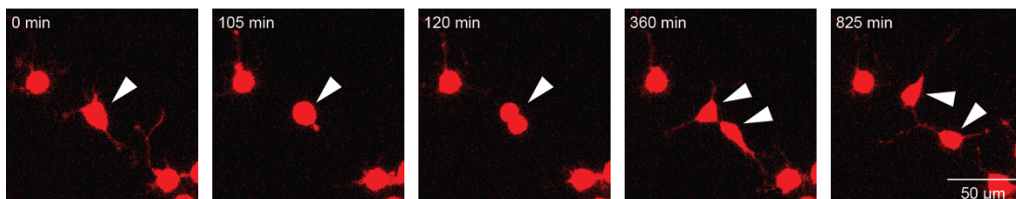


Figure R31: Representative image of tdTomato labeled neuroepithelial-like-2 ZFTA-RELA cell undergoing cell division in co-culture (arrowhead pointing to mother and daughter cells).

Changes made: Extended Data Figure 9g

#2-m3. Include additional, newer references for the role of neuronal activity in driving tumor progression

where applicable. For example, the authors should reference and discuss the recent findings on the role of neuronal signaling in EPN (Chen et al., Nature 2024).

We have included this citation in **Results**.

Results: Recent studies have revealed that neuronal activity can promote the proliferation and invasion of distinct molecularly defined tumor subpopulations through paracrine and synaptic signaling^{37,45–47}

Changes made: Results

#2-m4. Lines 74-75: Even if different subgroups harbor distinct malignant cell states, this does not necessarily imply that these states are driving different outcomes.

We have amended the phrasing.

Introduction: It remains unknown whether these subgroups have distinct cellular origins and composition of malignant cell states contributing to differences in outcomes and varying resistance to therapy.

Changes made: Introduction

#2-m5. Lines 78-80: This sentence suggests that cell states are sample-specific due to the cohort chosen, but this is more related to the methodology used to define cell states. Specifically, the study cited as reference 14 defined cell states within individual samples and then aggregated them to form metaprograms.

Reference 14 (Gojo et al., 2020) derived ST-EPN cell states using scRNA-seq data from three ZFTA-RELA fresh samples and one ST-YAP1 sample¹⁴. The method utilized by Gojo et al., 2020 and this manuscript defined cell states shared by patients by aggregating gene programs from individual samples. Gojo et al., 2020 also applied an independent method (batch correction and graph-based clustering) that **jointly analyzed** all samples to identify cell clusters. Both approaches were highly concordant in defining sample-specific cell states (Figure S4D in Gojo et al., 2020), indicating that the limited sample size - rather than the adopted methodology - generated sample/subgroup-specific cell states. Overall, these limitations highlight the need for a comprehensive study examining tumors across all subgroups of ST-EPNs.

#2-m6. Figure 1a: Basic information, such as age and sex (rather than gender), should be included

We have revised **Figure 1a** to include sex (rather than gender) (**Figure R32**). Age was previously included in the figure.

Figure R32

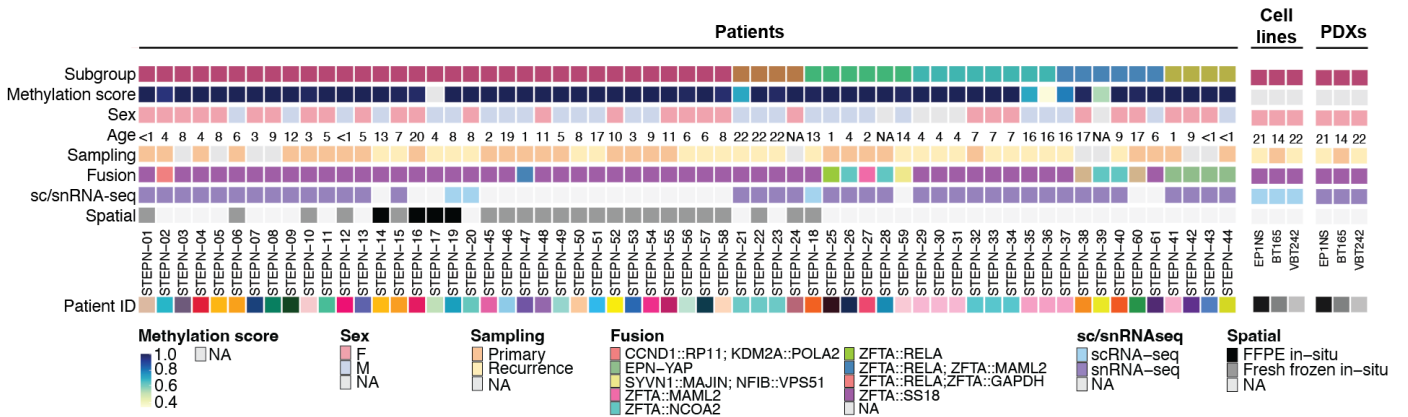


Figure R32. Oncoplot depicting ST-EPN samples profiled in this study.

Changes made: Figure 1a

#2-m7. Figure 1b-c: The authors mention 4 samples from the ZFTA cluster 1, but Figure 1b shows 5 such samples plotted. In Figure 1c, to what extent is the clustering driven by the sample of origin? The statement “we find that the molecular subgroup classification drives clustering of all malignant cells” is quite strong - typically, in most cancers, the sample of origin is what primarily drives clustering of malignant cells.

We thank the reviewer for these important observations. We have revised **Figure 1b** to include the correct number of samples that were profiled by DNA methylation sequencing (**Figure R33a**).

Regarding the reviewer’s second comment, our observation is that cells cluster **both** by samples and by subgroups. This can be seen by comparing the UMAP of malignant ST-EPN cells/nuclei profiled by sc/snRNAseq colored by subgroup (**Figure R33b**) and patient ID (**Figure R33c**), where cells both cluster by sample ID and cells of samples from the same subgroup generally cluster more closely than those of samples from other subgroups. To more clearly examine this point, we performed pseudobulk analysis of tumor cells by selecting 3000 highly variable genes (HVGs) across samples and performing principal component analysis (**Figure R33d**). We observed that samples cluster by subgroup (with some exceptions), suggesting that subgroup is a component in driving differential gene expression profiles. We have included the results for HVG=3000 in **Figure 1d** and amended the **Results** to more clearly explain this point.

Figure R33

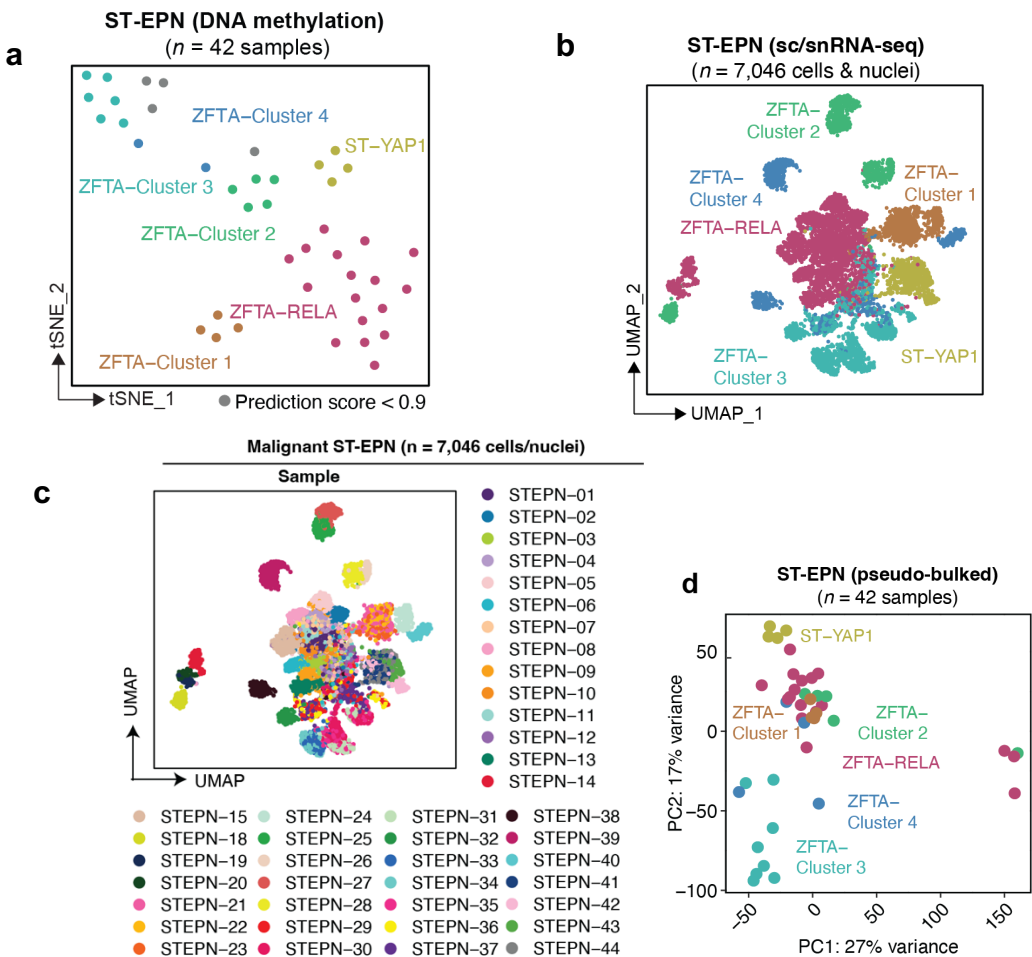


Figure R33. (a) tSNE plot ST-EPN samples profiled by DNA methylation profiles and colored by molecular subgroup. Samples with a methylation prediction score < 0.9 are indicated in grey. (b) UMAP plot of all ST-EPN patient cells/nuclei colored by sample of origin. (c) UMAP plot of all ST-EPN patient cells/nuclei colored by ST-EPN subgroup. (d) Principal

component analysis (PCA) plot of 3000 highly variable genes (HVGs) across samples, where each dot is a sample colored by molecular subgroup.

Results: As DNA methylation profiles utilized to categorize these tumors cluster subgroups differently (**Figure 1b**) and are considered epigenetic fingerprints of cellular origins in both development and tumorigenesis^{29,30}, we hypothesized that ST-EPN tumors would correspondingly be transcriptionally distinct. Accordingly, we found that the malignant cells cluster both by sample and molecular subgroup classification (**Figure 1c, Extended Data Figure 1e**), consistent with previous observations¹³. We further derived the top 3000 highly variable genes across samples and performed principal component analysis to demonstrate that samples cluster by their molecular subgroup classification (**Figure 1d**).

Changes made: Figure 1b, Figure 1d, Results

#2-m8. Figure 1f and line 133: The authors generated pseudo-bulk profiles using cells from different samples of the same subtype, which is not standard practice and is prone to batch effects. Why didn't they perform this analysis on a per-sample basis and demonstrate consistency across samples within the same subtype? They should score each sample against the different programs and provide statistical analysis.

We thank the reviewer for the suggestion and performed the projection analysis for each sample (**Figure R34**). Samples from the same subgroup generally demonstrate consistency in projection, with samples of ST-YAP1 projecting to choroid cells, samples of ZFTA-Clusters 1-4 to early neuroepithelial/mesenchymal cells, and ZFTA-RELA cells being broadly dispersed from neuroepithelial to neuronal and glial lineages. We have included these analyses to **Extended Data Figure 2a**.

Figure R34

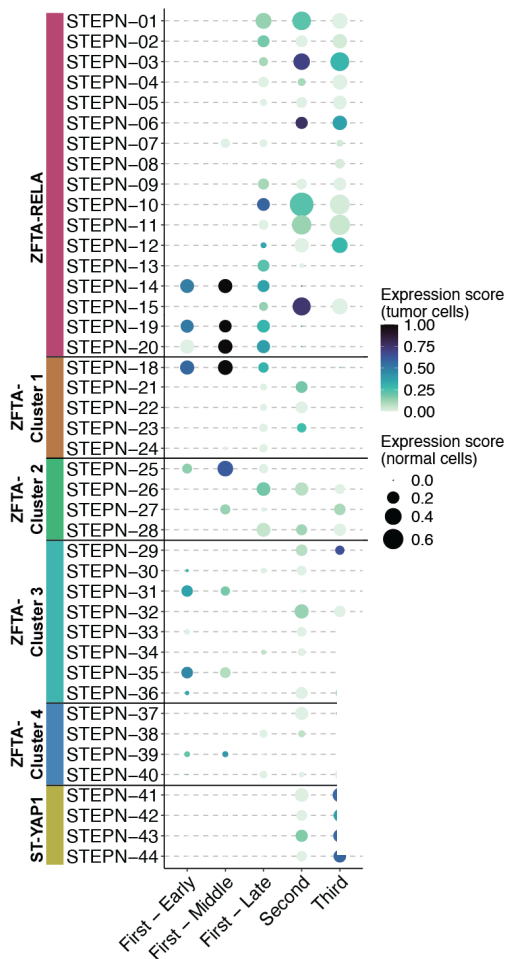


Figure R34. Projection of each of ST-EPN tumor samples (y-axis) onto developmental timepoints in the reference scRNA-seq dataset^{1,4} (First-Early: post-conception week (PCW) 4, First-Middle: PCW 5-6, First-Late: PCW 7-13).

Changes made: Extended Data Figure 2a

#2-m9. Additionally, in the caption (line 452), the term "tumor cell type gene expression" is not clearly defined—did the authors derive subtype-specific signatures? What genes are driving the observed similarity?

We thank the reviewer for pointing out this lack of clarity. We derived subgroup-specific signatures by first pseudo-bulking ST-EPN cells by subgroup and identifying differentially expressed genes using Seurat's *RunPrestoAll* function. Then, we scored the normal cells by these gene signatures. Thus, the "tumor cell type gene expression" refers to the score of these subgroup-specific gene signatures in the normal cells of the developmental reference dataset (Nowakowski et al., 2017 + Eze et al., 2024). We have amended the Figure 1 legends to "tumor subgroup-specific signature" and the labels in the dotplot to "Expression score (tumor profile in normal cells)". In addition, we have amended **Methods** to clarify this point.

Methods: Generation of developmental signatures. To identify developmental signatures enriched in ST-EPN subtypes or cell states, we performed projection of tumor cells (pseudo-bulked by ST-EPN subtype or cell state) onto normal cell type signatures. We first assembled a single-cell reference atlas of the developing human cerebral cortex from two published scRNA-seq datasets^{1,4}. We then subset the combined dataset to 500 cells (if available) per cell identity and used Seurat's *RunPrestoAll* function (keeping only genes expressed in at least 25% of all cells) to derive the top differentially expressed genes (adjusted *P*-value < 0.05) across cell types, which are reflective of developmental-specific cell type signatures. We then scored all ST-EPN malignant cells with these marker genes and annotated ST-EPN cells based on the maximum expression score for the respective signatures.

Changes made: Methods

#2-m10. Line 135: The authors only show expression scores, not correlations.

We thank the reviewer for this observation and have removed this statement from the manuscript.

#2-m11. Line 148: typo ('extent').

We have removed this statement from the manuscript.

#2-m12. Line 170 (Extended Data Figure 2c): The metaprograms are composed of lists of genes, so it's unclear what exactly the authors are correlating. Please clarify or consider using a different similarity metric.

We thank the reviewer for this constructive comment. We compare the metaprograms (MPs) identified in this study with MPs that were identified in previous work (Gojo et al., 2020)¹⁴, by performing Pearson's correlation and hierarchical clustering on the scores of metaprogram marker genes as follows:

- First, we compute single-cell/nucleus level expression scores for the respective MP signatures (each composed of a list of 30 genes). We do this by calculating for each cell *i* a score $SC_j(i)$, quantifying the average relative expression (*Er*) of the top 30 genes within each MP (*G_j*) and comparing the score to the average relative expression of a control gene (G_j^{cont}): $SC_j(i) = \text{average}[Er(G_j, i)] - \text{average}[Er(G_j^{cont}, i)]$. The control gene-set is defined by binning all genes into 30 bins of aggregate expression levels (*Ea*) and randomly selecting 100 genes from the same expression bin for each gene in the gene-set *G_j*.
- Second, after scoring each of our ST-EPN cells/nuclei for the MP signatures, we use pairwise correlation for each possible pair of MP, and show hierarchical clustering of the scores using a correlation-based distance metric (distance metric: 1 – Pearson correlation; linkage: Ward's linkage).

We have revised our methods section to provide a clearer explanation of this approach in **Methods**.

Methods: Program-wise hierarchical clustering. To compare sets of NMF programs or MPs, we first scored all ST-EPN cells/nuclei for the program signatures as described above (section: *Scoring gene sets*). After scoring, we performed pairwise correlation for each possible pair of programs and show hierarchical clustering of the scores using a correlation-based distance metric (distance metric: $1 - \text{Pearson correlation}$; linkage: Ward's linkage).

Changes made: Methods

#2-m13. Line 171: Figures 1f and 2a are addressing different aspects. Figure 2a derives patterns of variability within tumors, whereas Figure 1f claims that different subtypes globally resemble developmental processes. Therefore, the metaprograms do not corroborate the results shown in Figure 1f.

As also addressed in rebuttal for **Reviewer comment #2-3**, we agree with the reviewer that Figure 1f and 2a examine different aspects the tumors and utilize different analytical methods. We have removed this statement from the manuscript.

#2-m14. Figures 2c-d: These panels assume that each cell must be assigned to a metaprogram, which is generally not accurate - a cell's score should exceed a certain threshold for it to be assigned. This is evident in Figure 2e, where cells located in the middle of the cloud should not be assigned to any of the presented programs. Moreover, the varying proportions of metaprogram abundance in different tumor subtypes should be supported by more quantitative statistics. Finally, I do not understand (and could not find reference in the methods) whether the scoring was conducted per sample (i.e., with prior centering per sample) or after aggregating cells from all samples. The latter approach seems better suited for this analysis.

We thank the reviewer for these constructive comments and provide a point-by-point response here below.

These panels assume that each cell must be assigned to a metaprogram, which is generally not accurate - a cell's score should exceed a certain threshold for it to be assigned. This is evident in Figure 2e, where cells located in the middle of the cloud should not be assigned to any of the presented programs.

We agree with the reviewer that our approach relies on the underlying assumption that each cell belongs to a metaprogram, and that this concept has some intrinsic biological limitations, especially in the context of plasticity. This method of assigning certain cells to a metaprogram as previously described^{39,42,43,51}, is analogous to grouping cells into clusters and annotating cell clusters as cell types/states, which enables common differential cell composition analysis. As this grouping of cells by distinct groups is generally suboptimal in reflecting continuous processes like differentiation, we used a complementary method to present the metaprogram scores in a scatter plot in **Figure 2e** (hierarchy plot) to highlight the plasticity of tumor cells.

Moreover, the varying proportions of metaprogram abundance in different tumor subtypes should be supported by more quantitative statistics.

We thank the reviewer for the helpful suggestion. We have included supporting statistics to show the varying abundance of the metaprograms across different tumor subgroups (**Figure R35**). We used the propeller function in R package *speckle* (v1.8.0)⁵, a specialized method for performing statistical tests on compositional data, to determine statistical significance in cell proportions across subgroups. Briefly, this function fits a linear model for each cell type, with the groups as the explanatory variable to obtain false discovery rates (FDR) by performing Bayes shrinkage of the variances and correcting for multiple testing. Using this method, all metaprograms are significantly different across the subtypes, as indicated by the FDR values (**Figure R35**). We have included this analysis in **Extended Data Figure 3a**.

Figure R35

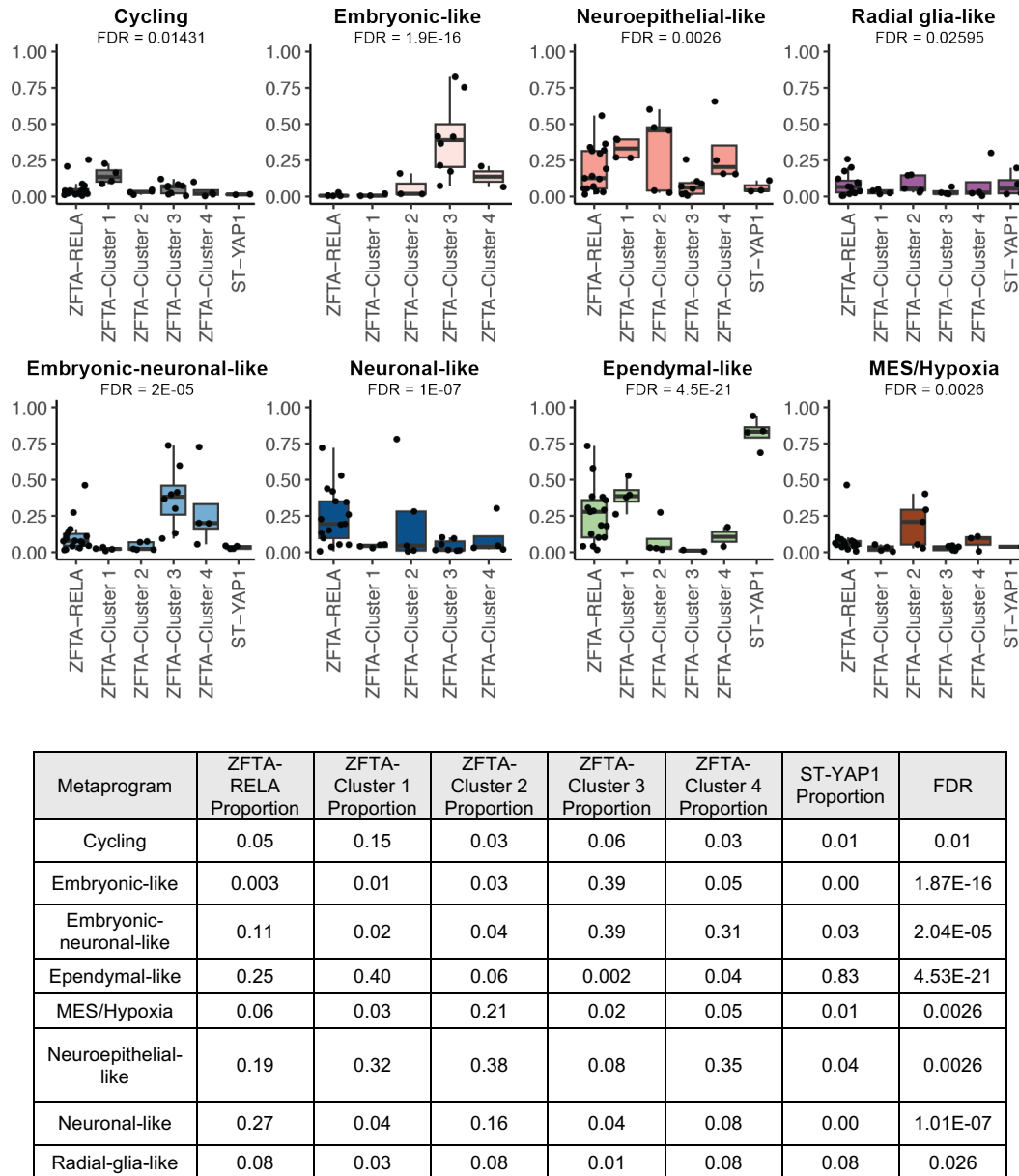


Figure R35. (top) Boxplots depicting proportion of each metaprogram (y-axis) across ST-EPN subgroups (x-axis). False discovery rate (FDR) labeled for each metaprogram. **(bottom)** Statistical analysis results of each metaprogram across subgroups.

Finally, I do not understand (and could not find reference in the methods) whether the scoring was conducted per sample (i.e., with prior centering per sample) or after aggregating cells from all samples. The latter approach seems better suited for this analysis.

The scoring was done after aggregating cells from all samples. We have included this in revised **Methods**.

Methods: Scoring gene sets. We first aggregated ST-EPN single cells/nuclei from all samples, and performed normalization and centering. Then, we scored these cells for NMF/developmental signatures as previously described^{14,39,42–44}. We did this by calculating for each cell i a score $SC_j(i)$, quantifying the average relative expression (Er) of the top 30 genes within each MP (G_j) and comparing the score to the average relative expression of a control gene (G_j^{cont}): $SC_j(i) = \text{average}[Er(G_j, i)] - \text{average}[Er(G_j^{cont}, i)]$. The control gene-set was defined by binning all genes into 30 bins of aggregate expression levels (Ea) and randomly selecting 100

genes from the same expression bin for each gene in the gene-set G_j . ST-EPN cells/nuclei were annotated based on the maximum expression score for the respective MP signatures after excluding the one related to cell cycle (“cycling”).

Changes made: Extended Data Figure 3a, Methods

#2-m15. The methods section (lines 659-670): The description relating to scoring and the cell state plot lacks appropriate references.

We included the appropriate references (Neftel et al., 2019) in **Methods**.

Methods: Cell state plot. We separated ST-EPN cells into neuroepithelial-like/embryonic-like versus neuronal-like/ependymal-like as previously described⁴³. In brief, we first calculated the y-axis value as $y = \max(SC_{neuroepithelial-like}, SC_{embryonic-like}) - \max(SC_{neuronal-like}, SC_{ependymal-like})$. For neuroepithelial-like/embryonic-like cells ($y > 0$), the x-axis value was defined as $x = \log_2(|SC_{ependymal-like} - SC_{neuronal-like}| + 1)$ and for neuronal-like/ependymal-like cells ($y < 0$), the x axis was defined as $\log_2(|SC_{embryonic-like} - SC_{neuroepithelial-like}|)$.

Changes made: Methods

#2-m16. In their important 2020 paper (Gojo et al.), the authors also sequenced ST-RELA samples and demonstrated that these lack the ependymal signature (main Figure 4 in that paper). However, here the ependymal-like program is strongly seen in these tumors (line 177). Additionally, in the 2020 study, the authors analyzed one ST-YAP1 sample, which they indicated is dominated by a program they term ST-YAP1 and lacks an ependymal signature (Figure 4 in that paper). This finding appears to conflict with their current results (line 198) - can the authors please clarify?

In Gojo et al. 2020, we profiled $n = 3$ ST-RELA and $n = 1$ ST-YAP1 fresh samples for scRNA-seq. This relatively low number of samples profiled in Gojo et al., 2020 did not allow us to derive robust metaprograms, and led us to the identification of subtype-specific metaprograms (e.g. ST-YAP1 metaprogram, or ST-RELA-Variable) without a clear link to a developmental cell type, as also addressed in **Reviewer comment #2-5**. This limitation sets the basis for our study, as we point out in our **Introduction**, and the significant expansion to $n = 42$ samples. Based on signatures derived from our bigger cohort of ZFTA-RELA samples ($n=17$ samples), we are now able to reason that ST-RELA (ZFTA-RELA) samples profiled in the Gojo et al., 2020 manuscript happened to express low levels of the ependymal signature. This is consistent with the variations observed in our current cohort in the current manuscript, in which several patients (STEPN-1, STEPN-13, STEPN-14) also exhibit low abundance of ependymal-like cells.

Additionally, based on the limitations of the study, in Gojo et al., 2020 study, the intratumoral heterogeneity was not well captured for the one ST-YAP1 sample, as the majority of the tumor was characterized as “ST-YAP1 metaprogram” (**Figure R36a**). We see that ST-YAP1 metaprogram in Gojo et al., 2020 correlates to both the STEPN-Ependymal-like and STEPN-Radial glia-like programs of our current manuscript. This also concurs with the cell states present in the new ST-YAP1 cohort, which is pre-dominantly ependymal-like and radial glia-like. Based on these findings, we reason that the ST-YAP1 metaprogram in Gojo et al, 2020 is a mixture of different cell states that is present in ST-YAP1 tumors, which we were able to characterize in a more granular manner in our new cohort.

Overall, we find that the findings from Gojo et al., 2020 and our current study are consistent and not conflicting. We thank the reviewer for this constructive discussion.

Figure R36

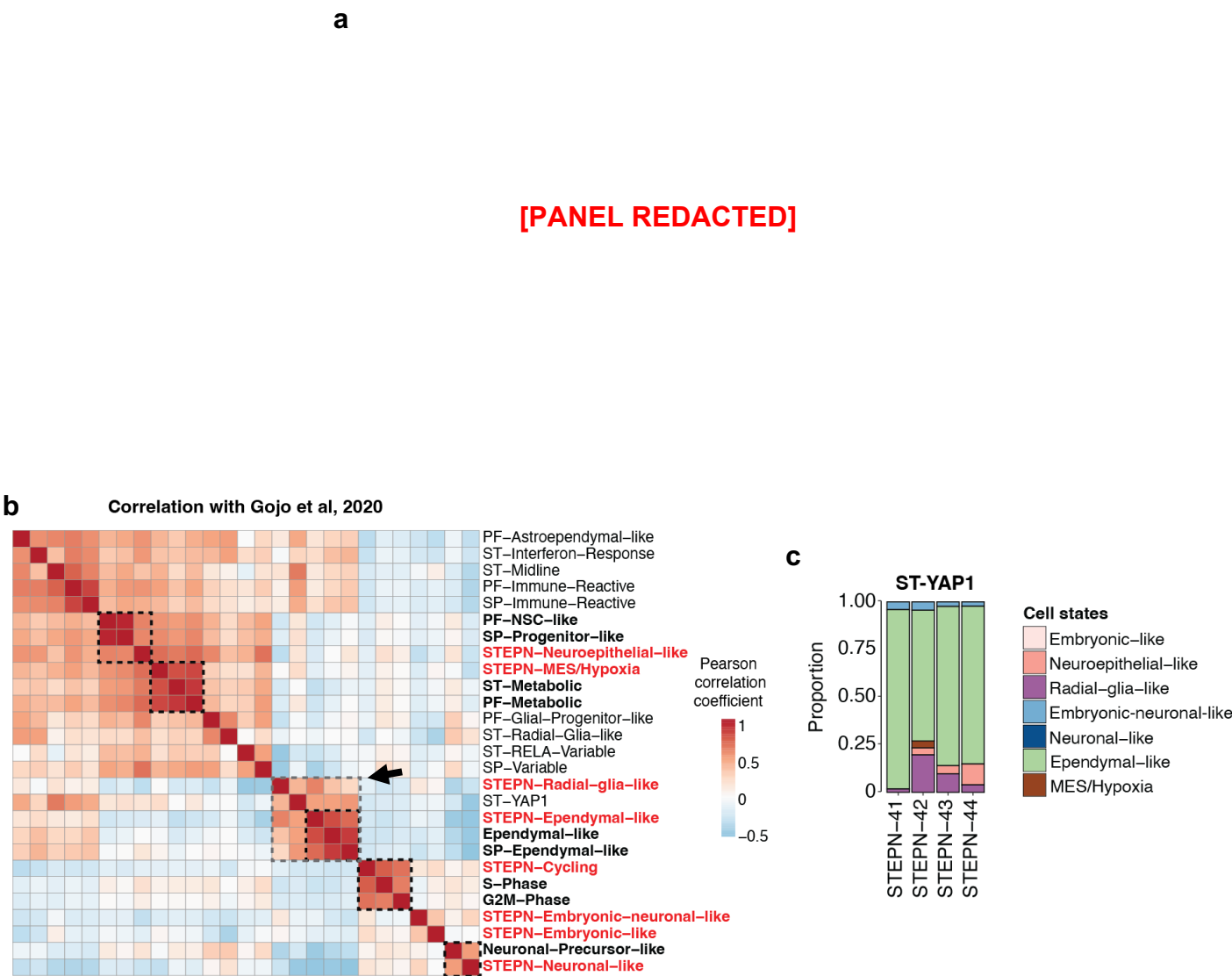


Figure R36. (a) [TEXT REDACTED] (b) Pairwise Pearson correlation between ST-EPN metaprograms identified in this study using top 30 marker genes of metaprograms in frozen tumor samples (red font) and previously defined metaprograms¹⁴ identified across different types of ependymoma (SP: spinal; ST: supratentorial; PF: posterior fossa). **(c)** Proportion of metaprograms (y-axis) in each sample (x-axis) in our current study.

#2-m17. Line 329: There is a missing word “and” in the phrase: “are highly invasive AND neuroepithelial-like cell states are the most proliferative.” Additionally, in Figure 2a, the neuroepithelial-like metaprogram and the cycling metaprogram are anti-correlated, while the cycling metaprogram is slightly positively correlated with the neuronal cell metaprogram, which the authors state in line 326 is the least proliferative. Can the authors please clarify this discrepancy?

We thank the reviewer for this comment. As shown in **Figure R20**, a certain proportion of each malignant cell state is proliferative. Hence, the overall cycling population is a mixture of different cell states. The heatmap of Figure 2a shows the correlation between the scores of the top 30 genes of each NMF program. As this is the case, we do not expect a specific correlation between neuroepithelial-like and cycling program in the heatmap. We have included these analyses in **Extended Data Figure 2e-f** to more clearly explain the proliferative rate of each cell state in patient tumors.

Figure R20

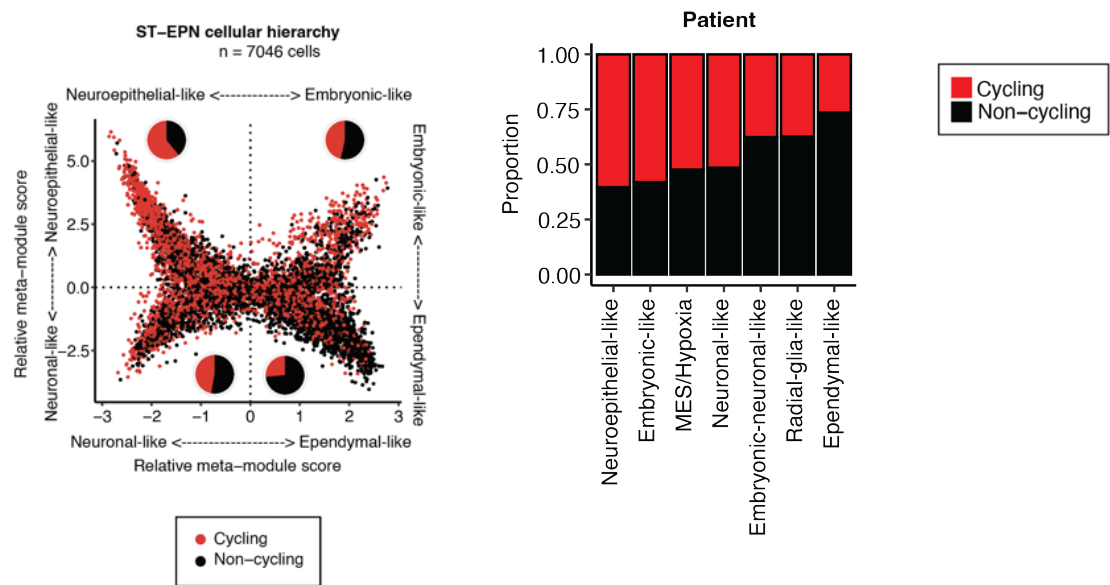


Figure R20. (left) Cell state plot colored by inferred cycling status. The frequency of cycling and non-cycling cell within each quadrant is displayed in the pie charts. **(right)** Barplot showing the proportion of cycling and non-cycling cells across ST-EPN cell states, ordered from the most (left) to the least (right) proliferative.

Results: Taken together, these results suggest that ZFTA-RELA transcriptomic cellular states have distinct morphological and behavioral attributes in the context of their native neuroglial microenvironment, in which neuronal-like cell states are both highly migratory and adopt distinct migratory patterns reminiscent of immature neurons during development.

Changes made: Extended Data Figure 2e-f, Results

Referee #3 (Remarks to the Author):

A. Summary of the key results

I want to thank the authors for presenting an interesting manuscript with a compelling story, I appreciate the effort that went into addressing this important topic. The manuscript is overall well written, and data clearly presented.

The authors make use of a multimodal framework of cutting-edge technologies to address questions regarding the tumor biology and origin of supratentorial ependymomas (ST-EPN). They employ single cell/single nuclei and spatial transcriptomics, and live cell imaging addressing several important questions. I've summarized the main findings below:

- ST-EPN subgroups align with temporally restricted cells across the human cortical development and have distinct developmental signatures.
- ST-EPN modes of organization where high order architecture ranging from disorganized to structured correlate with mesenchymal and hypoxia cell state signatures.
- The neural tumor microenvironment promotes plasticity towards a neural developmental trajectory, which affects the morphology and behavior of the ST-EPN cells.

I tentatively recommend accepting the manuscript with major revisions.

B. Originality and significance: if not novel, please include reference

The main novel findings are:

- ST-EPN subgroup characterization using normal developmental cortex to align the tumors data for temporally distinct subgroup signatures. Previous suggested cell type of origin has been radial glia cells, here the authors map ST-EPN tumors to two even earlier progenitor populations (caveat, needs revision)
- ST-EPN spatial investigation of the canonical ZFTA-RELA subgroup, finding higher order organization correlating with mesenchymal and hypoxia cell state signatures (also found in study authors reference 24, but here the authors examine a different type of tumor). (caveat, revision, that more cell states need to be analyzed).
- Using cell morphology, marker genes and live cell imaging elucidating the effect of the neural tumor microenvironment in promoting invasive behavior and developmental trajectory of ST-EPN cells. (caveat, needs revision)

Caveats: depending on revision results assessments may change.

C. Data & methodology: validity of approach, quality of data, quality of presentation

Dear Authors, I have tried to structure my comments using citations from the manuscript, specific line numbers and figure names. I have tried to explain as clearly as possible my reasoning and feedback. It was an interesting read, and I see the future potential of a beautiful work. However, there are many things I deem is needed to be revised. If anything is unclear, please contact the editor and I will do my best to explain.

[We thank the reviewer for their comprehensive and insightful evaluation of our manuscript. We have attempted to thoroughly address each point raised.](#)

Abstract

#3-1. The authors state “we integrated single cell- and spatial transcriptomics, as well as in vitro and in vivo

live-cell imaging”. However, reading the manuscript it does not include *in vivo* live-cell imaging, only *in vitro* live cell imaging, please rephrase.

We thank the reviewer for this point. We performed *in vivo* live-cell imaging to describe the morphologies of cell states, cell migratory patterns and tumor microtube dynamics *in vivo*, shown in **Figure 5a and h, Extended Data Figure 9d and f**. We have specified more clearly in the manuscript where *in vivo* live-cell imaging has been done in **Figure 5 legends and Extended Data Figure 9 legends**.

Figure 5a,h legends: **Figure 5. (a)** Morphological characterization of ZFTA-RELA cells (EPINS) in *in vivo* PDX (top row) and *in vitro* coculture model (bottom row). ZFTA-RELA cells express endogenous tdTomato. *In vivo* PDX images were acquired using in-vivo two-photon microscopy; blood vessels are shown in green. **(h)** *in vivo* time-lapse imaging of tumor microtube dynamics in a PDX model of ZFTA-RELA cells expressing tdTomato. Arrowheads point to the ends of tumor microtubes. Blood vessels are shown in green.

Extended Data Figure 9d,f legends: **Extended Data Figure 9. (d)** *in vivo* time-lapse imaging of PDX model of ZFTA-RELA (EPINS cells expressing tdTomato) acquired with two-photon microscopy. Tumor cells are indicated in red, and blood vessels in green. An invasive cell and its invasion route (dashed arrow, above) versus a stable cell (arrowhead, below) are shown. **(f)** Net TM turnover of ZFTA-RELA cells, including protrusion and retraction, over 4 hours for each morphological subtype in *in vitro* coculture (left) and *in vivo* (right).

Changes made: Figure 5 legends and Extended Data Figure 9 legends

#3-2. The authors state “We find that individual tumor subgroups harbor two distinct progenitor-like cell states reminiscent of early human brain development and diverge in the extent of neuronal or ependymal differentiation.”. However, in Figure 2e, four different cell states are displayed, and subgroups are mapped across these. Please clarify. E.g., was this specifically referring to the neuroepithelial-like and embryonic-like states?

We thank the reviewer for this question. We clarified this sentence that we are referring to the two progenitor-like populations neuroepithelial-like and embryonic-like states.

Abstract: We find that individual tumor subgroups harbor two distinct progenitor-like cell states-neuroepithelial-like and embryonic-like- which are reminiscent of early human brain development and diverge in the extent of neuronal or ependymal differentiation.

Changes made: Abstract

#3-3. The authors state “We further uncover several modes of spatial organization of these tumors”. However, only MES/Hypoxia is statistically significant and supported by the data analysis, and many of the state/cell types were not tested for, at least not in the provided material. See revision points.

We thank the reviewer for this important suggestion. We found no other cell states/types that significantly correlate with the global spatial coherence score of tumors (Figure R5a).

To determine further whether these findings were observed across more samples, we increased our spatial cohort by adding 33 new ZFTA-RELA 10X Xenium samples of 13 new patients (from $n=23$ to $n=56$ sections; $n=10$ to $n=23$ patients). We also amended the cell annotation method to include label projection of previously defined tumor cell states and/or non-malignant cell types from patient tumor sn/scRNA-seq dataset for more unbiased and time-efficient analysis across all samples (**Methods**).

With this increased spatial cohort and annotation methods, we confirm our previous findings that MES/Hypoxia is still significantly correlated with the global spatial coherence score across all samples ($R=0.68$, adjusted p -value = 0.0047 by Bonferroni correction) (**Figure R5a**), while no other metaprogram or

non-malignant cell type correlates significantly with the global spatial coherence score (Figure R5a). We also corrected the linear regression analysis by:

- Corrected P-values for multiple comparisons using Bonferroni correction
- Calculated mean value of samples that have multiple sections

We have included these new analyses to **Extended Data Figure 4e**, and have amended the manuscript to report the new number of samples in **Results**.

Figure R5

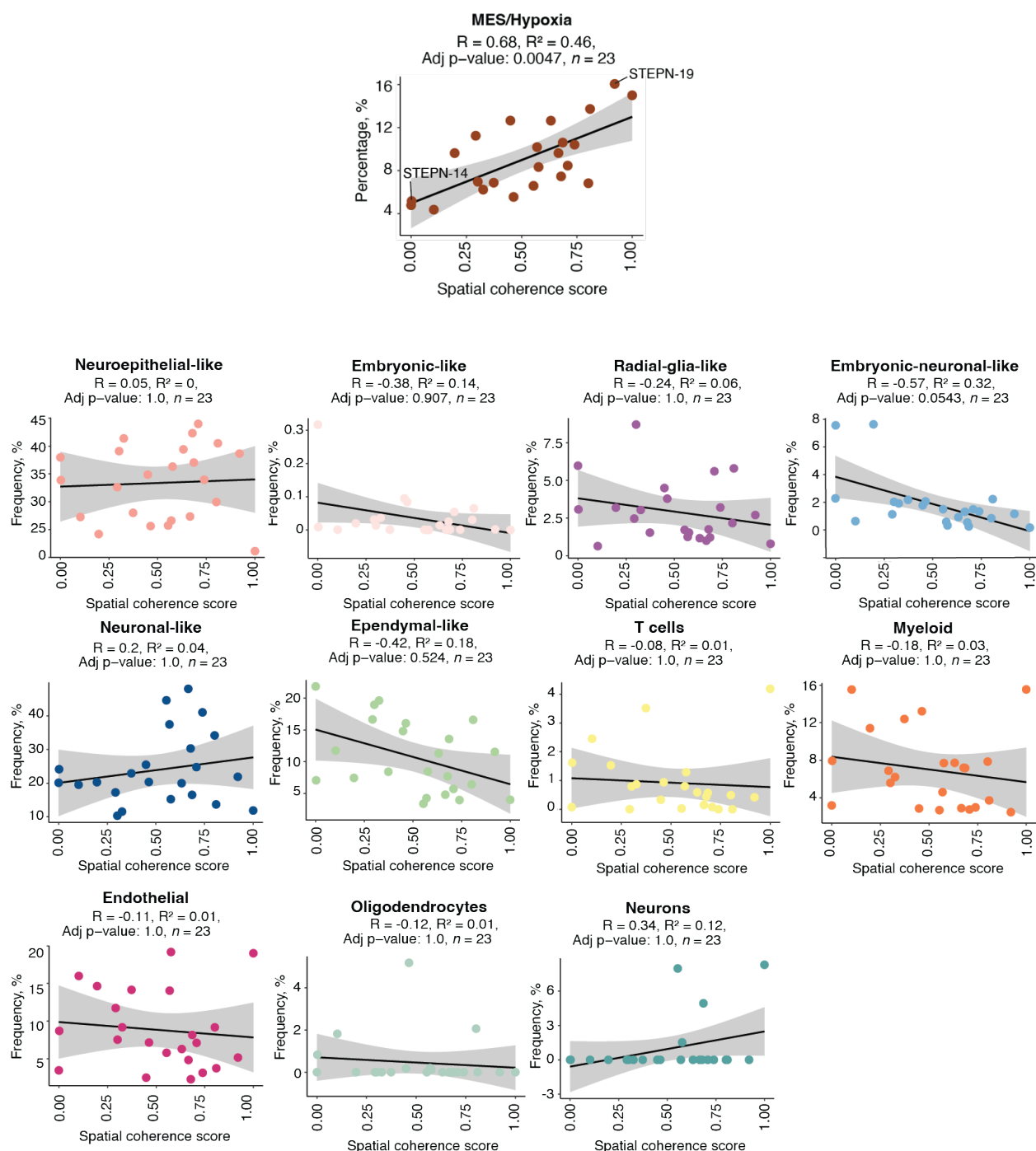


Figure R5. (a) Correlation between the spatial coherence score and the percentage of individual cell states/types in ZFTA-RELA tumor samples profiled by 10X Xenium. Data points (representing each tumor sample) are interpolated with a simple linear regression. For tumor samples with multiple sections, the average values are shown. The goodness of fit (R , R^2), Bonferroni-corrected P-values and total number of tumor sections (n) are displayed at the top.

Results: we performed 10X Genomics Xenium *in situ* transcriptomics on a total of 56 tumor sections from 23 ZFTA-RELA patients to allow for subcellular resolution mapping of 358 target genes (**Table S6**).

Methods: Xenium cell state/type cell assignment. To assign cell identities in Xenium data, we first performed unsupervised clustering using *Seurat* (v5.1.0), performing *RunUMAP* on the least number of PCs capturing 80% or more variance, shared-nearest-neighbor graph using *FindNeighbors*, and Louvain clustering (*FindClusters*) run at resolutions 0.1 to 1.0, in 0.1 increments. We then scored only for the normal cell types, by transferring labels from the single cell/nuc reference object from corresponding tumor subgroup using *FindTransferAnchors* and *TransferData* (dims = 20), yielding a label-transfer score for every cell-to-cell type pair. Independent module scores were calculated using *AddModuleScore_UCell* using marker genes present in the Xenium panel. For each cluster, average label-transfer score and module score were manually inspected and visualized through UMAP, and the clusters corresponding to normal cell types were identified in each sample. Cells belonging to the normal cell type annotation was removed, and malignant subset of each sample was annotated in the same way as normal cell types. To mitigate systemic score inflation, scores for each malignant cell type were mean centered across all cells: $centered_score_{ij} = score_{ij} - \overline{score}_{.j}$, where i corresponds to cells and j corresponds to cell types. Finally, each malignant cell was assigned with the cell type with highest mean centered score, and malignant and normal cells were combined to generate the final, fully annotated cells.

Changes made: Extended Data Figure 4e, Results, Methods

Introduction

Clearly written, describes the current knowledge gap, brings up the main findings, and frames the authors scientific work in a proper context.

We thank the reviewer for this kind comment.

#3-4. Lines 81-83: The authors use a reference to a review paper from 2012 stating that “it has recently become clear”. In my opinion ‘recently’ would be within the last ~ 5 years, I suggest the use of another term or removal of ‘recently’ in this context.

We removed “recently” from this sentence.

Introduction: In addition to the cell-intrinsic properties, it has become clear how cancer cell diversity is further influenced by their microenvironment, where interactions with adjacent cells and non-genetic factors such as deprivation of oxygen or nutrients augment cellular variability and malignant phenotypes.

Changes made: Introduction

#3-5. Lines 95-96: The authors claim to “find that malignant cells reflect two distinct temporally restricted progenitors [which are? Please name them] from early human embryonic cortical development”, please clarify here already by adding in parenthesis which these two progenitors are.

We have clarified the two progenitors in parenthesis.

Introduction: We first delineate the developmental correlates of different ST-EPN subgroups and find that malignant cells reflect two distinct temporally restricted progenitors-neuroepithelial-like and embryonic-like cells- from early human embryonic cortical development and exhibit varying degrees of neuronal or ependymal differentiation.

Changes made: Introduction

Results

The different sections of the results are overall clearly structured and easy to follow. Presenting the data and the authors investigation in a chronological order that makes sense and create an interesting story. However, there are multiple points that needs to be revised.

We thank the reviewer for this generous comment and constructive feedback.

#3-6. Line 110: The authors state “we compiled a patient cohort of 44 patient tumors”, but it is not immediately clear that several patients have been sampled multiple times. Furthermore, one patient has 2 samples in subgroup ZFTA Clusters 3 and 1 sample in subgroup ZFTA Clusters 4, please add an explanation somewhere in the paper of why the patient’s samples are spread across different subgroups. Although the authors illustrate this in Figure 1a, I suggest they further add in their main text when presenting the dataset that the 44 sample comes from 37 patients. If possible, it would also be of interest to know if the multiple samples from the same patients have been collected at different time-points or from different spatial areas of the tumors. Because sampling from different areas of a tumor could capture global heterogeneity and contain different genomic and transcriptomics profiles.

We thank the reviewer for this helpful suggestion. We have now clarified that our study contains **42 samples** from **34** patients included in the patient cohort for single-cell transcriptomics in **Results**.

In addition, we thank the reviewer for the interesting question regarding methylation classification across multiple samples in the same patient. In our patient cohort, Patient-29 has two samples in ZFTA-Cluster 3 and 1 sample in ZFTA-Cluster 4. While the first two samples were sampled from the supratentorial region of the brain, ZFTA-Cluster 4 sample (STEPN-37) was biopsied from metastatic lymph nodes. This suggests that microenvironmental factors may impact methylation classification. We have now also included tumor location in our updated supplementary table (**Table S1**).

While it is currently not well characterized how frequently and why ST-EPN tumors can shift in methylation classifications upon recurrence, studies in other tumors indicate a possible combination of both internal and external factors contributing to these shifts^{52,53}. Gliomas were observed to change methylation subclasses after recurrence, accompanied by increased immune cell infiltration within the tissue as well as a shift towards mesenchymal subclass⁴¹. In PF-EPNs, recurrent tumors that shifted methylation classification acquired new mutations, specifically in novel chromosome 1q gains and/or 6q losses⁴². As most patients received treatment such as radiotherapy and/or chemotherapy in between biopsies in our study, these results overall suggest that phenotypic evolution, microenvironmental changes and/or therapy pressures overall contribute to a change in methylation classification. Although we currently have limited paired primary and recurrent cases to thoroughly investigate this question in ST-EPN tumors, compiling a comprehensive cohort with patient-matched primary and recurrent samples would be highly valuable for future studies aimed at elucidating the mechanisms underlying shifts in methylation classification in ST-EPN tumors upon recurrence.

In addition, we unfortunately do not have the information regarding where the samples were taken from the specific areas in the tumor, which we have addressed in **Limitations of the Study**. We agree with the reviewer that future studies incorporating sampling from well-described different areas of the tumor of ST-EPN tumors could provide further insights into the heterogeneity across the entirety of the tumor mass.

Results: To comprehensively characterize the cellular tumor heterogeneity of supratentorial ependymomas (ST-EPN), we compiled a patient cohort of 42 patient tumor samples from 34 patients for single-cell transcriptomics, encompassing ZFTA-RELA ($n = 17$), ZFTA-Cluster 1 ($n = 4$), ZFTA-Cluster 2 ($n = 5$), ZFTA-Cluster 3 ($n = 8$), ZFTA-Cluster 4 ($n = 4$), and ST-YAP1 ($n = 4$) subgroups.

Limitations of the Study: **Limitations of the study.** In this study, we characterized tumor heterogeneity of patient samples diagnosed as supratentorial ependymomas. However, there may be a sample site bias in our data set due to random sampling of the whole tumor mass during surgery as well as dividing the taken tumor into pieces for sc/snRNAseq and spatial transcriptomics. Therefore, our findings are limited due to the sampling of tumors, and do not have a conclusive claim of the entire tumor mass

present in patients. In addition, due to limited availability of non-canonical ZFTA tumors and corresponding patient-derived cell lines, we have not included spatial transcriptomics and other functional live-cell imaging data of these tumor subgroups. We also did not include ST-EPN-SE subgroup tumor samples in our analyses. Finally, the mesenchymal and hypoxia cell state signatures are not a direct measurement of hypoxia itself, and rather act as proxy for potential effects or consequences of local and global hypoxic microenvironments.

Changes made: Results and Limitations of the Study

#3-7. Lines 110-111: The author's state: "we compiled a patient cohort of 44 patient tumors encompassing ZFTA-RELA (n = 20), ZFTA Clusters 1-4 (n = 20) and ST-YAP1 (n = 4) subgroups." However, in figure 1a the number for each subgroup is clearly stated. The authors also make a point of their study including more diverse subgroups than previous studies. I suggest that instead of grouping only ZFTA Clusters 1-4 here, that they be presented individually also in the text. E.g., ZFTA Clusters 1 (n = 4), ZFTA Clusters 2 (n = 4), ZFTA Clusters 3 (n = 8), and ZFTA Clusters 4 (n = 4).

We rephrased the sentence to this suggested phrasing.

Results: To comprehensively characterize the cellular tumor heterogeneity of supratentorial ependymomas (ST-EPN), we compiled a patient cohort of 42 patient tumor samples from 34 patients for single-cell transcriptomics, encompassing ZFTA-RELA (n = 17), ZFTA-Cluster 1 (n = 4), ZFTA-Cluster 2 (n = 5), ZFTA-Cluster 3 (n = 8), ZFTA-Cluster 4 (n = 4), and ST-YAP1 (n = 4) subgroups.

Changes made: Results

#3-8. Line 121: The authors state they have retain 6933 high-quality malignant cells, however in Extended Data Figure 1, subpanel 'Final cell type classification' the number is instead 6680 malignant cells. Please clarify where the 253 additional cells come from.

We thank the reviewer for highlighting this lack of clarity in our study. The inconsistencies were due to the different types of data described (scRNA-seq, snRNA-seq, or both datasets) and misreporting of the number of cells in the manuscript. The subpanel in our previous version of Extended Data Figure 1a included only nuclei isolated from frozen samples (profiled by snRNA-seq), while the text (line 122) referred to malignant cells/nuclei from frozen (snRNA-seq) and fresh (scRNA-seq) samples. However, the text also misreported this number of malignant cells/nuclei, where the correct value is 7,046 malignant cells/nuclei. We sincerely thank the reviewer for catching this typo, and we have now revised all figures and text to describe the dataset as all cells/nuclei (scRNA-seq and snRNA-seq) combined, for a total of 7,046 malignant cells/nuclei.

Figure R14

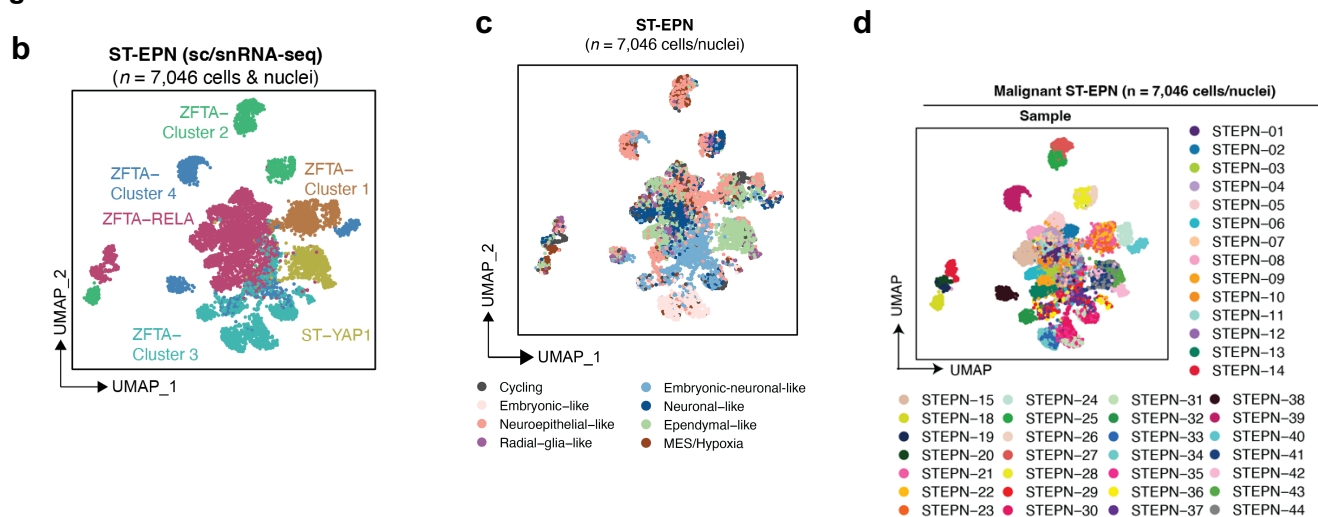


Figure R14. (b) UMAP plot of malignant ST-EPN cells profiled by sc/snRNA-seq and colored by molecular subgroup. **(c)** UMAP plot of malignant ST-EPN cells colored by identified cellular state. **(d)** UMAP plot of all ST-EPN patient cells/nuclei colored by sample of origin.

Results: We retained 7,046 high-quality malignant cells (mean of 5,511 genes per cell/nucleus; **Extended Data Figure 1e, Table S1**), which we processed uniformly and co-embedded into a shared low-dimensional space.

Changes made: *Figure 1c, Figure 2c, Extended Data Figure 1a, Extended Data Figure 1e, Results.*

#3-9. Line 127: The authors state that “subgroup classification drives clustering of all malignant cells (Figure 1b-c)”, however in the c panel there are distinct differences between DNA methylation and sc/sn RNA-seq data. I suggest the authors test coloring the sc/sn RNA-seq data UMAP also based on patient ID. Many of the subgroup clusters looks to be separate by the patients, not their subgroup classification. I suggest a different analysis for this. One can treat each patient sample as a pseudo-bulk sample, then select e.g., top 3000 variable genes across all pseudo-bulk samples, then do a Pearson correlation between all samples and display them in a heat map with subgroups as color legends. Look to see if subgroups are clustering together using such an approach instead.

We thank the reviewer for this helpful suggestion. We have included a UMAP plot colored by patient ID (**Figure R14c, Extended Data Figure 1e**). In addition, we performed the recommended pseudo-bulk RNA-seq analysis by selecting between 500 and 3,000 highly variable genes (HVGs) and performing Pearson correlation, as well as principal component analysis (**Figure R37**). We observed that samples generally cluster by subgroup, suggesting that subgroup is a component in driving differential gene expression profiles. We have included the results for HVG=3000 in **Figure 1d (Figure R33)** and amended the **Results** to more clearly explain this point.

Figure R37

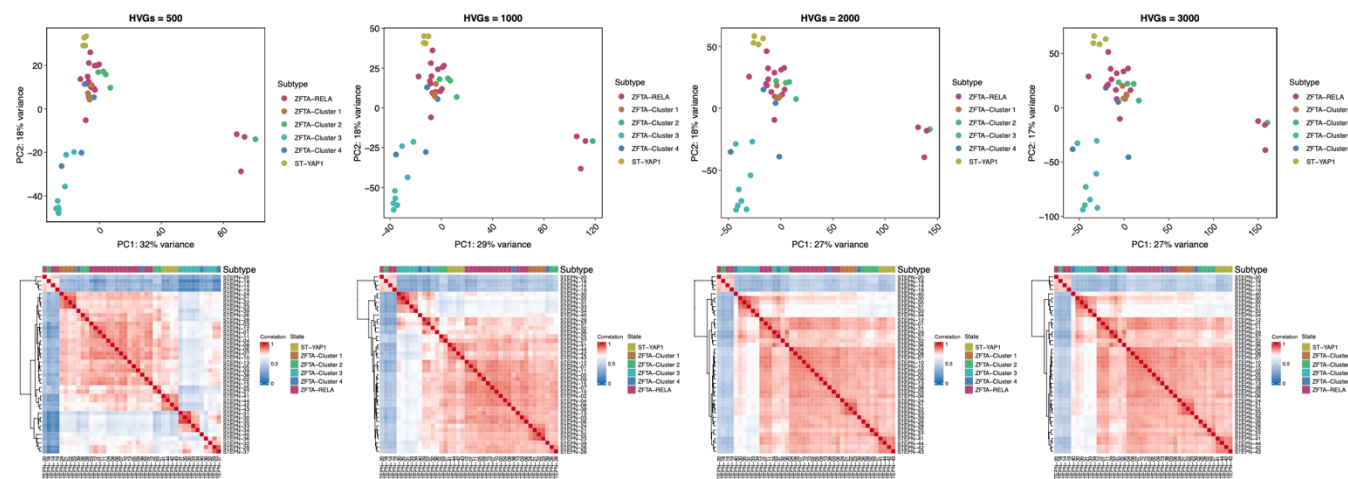


Figure R37. Principal component analysis (PCA) (top) and Pearson Correlation (bottom) of 500, 1000, 2000, or 3000 highly variable genes (HVGs) across samples. In PCA plots, each dot is a sample colored by molecular subgroup. In Pearson Correlation plots, each row and column is a sample.

Figure R33

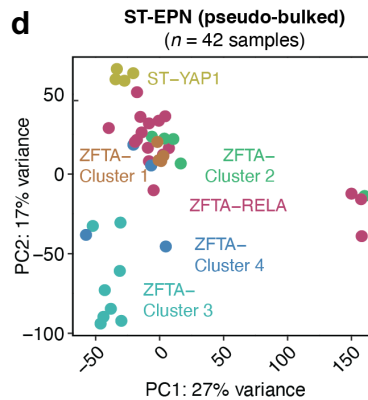


Figure R33. (d) Principal component analysis (PCA) plot of 3000 highly variable genes (HVGs) across samples, where each dot is a sample colored by molecular subgroup.

Results: As DNA methylation profiles utilized to categorize these tumors cluster subgroups differently (**Figure 1b**) and are considered epigenetic fingerprints of cellular origins in both development and tumorigenesis^{29,30}, we hypothesized that ST-EPN tumors would correspondingly be transcriptionally distinct. Accordingly, we found that the malignant cells cluster both by sample and molecular subgroup classification (**Figure 1c, Extended Data Figure 1e**), consistent with previous observations¹³. We further derived the top 3000 highly variable genes across samples and performed principal component analysis to demonstrate that samples cluster by their molecular subgroup classification (**Figure 1d**).

Changes made: Figure 1d, Results

#3-10. Line 139: The authors state that “ZFTA-Cluster 3 and 4 tumors mapped to mesenchymal cells found in the early first trimester”, which the figure supports, but for ZFTA-Cluster 3, in particular, it appears that cells map to OPC to an equal extent when considering cell types. Can the authors please explain why mesenchymal and not OPC is mentioned in their results interpretation?

We thank the reviewer for this question. As the reviewer pointed out, ZFTA-Cluster 3 also projects to OPC cells, but as we did not observe OPC-like malignant cells in our metaprogram analysis (Figure 2), we did not highlight these findings. Furthermore, we found generally some inconsistencies between the findings of Figure 1f and Figure 2D due to fundamental differences in analysis methods, as outlined below:

1. Figure 1F depicts correlation based on aggregate/“pseudobulk”. First, single cell counts of each tumor subgroup were aggregated (“pseudobulk”), which were then scored by the top 30 marker genes of cell types in the reference dataset (Nowakowski et al., 2017 + Eze et al., 2024). This is labeled “expression score (tumor cells)”. Similarly, single-cell counts of each cell type in the reference dataset were pseudobulked and scored by the top 30 differentially expressed genes of tumor subgroups in the tumor sc/snRNAseq data. This is labeled “expression score (normal cells)”. The dotplot in Figure 1F therefore displays correlations of samples aggregated and projected to the developmental dataset, representing **an aggregate** confounded by the heterogeneity of cell states of each patient per subgroup.
2. For Figure 2D, we performed NMF analysis of tumor sc/snRNAseq data and identified recurrent gene expression programs across samples (“metaprograms”). While each development-like metaprogram was annotated by the comparison of gene signatures to the developmental reference scRNA-seq dataset as well as gene ontology (GO) analysis, these are *de-novo* identified malignant programs in the tumor. As such, the plot in Fig. 2D more accurately reflects the true malignant cell composition across tumor samples.

Due to the confusing and sometimes conflicting findings of these two figures, we have **removed** the pseudobulk projection of tumor subgroups to cell types in the reference dataset and only retained the projection to time (**Figure R1**). We have also revised the manuscript results section to explain these discrepancies.

Figure R1

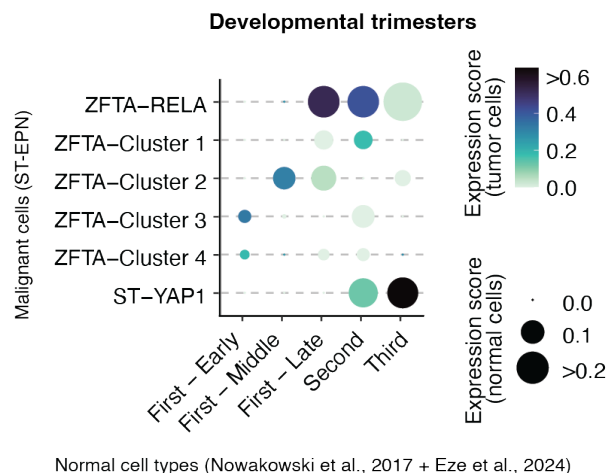


Figure R1. Projections of ST-EPN subgroups (y-axis) onto developmental time points found across the developing human cortex (x-axis). First-Early: post-conception week (PCW) 4, First-Middle: PCW 5-6, First-Late: PCW 7-12. PCW, post-conceptional week.

Results: Next, we projected gene expression profiles of non-malignant cells from the reference atlas^{27,28} onto individual ST-EPN cells (**Figure 1f, Extended Data Figure 2a**). While ZFTA-Cluster 3 and ZFTA-Cluster 4 tumors projected to the earliest time window of early and middle first trimesters, ZFTA-Cluster 2 and ZFTA-Cluster 1 tumors mapped onto the later windows of middle first and second trimesters. Finally, canonical ZFTA-RELA tumors projected broadly from the late first to second, and third trimester, while ST-YAP1 aligned with the latest timepoints. (**Figure 1f, Extended Data Figure 2a**). Overall, the correlation of ST-EPN subgroups to different developmental stages suggest divergent developmental patterns across subgroups.

Changes made: *Figure 1, Results*

#3-11. Lines: 155-169: The authors state that annotation of the metaprograms are based on the expression of known cell type-specific marker genes in Table S4. They also highlight specific examples in the text for all the 8 metaprograms. However, in Table S4, when checking if all the mentioned program specific genes are within this table, I can only find 7 out of 16. From the text the reader assumes that the annotation of the metaprograms is based on the marker genes in Table S4. This is confusing to me. Could the authors please explain why 9 of the mentioned markers are not in Table S4 among the metadata program specific top 200 genes? Also, if these 9 genes are below the top 200 genes for each metaprogram, please specify and explain the rational for a cut-off for which genes are deemed to be metaprogram specific enough to merit their use as annotation markers for the different metaprograms. Furthermore, this should also be taken into consideration for GO enrichment analysis, that a cutoff and number of top genes are consistent and that the marker genes used for annotation is included in this top set of genes used for GO enrichment analysis (Extended Data Figure 2b). Currently this does not appear to be the case, I suggest setting motivated cutoffs and re-running the GO analysis based on these. See further comments in section F.

We thank the reviewer for this thorough analysis. The reviewer is correct in that we annotated each metaprogram based on the expression of known cell type-specific marker genes, projection onto developing human cerebral cortex reference atlases, and functional analysis using Gene Ontology (GO) enrichment analysis. We have now revised all our graphs (**Figure R38: Figure 2b, Extended Figure 3c**), tables (**Table S4**

and Table S5), analyses to consistently use the [top 30 metaprogram marker genes](#) (and not 200, as previously indicated in parts of the manuscript). Of note, this is the same gene list that was used to score ST-EPN cells and assign them an identity (**Methods**).

Figure R38

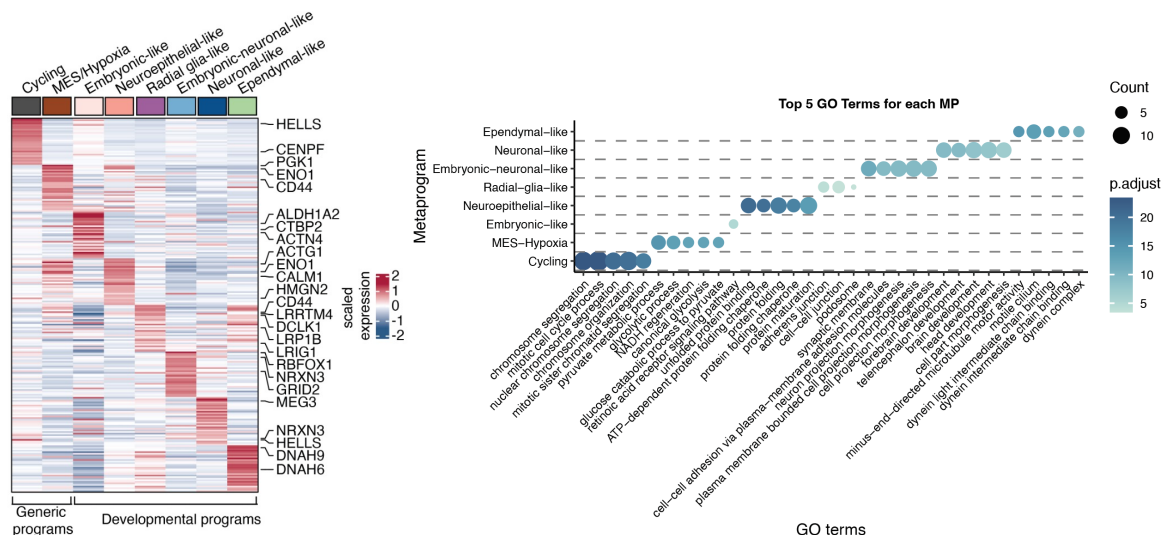


Figure R38. (left) Heatmap showing relative expression (z-score) of the top 30 genes for each metaprogram across all ST-EPN cells profiled in this study. Selected genes are highlighted. **(right)** Top five biological pathways enriched across the identified ST-EPN metaprograms by GO analysis of the top 30 NMF marker genes.

Methods: Scoring gene sets. We first aggregated ST-EPN single cells/nuclei from all samples, and performed normalization and centering. Then, we scored these cells for NMF/developmental signatures as previously described^{14,39,42–44}. We did this by calculating for each cell i a score $SC_j(i)$, quantifying the average relative expression (Er) of the top 30 genes within each MP (G_j) and comparing the score to the average relative expression of a control gene (G_j^{cont}): $SC_j(i) = \text{average}[Er(G_j, i)] - \text{average}[Er(G_j^{cont}, i)]$. The control gene-set was defined by binning all genes into 30 bins of aggregate expression levels (Ea) and randomly selecting 100 genes from the same expression bin for each gene in the gene-set G_j . ST-EPN cells/nuclei were annotated based on the maximum expression score for the respective MP signatures after excluding the one related to cell cycle (“cycling”). **Gene Ontology (GO) and Gene Set Enrichment (GSEA) analysis.** To identify biological terms that are over-represented in a specific gene set, we performed GO and GSEA analysis using *enrichGO*, respectively *gseGO* functions from the *clusterProfiler*¹³ package. As a database, we used the R package *org.Hs.eg.db* (v. 3.18.0), which is an organism annotation package for *Homo sapiens*, relying on the Ensembl version released on 2023-May10. Unless otherwise states, we used the following parameters: *ont* = “ALL”, *universe* = all genes retained for differential gene expression analysis, *minGSSize* = 3, *maxGSSize* = 800, *pvalueCutoff* = 0.05, *OrgDb* = *org.Hs.eg.db*, *pAdjustMethod* = “BH”.

We appreciate the increase in consistency upon the reviewer’s suggestions.

Changes made: Figure 2b, Extended Data Figure 3c

#3-12. Lines 173-175: The authors include GO results for 6 of 8 metaprograms in Extended Data Figure 3b. Please add all 8. If no significant results were obtained for the other 2 metaprogram, please state so.

We thank the reviewer for noticing this lack in consistency. We revised our GO results to include all 8 metaprograms (Figure R37 right, Extended Data Figure 3c).

Changes made: Extended Data Figure 3b

#3-13. Lines 177-178: The authors state “While ZFTA-RELA tumors were composed predominantly of neuronal-like and ependymal-like cells”, however looking at the subpanels for figure 3 d-e, I suggest they also

mention neuroepithelial-like cells. & 14. Line 178: The authors state “neuronal-like cells were absent in ZFTA-Cluster 1 and ST-YAP1 tumors”, however there are some proportions of neuronal-like cells in ZFTA-Cluster 1. I suggest adding “cells were few or absent”.

We have rephrased to address this point.

Results: While ZFTA-RELA tumors were composed predominantly of neuroepithelial-like, neuronal-like and ependymal-like cells, neuronal-like cells were few or absent in ZFTA-Cluster 1, ZFTA-Cluster 3, and ST-YAP1 tumors.

Changes made: Results

#3-14. Lines 217-225: This is great! The authors present high-level results from the spatial experiment and explain their rational for only going forward to analyze four non-malignant cell types and five malignant cell states.

We thank the reviewer for the kind enthusiasm.

#3-15. Lines 227-228: The authors state “was highly consistent across consecutive sections from individual patients, confirming high reproducibility of this technology (Extended Data Figure 6b).” Although the cell proportions are visualized in Extended Data Figure 6b the authors do not show any calculations supporting the specific claim that proportions were highly consistent across consecutive sections from individual patients. To make a claim for the consistency of proportions across consecutive sections, **please add calculations supporting this claim.**

We thank the reviewer for raising this point. We have now computed the Pearson’s correlation (according to cell state/type proportions) between sections belonging to the same sample and sections belonging to different samples, and then plotted these correlation values as violin plot (**Figure R39a**). The correlation values for sections from the same sample are higher than the ones for sections belonging to different samples, indicating the similarity in cell state/type compositions across the sections of the same sample. In addition, we took the samples from which there were at least 2 tumor sections, and for every possible pairing of sections belonging to the same sample we computed the correlation according to their cell state/type proportions. We then averaged the correlation values for all the section-pairs/sample (**Figure R40**). These values indicated that generally sections are highly correlated for each sample, except for a few exceptions such as STEP-47 and STEP-48. While we do overall see consistency across sections for each sample, we have removed this claim as highly consistent cell states/types identified across consecutive sections is also reflective of how the tumor sections were cut, not only the reproducibility of the technology.

Figure R39

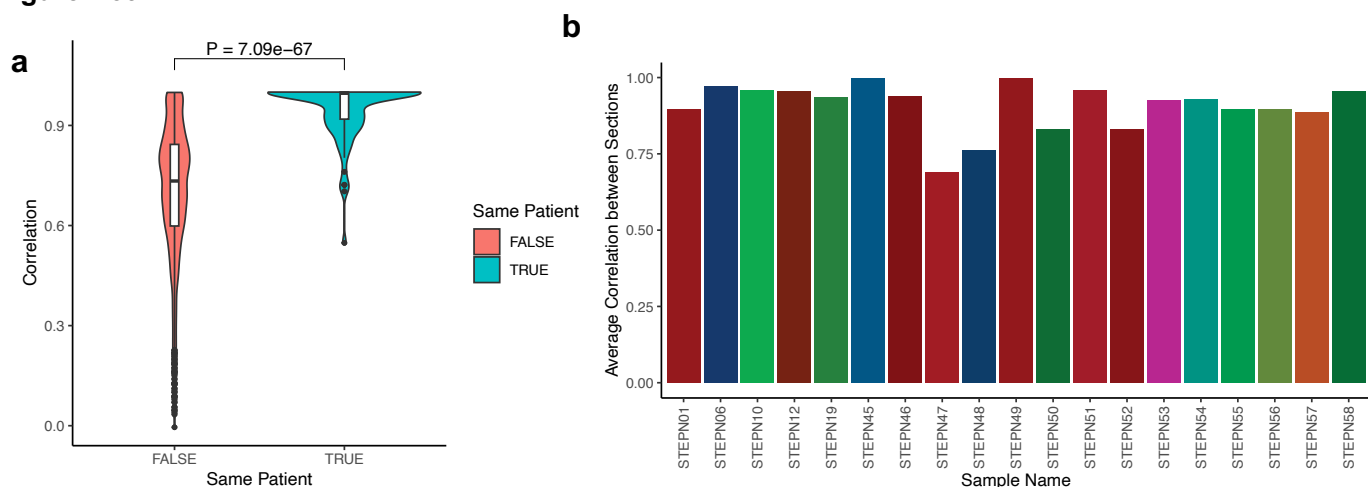


Figure R39. (a) Violin plot of Pearson correlation values (x-axis) of sections that are from the same or different samples (y-axis). **(b)** Average correlation between sections (y-axis) for each sample that has more than one section (x-axis).

#3-16. Lines 231-234: The authors state “After annotating spatially distributed cell states, we observed that tumors segregated into two global organization patterns – (1) high compartmentalization of cell states, which we term structured tumors, and (2) cell states/types scattered throughout the section, which we term disorganized tumors. (Extended Data Figure 5a).” However, there is no indication for case (1) and (2) in Extended Data Figure 5a. It would be valuable to have something similar to what is shown in Figure 3b. I suggest that the authors also reference Figure 3b here and add spatial coherence score visualizations to Extended Data Figure 6b.

We thank the reviewer for the helpful suggestions. We have modified the figure (**Figure R40a**) to include spatial coherence scores for each section, as similar representation to **Figure 3b**. Tumors are ordered from left to right by increased spatial coherence scores, meaning that tumor sections on the far right show high compartmentalization of cell states (case (1), whereas tumors on the far right have cell states/types scattered throughout the section (case (2)). We have also included image of each section in for the disorganized and organized sections. Finally, we have amended **Figure 3b** to include a graphical representation of how spatial coherence score is calculated for an example section (**Figure R40b**).

Figure R40

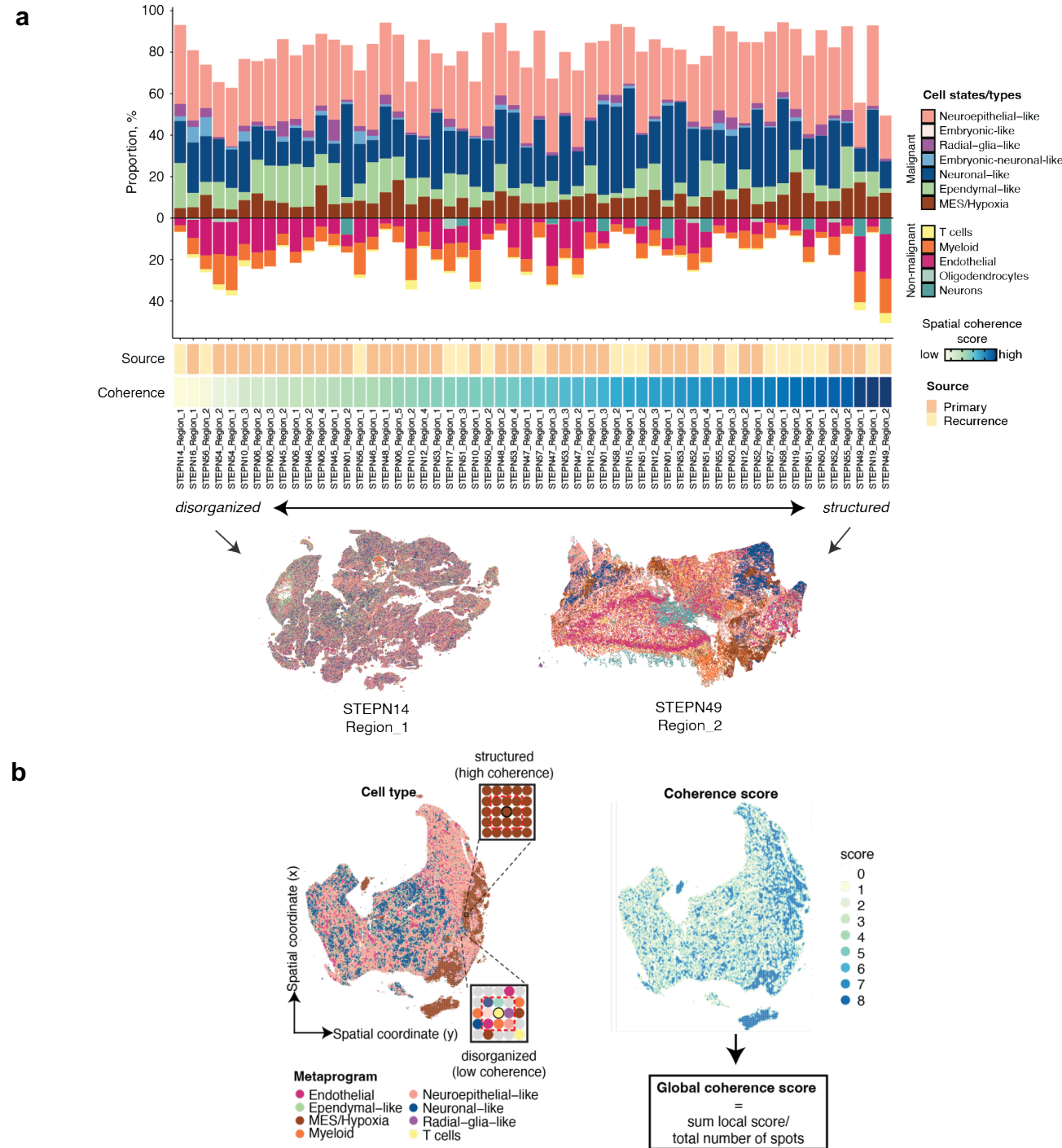


Figure R40. (a) Proportion of tumor metaprograms and non-malignant cell types (y-axis) in all ZFTA-RELA tumor sections profiled by 10X Xenium (x-axis). Example spatial maps of two sections, demonstrating differences in spatial organization between disorganized (STEPN14-Region-1) and structured (STEPN49-Region-2) tumors. **(b)** Schematics illustrating the calculation of the spatial coherence score, depicting two regions with high, respectively low spatial coherence scores (**left**), and the calculated coherence score for the example section (**right**).

Changes made: Figure 3b, Extended Data Figure 5d

#3-17. Lines 242-245: The authors state “we performed linear regression analyses between the spatial coherence score of each section and the frequency of cells observed for each state/type.” By formulating this

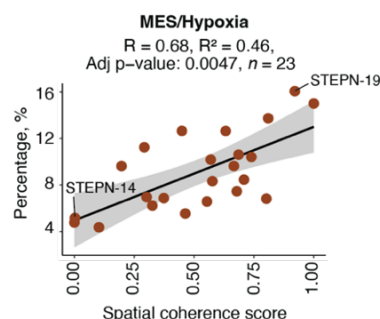
explorative hypothesis, one expected the authors to investigate all 9 cell states/types (4 non-malignant, 5 malignant). This should then result in 9 linear regression calculations and 9 p-values which then multiple correction is applied for, e.g., preferably using Bonferroni correction given the small number of p-values to correct for. However, the authors only show linear regression calculations and p-values for MES/Hypoxia and vascular cells (no mention of multiple correction). This needs to be addressed. Given 9 test, the adjusted p-value for MES/Hypoxia would be $p_{\text{adjusted}} = 0.018$ (still significant). The authors draw the conclusion that the MES/Hypoxia state is especially important for spatial organization without investigate and/or show results from the other cell types and state. Please address this, which could have implications for how the results are interpreted and lead to major revision of this section of the manuscript. See further comments in section F.

&

#3-18. Lines 246-249: The authors hypothesize “that vascular perfusion might decrease hypoxia, and therefore revert this structural organization”. In my opinion, this is too early a hypothesis to form from this data. First the previous comment needs to be addressed, there could be more states than MES/Hypoxia that correlate with structural organization. In addition, in Figure 3c, the p-value (appearing to be non-adjusted for multiple testing) for linear regression between structural organization and vascular cells is not significant, i.e., the data and analysis does not currently support this line of reasoning and the authors conclusions. My suggestion is to remove this from the manuscript.

We thank the reviewer for these important points. As also addressed in **Reviewer comment #3-3**, we performed linear regression analyses for all cell states/types with an increased sample size. In addition, as the reviewer suggested we corrected for the p-values by the number of statistical tests performed using Bonferroni correction (**Figure R5a**). In addition, as the reviewer also had advised in **Reviewer comment #3-19**, we calculated the mean value of samples that have multiple sections. With these corrections, we find that the correlation between spatial coherence score and MES/Hypoxia continues to be statistically significant ($R = 0.68$; adjusted P value = 0.0047). We find that no other cell state/type is significantly correlated/anti-correlated with the global spatial coherence, including the vascular cells. As this is the case, we have amended **Figure 3**, included correlation with other cell states/types in **Extended Data Figure 4e**, and removed previous conclusions regarding vascular cells and structural organization in the relevant lines of **Results** and **Discussion**.

Figure R5



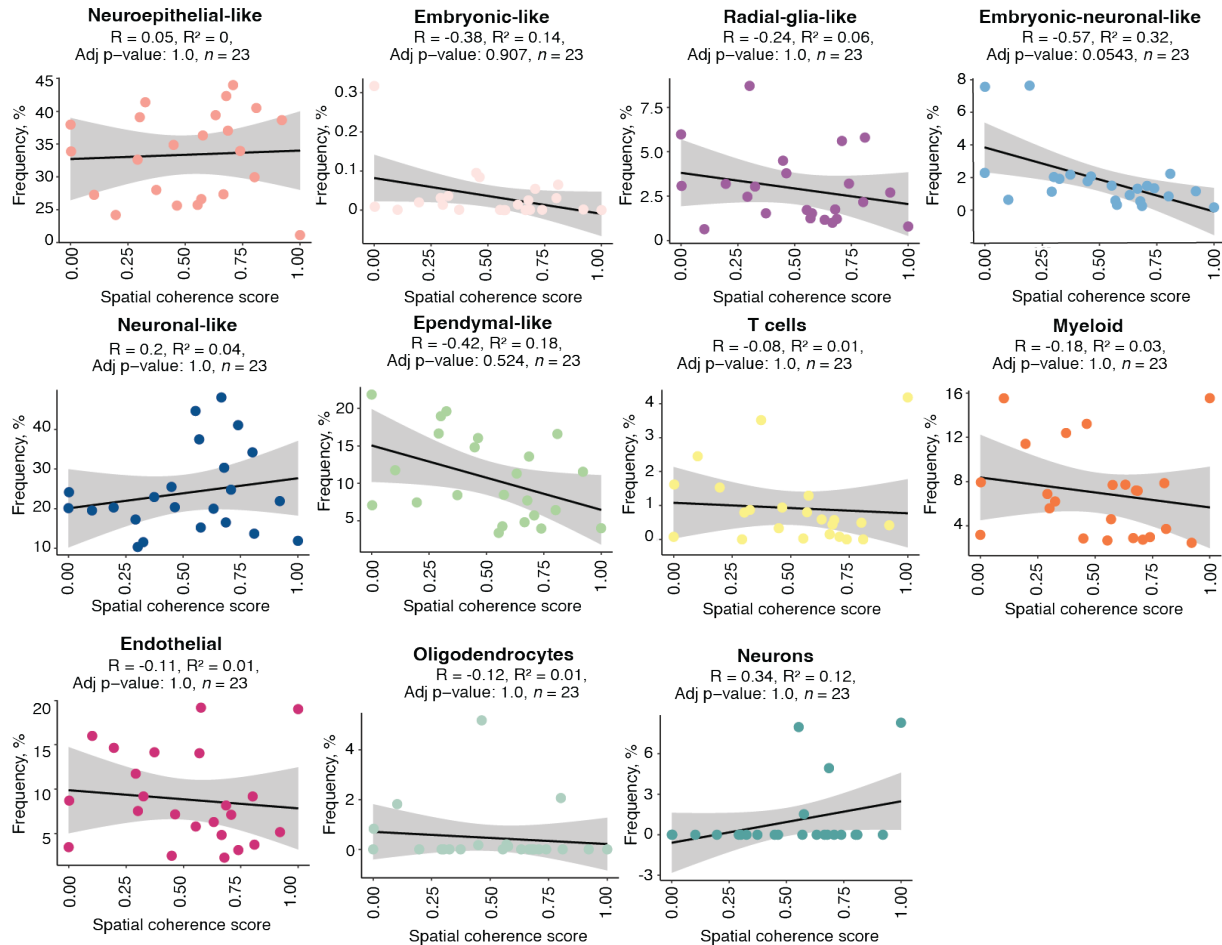


Figure R5. (a) Correlation between the spatial coherence score and the percentage of individual cell states/types in ZFTA-RELA tumor samples profiled by 10X Xenium. Data points (representing each tumor sample) are interpolated with a simple linear regression. For tumor samples with multiple sections, the average values are shown. The goodness of fit (R , R^2), Bonferroni-corrected P-values and total number of tumor sections (n) are displayed at the top.

Changes made: Extended Data Figure 4e, Results, Discussion

#3-19. Lines 251-254: The authors state “This indicates that higher spatial organization of ZFTA-RELA tumors correlates with more aggressive tumors.” However, there is no mention of how this is calculated in the material and methods. The statistical test (mentioned in the figure legend) in Extended Data Figure 6f-g requires a normal distribution, which has not been assessed by the authors, please re-evaluate and used appropriate motivated statistics. The p-values are not significant, thus the claims and reasoning by the authors for this part is not supported by the data and data analysis. For this data, care needs to be taken here, because some tumors have been sampled from multiple regions (perhaps use an average measurement from each tumor?). Also, it is unclear if patients with primary tumors have had recurrent disease at a later time-point. My assessment is that this part is too speculative and the sample size in this cohort is too small to draw any conclusions here. I suggest removing this from the manuscript.

We thank the reviewer for these important points and agree that the comparison of primary versus recurrent tumors is too speculative based on small cohort size. As also addressed in **Reviewer comment #3-3** and **#3-17~18**, we increased the spatial cohort by adding 33 new ZFTA-RELA 10X Xenium samples of 13 new patients (from $n=23$ to $n=56$ sections; $n=10$ to $n=23$ patients). In this increased cohort, we first calculated the mean spatial coherence values for samples that had multiple sections, and computed whether the spatial coherence scores of primary versus recurrent tumors were normally distributed by Shapiro Wilk’s test. Both

primary and recurrent tumor scores were normally distributed (Shapiro Wilk's test p-value for Primary: 0.978, p-value for Recurrence: 0.116), and therefore we performed a student's t-test (**Figure R17a**). We did not see a statistical difference in spatial coherence between primary and recurrent tumors. As this is the case, we have removed this part from the manuscript and removed conclusions in the manuscript pertaining to primary versus recurrent samples, as well as statements regarding spatial organization correlating with more aggressive tumors.

Figure R17

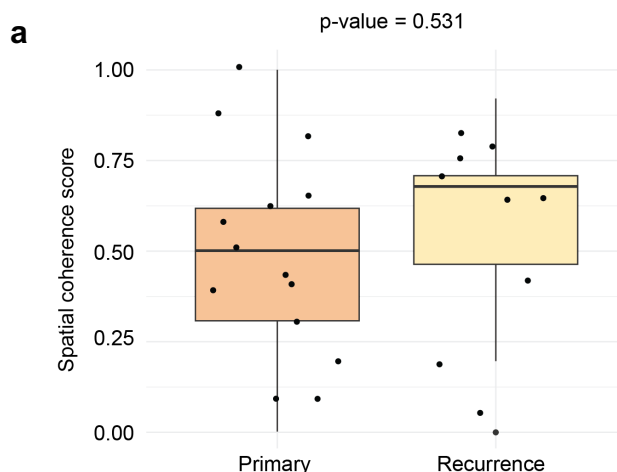


Figure R17. (a) Scaled spatial coherence score across ZFTA-RELA derived from primary or recurrent patients. Each dot depicts the average value of sections for each sample. P-value calculated by two-sided student t-test.

#3-20. Lines 260-268: The authors describe the spatial Xenium data by dividing different regions into different spatial niches. However, to me as a reader it gets confusing. There are first niches 1-4 for each section, but results are only shown for one subregion from two sections. I'm not certain why the authors selected to only look for 4 niches across all tumors (I would like to see an extended data figure showing spatial niches for all samples), which also have different degrees of spatial organization, which could affect how many niches there could be. Please motivate why 4 niches were chosen. E.g., why not pick a higher number of niches, e.g., 9 that was investigated before (4 non-malignant, 5 malignant)? Especially since the authors do make observations that indicate that transcriptional state and cell types are directly correlated with morphology (Figure 4a). Please provide a clear motivation for this choice, and additional supplementary figures showing the results from the analysis.

&

#3-21. Lines 269-274: Here the authors define recurrent spatial niches (methods lines 769-773) based on the frequency of cell types, reading the manuscript and looking at the corresponding subpanels and pie chart plots (Figure 4), I interpret this as also including the different cell states. The authors then conclude that "Most recurrent niches were predominantly composed of one malignant cell state (Figure 4c), which indicates that malignant cell states preferentially co-localize with themselves." But according to the current definition of these recurrent cell types, this argument becomes circular (since recurrent spatial niches were also defined by frequency of cell types). Thus, in my opinion, the conclusion that "we uncovered local spatial patterns of cell states across ZFTA-RELA tumors, with a common observation of malignant cell states preferentially co-localizing with themselves.", is not supported by this data analysis. See section F for suggestion of re-analysis.

We thank the reviewer for these constructive questions. Overall, we see that cellular composition, morphology and spatial arrangements can vary widely within and across spatial regions. As such, the "correct" number of spatial niches is fundamentally elusive, where higher or lower number of clusters can reveal different aspects of underlying biology with each approach offering different insights and limitations. This challenge is conceptually very similar to determining the optimal number of clusters in scRNA-seq data

analysis, where in both cases the scientist attempts to balance resolution and interpretability to reach some biological truths⁵⁴. While for scRNA-seq the field has grown in generating various algorithms for clustering^{55–57}, spatial transcriptomics field still requires more systematic studies comparing clustering methods to reliably identify biologically meaningful local spatial clusters in samples.

With these conceptual limitations, we addressed this question using several approaches. We first compared cell states/types using different number of niches (niches.k), to determine whether results were consistent across the different numbers of k. (**Figure R41**). According to the reviewer's recommendation in **Section F (Reviewer comment #3-X)**, we compared the results of niches.k=4 and 5, as well as a higher value of 6, and found that these parameters resulted in similar identification of metaniches (**Figure R41**). We further examined what each niche for the niches.k selected (4-6) captures on the tissue (**Figure R42**). These examinations showed us that niches.k = 6 captures more histologically discrete regions within the tissue than niches.k = 4, such as ependymal rosettes (**Figure R43**). We also compared the results of niches.k=9 as the reviewer had also suggested, and found that this value broke apart the ependymal rosettes (**Figure R43**). As this was the case, we chose to include niches.k = 6 results for all revised figures.

As the niche/metaniche analysis fundamentally does not have a metric to quantitatively compare the different values of niches, we further aimed to validate the results from metaniches by using an orthogonal analysis called *CellCharter*⁵⁸. Briefly, this method aims to simultaneously determine spatial patterns across samples, with built in algorithms to determine the “most stable” cluster assignments across multiple clustering parameters. We reasoned that this approach could address **Reviewer comment #3-20** regarding selection of optimal clusters, as well as **Reviewer comment #3-21** regarding the possibly circular nature of metaniche analysis.

Using CellCharter, we aggregated all cells and their features, and examined reproducibility/stability of the number of clusters (stability defined as cluster assignment reproducible across multiple clustering runs) from $n = 3$ to 30 (**Figure R44a**). We found that $n = 4$ and $n = 26$ had the highest cluster stability. As $n = 4$ mainly grouped cells into low-resolution clusters enriched with either malignant or normal cells (**Figure R44b**), we chose the next highest $n = 26$ local patterns, or *spatial clusters*, for downstream analysis. We compared cell states/types composition across spatial clusters (**Figure R44c**), and identified the most abundantly present cell state/type per spatial cluster that we refer to as “enriched”. We found that the categories of cell state/type-enriched spatial clusters derived with this orthogonal method indeed resembled the dominant cell states/types of metaniches identified previously (**Figure R44d-e**). While the metaniche analysis is based on correlations of the proportion of cell state/types for each niche in each section, potentially making the statement that ‘malignant cell state preferentially co-localize with themselves’ circular, CellCharter spatial clusters are identified through integration of spatial coordinates and gene expression profiles in an unbiased method across all sections simultaneously and thus is not limited by this issue. Comparable results from metaniche and spatial cluster analyses reinforced our confidence in the validity of our findings was reinforced. Furthermore, CellCharter analysis revealed that within each cell state/type-enriched category, there were variations in the different cell types present in each spatial cluster. This uncovers greater complexity in local cell composition across samples. Finally, we were able to designate which spatial clusters were found across patients versus which clusters were patient-specific, further highlighting the heterogeneity of spatial organization in ZFTA-RELA tumors.

In conclusion, we aimed to address the reviewer's comments by both clarifying the conceptual limitations and providing an orthogonal method to validate our previous findings. We thank the reviewer for their questions which enriched our study, and we have included these analyses to **Figure 4** and **Extended Data Figure 6**, as well as **Results** section “**ZFTA-RELA tumors are locally organized into recurrent spatial clusters**”.

Figure R41

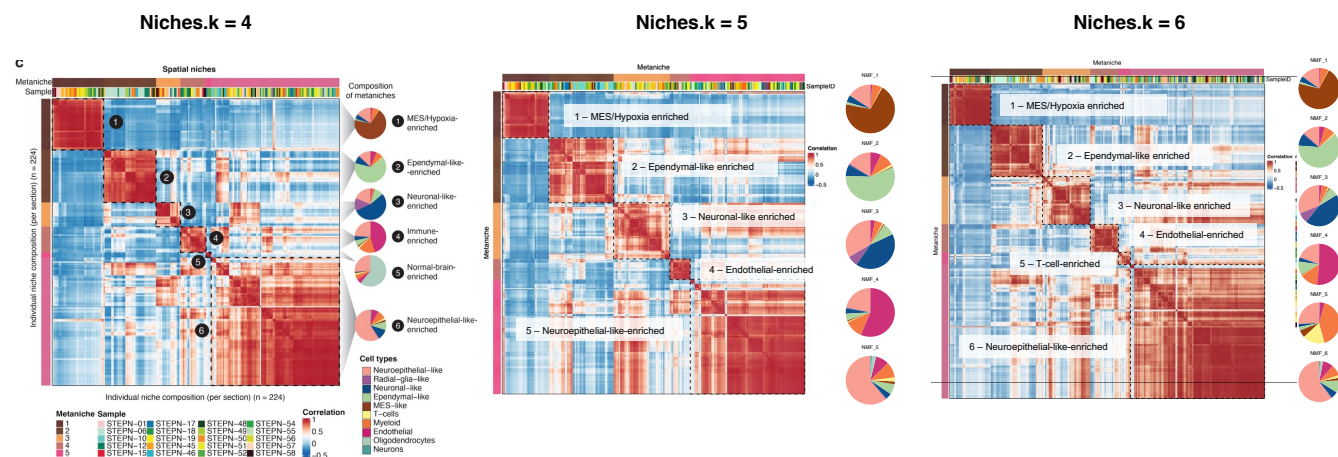
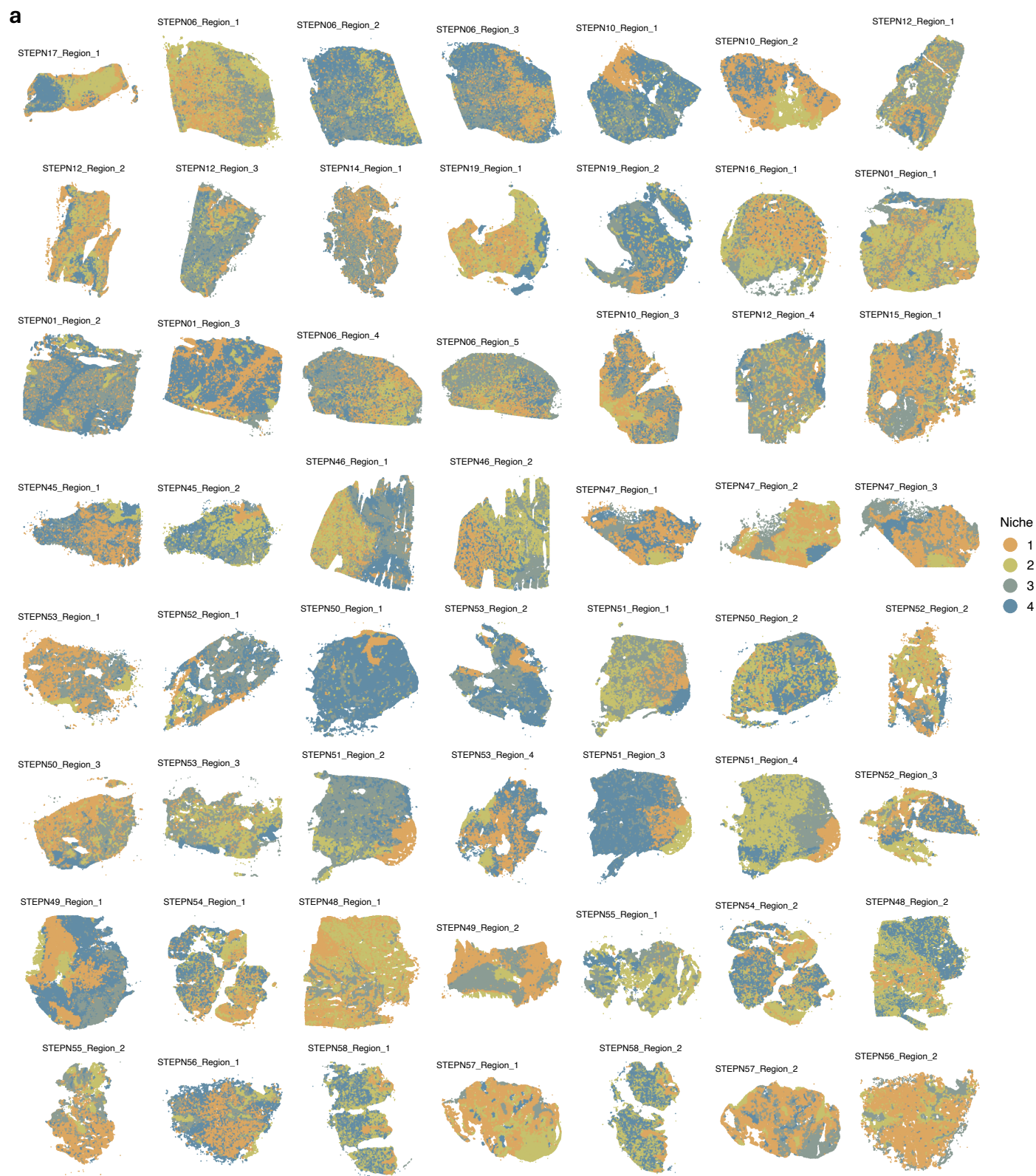
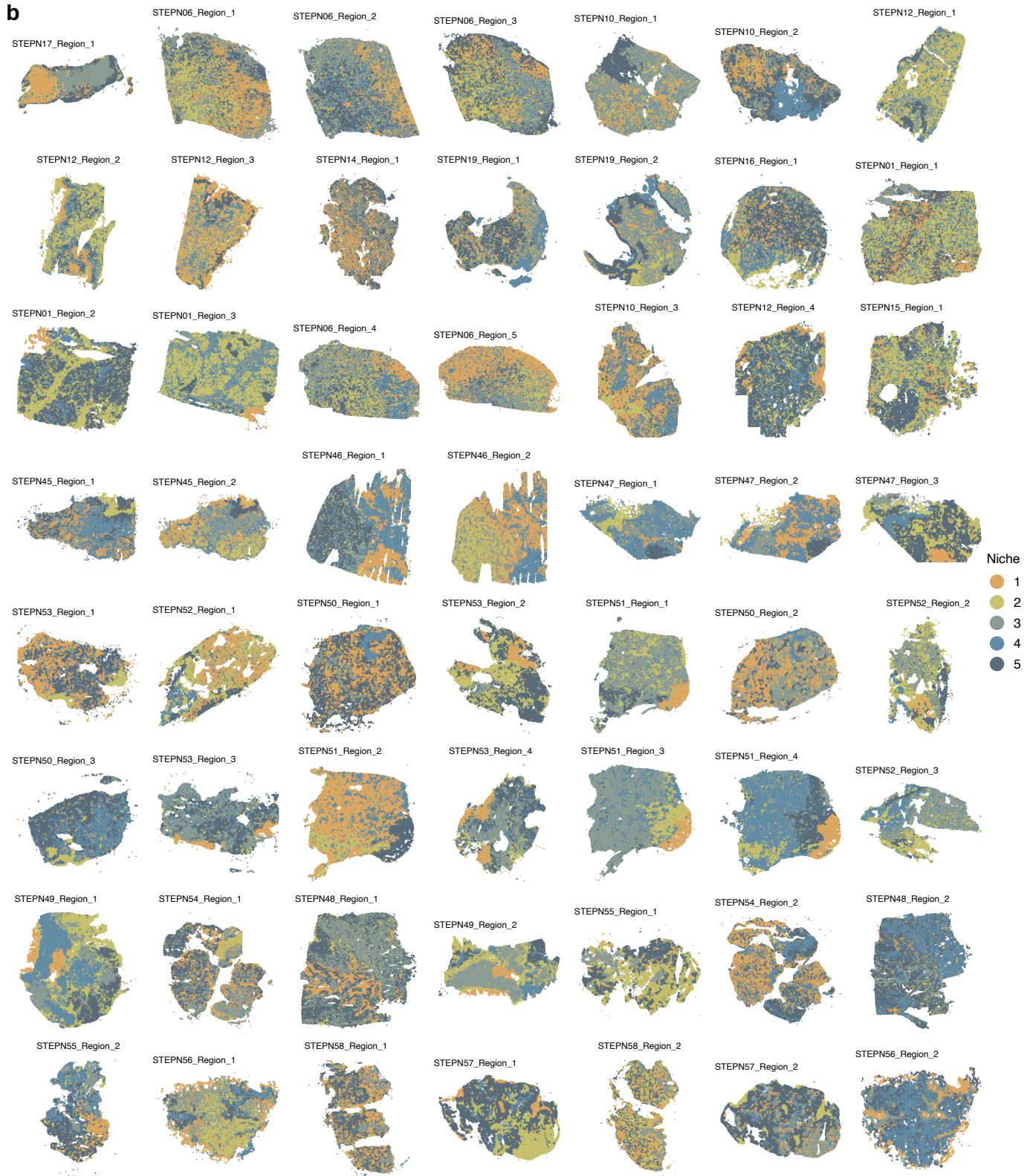


Figure R41. Pairwise Pearson correlation (color scale) representing similarity in cell type composition of the spatial niches identified in 56 ZFTA-RELA tumor sections analyzed by 10X Xenium. Each row and column represent one individual spatial niche identified in one tumor section, summarized by its relative frequency of cell states/types. Recurrent niches (metaniches) are highlighted by dashed line boxes. Niches are annotated based on the most frequent cell types (displayed on the right side in the pie charts). **(left)** niches.k=4, **(middle)** niches.k=5, **(right)** niches.k=6.

Figure R42





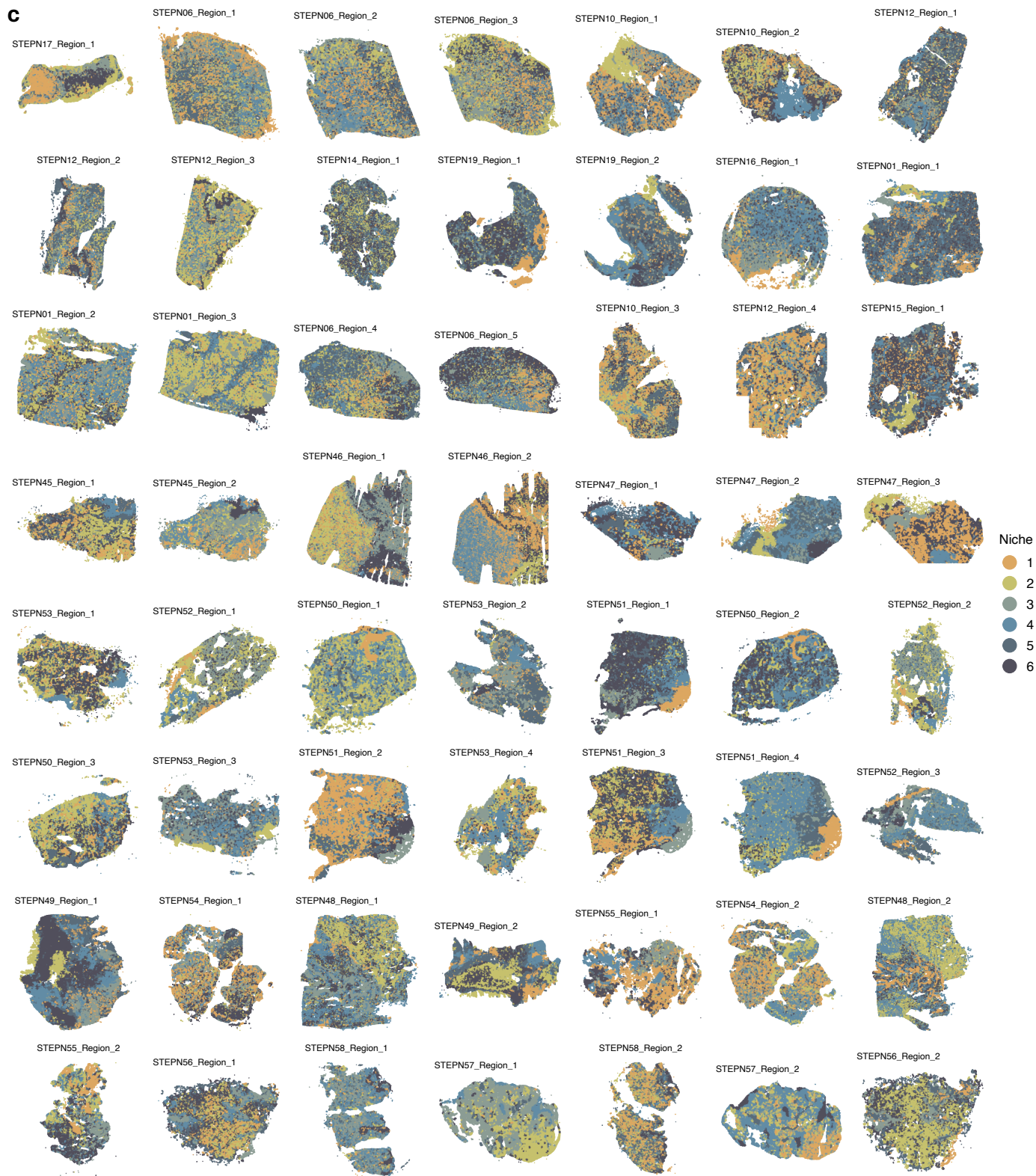


Figure R42. Spatial maps of 56 ZFTA-RELA tumor sections profiled by 10X Xenium, depicting distribution of niches when (a) niches.k = 4, (b) niches.k = 5, (c) niches.k=6 are identified for each section.

Figure R43

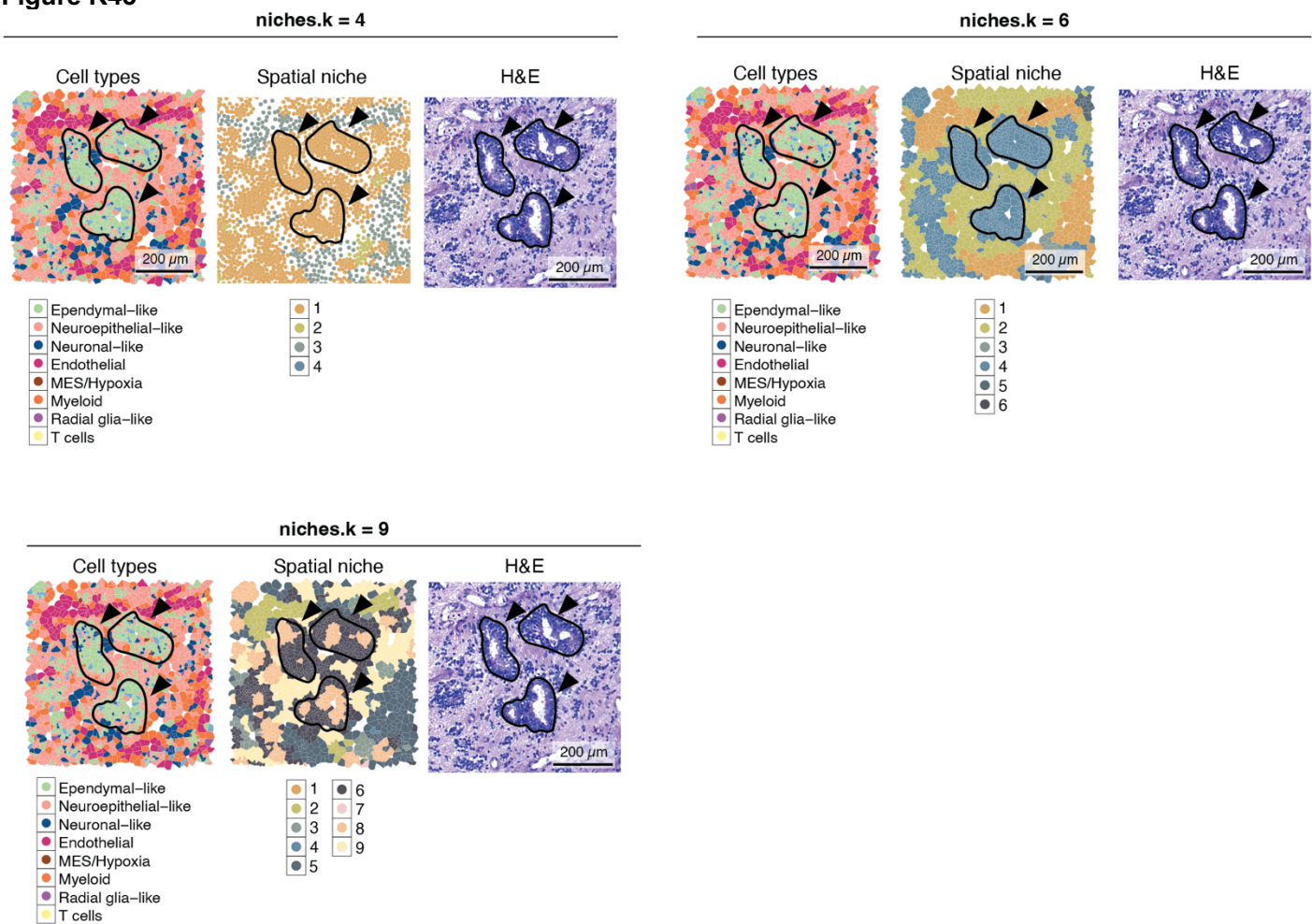


Figure R43. Overlay of 10X Xenium metaprogram assignment, CellCharter spatial clusters, and H&E staining for ependymal-rosettes (circled) in ZFTA-RELA tumor section (STEPN-12), showing correspondence of morphologically distinct local tumor region with metaprogram annotation and spatial niche assignment. **(left)** niches.k = 4, **(right)** niches.k = 6.

Figure R44

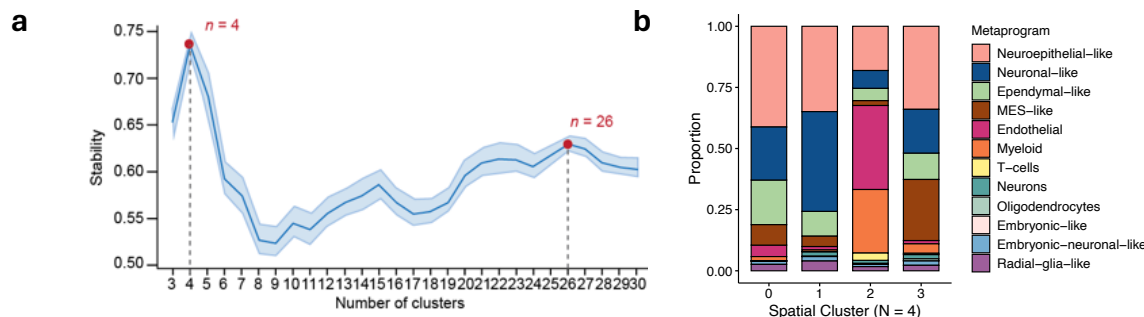


Figure R44. (a) CellCharter spatial cluster stability (y-axis) across different number of clusters (x-axis). Most stable clusters are marked. All depicted data are mean values with 95% confidence interval. **(b)** Proportion of cell states/types (x-axis) across spatial clusters (y-axis) at spatial cluster = 4. **(c)** Proportion of cell states/types (x-axis) across 26 spatial clusters (y-axis) **(left)**. Spatial clusters are grouped by similarity in the cell state/type most dominantly present in each cluster. Representative spatial maps of myeloid-enriched, MES/Hypoxia-enriched, and endothelial-enriched spatial clusters with their corresponding metaprogram annotations **(right)**. **(d)** Comparison of spatial niche assignment **(top)** to CellCharter spatial clusters grouped by similarity in the cell state/type most dominantly present in each cluster **(bottom)**. **(e)** Correlation of cell type proportion in CellCharter spatial clusters (y-axis) and in metaniches (x-axis). Spatial clusters identified by CellCharter are grouped based on the most abundant cell types, and cells assigned to spatial clusters of the same group are pooled to compute proportions of cell types. Cells assigned to each metaniche are also pooled to compute proportions of cell types. The correlation in cell type proportions is assessed by linear regression. Each data point is a cell type. **(f)** Heatmap depicting spatial clusters (rows) present across 56 tumor sections (columns). Rows are ordered by entropy of spatial cluster, with top half annotated as sample-specific (present in 1-2 samples) and bottom half annotated as sample-shared (present in greater than 3 samples). Top row is colored by sample ID.

Changes made: Figure 4, Extended Data Figure 6, Results section “ZFTA-RELA tumors are locally organized into recurrent spatial clusters”

#3-22. Lines 310-311: The authors states “we characterized subcellular and cellular tumor cell behaviour with longitudinal live-cell imaging”, to me the use of ‘longitudinal’ is more used when studies span months and years, I suggest the use of ‘temporal’ instead.

We rephrased **Results** to include this point.

Results: We next investigated whether cellular states with distinct morphologies also have unique cellular behaviors. To this end, we characterized subcellular and cellular tumor cell behavior with temporal live-cell imaging.

#3-23. The authors uses the word ‘somatokinesis’ to describe what I believe is cell movement. I can only find one mention of this word in ncbi. Is there some other word that is more commonly and broadly used? Please define this term clearly in the text for the reader.

We thank the reviewer for this helpful feedback. As the reviewer correctly stated, somatokinesis describes cell soma movement. To use a term more broadly usable, we have changed to “**cell migration**” or “**cell migration speed**” for the text in manuscript and relevant figures.

#3-24. The authors use the definitions ‘continuous invasion mode’ and ‘saltatory invasion patterns’. Please clearly describe and define to the reader what these entails. E.g., in the main text or more detailed in material and methods.

We thank the reviewer for this helpful comment. Firstly, as we did not explicitly examine invasion in our assays we changed “continuous invasion mode” to “continuous migratory mode”, and “saltatory invasion patterns” to “saltatory migratory patterns”. Accordingly, the “continuous migratory mode” of cells involves a steady motion across time, whereas in a saltatory pattern cells move in rapid bursts of motion that are separated by stationary pauses, progressing in a discontinuous, “jumping” manner. “Saltatory movement” has been used to describe the migration patterns of immature neurons during development^{59–62}. We chose to use this term to describe the movement of neuronal-like tumor cells, as they closely follow the movement patterns of these immature neurons. We have described these terms in detail in **Results**.

Results: Moreover, we noted distinct modes of migration across cell states, including a continuous migratory mode where cells undergo a steady motion across time, and saltatory migratory mode where cells move in intermittent bursts of rapid, jumping motion.

Changes made: Results

#3-25. The authors use different terms to describe invasive cell behavior, such as: protrusion, retraction, branching, branching migration, locomotion and translocation. I suggest adding clear and descriptive definitions by the authors in the methods section of what these are and entail.

We thank the reviewer for this feedback. “Protrusion” is a form of tumor microtube (TM) dynamics that entails an outward growth of the TM, while “retraction” entails a shrinkage of the TM, or “retreating” growth of the TM. In addition, “branching” is a form of TM dynamics that entails growth of a secondary TM that protrudes out of the pre-existing TM. We have explained these terms in more detail in **Methods**.

Methods: Measurement of TM dynamics. All TM measurements was performed in Fiji⁶³. For both *in vivo* and *in vitro* experiments, TM protrusion was defined as TM outgrowth, retraction as decreased growth, and branching as TM dynamics where secondary TM protrudes out pre-existing TM as previously described^{32,33}. Net TM turnover was measured by measuring the length of all TMs at the starting time point and after 4h.

Changes made: Methods

#3-26. Line 320: The authors mention ‘EPN-glial-like’ cell for the first time in the manuscript. It is not clear to the reader which meta program or cell state these are related to. Please clarify. Is it Radial glial-like? If yes, then why in Extended Data Figure 10d does it appear to be more of this in monoculture but on line 341 is says that it is detected less?

We thank the reviewer for this important question. We agree with the reviewer that “EPN-glial-like” is a new term that is disconnected from the rest of the manuscript. To address this point, we aimed to further characterize the “EPN-glial-like” (tumor cells with greater than 3 short TMs) by performing additional immunofluorescence (IF) and 10X Xenium on all tumor morphological subtypes. We first amended the methods for IF staining and increased the panel of markers to increase consistency and rigor of the study. For example, we improved the detection of tumor cells by amplifying the tdTomato signal using anti-tdTomato antibody. This allowed us to more clearly distinguish the tumor cells and helped us identify the previous background signals that led to misleading results. In addition, we improved our method of quantifying immunofluorescence signal to a semi-automated methodology, by performing segmentation of tdTomato signal of each cell to capture the morphology and area of the cell. Consequently, we found that the tumor cells with more than 3 short TMs are **negative** for Vimentin and GFAP, and **positive** for S100B as well as **CCDC40** (**Figure R9a**), a protein specifically expressed in ependymal motile cilia of the brain ventricles^{17,18}. We confirmed that *CCDC40* is specifically expressed in ependymal-like cells of patient tumor samples (**Figure R9b**). Furthermore, by 10X Xenium of tumor cells in coculture with rat E19 cortical cells, we overlaid cell morphology to a panel of 366 transcripts. This allowed us to expand the number of markers examined the for these cells. From these analyses, we found that tumor cells with more than 3 short TMs are positive for *MAP3K19* and *DNER* (**Figure R9c**), which are also highly expressed by ependymal-like cells in patient tumor samples (**Figure R9b**). These cells also express other progenitor-like markers, including *SOX10* (**Figure R9c**), which is not expressed in the ependymal-like cells of patient tumors (**Figure R9b**).

To further clarify the difference between tumor cells in patient tumors and tumor cells in coculture, we analyzed the 10X scRNAseq data across patients and patient-derived models (monoculture, coculture, PDX) (**Figure R10**). Tumor cells in adherent and spheroid monoculture conditions were predominantly neuroepithelial-like, while subpopulations in patient and PDX were distinctly neuronal-like (bottom left quadrant in **Figure R10c-d**) and ependymal-like (bottom right quadrant) cells. Finally, tumor cells in coculture with rat E19 cortical cells were somewhere in between, with transitional cells midway progenitor-like and mature-like gene signatures. These findings suggest that tumor cells in coculture are less differentiated than cells in patient tumors. In light of these analyses and that tumor cells with more than 3 short TMs are positive for both ependymal-like and progenitor-like markers, we reasoned that these cells are likely more immature ependymal-like cells. Accordingly, we amended the nomenclature of these cells to “**immature ependymal-**

like” cells in manuscript, referring to cells that are simultaneously enriched with immature progenitor-like as well as tumor ependymal-like metaprogram gene signatures.

Figure R9

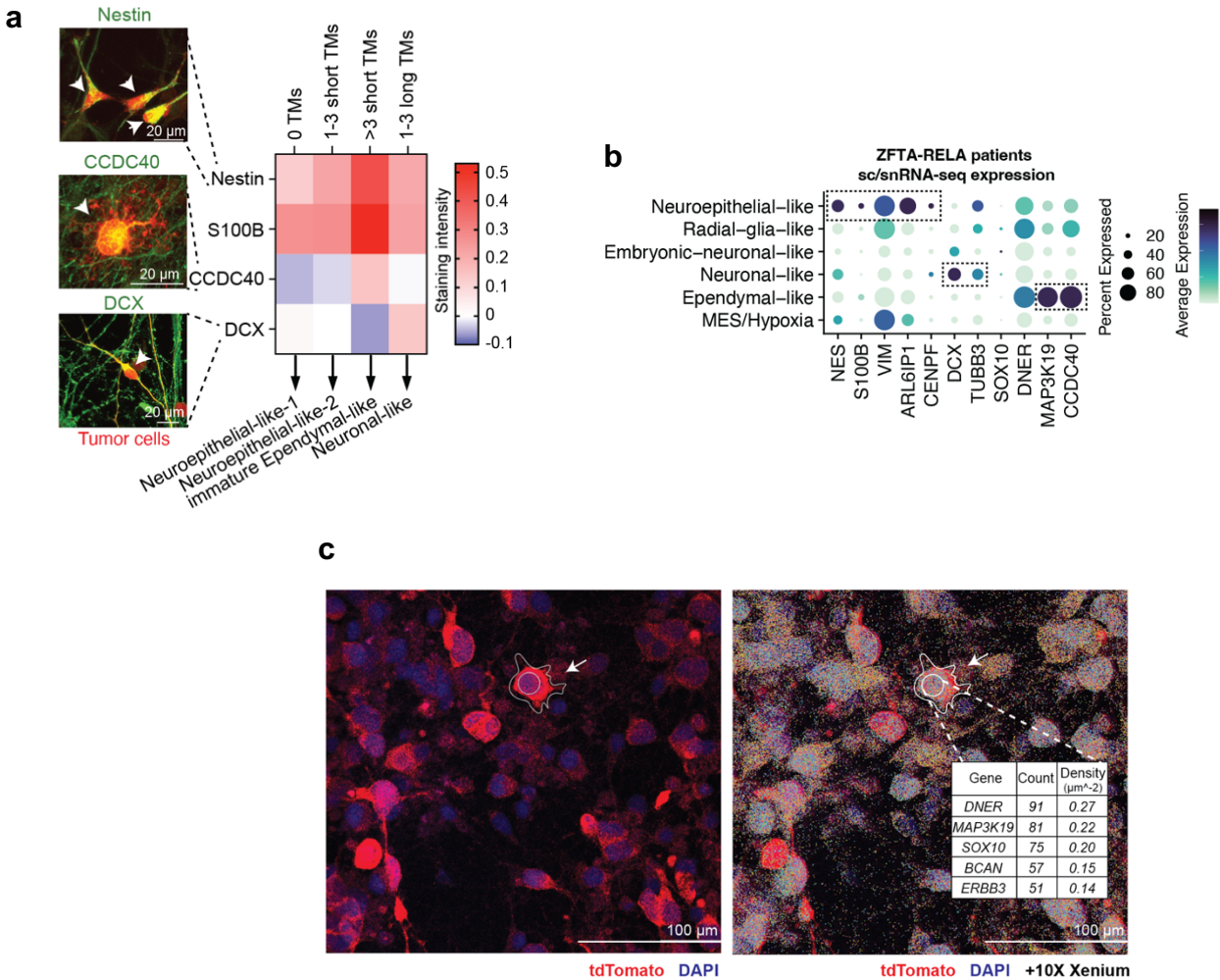


Figure R9. (a) Representative immunofluorescence images and mean intensity of signal of an *in vitro* co-culture model of ZFTA-RELA cells expressing tdTomato, categorized by morphological cell subtypes. Data represent the mean intensity value of $n = 155$ cells acquired from $n = 3$ independent biological replicates per marker. **(b)** Dotplot showing scaled mRNA expression of marker genes for each metaprogram in ZFTA-RELA patient tumors profiled by sc/snRNAseq. Dot size represents percent expressed, while color scale represents the average expression. **(c) (left)** Immunofluorescence image of ZFTA-RELA (EP1NS) cells expressing tdTomato) cocultured with rat E19 cortical cells, stained for anti-tdTomato and DAPI. **(right)** Immunofluorescence image overlaid with 10X Xenium transcripts. Each dot represents a detected transcript from the custom ($n= 103$) and base panel ($n= 255$) probes used for patient tumor sections. Table indicates top 5 enriched markers in demarcated region (white).

Figure R10

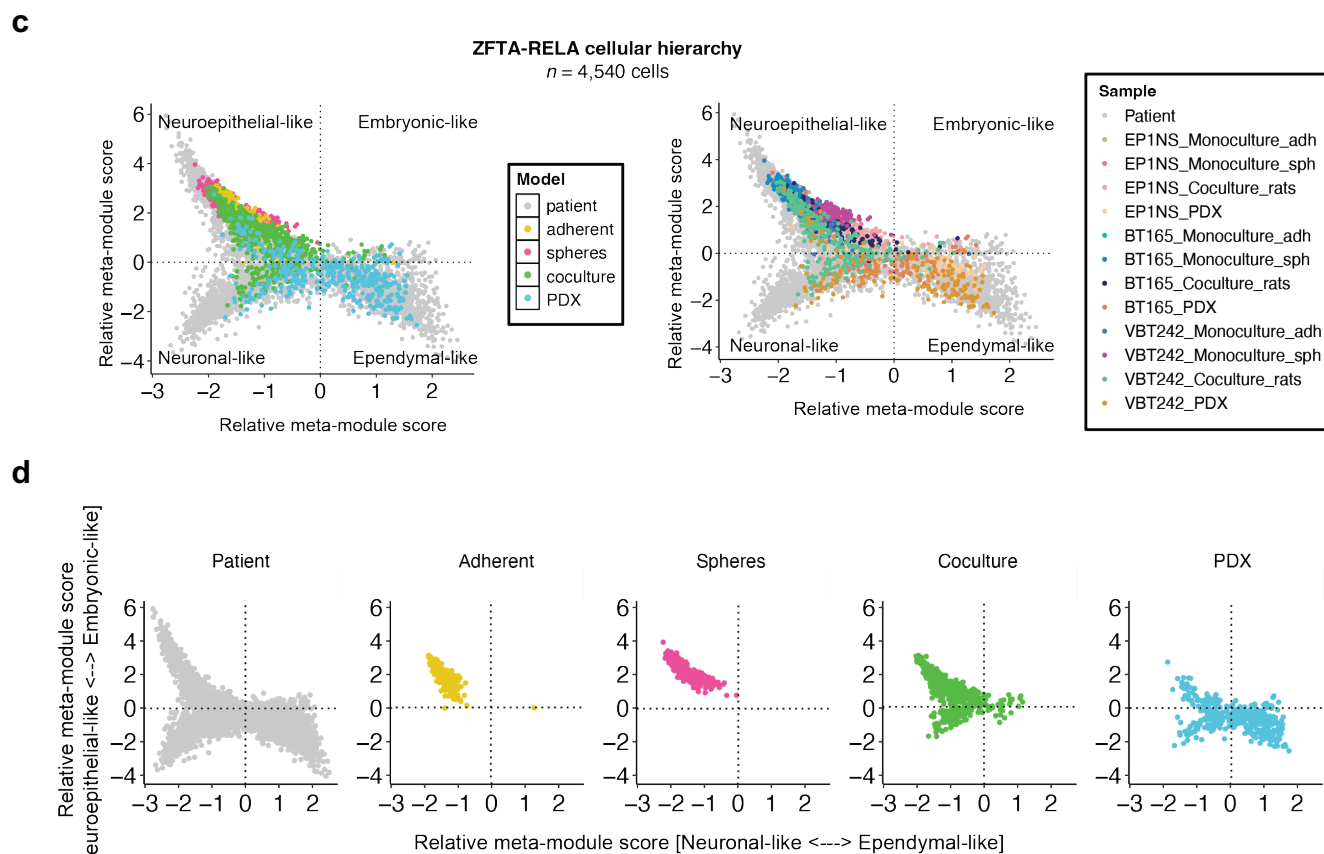


Figure R10. (c) Cell state plots of ZFTA-RELA patient tumor cells profiled by snRNAseq overlayed with single cells derived from PDX ($n = 3$), adherent ($n = 3$), neurosphere ($n = 3$) and coculture with rat E19 cortical cells ($n = 3$) models of EP1NS, BT165, and VBT242. Plots are colored by model (left) and by sample (right). **(d)** Cell state plots, separated by ZFTA-RELA patient tumors profiled by snRNAseq, adherent monoculture, spheroid monoculture, coculture with rat E19 cortical cells, and PDX profiled by sc/snRNAseq.

Changes made: Figure 5, Extended Data Figure 9

#3-27. Lines 338-340: The authors state that “in coculture versus monoculture and found significant enrichment of neuronal-like GO terms and gene signatures only when tumor cells are grown in coculture, as well as a decrease in neuroepithelial-like signature”, however only GO analysis is shown for coculture. Please include how this analysis was done in the methods section, also include how many genes were used for the GO analysis and do the same for the monoculture. Display these top GO enrichment terms in plots side by side in Extended Data Figure 10, to show that/if there is a difference among GO terms between monoculture and coculture.

We thank the reviewer for this feedback. As we have now additionally performed coculture in 3 different ZFTA-RELA cell lines (EP1NS, BT165, VBT242), we performed GO analysis of genes enriched in monoculture and coculture of the 3 cell lines to **Extended Data Figure 10f-g (Figure R45)**. We have added information on how this analysis was done and how many genes were used in the corresponding legend and/or method sections.

Figure R45

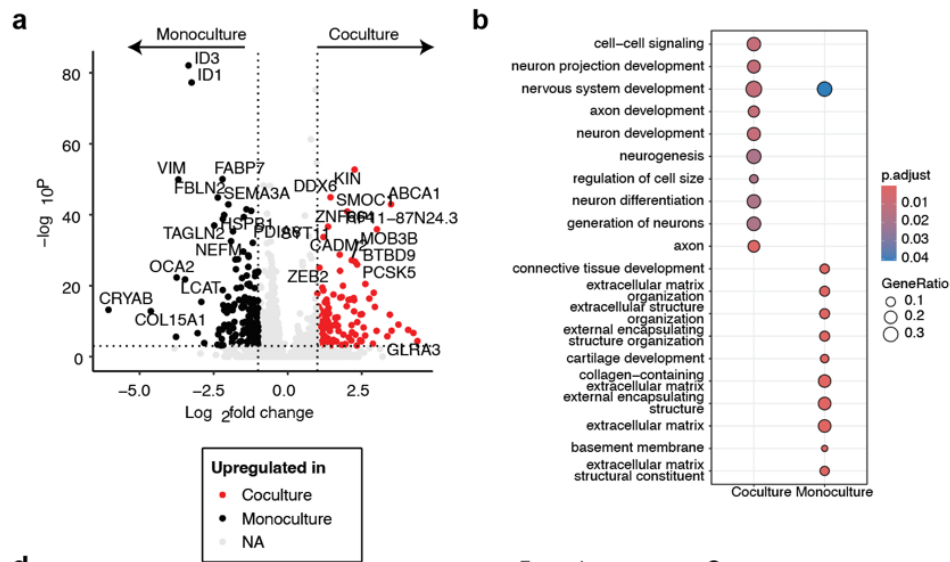


Figure R45. (a) Volcano plot showing genes differentially expressed between ZFTA-RELA cells (EP1NS, BT165, VBT242) in coculture versus monoculture based on Seurat’s *FindAllMarkers* function. Genes with $|\log_2FC| > 1$ and adjusted P-value < 0.001 are highlighted in color. **(b)** Top 10 gene ontology (GO) terms enriched in genes differentially expressed (input for GO: $|\log_2FC| > 1$ and adjusted P-value < 0.001) between coculture and monoculture ZFTA-RELA cells based on the genes identified in (a).

Methods: Differential gene expression analysis. Single-cell. To identify genes that are differentially expressed between two conditions (monoculture vs coculture) or cell states (neuroepithelial-like vs embryonic-like), we first used Seurat’s *RunPrestoAll* function (only genes minimally detected in 15% cells). For downstream gene ontology (GO) analysis we used the top 30 genes (ordered by average $\log_2FC$) with an average $|\log_2FC| > 1$ and adjusted P-value threshold as indicated in the corresponding figure legend, for gene set enrichment analysis (GSEA), we used all genes with an adjusted P-value threshold of 1^{-05} . **Gene Ontology (GO) and Gene Set Enrichment (GSEA) analysis.** To identify biological terms that are over-represented in a specific gene set, we performed GO and GSEA analysis using *enrichGO*, respectively *gseGO* functions from the *clusterProfiler*¹³ package. As a database, we used the R package *org.Hs.eg.db* (v. 3.18.0), which is an organism annotation package for *Homo sapiens*, relying on the Ensembl version released on 2023-May10. Unless otherwise states, we used the following parameters: *ont* = "ALL", *universe* = all genes retained for differential gene expression analysis, *minGSSize* = 3, *maxGSSize* = 800, *pvalueCutoff* = 0.05, *OrgDb* = *org.Hs.eg.db*, *pAdjustMethod* = "BH".

Changes made: Extended Data Figure 10, Methods

#3-28. Line 345-346: The authors state that “only ZFTA-RELA cells grown *in vivo* and cocultures are enriched in neuronal-like and ependymal-like states,” Extended Data Figure 10f. However, the used of the word ‘enriched’ is not supported by any calculations or statistical test. From the plot in Extended Data Figure 10f it does not look like ZFTA-RELA cells grown *in vivo* are present only in neuronal-like and ependymal-like states, it looks like they are spread out a bit everywhere. Please include a statistical test, if the p-value is not significant, then remove this statement and conclusion from the manuscript, since it would then not be supported by the data. E.g., one could divide the data into 2 categories i) Neuronal-like + Ependymal-like and ii) Neuroepithelial-like and Embryonic-like, based on the plot. Then perform a Chi-squared test.

We thank the reviewer for pointing out the importance of a statistical test supporting our statement that “only ZFTA-RELA cells grown *in vivo* and cocultures are enriched in neuronal-like and ependymal-like states,”. We have included additional scRNA-seq data of the 3 different cell lines in coculture with rat E19 cortical cells (EP1NS, BT165, VBT242) to increase rigor of the study and performed the recommended analysis (dividing the data into 2 categories i) **differentiated** = Neuronal-like + Ependymal-like and ii) **progenitor** = Neuroepithelial-like and Embryonic-like) and performed a Chi-squared test, and the **P-value for**

the Chi-squared test is $P = 8.458^{296}$. (Figure R46). We can therefore conclude that there is a significant relationship between the tumor model (spheres, adherent, coculture, PDX or patient) and the frequency of progenitor (Neuroepithelial-like and Embryonic-like) and differentiated (Neuronal-like + Ependymal-like) states. We have included these results to Extended Data Figure 7e.

Figure R46

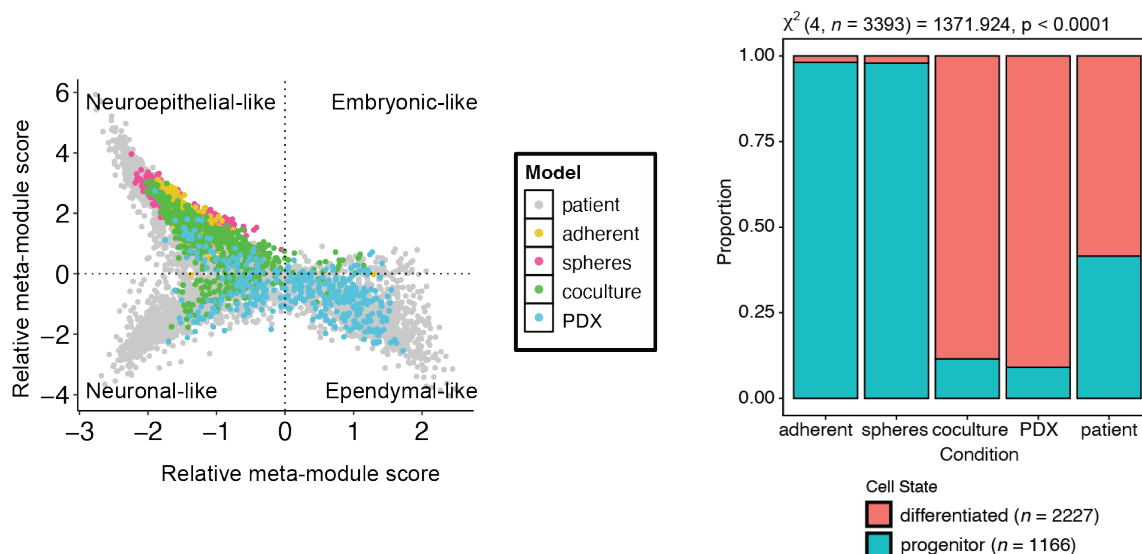


Figure R46. (left) Cell state plot of ZFTA-RELA patient tumor cells profiled by sc/snRNA-seq overlaid with single cells derived from adherent monoculture ($n = 3$), neurosphere monoculture ($n = 3$), coculture with rat E19 cortical cells ($n = 3$), and PDX ($n = 3$) models of EP1NS, BT165 and VBT242. **(right)** Proportion barplot showing distribution of progenitor-like (neuroepithelial-like + embryonic-like) and differentiated (neuronal-like + ependymal-like) cell states across patient tumor and models (adherent monoculture, neurosphere monoculture, coculture with rat E19 cortical cells, and PDX). Chi-square test of independence showing significant association between cell state and patient tumor and models ($p < 0.0001$).

Changes made: Extended Data Figure 7e

#3-29. Line 350-351: The authors state “Overall, these results highlight how the neural microenvironment promotes plasticity towards a neuronal trajectory, affecting the morphology and behavior of ZFTA-RELA cells.”. At this point I would say that the results partly do. E.g., Figure 5j does. But, as mentioned in the comment for lines 345-346, calculations and a statistical test is needed for this to also be a supported statement for Extended Data Figure 10f.

We thank the reviewer for the helpful suggestions to further validate our findings in **Reviewer comments #3-28**. As addressed previously, we performed the suggested statistical test (chi-squared test) and found that the comparison between progenitor and differentiated population is statistically significant (chi-squared test $P = 8.458^{296}$, Figure R46). Therefore, we conclude that the statement “Overall, these results highlight how the neural microenvironment promotes plasticity towards a neuronal trajectory, affecting the morphology and behavior of ZFTA-RELA cells” is supported.

#3-30. Figure 5j-k. Check for normal distribution and then decide on appropriate test to use, explain in material and methods.

We used the Shapiro-wilk test to determine that metaprogram module scores in scRNAseq data of monoculture and coculture were not normally distributed, and we thus used the Wilcoxon rank sum test (Figure R47). As we have amended Figure 5 and Results to include results of monoculture and coculture across 3

different ZFTA-RELA cell lines as also addressed in [Reviewer comment #3-28](#), we have removed this analysis from the manuscript.

Figure R47

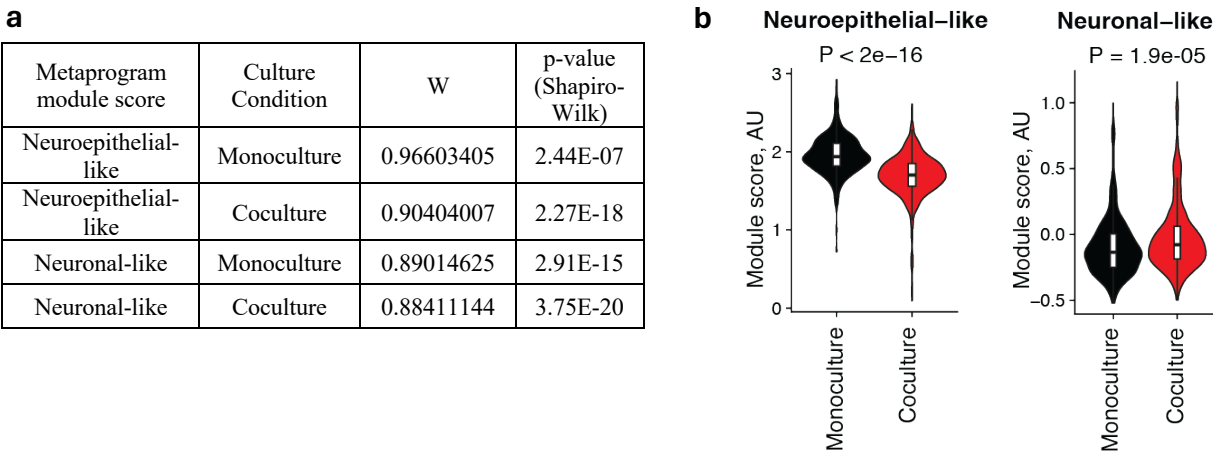


Figure R47. (a) Shapiro-Wilk normality test for distribution of metaprogram module score in monoculture and coculture conditions. **(b)** Violin plots showing metaprogram scores for selected metaprograms in ZFTA-RELA cells cultured in monoculture or coculture. Boxplots denote Tukey’s whiskers (25–75 percentile represented by minima-maxima; statistical median as center). P values calculated by Wilcoxon sum rank test.

#3-31. Figure 1a: For methylation data legend, it seems like there is also a grey color, e.g., for 3 of the ZFTA-RELA samples, cell lines and PDXs. If this means Not Applicable (NA), please add a grey box denoting NA at the bottom for the Methylation score legend.

We revised [Figure 1a](#) to include Not Applicable (NA) values as grey box in the Methylation score legend ([Figure R32](#)).

Figure R32

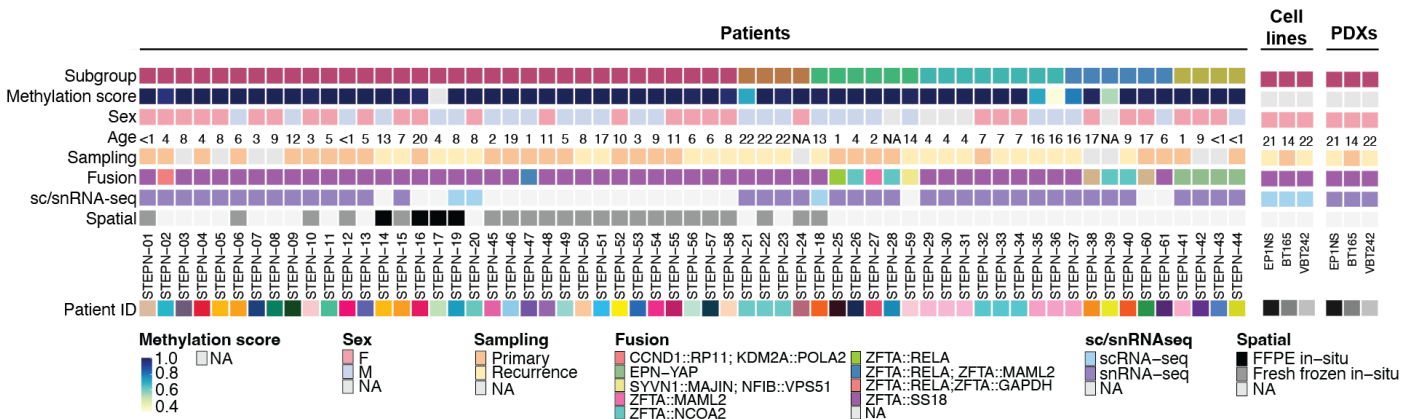


Figure R32. Oncoplot depicting ST-EPN samples profiled in this study.

Changes made: Figure 1a

#3-32. Figure 1e: Please write out the abbreviations in the figure legend for CP, OSVZ and VZ. Also indicate with a text label that the green cells on-top are neurons I believe. Currently only the newborn and maturing neurons are text labeled.

&

#3-33. Figure 1e: Would it be possible to also color code layers within the rectangle of the cortex? The cell type illustration has some background colors that seems to indicate different layers.

We revised the figure legends for **Figure 1e (Figure R48)** to include all the abbreviations present in the figure including for CP, OSVZ, and VZ. In addition, we included neurons to the text labels, and color coded the layers within the rectangle of the cortex. Finally, we removed “Choroid Plexus” and “MGE Inhibitory Neuron Lineage” annotations as these are not directly relevant regions for the manuscript and amended the annotations for developmental timepoints on the bottom part of the schematic to match the developmental timepoints used in the tumor to development projection in **Figure 1f**.

Figure R48

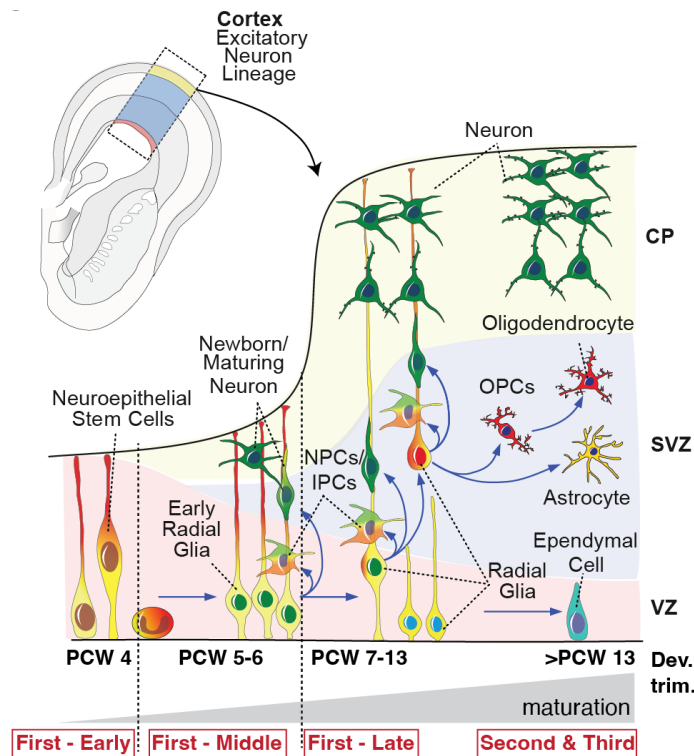


Figure R48. Top part: Simplified schematic of cell types present during human cortical development. Dev. Trim; developmental trimester; PCW, post conception week; IPCs, intermediate progenitor cells; NPCs, neural progenitor cells; MGE, medial ganglionic eminence; OPC, oligodendrocyte precursor cells; CP, Cortical plate; SVZ, subventricular zone; VZ, ventricular zone. Bottom part: Abundance of cell types (x-axis) present during the first, second and third trimester of human cortical development (y-axis). Data reanalyzed from two published scRNA-seq datasets of the human developing cortex^{1,4}

Figure 1e legends: (e) Simplified schematic of cell types present during human cortical development. Dev. Trim; developmental trimester; PCW, post conception week; IPCs, intermediate progenitor cells; NPCs, neural progenitor cells; MGE, medial ganglionic eminence; OPC, oligodendrocyte precursor cells; CP, Cortical plate; SVZ, subventricular zone; VZ, ventricular zone. Bottom part: Abundance of cell types (x-axis) present during the first, second and third trimester of human cortical development (y-axis). Data reanalyzed from two published scRNA-seq datasets of the human developing cortex^{1,4}

Changes made: Figure 1e, Figure 1e legends

#3-34. Figure 1f and Extended data figure 2a: I found the different expression score labels to be a bit confusing. Although it is explained in the figure legend, I would find it helpful with an addition to the labels. E.g., instead of writing (tumor cells), write (normal profile in tumor cells), to indicate that normal cell profiles have been

projected in tumor cell data. For (normal cells), it would be (tumor profile in normal cells). I've seen this somewhere also in the Extended Data, would be great to have it updated there to.

We have changed the labels of **Figure 1e (Figure R1)** and **Extended data Figure 2b (Figure R49)**.

Figure R1

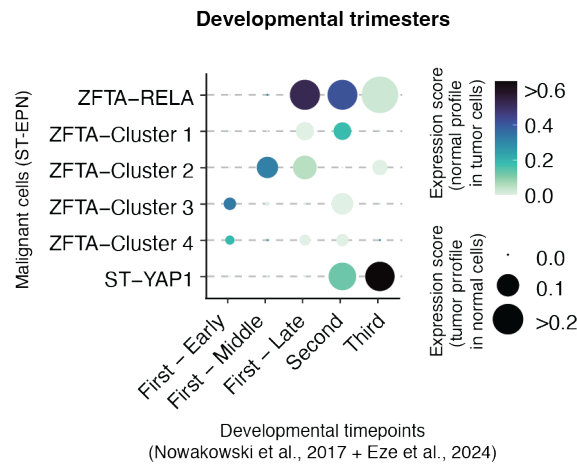


Figure R1. Projections of ST-EPN subgroups (y-axis) onto developmental time points found across the developing human cortex (x-axis). First-Early: post-conception week (PCW) 4, First-Middle: PCW 5-6, First-Late: PCW 7-13, Second: PCW 14-27, Third: PCW 28-37. PCW, post-conceptional week; dev. trim, developmental trimesters.

Figure R49

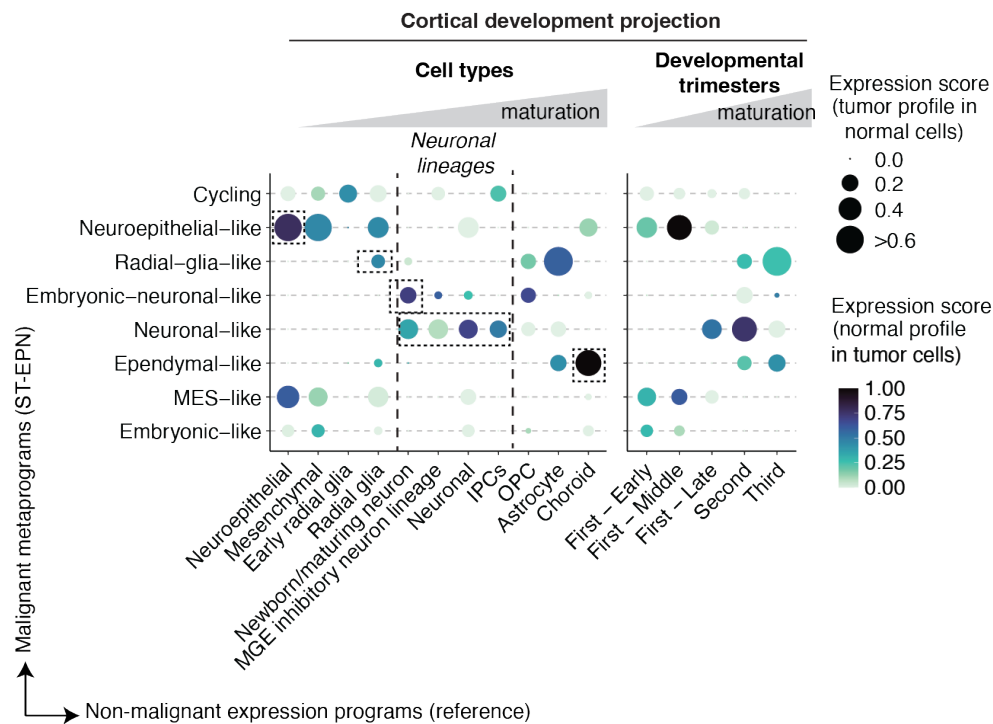


Figure R49. Projections of each ST-EPN metaprogram (y-axis) onto cell types present during the first, second and third trimester of human cortical development. First-Early: post-conception week (PCW) 4, First-Middle: PCW 5-6, First-Late: PCW 7-13. Color scale depicts expression score of normal cell type gene expression in tumor cells, and dot size represents expression score of tumor cell type gene expression in normal cells.

Figure 2e: Just wanted to say that this representation looks great and makes it a lot easier to interpret the cell proportions from Figure 2d!

We thank the reviewer for this kind comment.

#3-35. Extended Data Figure 1d: In general, it is difficult to see the placement of all the cell types in these UMAPs. Perhaps using different colors and/or transparency, or dividing the visualization into more plots? E.g., The third developmental trimester is not at all visible.

We agree with the reviewer that it is hard to see placement of all cell types in the previous version of this plot. One of the major limitations is that certain cell types that are present exclusively in Nowakowski et al. 2017 are present at low numbers compared to the populations of Eze et al. 2024. We subset each cell type to a max number of $n = 500$ cells (when available; if not, the total number of cells for that specific type was used). We have revised the **Extended Data Figure 1f** to aid with visibility of cell types in the UMAPs. We have changed the transparency and highlighted certain cell types by dotted circles.

Figure R15

b

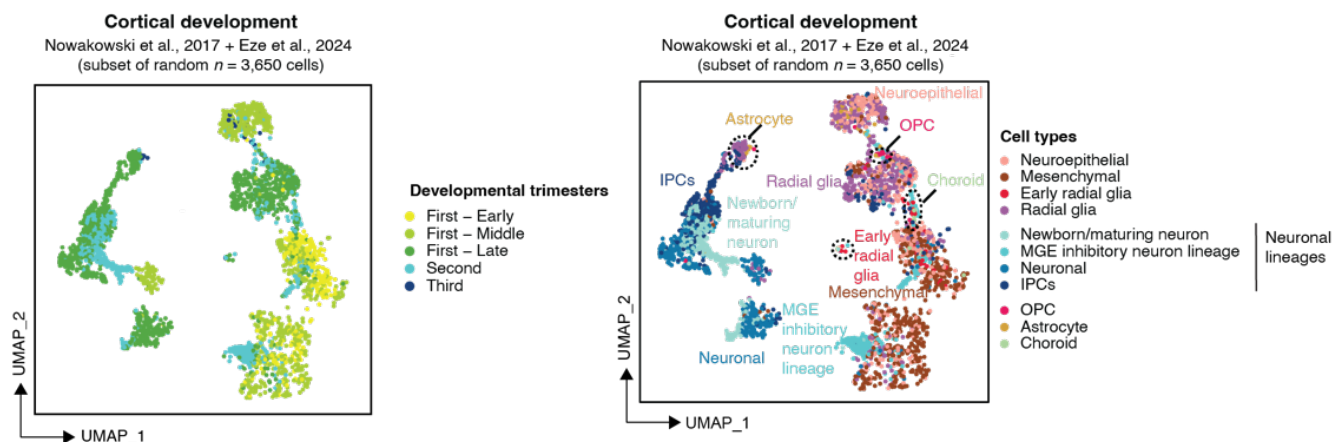


Figure R15. (b) UMAP plot of cell types present during the first, second and third trimester of human cortical development [refs], colored by developmental trimester (left) and cell types (right). Original publications listed for each plot.

#3-36. Extended Data Figure 4a: I find the labeling a bit confusing. Which is the case and control? From the y-axis labeling I would assume that ZFTA-RELA is the case and ZFTA-Cluster 3 is the control. But in the plot the upregulated genes are shown for ZFTA-Cluster 3, I would expect the plot to show upregulated genes for the case. I suggest doing something similar to what was done by the authors in Extended Data Figure 10a instead.

We thank the reviewer for pointing out this lack of clarity in the figure. We have revised the plot to show the results of differential gene expression analysis between ZFTA-Cluster 3 (on the left side, same color scheme as in the manuscript) and ZFTA-RELA tumors (displayed on the right, with the same color scheme as in the manuscript). (**Figure R30, Extended Data Figure 3b-d**).

Figure R30

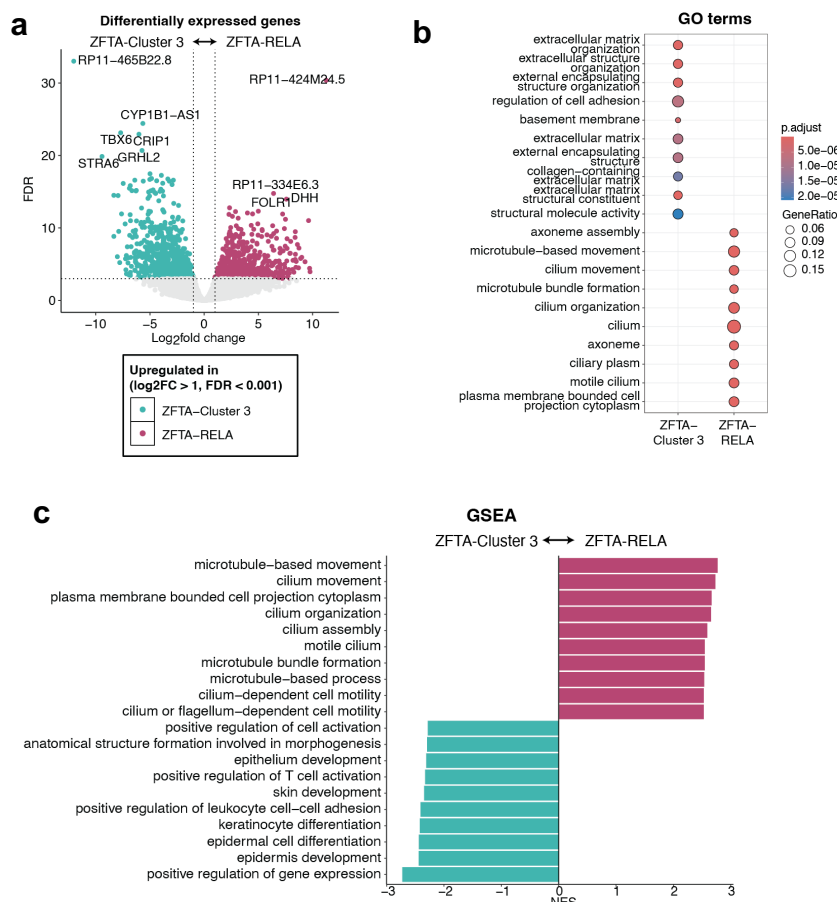


Figure R30. (a) Volcano plot showing genes differentially expressed between ZFTA-RELA and ZFTA-Cluster 3 tumors. Cells were pseudobulked by sample, and differential gene expression analysis was run using *edgeR*. Genes with $|\log_2FC| > 1$ and $FDR < 0.001$ are highlighted in color. **(b and c)** Top 10 gene ontology (GO) terms **(b)** and pathways identified by GSEA analysis **(c)** comparing genes differentially expressed (input for GO: $|\log_2FC| > 1$ and $FDR < 0.001$, input for GSEA: $FDR < 0.001$) between ZFTA-RELA and ZFTA-Cluster 3 tumors based on the genes identified in **(a)**.

#3-37. Figure 4 and Extended Data Figure 7: Please use a different color scale for the heatmaps, it is difficult to interpret when the scale is from 0-60%, but the color range go from blue-white-red. Use e.g., white-yellow-red, something that steadily increases. The blue-white-red is appropriate when looking at number ranges e.g., between -1 to 1. Furthermore, for these heatmaps, please include all cells and states in the color legend. I suggest having two color legends, keep the one with niches A-E but add one including all cells and cell types, e.g., more annotations are shown in the pie charts. Extended Data Figure 10a: This looks really nice, something similar could be done for easier interpretation of Extended Data Figure 4a.

We thank the reviewer for the detailed feedback. As also addressed in [Reviewer comment #3-20~21](#), we amended the metaniche analysis with CellCharter, and as a result, we had more figures to include in the main and extended data figure. As the niche heatmap only provides information regarding one section, we have removed this plot from the manuscript.

#3-38. Figure 5h: The illustration showcases monocultures and cocultures as if they are performed in the same Petri dish with some form of separation wall. However, from reading the manuscript this does not appear to be the case. Please revise the image to reflect the experimental design.

We have revised the schematic to reflect the experimental design (**Figure R50 left**). In addition, because we have included scRNA-seq data of more cell lines in coculture as also addressed in **Reviewer comment #3-28**, we have removed this schematic from the manuscript and replaced it with a schematic outlining the different patient-derived models analyzed by scRNA-seq (**Figure R50 right**).

Figure R50

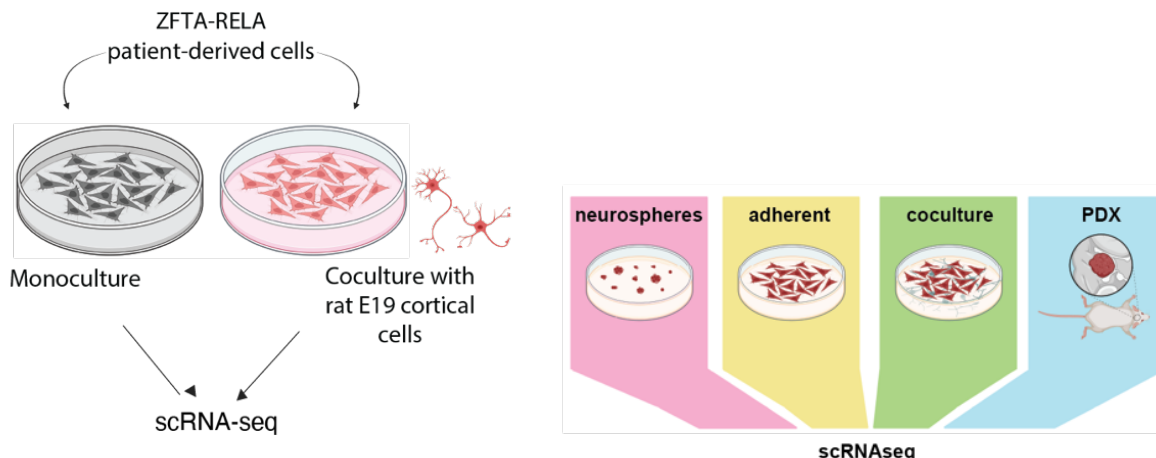


Figure R50. (left) Experimental design to test the compositional changes of ZFTA-RELA cells in monoculture versus coculture with rat E19 cortical cells. **(right)** Graphical summary of ZFTA-RELA patient tumor cells profiled by sc/snRNA-seq derived from neurosphere ($n = 3$), adherent ($n = 3$), coculture with rat E19 cortical cells ($n = 3$) and PDX ($n = 3$) models of EP1NS, BT165 and VBT242.

Changes made: Figure 5h

Methods

#3-39. I'm unable to find anything in the methods related to the methylation analysis in Figure 1b. Please add experimental protocol procedures and data analysis.

We have modified **Methods** to include the experimental protocol and data analysis methods for methylation classification of tumors.

Methods: Molecular classification of tumors. Molecular classification via DNA-methylation profiling was performed using the Heidelberg Brain Tumor Methylation Classifier v12.4 (www.molecularneuropathology.org) as previously described⁶⁴. Genome-wide DNA methylation profiling was performed using the Infinium HumanMethylation 450k or EPIC Kit according to the manufacturer's instructions (Illumina, San Diego, CA, USA). Raw signal intensities from the resulting idat-files were calculated using the minfi Bioconductor package v.1.24.0⁶⁵. The normalization and batch effect correction were performed using the limma package v3.34.5⁶⁶. Afterwards, beta-values were calculated before selecting CpG-probes for downstream analysis. Unsupervised nonlinear dimension reduction was performed by using the 1-variance weighted Pearson correlation between the samples to construct the distant matrix, which was subsequently used as input for a t-SNE visualization. Classification was performed by using the established random forest algorithm of the Heidelberg Brain Tumor Methylation Classifier v12.4, which included all molecular EPN groups, including the newly established ST-EPN clusters derived in previous studies^{23,26,64}.

#3-40. I'm unable to find any information regarding how the samples fusion gene statuses were determined, shown in Figure 1a. If this was provided from the hospital as metadata, I suggest stating which type of metadata was provided by the hospital in the material and methods section regarding the samples.

We have modified the **Methods** to include how sample gene fusions were determined.

Methods: Detection of fusion genes in tumors. Gene fusion status for most cases was derived from earlier studies^{23,26} using either RNA-sequencing or RT-PCR as described in the respective studies. The gene fusions and their respective isoforms were derived using two independent methods, InFusion and Arriba^{67,68}.

#3-41. Line 533: The authors have included ethical approval information for the human subjects, but not for the animal studies. This information appears to be included in the Reporting Summary document, please also include this information in the main manuscript ethical approval section.

We have included the animal studies in **Methods**.

Methods: In vivo imaging of ZFTA-RELA patient derived xenografts. For all *in vivo* imaging experiments, male NMRI nude mice were used. All animal procedures were performed in accordance with the European Directive on animal experimentation (2010/63/EU) and institutional laboratory animal research guidelines after approval of the Regierungspräsidium Karlsruhe, Germany. Efforts were made to minimize animal suffering and to reduce the number of animals used according to the 3R principles. Mice were clinically scored, and experiments were terminated if they showed marked neurological symptoms or weight loss exceeding 20%.

#3-42. Line 637: The authors mention the use of the InferCNV R package to identify malignant and non-malignant cells. However, they do not state how InferCNV was run or if a reference was used. This is important to state in the methods, e.g., if no reference is used then amplification and deletions that are broadly present in the dataset will be missed.

We have included additional details to the identification of malignant and non-malignant cells in **Methods**.

Methods: Identification of malignant and non-malignant cells. To discriminate between malignant and non-malignant cells, we used: (1) Seurat clustering, (2) unbiased cell type annotation using the automated annotation package SingleR (v1.6.1) using the Human Primary Cell Atlas⁴¹ data as a reference, and (3) inference of copy-number variation using the InferCNV R package (v1.8.0). We began by performing an initial annotation of cell lineages through dimensionality reduction using UMAP and unsupervised Louvain clustering. To refine our classifications, we analyzed clustering patterns and the expression of established immune, endothelial, and oligodendrocyte gene markers. We validated these initial lineage assignments using the automated annotation package SingleR (v1.6.1), which assigns cell identities based on their similarity to a reference set of bulk transcriptomes from healthy cells (Human Primary Cell Atlas data⁴¹). We then inferred genome-wide evidence for somatic chromosomal copy number variation (CNVs) using the *inferCNV* R package as previously described^{14,39,42–44}. We spiked-in non-malignant immune and oligodendrocyte nuclei from previous publications^{14,39,44} as reference normal karyotype for CNV inference, and used hierarchical clustering of the single-cell copy-number profiles within each sample with the reference non-malignant copy-number profiles to classify the presence of CNVs. Due to presence of false positive/negative rates of inferring CNVs with nuc-seq data from frozen samples, CNV results were used as secondary validations. CNVs should be detected in cell cluster classified as malignant cells by transcriptional clustering. In contrast, CNVs should be absent from the cell cluster classified a normal cell type. Reassuringly, we observed that nuclei lacking CNVs typically clustered together and were identified as macrophage, T cell, or endothelial cell nuclei based on their high expression of myeloid, T cell, or endothelial markers, respectively. All other clusters consistently exhibited CNVs and were classified as malignant tumor cells. Only cells identified as malignant were selected for subsequent NMF analysis.

#3-43. Lines 684-688: The authors describe ‘Correlation to reference datasets’, which in the main text (line 133-135) is described and referred to as a way to project the gene expression from the different subgroups onto reference cell atlases. It was not immediately clear to me as a reader that this part of the methods corresponds to this part of the results. I suggest some clarification in the methods description to help the reader navigate. Furthermore, I want to see a more detailed description of how this was done. Currently, it is explained that marker genes are found for each cell types in the reference datasets, and “within ST-EPN molecular subgroups or cellular states”, however in Figure 2d only cellular states seems to be presented. How was the ST-EPN molecular subgroups score used? Then the authors state “We then scored all ST-EPN malignant cells with marker genes identified across the developing human cortex, and other way around”, however, there is no mention of what score cut-off values or ranking were considered for the annotation itself, please add this to the methods section.

We have amended **Methods** to more clearly explain our methods.

Methods: Generation of developmental signatures. To identify developmental signatures enriched in ST-EPN subtypes or cell states, we performed projection of tumor cells (pseudo-bulked by ST-EPN subtype or cell state) onto normal cell type signatures. We first assembled a single-cell reference atlas of the developing human cerebral cortex from two published scRNA-seq datasets^{1,4}. We then subset the combined dataset to 500 cells (if available) per cell identity and used Seurat's *RunPrestoAll* function (keeping only genes expressed in at least 25% of all cells) to derive the top differentially expressed genes (adjusted *P*-value < 0.05) across cell types, which are reflective of developmental-specific cell type signatures. We then scored all ST-EPN malignant cells with these marker genes and annotated ST-EPN cells based on the maximum expression score for the respective signatures.

#3-44. Please add a methods part describing the correlation between Gojo et al 2020. It is mentioned in the text (line 170-171), but I can't find a description in the methods part. E.g., from what values were these correlations calculated? Were the top 200 genes from the metaprograms used to calculate a score (how was that score calculated) and then how many top genes for the different gene programs from Gojo et al 2020 was used to calculate scores that were then used in the correlation analysis?

We thank the reviewer for pointing out this lack of clarity in the description of our approach. From our study, we selected the top 30 genes for each metaprogram. We have included this in **Extended Data Figure 2d legends**. From Gojo et al. 2020, we used the published metaprogram genes (Table S2), which include the top 30 genes for all metaprograms, except for the following programs (neuronal-precursor-like, S-phase, G2M-phase, ependymal-like) that have 60 genes. We also amended **Methods**.

Extended Data Figure 2d legends: (d) Pairwise Pearson correlation between ST-EPN metaprograms (using the top 30 NMF marker genes) identified in this study in frozen tumor samples (red font) and previously defined metaprograms¹⁴ identified across different types of ependymoma (SP: spinal; ST: supratentorial; PF: posterior fossa).

Methods: Scoring gene sets. We first aggregated ST-EPN single cells/nuclei from all samples, and performed normalization and centering. Then, we scored these cells for NMF/developmental signatures as previously described^{14,39,42-44}. We did this by calculating for each cell *i* a score $SC_j(i)$, quantifying the average relative expression (*Er*) of the top 30 genes within each MP (*G_j*) and comparing the score to the average relative expression of a control gene (*G_j^{cont}*): $SC_j(i) = \text{average}[Er(G_{j,i})] - \text{average}[Er(G_{j^{cont}}^{cont}, i)]$. The control gene-set was defined by binning all genes into 30 bins of aggregate expression levels (*Ea*) and randomly selecting 100 genes from the same expression bin for each gene in the gene-set *G_j*. ST-EPN cells/nuclei were annotated based on the maximum expression score for the respective MP signatures after excluding the one related to cell cycle ("cycling"). **Program-wise hierarchical clustering.** To compare sets of NMF programs or MPs, we first scored all ST-EPN cells/nuclei for the program signatures as described in section: *Scoring gene sets*. After scoring, we performed pairwise correlation for each possible pair of programs and show hierarchical clustering of the scores using a correlation-based distance metric (distance metric: 1 – Pearson correlation; linkage: Ward's linkage).

Changes made: Extended Data Figure 2d legends, Methods

#3-45. Please add differential gene expression analysis and GO enrichment analysis shown in Extended Data Figure 4a to the methods section.

We have included this information to **Methods**.

Methods: Differential gene expression analysis. Single-cell. To identify genes that are differentially expressed between two conditions (monoculture vs coculture) or cell states (neuroepithelial-like vs embryonic-like), we first used Seurat's *RunPrestoAll* function (only genes minimally detected in 15% cells). For downstream gene ontology (GO) analysis we used the top 500 genes (ordered by average log₂FC) with an average |log₂FC| > 1 and adjusted P-value threshold as indicated in the corresponding figure legend, for gene set enrichment analysis (GSEA), we used all genes with an adjusted P-value threshold of 1⁻⁰⁵. **Pseudobulk.** We performed pseudo-bulk differential expression analysis to identify transcriptional programs that distinguish ZFTA-RELA from ZFTA-Cluster 3 tumors. We first extracted raw count matrices from each tumor sample, and calculated pseudo-bulked gene expression counts using the SingleCellExperiment's *aggregateAcrossCells* function. We then conducted differential gene expression analysis using edgeR. We filtered out genes with low expression counts using the *filterByExpr* function, based on our grouping variable (tumor subtype). We then normalized the counts using the *calcNormFactors* function. We constructed a design matrix to include tumor subtype, estimated negative binomial dispersions using the *estimateDisp* function, and fitted a generalized linear model using *glmFit*, followed by hypothesis testing with *glmLRT* for differential expression. We considered significant genes

based on a false discovery rate (FDR) threshold of <0.001. For downstream gene ontology (GO) analysis we used the top 500 genes (ordered by $\log_2FC$) with $FDR < 0.001$ and $|\log_2FC| > 1$, for gene set enrichment analysis (GSEA), we used all genes with $FDR < 0.001$. **Gene Ontology (GO) and Gene Set Enrichment (GSEA) analysis.** To identify biological terms that are over-represented in a specific gene set, we performed GO and GSEA analysis using *enrichGO*, respectively *gseGO* functions from the *clusterProfiler*¹³ package. As a database, we used the R package *org.Hs.eg.db* (v. 3.18.0), which is an organism annotation package for *Homo sapiens*, relying on the Ensembl version released on 2023-May10. Unless otherwise states, we used the following parameters: *ont* = "ALL", *universe* = all genes retained for differential gene expression analysis, *minGSSize* = 3, *maxGSSize* = 800, *pvalueCutoff* = 0.05, *OrgDb* = *org.Hs.eg.db*, *pAdjustMethod* = "BH".

#3-46. In general, for all GO enrichment analysis, please add information regarding version control, these databases update regularly and using the same list of genes for a search in a few months might yield slightly different results. E.g., if using g:profiler for GO one can specify which Ensembl version and Ensembl Genomes version the database was built on (<https://biit.cs.ut.ee/gprofiler/page/archives>). Does clusterProfiler used here have something similar? Please also indicate which settings the enrichGO function was run under, e.g., if there were some non-default settings. Also, regardless mention the p-value cut-off and multiple testing method selected even if default was used. Also mentioned if ranked or non-ranked gene lists were used or a possible choice with the enrichGO function search.

We thank the reviewer for the helpful suggestion. We used the library *org.Hs.eg.db* as annotation database, which is based on ENSOURCEDATE of 2023-May-10. While the package does not explicitly state the Ensembl release number, based on the Ensource date, it's likely that the package utilizes data from Ensembl release 109, which was current around that date. We have added the specific settings, P-value cut-off and multiple testing method selected under which we ran enrichGO in **Methods**.

Methods: Gene Ontology (GO) and Gene Set Enrichment (GSEA) analysis. To identify biological terms that are over-represented in a specific gene set, we performed GO and GSEA analysis using *enrichGO*, respectively *gseGO* functions from the *clusterProfiler*¹³ package. As a database, we used the R package *org.Hs.eg.db* (v. 3.18.0), which is an organism annotation package for *Homo sapiens*, relying on the Ensembl version released on 2023-May10. Unless otherwise states, we used the following parameters: *ont* = "ALL", *universe* = all genes retained for differential gene expression analysis, *minGSSize* = 3, *maxGSSize* = 800, *pvalueCutoff* = 0.05, *OrgDb* = *org.Hs.eg.db*, *pAdjustMethod* = "BH".

#3-47. Line 740 Metaprogram cell assignment. Which figure and data is this paired with? As a reader I'm confused, is this for Figure 4 a subpanel 'Cell types' from the Xenium spatial data? Please clarify.

We revised **Methods** to clarify that the Methods section refers to "Xenium cell state/type cell assignment".

Methods: Xenium cell state/type cell assignment. To assign cell identities in Xenium data, we first performed unsupervised clustering using *Seurat* (v5.1.0), performing *RunUMAP* on the least number of PCs capturing 80% or more variance, shared-nearest-neighbor graph using *FindNeighbors*, and Louvain clustering (*FindClusters*) run at resolutions 0.1 to 1.0, in 0.1 increments. We then scored only for the normal cell types, by transferring labels from the single cell/nuc reference object from corresponding tumor subgroup using *FindTransferAnchors* and *TransferData* (dims = 20), yielding a label-transfer score for every cell-to-cell type pair. Independent module scores were calculated using *AddModuleScore_UCell* using marker genes present in the Xenium panel. For each cluster, average label-transfer score and module score were manually inspected and visualized through UMAP, and the clusters corresponding to normal cell types were identified in each sample. Cells belonging to the normal cell type annotation was removed, and malignant subset of each sample was annotated in the same way as normal cell types. To mitigate systemic score inflation, scores for each malignant cell type were mean centered across all cells: $centered_score_{ij} = score_{ij} - \overline{score}_{.j}$, where *i* corresponds to cells and *j* corresponds to cell types. Finally, each malignant cell was assigned with the cell type with highest mean centered score, and malignant and normal cells were combined to generate the final, fully annotated cells.

#3-48. Line 764 Spatial niche. After reading the paper and this methods part I'm still unsure as a reader how the different niches are defined. I think it is because the authors both refer to 'spatial niches' which are mentioned as niche 1, 2, 3, and 4. Then the authors define recurrent niches A, B, C, D and E.

We included more detailed explanation regarding how spatial niche and metaniches (“recurrent niches”) were identified **Methods**. In addition, we have assigned numerical labels (1–6) to metaniches (recurrent niches). We also included additional description regarding methods to derive CellCharter spatial clusters in **Methods**.

Methods: Spatial niche and metaniche (MN) identification. Spatial niche analysis was performed for each tumor section in R using the *BuildNicheAssay*’s function from Seurat (v. 5), that demarcates regions of tissue defined by a similar composition of spatially adjacent cell types. This function, which requires a pre-established number of neighbors (*neighbors.k* set to 20) and expected niches (*niches.k* set to 6), considers the $n = \text{neighbors.k}$ spatially closest neighbors to each cell, counts the occurrences of each cell state/type in this area, and uses k-mean clustering to group cells with similar neighborhoods together to resolve $n = \text{niches.k}$ spatial niches for each tumor section. We applied *BuildNicheAssay* to our spatial cohort using *niches.k* = 4-6, and thereby resolved 4, 5 and respectively 6 niches within each tumor section. To identify niches that were consistently detected across multiple tumor sections [termed metaniches (MNs)], we first calculated the relative proportion of each cell state/type within a given niche-section. We then computed correlation matrix between the different niches by hierarchical clustering as explained in section: Program-wise hierarchical clustering, to understand how similar the distributions of cell states are across different spatial niches. Finally, we used the function *cutree* from the ‘stast’ R-package to automatically extract $n = 5-6$ highly correlated MNs.

Spatial clusters identification by CellCharter. To simultaneously uncover local spatial patterns across multiple samples and provide an orthogonal validation to our niche analysis, we used CellCharter⁵⁸ to identify spatial clusters. To enable joint analysis among different samples, CellCharter applied dimension reduction and batch correction through a variational autoencoder method, scVI. Next CellCharter encoded spatial information into a spatial network based on spatial proximity among cells. To enable neighborhood aggregation, CellCharter concatenated features from neighbor cells. We then performed Gaussian mixture model (GMM) based clustering on these concatenated and aggregated features to determine spatial clusters. We ran GMMs on a wide range of cluster numbers (N) to identify stable cluster number for final GMM and downstream analysis. More specifically, N clusters were deemed stable if cluster assignments are concordant across multiple runs of GMMs for N, N-1, and N+1 based on the Fowlkes-Mallows Index (FMI).

#3-49. Lines 354-356: The authors claim that “Our work uncovers different developmental signatures as well as malignant cell states across ST-EPN subgroups including previously uncharacterized non-canonical ST-ZFTA tumors, and demonstrates that these cell states have distinct spatial, morphological, and behavioral characteristics.” However, this framing is misleading, because Xenium spatial data analysis and cell experiments were only performed in canonical ST-ZFTA tumors. Please reframe and clarify.

We have rephrased this claim in the discussion to clarify that the spatial transcriptomics and live-cell imaging were only done for ZFTA-RELA (canonical) tumors.

Discussion: Our work uncovers different malignant cell states as well as developmental signatures across molecular subgroups of ST-EPN tumors, including previously uncharacterized non-canonical ST-ZFTA tumors¹³. In addition, we demonstrate that individual cell states in ZFTA-RELA tumors have distinct spatial, morphological, and behavioral characteristics.

#3-50. Lines 382-383: The authors state “In addition, the strong ependymal signatures in ST-YAP1 corroborates previous work demonstrating ciliogenic marker expression in these tumors^{14,55}”, referencing previous work. But the dataset presented in this manuscript is new, please check for high ciliogenic marker expression in this ST-YAP1 cohort data. E.g., add a plot in the Extended Data.

We thank the reviewer for this comment. ST-YAP1 samples in this cohort mainly express the Ependymal-like metaprogram. The top 30 genes of the Ependymal-like metaprogram are significantly enriched in members of the GO-term “cilium movement” (p-value = 7.614×10^{-7} , Hypergeometric test), including *DNAH9*, *DNAH11*, *PACRG*, *ARMC3*, *DNAH6*, *DAW1*, *DNAH7*, *EFHC1*, *NEK10* (**Figure R51**). We highlight *DNAH6* and *DNAH9* in **Figure 2b** of the manuscript, and concurrently removed this sentence from the manuscript to make the discussion section more succinct.

Figure R51

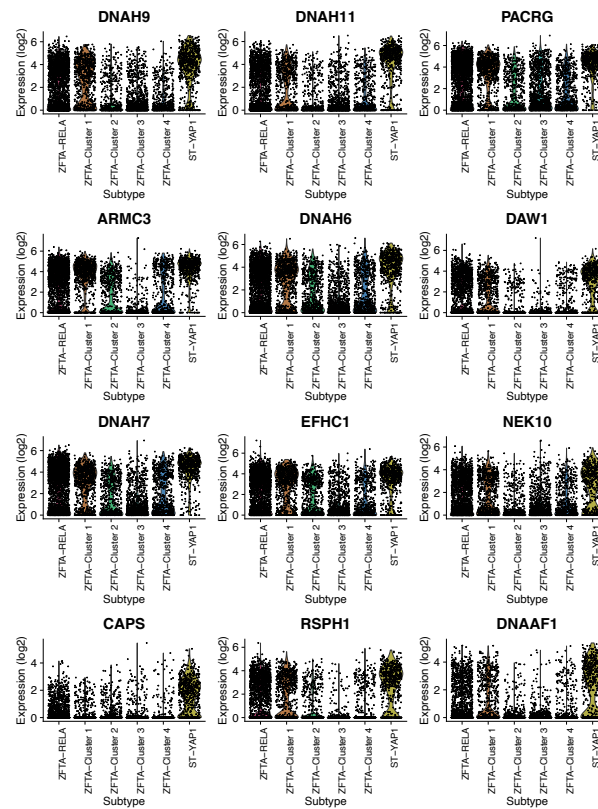


Figure R51. Violin plots of ciliogenic gene expression across ST-EPN subgroup patient tumors.

#3-51. Lines 396-397: Part about relapses was not supported by the data, please remove.

As also addressed in rebuttal for **Reviewer comment #3-19**, we removed this statement from the manuscript and relevant statements regarding the significance of comparison between primary versus recurrent tumors.

#3-52. Please include a section in the discussion bringing up limitations of the study. Some suggestions for this are:

- tumor global heterogeneity and sample site bias.
- Too few primary/relapse data points to establish if relapses generally are more structured than primary tumors.
- That spatial data was only generated on the canonical subtype and not for the non-canonical.
- That the ST-SE subgroup was not included in the study (authors ref. 13).
- That mesenchymal and hypoxia cell state signatures are not direct measurements of e.g., hypoxia and acts as a proxy.

We thank the reviewer for this helpful suggestion and have included limitations of the study addressing several points including the reviewer's comments.

Limitations of the Study: **Limitations of the study.** In this study, we characterized tumor heterogeneity of patient samples diagnosed as supratentorial ependymomas. However, there may be a sample site bias in our data set due to random sampling of the whole tumor mass during surgery as well as dividing the taken tumor into pieces for sc/snRNAseq and spatial transcriptomics. Therefore, our findings are limited due to the sampling of tumors, and do not have a conclusive claim of the entire tumor mass present in patients. In addition, due to limited availability of non-canonical ZFTA tumors and corresponding patient-derived cell lines, we have not included spatial transcriptomics and other functional live-cell imaging data of these tumor subgroups. We also did not include ST-EPN-SE subgroup tumor samples in our analyses. Finally, the mesenchymal and hypoxia cell state signatures are not

a direct measurement of hypoxia itself, and rather act as proxy for potential effects or consequences of local and global hypoxic microenvironments.

D. Appropriate use of statistics and treatment of uncertainties

#3-53. General comment on statistics. Please include a section of this in the material and methods part. All distributions tested needs to first be assessed if they are normally distributed (e.g., using a Shapiro-Wilk test). Then select appropriate parametric or non-parametric tests. Make clear how multiple testing has been considered and performed. For test with few p-values, it is most robust to use Bonferroni correction.

We have assessed the normality of all our data and performed downstream statistical comparison tests accordingly. This information is now included in **Methods**.

Methods: Statistical Analysis. All statistical analyses were conducted using R (version 4.3.3). To determine statistical significance, we first assessed data normality using the Shapiro-Wilk test (shapiro.test). If all groups followed a normal distribution (Shapiro-Wilk P-value > 0.05), we performed Student's t-test. If one or more groups did not follow a normal distribution (Shapiro-Wilk P-value < 0.05), we used the non-parametric Kruskal-Wallis H test for comparisons involving more than two groups, followed by the Wilcoxon rank sum test. Multiple comparisons were corrected using the Bonferroni test as appropriate. Only statistically significant P-values are reported. Data points are presented as mean $\pm$ SD or SEM, depending on the experiment, as specified in the figure legends. Boxplots follow the Tukey method, with bars extending from the 25th to 75th percentiles and the center bar indicating the statistical median.

#3-54. Please add the statistical test used in Extended Data Figure 4c to the methods part (line 695). Also, test if the data is normal distributed, which is an assumption required for the use of the student's t-test. If not normally distributed, please re-calculate using appropriate statistical test, e.g., Wilcoxon rank sum test. Also, because 3 tests are made, use multiple testing correction of the 3 p-values, describe in the methods section (and figure legend) which multiple correction method was used.

We thank the reviewer for this constructive feedback. According to the reviewer's recommendations, we have:

- Checked if the data are normally distributed using the Shapiro-Wilk Test. Neuroepithelial-like and neuronal-like scores are normally distributed (Shapiro P-value > 0.05), whereas ependymal-like scores in ZFTA-RELA do not follow a normal distribution. Therefore, to keep uniformity, we proceeded with a Wilcoxon rank sum test to compare the scores for the cell types across ST-EPN subtypes. (**Figure R52**).
- We applied multiple testing Bonferroni correction, given that we made 3 tests.
- Our revised figure (**Extended Data Figure 4h**) now includes information on statistical tests in the legend.

Figure R52

Score	Subgroup	Shapiro W	Shapiro P	Comments
Neuroepithelial-like	ST-YAP1	0.88	0.155	Normally distributed
Neuroepithelial-like	ZFTA-RELA	0.955	0.125	Normally distributed
Neuronal-like	ST-YAP1	0.94	0.577	Normally distributed
Neuronal-like	ZFTA-RELA	0.979	0.676	Normally distributed
Ependymal-like	ST-YAP1	0.984	0.98	Normally distributed
Ependymal-like	ZFTA-RELA	0.935	0.0265	Not normally distributed

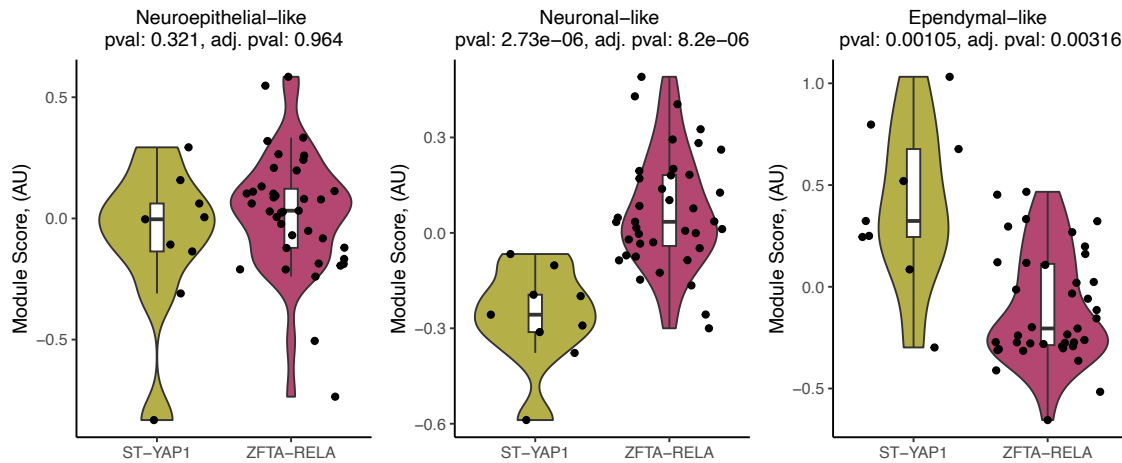


Figure R52. (top) Summary of results from Shapiro-Wilk normality test of gene expression module scores in ZFTA-RELA and ST-YAP1 samples profiled by bulkRNA-seq in an external cohort¹³. **(bottom)** Box plots showing gene expression module scores of cell state signatures across ST-EPN samples profiled by bulkRNA-seq in an external cohort¹³. Bonferroni-corrected significant P-values calculated by Wilcoxon-rank sum test are shown. AU, arbitrary unit.

#3-55. Figure 3c: Linear regression between structural organization and vascular cells is not significant, but results are interpreted as if they were. This needs to be removed from the manuscript in the places it occurs.

As also addressed in [Reviewer comment #3-17~18](#), as this correlation is not significant, we have removed related statements in the manuscript.

#3-56. Extended Data Figure 6f-g: The tests between primary and relapses are not significant, but results are interpreted as if they were. This needs to be removed from the manuscript in the places it occurs.

As also addressed in [Reviewer comment #3-19](#), we removed this statement from the manuscript and relevant statements regarding the significance of comparison between primary versus recurrent tumors.

#3-57. Please add for Extended Data Figure 6a how R2 and the p-value was calculated in the methods section. Also explain the rationale for why “For the smart-seq2 dataset, neuronal-like and embryonic-neuronal like metaprograms were combined (“neuronal-like”), as well as neuroepithelial-like and embryonic-like (“neuroepithelial-like).” i.e., motivate/explain more clearly why combine and not keep them separate. This division and grouping of metaprograms does not appear to make sense, especially since the authors in the main text state, lines 184-187, “To summarize the distribution of metaprograms in each ST-EPN subgroup, we scored all single cells for the two most immature (neuroepithelial-like and embryonic-like) and the two more mature (neuronal-like and ependymal-like) lineage-related expression programs and visualized the results as a cell state plot (Figure 2e).”

We thank the reviewer for this helpful feedback. Firstly, we modeled the relationship between the two variables using the *lm* function from the R *stats* package. We extracted the goodness of fit (R, R² and P value) using the *summary()* function. We have added this information to **Methods**.

Secondly, we now show the correlation between each cell state/type identified by Smart-seq2 and by Xenium analysis of the new spatial cohort, including embryonic-neuronal-like and embryonic-like (**Figure R53**) and have included these results in **Extended Data Figure 4b**.

Figure R53

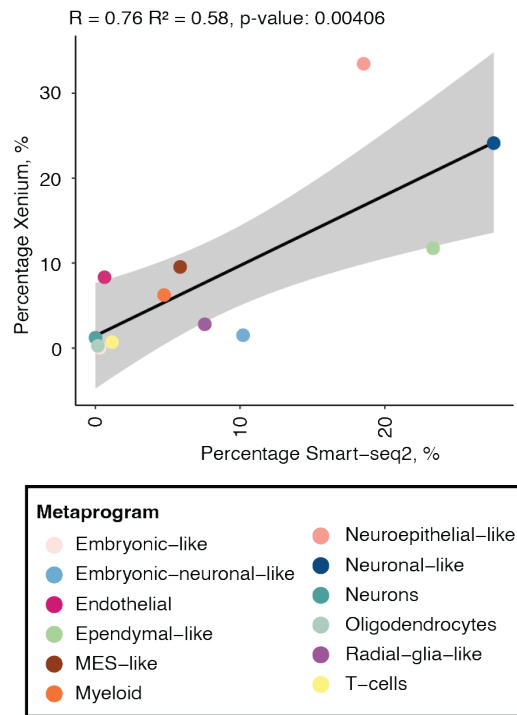


Figure R53. Scatterplot showing the percentage of cells detected in ZFTA-RELA tumors profiled by snRNA-seq and by 10X Xenium spatial transcriptomics. The goodness of fit (R , R^2) and P-value are displayed at the top.

Methods: Linear regression. We used simple linear regression to assess whether the abundance of a particular cell state/type is associated with increased spatial coherence and to evaluate the concordance between cell state/type abundance identified by sn/scRNA-seq and spatial Xenium analyses. We modeled the relationship between the two variables using the *lm* function from the R *stats* package. We extracted the goodness of fit (R , R^2 and P value) using the *summary()* function. We adjusted P-values for multiple comparisons using Bonferroni correction.

Changes made: Extended Data Figure 4b, Methods

#3-58. General comment, please add a section in methods describing in detail how and using what programs the linear regressions were done in the paper, currently there is only some description in the figure legends.

We have revised **Methods**.

Methods: Linear regression. We used simple linear regression to assess whether the abundance of a particular cell state/type is associated with increased spatial coherence and to evaluate the concordance between cell state/type abundance identified by sn/scRNA-seq and spatial Xenium analyses. We modeled the relationship between the two variables using the *lm* function from the R *stats* package. We extracted the goodness of fit (R , R^2 and P value) using the *summary()* function. We adjusted P-values for multiple comparisons using Bonferroni correction.

#3-59. Extended Data Figure 10c: The statistical test required normal distribution, please test for normal distribution and use appropriate test and describe rational in methods section.

We tested for normal distribution of this data, and as they were not normally distributed, we used the Wilcoxon rank sum test (**Figure R54**). However, as also addressed in **Reviewer comment #3-26,28,38**, because we have now included additional scRNA-seq data for 3 different cell lines, we had more figures to include in the manuscript. As this is the case, we have removed these plots from the manuscript.

Figure R54

Gene	Culture condition	W	p-value (Shapiro-Wilk)
VIM	Monoculture	0.89990898	1.63E-14
VIM	Coculture	0.64466566	1.39E-32
DCX	Monoculture	0.91358367	2.24E-13
DCX	Coculture	0.89873526	7.20E-19
GFAP	Monoculture	0.14289106	4.43E-37
GFAP	Coculture	0.2124527	2.48E-43

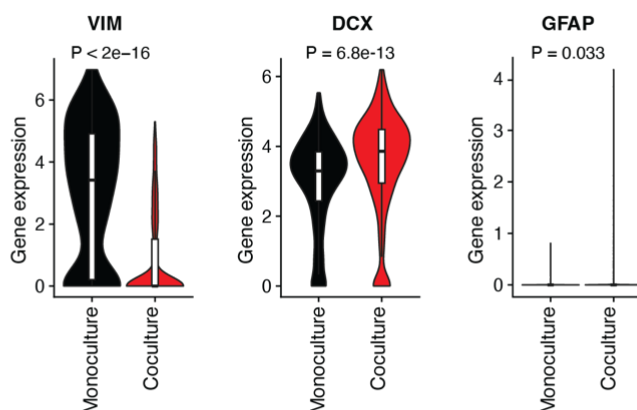


Figure R54. (left) Shapiro-Wilk normality test results of gene expression of *VIM*, *DCX*, and *GFAP* in ZFTA-RELA cells grown in monoculture or coculture conditions. **(right)** Violin plots of gene expression of *VIM*, *DCX*, and *GFAP* in ZFTA-RELA cells grown in monoculture or coculture conditions.

#3-60. Please include description of the statistical test with rational for selecting the test in Figure 5c, and Extended Data Figure 9e, in the methods section.

We thank the reviewer for this feedback. We tested for normality of somatokinesis (or cell migration speed) by Shapiro-Wilk normality test. As the data is not normally distributed, we tested for statistical significance using Kruskal-Wallis test for multiple comparisons. (**Figure R55**). We have amended **Methods**.

Figure R55

	NE1	NE2	Neuronal	Immature ependymal-like
W	0.6896	0.7506	0.7623	0.7476
p-value	<0.0001	<0.0001	<0.0001	<0.0001
passed normality test?	No	No	No	No

Comparison	Mean rank diff.	Significant?	Summary	Adjusted P Value
Neuroepithelial-like 1 vs. Neuroepithelial-like 2	-59.76	Yes	**	0.0046
Neuroepithelial-like 1 vs. Immature ependymal-like	33.77	No	ns	0.2864
Neuroepithelial-like 1 vs. Neuronal-like	-70.52	Yes	***	0.0006
Neuroepithelial-like 2 vs. Immature ependymal-like	93.52	Yes	****	<0.0001
Neuroepithelial-like 2 vs. Neuronal-like	-10.77	No	ns	>0.9999
Ependymal-like vs. Neuronal-like	-104.3	Yes	****	<0.0001

Figure R55. (top) Shapiro-Wilk normality test results for cell migration speed values of different cell states. **(bottom)** Kruskal-Wallis test results for comparisons between cell migration speed values of different cell states

Methods: Statistical Analysis. All statistical analyses were conducted using R (version 4.3.3). To determine statistical significance, we first assessed data normality using the Shapiro-Wilk test (shapiro.test). If all groups followed a normal distribution (Shapiro-Wilk P-value > 0.05), we performed Student's t-test. If one or more groups did not follow a normal distribution (Shapiro-

#3-62. The authors claim that high order architecture is influenced by MES/Hypoxia is supported by the data and analysis. However, this analysis is incomplete until test also have been done for all 9 cell types and states. There could be more cell types and states contributing to this higher order architecture. See comment about Lines 246-249.

We thank the reviewer for these important comments in **Reviewer comment #3-17** regarding these statements. As also addressed in **Reviewer comment #3-3** and **#3-17~18**, we performed linear regression analyses for all cell states/types with an increased sample size and confirmed that MES/Hypoxia is the **only** cell state that is significantly correlated with global spatial coherence score ($R = 0.68$; adjusted P value = 0.0047) (**Figure R5a**).

#3-63. The authors claim that “This indicates that higher spatial organization of ZFTA-RELA tumors correlates with more aggressive tumors.”, is not supported by the data, see comments for lines 251-254.

We thank the reviewer for these comments. As also addressed in rebuttal for **Reviewer comment #3-19**, we performed extensive statistical analyses of both the spatial coherence score and cell state composition of primary versus recurrent tumors in both our sc/snRNAseq and spatial cohort to confirm that there is no a statistical significance. We have removed these claims in the **Discussion** section.

#3-64. The section of the manuscript ‘ZFTA-RELA tumors are locally organized into recurrent tumor niches’ needs to be revised due to what appears to be circular arguments.

We thank the reviewer for this important feedback regarding the local neighborhood analyses of patient tumors. As also addressed in **Reviewer comment #3-20 & 21**, we have included additional analysis using CellCharter to overcome the limitations of the metaniche analysis and validate our findings.

F. Suggested improvements: experiments, data for possible revision

I have tried to collect the main suggested improvements here, but some are also under other sections.

#3-65. Please include an additional inferCNV analysis using normal cells as reference to see if there are any potentially different CNV patterns between and within subgroups. This could be relevant supporting information when investigating subgroup and inter-tumoral differences of the cohort.

We have included inferCNV analysis of patient tumors using normal cells as reference, as also addressed in **Reviewer comment #3-42 (Figure R26)**. We have included these analyses to **Extended Data Figure 1d**.

Figure R26

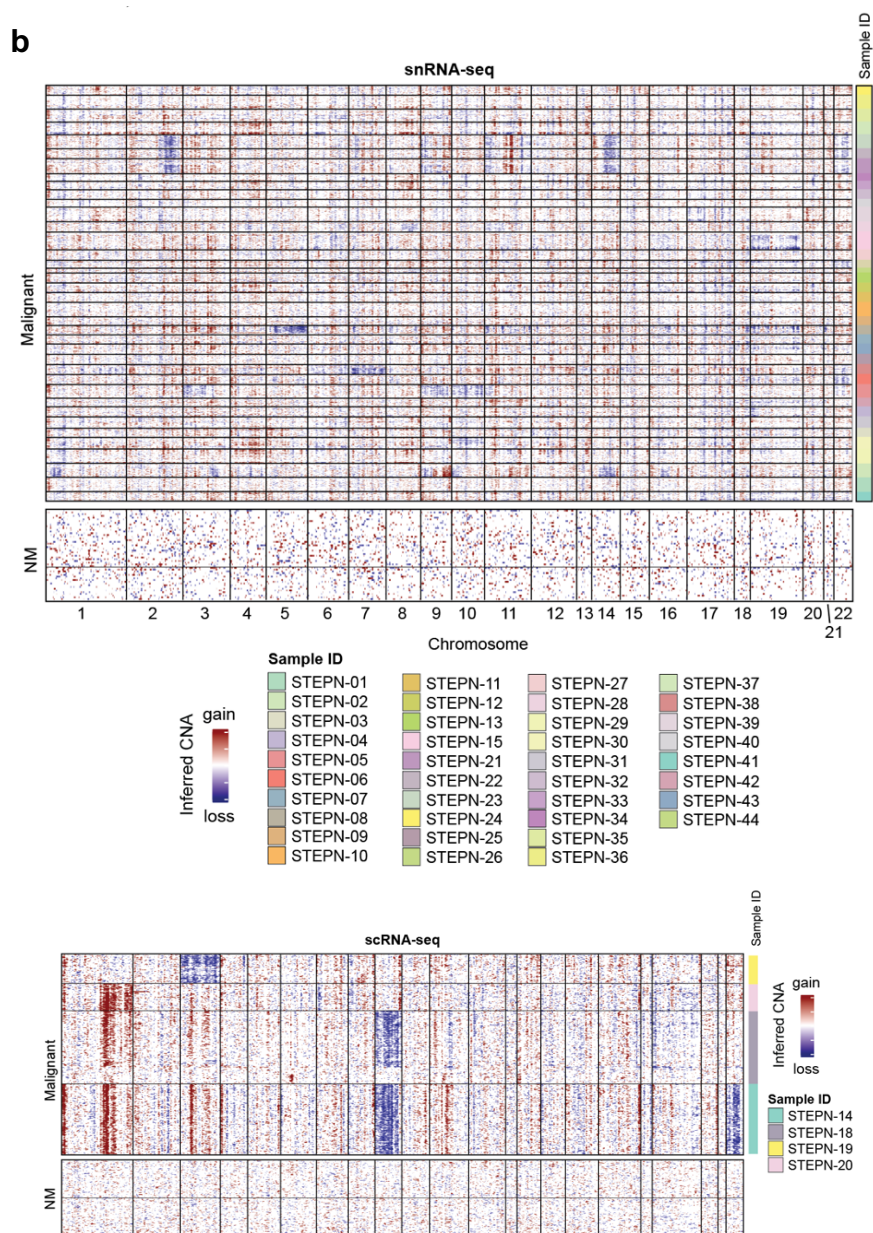


Figure R26. (a) UMAP plot of all ST-EPN patient cells/nuclei ($n = 7,840$) profiled by sc/snRNA-seq. Plots are colored by: inferred malignant status based on copy number variation (CNV, left panel) and final cell type classification (right panel). **(b)** CNV profiles of each cell (y-axis) across chromosomal location (x-axis) inferred from sc/snRNA-seq data. Cells are colored by tumor subtype, and ordered according to hierarchical clustering. Cells are split by inferred malignant/non-malignant origin. Representative fresh spike-in nonmalignant cells are shown at the bottom.

Changes made: Extended Data Figure 1d

#3-66. Line 127: Try a different approach too comparing the transcriptomes of the different subgroups. Now they do not cluster based on subgroups but look to cluster based on patient ID. See comments in section C for more details.

We thank the reviewer for this comment and helpful suggestions. As also addressed **Reviewer comment #3-9**, we visualized the cells on UMAP by patient ID, and also selected between 500 and 3000 highly variable genes (HVGs) from each patient sample scRNAseq data and performed Pearson correlation

and principle component analysis (**Figure R14c, R34, R37**). We observed that samples generally cluster by sample and subgroup, suggesting that subgroup is a component in driving differential gene expression profiles. We have included the results for HVG=3000 in **Figure 1d (Figure R34)** and amended the **Results** to more clearly explain this point.

Results: As DNA methylation profiles utilized to categorize these tumors cluster subgroups differently (**Figure 1b**) and are considered epigenetic fingerprints of cellular origins in both development and tumorigenesis^{48,49}, we hypothesized that ST-EPN tumors would correspondingly be transcriptionally distinct. Accordingly, we find that the malignant cells cluster both by sample and molecular subgroup classification (**Figure 1c, Extended Data Figure 1e**), consistent with previous observations⁵⁰. We further derived the top 3000 highly variable genes across samples and performed principal component analysis to demonstrate that samples cluster by their molecular subgroup classification (**Figure 1d**).

#3-67. Line 157-158: In addition to marker genes. The authors also use GO terms to annotate the metaprograms (Extended Data Figure 2b). From the methods section (line 673) it states that top 30 genes from each metaprogram is used. E.g., in Table S4 genes MKI67 and TOP2A (used to annotate the cycling metaprogram) are not within top 30. Neither is DCX for neuronal-like. I would suggest to re-run the GO analysis using all the top 200 genes for each metaprogram. Although, in the Extended Data file at the bottom there is a list of supplementary materials, stating that Table S5 “Gene ontology terms enriched in NMF metaprograms identified in ST-EPN patients (top 200 marker genes/metaprogram selected; GO terms with P-value < 0.01)”. This is inconsistent with the main material and methods part. Please clarify and update to correct and use consistent information.

As also addressed in **Reviewer comment #3-11 and #3-61**, we have revised all graphs and tables to use the top 30 metaprogram genes, which are the same genes used to score ST-EPN patient tumor cells (**Methods**). We appreciate the increase in consistency upon the reviewer’s suggestions.

#3-68. Lines 242-245: Given that the probe panel was design to capture cells from all eight metaprograms and multiple cell types, and that the authors have stated that their intention was to investigate 9 of these after data collection. I suggest additional analysis to be made (with multiple testing corrections), including all 9 cell types and metaprograms in separate linear regression analysis. E.g., included in an additional Extended Data Figure. There might be additional interesting biology not explored here that could have an impact on the higher order architecture of the tumor organization. This could substantially change the story of the paper in this section. Especially since the current data and analysis does not support the authors investigation into vascular cells as an explanation for disorganization (Figure 3 c and d), (p-value is not significant, could be study is underpowered, but can’t draw these conclusions with this data, please discuss in a revised discussion).

We thank the reviewer for the helpful suggestion and comments. We have revised the **Discussion** section to incorporate all new analyses regarding linear regression of other cell states and global spatial coherence scores, and have removed statements regarding the correlation of vascular cells to spatial coherence.

#3-69. For the section of the manuscript ‘ZFTA-RELA tumors are locally organized into recurrent tumor niches’. Consider characterizing recurrent niches using an unsupervised approach to avoid a circular argument. One alternative could also be to set out to investigate the recurrence of the 5 different malignant states captured by the Xenium data. Do the same analysis using BuildNicheAssay’s function from Seurat (v. 5), setting k=5. Then annotate the 5 different spatial niches for each section according to the same cell types/states as in Figure 4c.

We thank the reviewer for this helpful suggestion. As also addressed in **Reviewer comment #3-20**, we aimed to amend the limitations of metaniche analysis by using CellCharter that allows identification of spatial clusters simultaneously across all sections. We have included these new analyses to **Figure 4 and Extended Data Figure 6**.

#3-70. Line 345-346: Perform calculations and statistical test to support statement of enrichment. See section C for more details.

We thank the reviewer for the helpful suggestions outlined in **Reviewer comment #3-28**. We performed the suggested analysis and can conclude that there is a statistically significant relationship between the tumor model (spheres, adherent, coculture, PDX or patient) and the frequency of progenitor (Neuroepithelial-like and Embryonic-like) and differentiated (Neuronal-like + Ependymal-like) states (**Figure R45**).

Changes made: Extended Data Figure 9e

G. References: appropriate credit to previous work?

#3-71. Lines 67-68: the authors state “differences across species underscore the challenge in extrapolating findings from these studies in mice.”, please provide/mark a reference for this statement.

We have included the references for **Introduction**.

Introduction: differences across species underscore the challenge in extrapolating findings from these studies in mice⁶⁹.

#3-72. Lines 160-169: the authors mention genes associated with different metaprograms, please add literature references where missing.

We have included the references for **Results**.

Results: Two metaprograms were associated with generic cancer cell processes, including a cycling metaprogram (*HELLS*, *CENPF*)^{70,71} which is enriched in genes associated with cell proliferation, and a mesenchymal (MES)/Hypoxia (*PGK1*, *ENO1*, *CD44*)^{25,35} metaprogram which highly expresses hypoxia-related genes (**Figure 2b**, **Table S4**). The remaining six metaprograms were distinctly related to developmental programs, including: (1) embryonic-like (*ALDH1A2*, *CTBP2*, *ACTN4*) which expresses genes essential for formation and patterning of the embryo⁷²⁻⁷⁴ and is enriched in GO terms related to retinoic acid receptor signaling, not matching to a particular cell type found in the developing human brain; (2) neuroepithelial-like that strongly project to neuroepithelial stem cells of the developing human cortex; (3) radial glia-like that expresses radial glia genes (*LRIG1*, *DCLK1*)^{75,76}; (4) embryonic-neuronal-like (*RBFOX1*, *CNTNAP2*) enriched with immature neuronal genes^{77,78}; (5) neuronal-like (*KIRREL3*, *NRXN3*) enriched in neuronal differentiation-related genes^{79,80}; (6) ependymal-like (*DNAH9*, *DNAH6*) enriched with cilia-related genes⁸¹ (**Figure 2b**, **Extended Data Figure 2b-c**, **Table S4-5**).

#3-73. Line 230: In this section the author could use their ref 24 (Greenwald, A. C. et al. Integrative spatial analysis reveals a multi-layered organization of glioblastoma. Cell 187, 2485-2501.e26 (2024), to high-light previous work done by others looking at spatial organization and hypoxia.

We have included the references for **Results**.

Results: mesenchymal (MES)/Hypoxia (*PGK1*, *ENO1*, *CD44*) metaprogram, which highly expresses hypoxia-related genes^{15,82}

H. Clarity and context: lucidity of abstract/summary, appropriateness of abstract, introduction and conclusions This is described under C.

I. Comments on the Data repository

#3-74. Please add a separate file about the metadata on GEO for the spatial data. The current files are large and having a separate small metadata file makes the data more accessible and more in line with FAIR data principles. E.g., a version of the authors supplementary table 1 could be used here.

We thank the reviewer for this helpful suggestion. We included the metadata file to GEO that outlines relevant details about the spatial data (patient ID, age, sex and others).

#3-75. In the manuscript the authors state that they are analyzing 23 tissue sections using the spatial transcriptomics method Xenium. However, in the GEO database only 22 samples are provided. Please update and add all samples to the GEO database, or if it only was 22 sections, please update and correct this number in the manuscript.

We have expanded the spatial cohort to include **56** tissue sections and have included all sections into the GEO database.

#3-76. For the single cell and single nuclei data in the GEO, please add more metadata information to facilitate easier re-use of data. I suggest an additional supplementary file that helps the reader navigate the 42 sample GEO upload. E.g., 3 columns: 1) GSM number (GSM8115764), 2) Sample group (1), 3) Description (scRNAseq? Xenograft ZFTA-RELA?). From the GEO database entry, I'm currently having difficulties to easily understand what these samples are, I think that adding such a Description would greatly increase the data's adherence to FAIR principles. Due to this I'm unable to ascertain if all data has been uploaded in accordance with the authors figure 1a. However, I do note that in Figure 1a it looks like there should be 47 samples for sc/snRNAseq, but only 42 are added to the GEO database, please double check that all data has been deposited.

We thank the reviewer for this constructive feedback. We have included in the GEO supplementary file an additional column that includes GSM number, and description. In addition, we have confirmed that there are 42 sc/snRNA-seq patient data samples and **56** 10X Xenium data deposited in GEO.

Referee #3 (Remarks on code availability):

The authors provide extensive and descriptive annotation of the code for their work. I have not tested the code.

We have added extensive description and notebooks to our github repository to ensure reproducibility of our analyses.

References

1. Nowakowski, T. J. *et al.* Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science* **358**, 1318–1323 (2017).
2. Nishino, J., Kim, S., Zhu, Y., Zhu, H. & Morrison, S. J. A network of heterochronic genes including *Imp1* regulates temporal changes in stem cell properties. *Elife* **2**, e00924 (2013).
3. Cadwell, C. R., Bhaduri, A., Mostajo-Radji, M. A., Keefe, M. G. & Nowakowski, T. J. Development and Arealization of the Cerebral Cortex. *Neuron* vol. 103 Preprint at <https://doi.org/10.1016/j.neuron.2019.07.009> (2019).
4. Eze, U. C., Bhaduri, A., Haeussler, M., Nowakowski, T. J. & Kriegstein, A. R. Single-cell atlas of early human brain development highlights heterogeneity of human neuroepithelial cells and early radial glia. *Nat Neurosci* **24**, 584–594 (2021).
5. Phipson, B. *et al.* propeller: testing for differences in cell type proportions in single cell data. *Bioinformatics* **38**, 4720–4726 (2022).
6. Subramanian, L., Bershteyn, M., Paredes, M. F. & Kriegstein, A. R. Dynamic behaviour of human neuroepithelial cells in the developing forebrain. *Nat Commun* **8**, 14167 (2017).
7. Akula, S. K., Exposito-Alonso, D. & Walsh, C. A. Shaping the brain: The emergence of cortical structure and folding. *Dev Cell* **58**, 2836–2849 (2023).
8. Götz, M. & Huttner, W. B. The cell biology of neurogenesis. *Nat Rev Mol Cell Biol* **6**, 777–788 (2005).
9. Manno, G. La *et al.* RNA velocity of single cells. *Nature* 2018 560:7719 **560**, 494–498 (2018).
10. Akiba, C., Takezawa, A., Tsai, Y., Hirose, M. & Suzuki, T. bHLH family proteins control the timing and completion of transition from neuroepithelial cells into neural stem cells. *Development* **151**, dev202630 (2024).
11. Subramanian, L., Bershteyn, M., Paredes, M. F. & Kriegstein, A. R. Dynamic behaviour of human neuroepithelial cells in the developing forebrain. *Nat Commun* **8**, (2017).
12. Tritschler, S. *et al.* Concepts and limitations for learning developmental trajectories from single cell genomics. *Development* **146**, dev170506 (2019).
13. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation* **2**, 100141 (2021).
14. Gojo, J. *et al.* Single-Cell RNA-Seq Reveals Cellular Hierarchies and Impaired Developmental Trajectories in Pediatric Ependymoma. *Cancer Cell* **38**, 44–59 (2020).
15. Greenwald, A. C. *et al.* Integrative spatial analysis reveals a multi-layered organization of glioblastoma. *Cell* **187**, 2485–2501.e26 (2024).
16. Zhang, Y. *et al.* Rapid Single-Step Induction of Functional Neurons from Human Pluripotent Stem Cells. *Neuron* **78**, 785–798 (2013).
17. Becker-Heck, A. *et al.* The coiled-coil domain containing protein CCDC40 is essential for motile cilia function and left-right axis formation. *Nat Genet* **43**, 79–84 (2011).
18. Uhlén, M. *et al.* Tissue-based map of the human proteome. *Science (1979)* **347**, 1260419 (2015).
19. Mangena, V. *et al.* Glioblastoma Cortical Organoids Recapitulate Cell-State Heterogeneity and Intercellular Transfer. *Cancer Discov* **15**, 299–315 (2025).
20. Winter, J. M., Yadav, T. & Rutter, J. Stressed to death: Mitochondrial stress responses connect respiration and apoptosis in cancer. *Mol Cell* **82**, 3321–3332 (2022).
21. Vind, A. C., Zhong, F. L. & Bekker-Jensen, S. Death by ribosome. *Trends Cell Biol* **35**, 615–626 (2025).
22. Kang, J. *et al.* Ribosomal proteins and human diseases: molecular mechanisms and targeted therapy. *Signal Transduct Target Ther* **6**, 323 (2021).
23. Pajtler, K. W. *et al.* Molecular Classification of Ependymal Tumors across All CNS Compartments, Histopathological Grades, and Age Groups. *Cancer Cell* **27**, 728–743 (2015).

24. Tauziède-Espariat, A. *et al.* Supratentorial non-RELA, ZFTA-fused ependymomas: a comprehensive phenotype genotype correlation highlighting the number of zinc fingers in ZFTA-NCOA1/2 fusions. *Acta Neuropathol Commun* **9**, 135 (2021).
25. Ghasemi, D. R. *et al.* EPEN-14. MOLECULAR MARKERS AND PREDICTORS OF OUTCOME FOR PAEDIATRIC INTRACRANIAL EPENDYMOMA IN THE PROSPECTIVE, MULTICENTRE E-HIT2000 TRIAL AND SUBSEQUENT HIT-REGISTRIES: A POOLED ANALYSIS OF 244 PATIENTS. *Neuro Oncol* **26**, 0 (2024).
26. Zheng, T. *et al.* Cross-Species Genomics Reveals Oncogenic Dependencies in ZFTA/C11orf95 Fusion-Positive Supratentorial Ependymomas. *Cancer Discov* **11**, 2230–2247 (2021).
27. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nature Methods* **16**, 1289–1296 (2019).
28. Debanne, D., Campanac, E., Bialowas, A., Carlier, E. & Alcaraz, G. Axon Physiology. *Physiol Rev* **91**, 555–602 (2011).
29. Rubio-Teves, M. *et al.* Benchmarking of tools for axon length measurement in individually-labeled projection neurons. *PLoS Comput Biol* **17**, e1009051- (2021).
30. Osswald, M. *et al.* Brain tumour cells interconnect to a functional and resistant network. *Nature* **528**, 93–98 (2015).
31. Jung, E. *et al.* Tweak-Homolog 1 Drives Brain Colonization of Gliomas. *The Journal of Neuroscience* **37**, 6837 (2017).
32. Venkataramani, V. *et al.* Glioblastoma hijacks neuronal mechanisms for brain invasion. *Cell* **185**, 2899-2917.e31 (2022).
33. Tetzlaff, S. K. *et al.* Characterizing and targeting glioblastoma neuron-tumor networks with retrograde tracing. *Cell* **188**, 390-411.e36 (2025).
34. Liewald, D., Miller, R., Logothetis, N., Wagner, H.-J. & Schüz, A. Distribution of axon diameters in cortical white matter: an electron-microscopic study on three human brains and a macaque. *Biol Cybern* **108**, 541–557 (2014).
35. Yamada, K. M., Spooner, B. S. & Wessells, N. K. ULTRASTRUCTURE AND FUNCTION OF GROWTH CONES AND AXONS OF CULTURED NERVE CELLS. *Journal of Cell Biology* **49**, 614–635 (1971).
36. Dent, E. W., Gupton, S. L. & Gertler, F. B. The growth cone cytoskeleton in axon outgrowth and guidance. *Cold Spring Harb Perspect Biol* **3**, (2011).
37. Venkatesh, H. S. *et al.* Neuronal activity promotes glioma growth through neuroligin-3 secretion. *Cell* **161**, 803–816 (2015).
38. Aran, D. *et al.* Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol* **20**, 163–172 (2019).
39. Liu, I. *et al.* GABAergic neuronal lineage development determines clinically actionable targets in diffuse hemispheric glioma, H3G34-mutant. *Cancer Cell* **42**, 1528-1548.e17 (2024).
40. Liu, I. *et al.* The landscape of tumor cell states and spatial organization in H3-K27M mutant diffuse midline glioma across age and location. *Nat Genet* **54**, (2022).
41. Regev, A. *et al.* The human cell atlas. *Elife* **6**, (2017).
42. Filbin, M. G. *et al.* Developmental and oncogenic programs in H3K27M gliomas dissected by single-cell RNA-seq. *Science* (1979) **360**, 331–335 (2018).
43. Neftel, C. *et al.* An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. *Cell* **178**, 835-849.e21 (2019).
44. Liu, I. *et al.* The landscape of tumor cell states and spatial organization in H3-K27M mutant diffuse midline glioma across age and location. *Nat Genet* **54**, 1881–1894 (2022).
45. Venkatesh, H. S. *et al.* Targeting neuronal activity-regulated neuroligin-3 dependency in high-grade glioma. *Nature* **549**, (2017).
46. Venkatesh, H. S. *et al.* Electrical and synaptic integration of glioma into neural circuits. *Nature* **573**, 539–545 (2019).

47. Chen, H.-C. *et al.* Histone serotonylation regulates ependymoma tumorigenesis. *Nature* **632**, 903–910 (2024).
48. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev* **16**, 6–21 (2002).
49. Shen, H. & Laird, P. W. Interplay between the Cancer Genome and Epigenome. *Cell* **153**, 38–55 (2013).
50. Zheng, T. *et al.* Cross-Species Genomics Reveals Oncogenic Dependencies in ZFTA/C11orf95 Fusion-Positive Supratentorial Ependymomas. *Cancer Discov* **11**, 2230–2247 (2021).
51. Hovestadt, V. *et al.* Resolving medulloblastoma cellular architecture by single-cell genomics. *Nature* **572**, 74–79 (2019).
52. Drexler, R. *et al.* Temporal change of DNA methylation subclasses between matched newly diagnosed and recurrent glioblastoma. *Acta Neuropathol* **147**, 21 (2024).
53. Donson, A. M. *et al.* Significant increase of high-risk chromosome 1q gain and 6q loss at recurrence in posterior fossa group A ependymoma: A multicenter study. *Neuro Oncol* **25**, 1854–1867 (2023).
54. Grabski, I. N., Street, K. & Irizarry, R. A. Significance analysis for clustering with single-cell RNA-sequencing data. *Nat Methods* **20**, 1196–1202 (2023).
55. Yu, L., Cao, Y., Yang, J. Y. H. & Yang, P. Benchmarking clustering algorithms on estimating the number of cell types from single-cell RNA-sequencing data. *Genome Biol* **23**, 49 (2022).
56. Zhang, S., Li, X., Lin, J., Lin, Q. & Wong, K.-C. Review of single-cell RNA-seq data clustering for cell-type identification and characterization. *RNA* **29**, 517–530 (2023).
57. Liang, X. *et al.* A critical assessment of clustering algorithms to improve cell clustering and identification in single-cell transcriptome study. *Brief Bioinform* **25**, bbad497 (2024).
58. Varrone, M., Tavernari, D., Santamaria-Martínez, A., Walsh, L. A. & Ciriello, G. CellCharter reveals spatial cell niches associated with tissue remodeling and cell plasticity. *Nat Genet* **56**, 74–84 (2024).
59. Itoh, Y. *et al.* PDK1–Akt pathway regulates radial neuronal migration and microtubules in the developing mouse neocortex. *Proceedings of the National Academy of Sciences* **113**, E2955–E2964 (2016).
60. Marín, O., Valiente, M., Ge, X. & Tsai, L. H. Guiding neuronal cell migrations. *Cold Spring Harb Perspect Biol* **2**, (2010).
61. Evsyukova, I., Plestant, C. & Anton, E. S. Integrative mechanisms of oriented neuronal migration in the developing brain. *Annu Rev Cell Dev Biol* **29**, 299–353 (2013).
62. Cooper, J. A. Mechanisms of cell migration in the nervous system. *Journal of Cell Biology* **202**, 725–734 (2013).
63. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676–682 (2012).
64. Capper, D. *et al.* DNA methylation-based classification of central nervous system tumours. *Nature* **555**, 469–474 (2018).
65. Aryee, M. J. *et al.* Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363–1369 (2014).
66. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47–e47 (2015).
67. Okonechnikov, K., Conesa, A. & García-Alcalde, F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* **32**, 292–294 (2016).
68. Uhrig, S. *et al.* Accurate and efficient detection of gene fusions from RNA sequencing data. *Genome Res* **31**, 448–460 (2021).
69. Breschi, A., Gingeras, T. R. & Guigó, R. Comparative transcriptomics in human and mouse. *Nat Rev Genet* **18**, 425–440 (2017).
70. Song, R. *et al.* CENPF overexpression in bladder cancer cells enhances proliferation, migration, invasion, and apoptosis. *Sci Rep* **15**, 25398 (2025).
71. von Eyss, B. *et al.* The SNF2-like helicase HELLS mediates E2F3-dependent transcription and cellular transformation. *EMBO J* **31**, 972–985–985 (2012).

72. Lee, H.-T. *et al.* Phosphorylation-mediated disassembly of C-terminal binding protein 2 tetramer impedes epigenetic silencing of pluripotency in mouse embryonic stem cells. *Nucleic Acids Res* **52**, 13706–13722 (2024).
73. Li, C. *et al.* Unraveling the role of the WIPF1/ACTN4 complex in podosome formation of human placental EVTs: Insights into recurrent spontaneous abortion. *Genes Dis* **12**, 101665 (2025).
74. Leon, E., Nde, C., Ray, R. S., Preciado, D. & Zohn, I. E. ALDH1A2-related disorder: A new genetic syndrome due to alteration of the retinoic acid pathway. *Am J Med Genet A* **191**, 90–99 (2023).
75. De Vincenti, A. P., Bonafina, A., Ledda, F. & Paratcha, G. Lrig1 regulates cell fate specification of glutamatergic neurons via FGF-driven Jak2/Stat3 signaling in cortical progenitors. *Development* **151**, dev202879 (2024).
76. Shu, T. *et al.* Doublecortin-like Kinase Controls Neurogenesis by Regulating Mitotic Spindles and M Phase Progression. *Neuron* **49**, 25–39 (2006).
77. Jacko, M. *et al.* Rbfox Splicing Factors Promote Neuronal Maturation and Axon Initial Segment Assembly. *Neuron* **97**, 853–868.e6 (2018).
78. Canali, G. *et al.* Genetic variants in autism-related CNTNAP2 impair axonal growth of cortical neurons. *Hum Mol Genet* **27**, 1941–1954 (2018).
79. Martin, E. A. *et al.* The intellectual disability gene Kirrel3 regulates target-specific mossy fiber synapse development in the hippocampus. *Elife* **4**, e09395 (2015).
80. Boxer, E. E. *et al.* Neurexin-3 defines synapse- and sex-dependent diversity of GABAergic inhibition in ventral subiculum. *Cell Rep* **37**, 110098 (2021).
81. Hawrylycz, M. J. *et al.* An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, (2012).
82. Kim, J. W., Tchernyshyov, I., Semenza, G. L. & Dang, C. V. HIF-1-mediated expression of pyruvate dehydrogenase kinase: A metabolic switch required for cellular adaptation to hypoxia. *Cell Metab* **3**, (2006).

RE: Nature manuscript 2024-08-16431

“Single-cell multidimensional profiling of tumor cell heterogeneity in supratentorial ependymomas”

Point-by-point response to reviewer comments

We thank the reviewers for their generous and thoughtful feedback to the manuscript. Please see below for a point-by-point response to Reviewer #3’s comments.

List of Figures included in this rebuttal letter:

Figure R1. Proportion of metaprograms in primary and recurrent patients in snRNA-seq and Xenium data

Figure R2. Revised Extended Data Figure 2e

Figure R3. Revised Figure 2a

Figure R4. Cell cycle genes in neuroepithelial-like cells proximal or distant from MES/Hypoxia regions

Figure R5. Single and multiple linear regression of transformed data of cell states/types with spatial coherence

Figure R6. Revised Extended Data Figure 6f

Figure R7. Revised Figure 4d and Extended Data Figure 6g

Figure R8. Mean proportion of progenitor-like and differentiated cells across models

Figure R9. Revised Figure 2c

Figure R10. Kruskal-Wallis test and Dunn’s multiple comparison test of cell migration speed of cell states

Figure R11. Revised Extended Data Figure 1d

Referee #1 (Remarks to the Author):

The authors have truly added a great amount of content to the manuscript and have largely addressed many of the initial concerns (not only by this referee, but the others as well). I would like to commend them for this impressive work, and do feel the manuscript is suitable for publication at this point.

Referee #2 (Remarks to the Author):

The authors have addressed all my concerns in this comprehensive revision. It is an important contribution to the literature.

Referee #3 (Remarks to the Author):

Dear Authors, thank you for the revised version of the manuscript. I would like to acknowledge the effort, care, and attention to detail that has gone into improving the work, as well as the clear and extensive communication about the revisions. After reading the updated manuscript, I still have a few comments and questions, outlined below.

[We thank the reviewer for the constructive and thoughtful comments to improve the manuscript.](#)

Main

1. Lines 92–94: The authors state: “However, these studies were limited by patient number and did not include the non-canonical ST-ZFTA subgroups, which led to many cell states being sample- or subgroup-specific.” In this new cohort—2–3 times larger—the authors still observe (lines 151–152) that “the malignant cells cluster both by sample and molecular subgroup classification.” This suggests that similar patterns emerge despite the increased sample size. I would recommend adjusting the wording around lines 92–94 to something like: “However, these studies were limited by patient number and did not include the non-canonical ST-ZFTA subgroups, which could have led to many cell states appearing sample- or subgroup-specific.” If this change is made, I also suggest including a comment elsewhere in the manuscript to note that, despite the increased sample size, global similarities remain evident at both the sample and subgroup level. Alternatively, you might consider avoiding patient number as a differentiating factor altogether, e.g.: “However, these studies did not include the non-canonical ST-ZFTA subgroups.” The current wording conveys certainty about sample size as the cause of observed differences. Do we actually know that an increased cohort would change the sample- or subgroup-specific clustering, or could this reflect true underlying heterogeneity?

[We thank the reviewer for this point. We agree with the reviewer that while sample size may be a factor, we do not conclusively know if the sample size is the primary driving cause for our observations in lines 92-94. As the reviewer suggested, a clear difference is the lack of representation of non-canonical ST-ZFTA subgroups included in the previous studies. We have revised the language accordingly.](#)

Introduction: [However, these studies did not include the non-canonical ST-ZFTA subgroups, which led to many cell states being sample- or subgroup-specific.](#)

Changes made: Introduction

2. Line 118: The authors write “... with neuronal-like cancer cells being the most migratory...”. Would it be possible, based on the available metadata, to explore whether patients with a higher proportion of neuronal-like metaprograms show a greater tendency toward relapse, compared with other patients—assuming the number of relevant cases permits such an analysis?

We thank the reviewer for this interesting point. We compared the metaprogram proportion of primary versus recurrent tumors in both sc/snRNA-seq (**Figure R1 top**) and spatial ZFTA-RELA data (**Figure R1 bottom**). We used the propeller function in R package *speckle* (v1.8.0)¹, a specialized method for performing statistical tests on compositional data, to determine statistical significance in cell proportions across subgroups. Briefly, this function fits a linear model for each cell type, with the groups as the explanatory variable to obtain false discovery rates (FDR) by performing Bayes shrinkage of the variances and correcting for multiple testing. Using this method, **we did not see any statistically significant change in proportion of the neuronal-like metaprogram nor any other malignant metaprograms in primary versus recurrent tumors.**

Figure R1



Figure R1. (top) Proportion of metaprograms (y-axis) in primary and recurrent patient samples (x-axis) in ZFTA-RELA subgroup analyzed by sc/snRNAseq. Statistical significance of cell proportions by primary and recurrent tumors in ZFTA-

RELA using the propeller function from the R package *speckle* (right). **(bottom)** Proportion of metaprograms (y-axis) in primary and recurrent patient samples (x-axis) in ZFTA-RELA subgroup analyzed by 10X Xenium. Statistical significance of cell proportions by primary and recurrent tumors in ZFTA-RELA using the propeller function from the R package *speckle* (right).

3. Line 189: The authors provide example genes for 7 out of 8 metaprograms among the top 30 genes used to define these metaprograms. Could the authors clarify why this was not done for (2) the neuroepithelial-like metaprogram? For consistency, I suggest either including representative gene examples for this metaprogram in Table S4 or briefly explaining why this was not possible.

Results: neuroepithelial-like (*TUBA1A*, *ACTG1*)^{2,3} that strongly project to neuroepithelial stem cells of the developing human cortex;

4. Lines 196–199: The authors state that neuroepithelial-like and embryonic-like cells are the most proliferative, based on Extended Data Figures 2e–f. This is consistent with Extended Data Figure 2f; however, in Extended Data Figure 2e the pie charts suggest a much larger difference between neuroepithelial-like and embryonic-like cells. Could the authors clarify this discrepancy? One possible explanation is that Extended Data Figure 2e includes all cells, so that the cell-cycle signal is measured not only in neuroepithelial-like cells but also in others that fall within the neuroepithelial-like square of the plot. If this is the case, I suggest reconsidering whether the pie charts in Extended Data Figure 2e should be included, as they can be confusing when compared to Extended Data Figure 2f. For example, in 2e the difference between neuroepithelial-like and embryonic-like appears large, while in 2f it is small; conversely, in 2f the difference between embryonic-like and neuronal-like is clear, but in 2e they appear identical.

Figure R2

Figure R2. Revised Extended Data Figure 2e

Changes made: Extended Data Figure 2e

5. Lines 229–231: The authors state: “While there was a continuum of gene expression to both neuronal and ependymal-like cells in ZFTA-RELA tumors, the other ST-EPN subgroups only harbored one of the mature cell states.” However, for ZFTA-Cluster 4 this does not appear to be the case. It seems to derive from a neuroepithelial-like continuum but remain “stuck” between the two mature metaprograms. In my reading, it looks to have roughly equal numbers of cells in the neuronal-like, ependymal-like, and embryonal-like regions. Could the authors clarify whether this interpretation is correct? If so, I suggest revising the sentence to reflect this distinction. For example: “While there was a continuum of gene expression to both neuronal and ependymal-like cells in ZFTA-RELA tumors, the other ST-EPN subgroups showed unique cell state signatures, with ZFTA-Clusters 1–3 harboring only one mature cell state, and ZFTA-Cluster 4 appearing to remain in a premature state.”

We agree with the reviewer’s interpretation and have amended the sentence to highlight more clearly the states observed in ZFTA-Cluster 4.

Results: While there was a continuum of gene expression to both neuronal and ependymal-like cells in ZFTA-RELA tumors, the other ST-EPN subgroups showed unique cell state signatures, with ZFTA-Clusters 1–3 harboring only one mature cell state, and ZFTA-Cluster 4 appearing to remain in a premature cell state (**Figure 2e**).

Changes made: Results “The cellular hierarchy of ST-EPN tumors mirrors human cortical development” section

6. Figure 2a: It looks like the heatmap color bars for the cycling metaprogram is dark-purple on the y-axis (to the middle-bottom-right) but dark-grey in all other places in the manuscript. Color-typo? Please update for consistency.

We thank the reviewer for pointing out this discrepancy. We have revised Figure 2a.

Figure R3

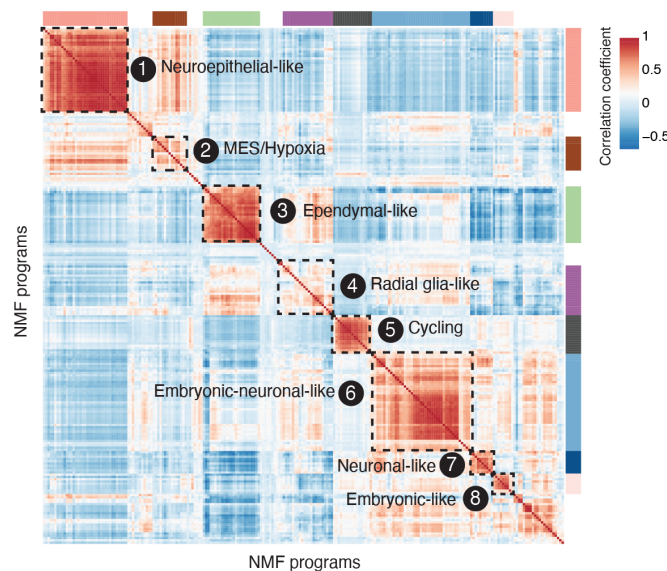


Figure R3. Revised Figure 2a

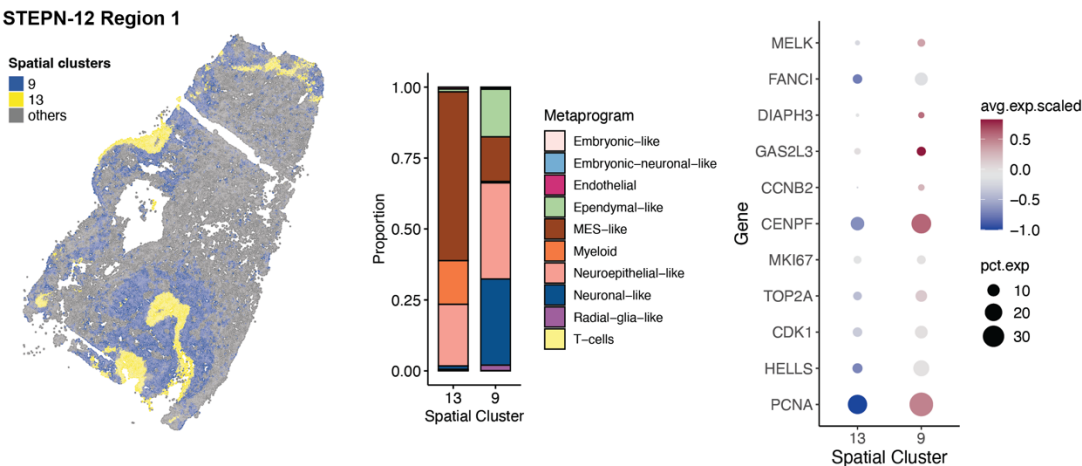
Changes made: Figure 2a

Spatial transcriptomics section:

7. I suggest including an analysis of the cycling metaprogram in the spatial samples. Since this metaprogram was identified in the sc/snRNA-seq data and there are at least four probes related to cycling in the Xenium panel, exploring its spatial distribution could provide valuable insights into tumor growth dynamics. For example, it would be interesting to see whether the cycling metaprogram is anti-correlated with the MES/Hypoxia metaprogram in a gradient-like manner. This could help understand whether dividing cells are spatially excluded from hypoxic regions, and whether slow-cycling or quiescent cells near hypoxic regions might contribute to recurrent disease. If the number of matched primary and relapse samples is too small for detailed analysis, a preliminary profiling of the cycling metaprogram could still be informative. At minimum, the authors could provide a rationale if this analysis is not technically feasible given the panel design. Including this would strengthen the spatial characterization of the tumor landscape and complement the MES/Hypoxia findings.

We thank the reviewer for this interesting question. We did not include the cycling metaprogram in the annotation of the spatial dataset as these markers were not specific enough to isolate out a specific population. However, to determine whether cycling is anti-correlated with MES/Hypoxia in a gradient-like manner, we queried the expression of cell cycle genes in non-MES/Hypoxia malignant cells proximal or distant from MES/Hypoxia cells. We first revisited the gene panel and identified 11 genes that are cell cycle related. Next, we identified two neighbor spatial clusters with distinct proportions of MES/Hypoxia cells-spatial clusters 9 and 13-which are present in 3 samples (**Figure R4**). Spatial cluster 13 encompass a high proportion of MES/Hypoxia cells, while cluster 9 are composed of a low proportion of MES/Hypoxia cells, but in proximity to cluster 13. Both spatial clusters contain at least 10% neuroepithelial-like cells, and thus we were able to compare expression of cell cycle genes in neuroepithelial-like cells proximal or distant from MES/Hypoxia. Neuroepithelial-like cells in cluster 9, which are depleted of MES/Hypoxia cells, show substantially lower expressions of all cell cycle genes (**Figure R4**). This is consistent with the reviewer’s and our hypothesis that proliferating cells are generally excluded from hypoxic regions. The limited number of samples with both clusters 9 and 13 prevents us from further exploring the possibility of enrichment of quiescent malignant cells in relapsed samples. Overall, we appreciate the reviewer’s astute suggestion and believe that this could be a direction to explore for future studies.

Figure R4



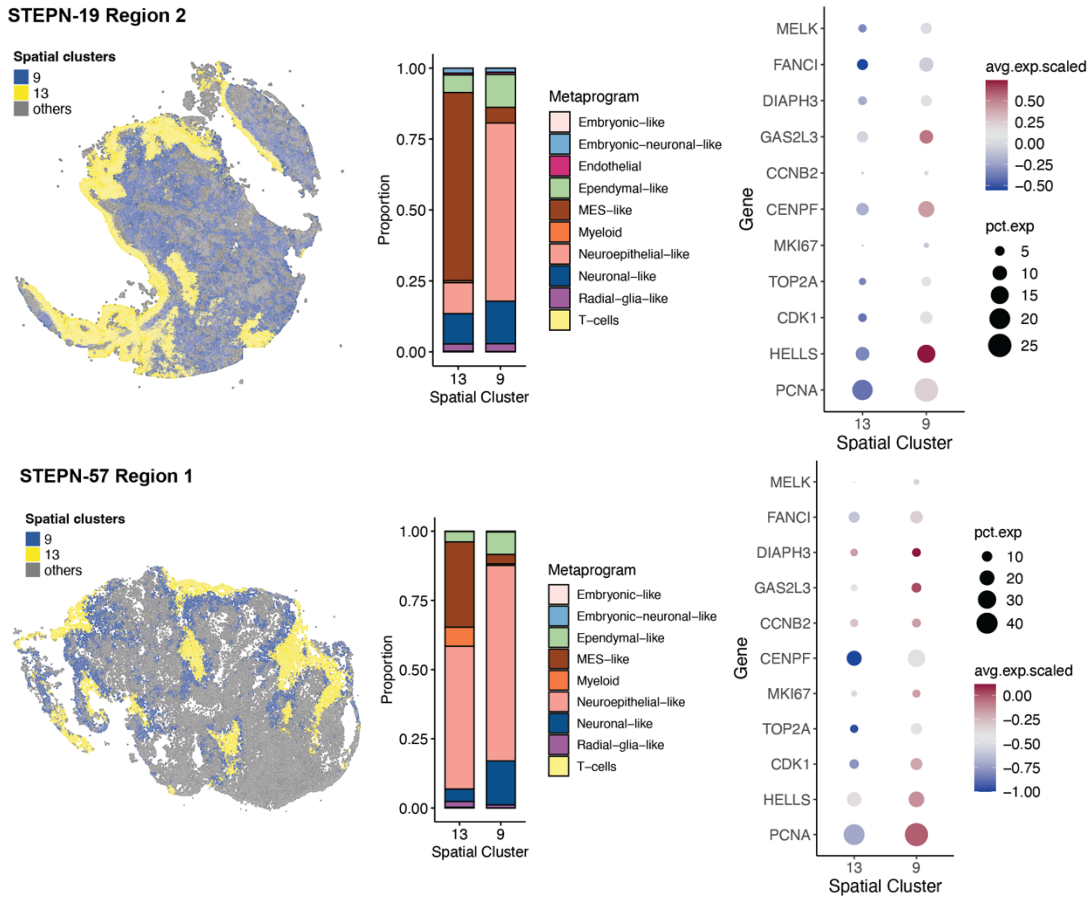


Figure R4. Comparison of spatial cluster 13 and 9 across three samples – STEP-12 Region 1 (top row), STEP-19 Region 2 (middle row) and STEP-59 Region 1 (bottom row). (left) Spatial map of spatial clusters 13, 9, and others (in grey). (middle) proportion barplot of cell states/types in spatial cluster 13 and 9. (right) Expression of cycling-related genes (y-axis) in spatial clusters 13 and 9 (x-axis). Size of dot depicts the percentage of cells expressing the gene, and color scale is scaled average expression.

8. Figure 3c–d: The authors provide a clear overview of the data, including helpful visual examples. For consistency, I suggest ensuring that the same samples shown visually in 3c are also highlighted in 3e. Additionally, in 3d the y-axis is labeled “percentage %,” whereas in Extended Data Figure 4e it is labeled “Frequency %.” I recommend standardizing this terminology for consistency across figures.

We have now revised Figure 3d to highlight the same samples as Figure 3c. In addition, we have standardized the terminology of the y-axis in Figure 3d and Extended Data Figure 4e to “transformed proportion”, which is further addressed in our response to Reviewer comment #9.

Changes made: Figure 3d, Extended Figure 4e

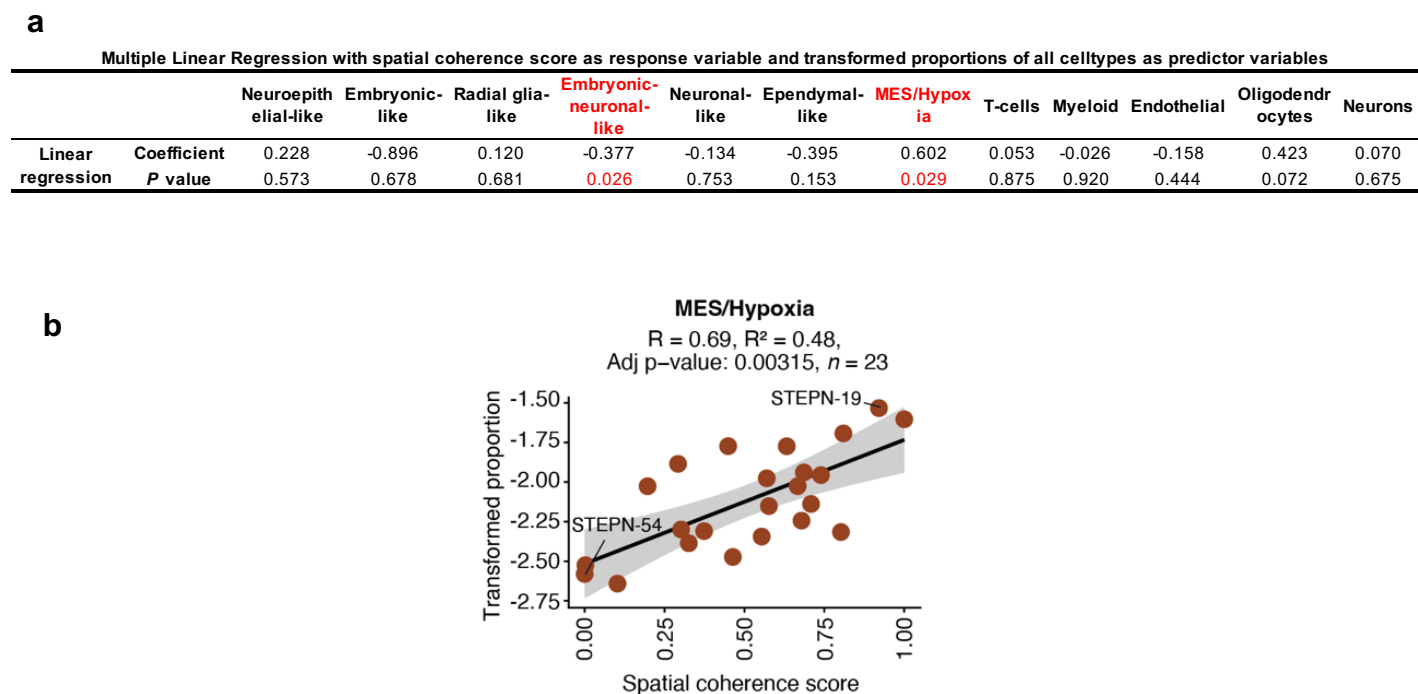
9. Figure 3d and Extended Data Figure 4e: The authors’ initial approach of correlating each cell type proportion with the global spatial coherence score and applying a Bonferroni correction is acceptable for exploratory analysis, and the plots are informative for visualizing these relationships. However, because the follow-up test—removing MES/Hypoxia cells and recalculating the spatial coherence score—is conditional on MES/Hypoxia cells being significant, the resulting p-value is not fully independent and should be interpreted as a post hoc test rather than a confirmatory test. A more rigorous approach would be to perform a multiple linear regression including all cell types as predictors of the spatial coherence score. This would provide a direct estimate of each cell type’s unique contribution while adjusting for correlations among them. The authors could retain their original analysis and plots for descriptive purposes, while using the multiple

regression to confirm whether MES/Hypoxia cells remain significantly associated with spatial coherence after accounting for other cell types. This approach would give a statistically more robust assessment of each cell type's independent effect and reduce the risk of inflated type I error from sequential testing.

We thank the reviewer for this point. We agree that multiple linear regression is a robust assessment to determine the effect of multiple variables. In our study, the cell states/ types are **proportions** that add up to one, and as such they are inherently dependent on one another. Thus, for multiple linear regression that accounts for all cell states/types together, we first performed logit transformation followed by multiple linear regression (**Figure R5a**). From this analysis, MES/Hypoxia was still the most significant, as well as embryonic-neuronal-like. After this analysis, we decided that it is generally best practice to transform compositional data for both simple and multiple linear regression, and thus we re-performed simple linear regression with Bonferroni correction with the transformed data. We found that MES/Hypoxia is still significantly positively correlated with spatial coherence, and embryonic-neuronal-like cells are inversely correlated with spatial coherence (**Figure R5b**).

Overall, our conclusion is that MES/Hypoxia is still significantly correlated with spatial coherence score after multiple linear regression. In addition, we uncover embryonic-neuronal-like as a significant inverse correlation that is more evident by logit transformation of the data. We included these new findings into **Results, Discussion, Extended Data Figure 4e** and **Supplementary Table 7**. In addition, we have removed the analysis that removed MES/Hypoxia cells to re-calculate the spatial coherence score as the cell states are dependent on one another, and likely is not the most rigorous test. We thank the reviewer for the very constructive feedback which allowed us to more rigorously examine our methods and consequently uncover further interesting findings related to the spatial organization of ZFTA-RELA tumors.

Figure R5



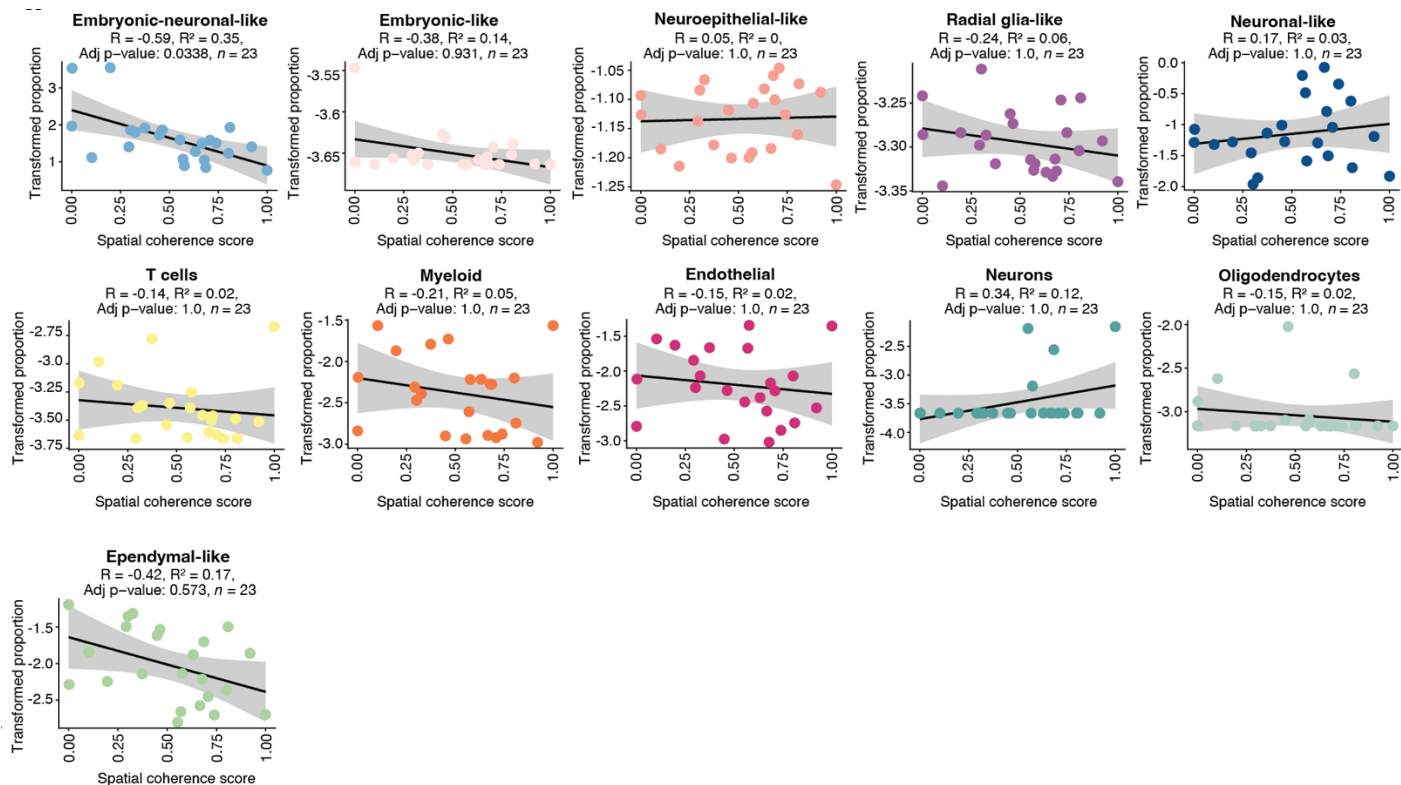


Figure R5. (a) Multiple linear regression of logit transformed proportion of cell states/types to spatial coherence scores. **(b)** Simple linear regression of logit transformed proportion of cell states/types to spatial coherence scores. P-values adjusted by Bonferroni correction.

Results: This analysis revealed that higher spatial coherence scores were positively correlated with the logit transformed proportion of MES/Hypoxia cells ($R = 0.69$, Bonferroni-corrected P-value = 0.00315, $n = 23$, **Figure 3d, Table S7**), while inversely correlated with the logit transformed proportion of embryonic-neuronal-like cells ($R = -0.59$, Bonferroni-corrected P-value = 0.0338, $n = 23$, **Figure 3d, Table S7**). These findings were also corroborated by multiple linear regression analysis (**Table S7**), suggesting that the presence of MES/Hypoxia and embryonic-neuronal-like cell states is specifically associated with *structured* or *disorganized* tumors, respectively. We next calculated the spatial coherence for each cell state/type to determine whether certain cell states/types are more likely to co-localize with themselves (**Extended Data Figure 4f**). MES/Hypoxia cells showed high spatial coherence, suggesting that hypoxia both correlates with global tumor spatial coherence as well as co-localization of MES/Hypoxia cells in distinct local neighborhoods. The inverse was observed for embryonic-neuronal-like cells, suggesting that these cells are more likely to be scattered throughout the tissue. Overall, our results suggest that ZFTA-RELA tumors have distinct global spatial organization patterns, with more organization being highly correlated with the presence of MES/Hypoxia cells, while more disorganization being correlated with embryonic-neuronal-like cells.

Discussion: Despite hypoxia being described as critical for metabolism and epigenetic regulation in posterior fossa ependymomas⁷⁵, its role in ST-EPNs has been underappreciated. It would therefore be of interest for future studies to examine hypoxia-driven tissue structural changes and their functional implications directly in ST-EPN tumor models. In addition, we highlight a novel inverse relationship of embryonic-neuronal-like to global spatial organization, opening new questions regarding the extent of tumor microenvironment re-modeling in the presence of this cell state and its therapeutic significance.

Changes made: Results, Discussion, Extended Data Figure 4e, Supplementary table 7

10. Lines 309–310: The authors state: “We overall identified six recurrent niches, each enriched in distinct malignant cell states or non-malignant cells (Extended Data Figure 6b).” The term “enriched” implies that a

statistical test has been applied, which is not the case here. I suggest using “enhanced” instead when describing the meta-niches to avoid implying statistical significance.

We have amended the language to the reviewer’s suggestion.

Results: We overall identified six recurrent niches, each enhanced in distinct malignant cell states or non-malignant cells (**Extended Data Figure 6b**).

Changes made: Results “ZFTA-RELA tumors are locally organized into recurrent tumor spatial clusters” Section

11. Lines 321–322: Similarly, please update the wording to describe the clusters as “enhanced” rather than “enriched” if no statistical test has been performed. This recommendation applies to all downstream analyses of these clusters.

We thank the reviewer for this suggestion. We have revised “enriched” to “enhanced”, “high proportion of”, or “prevalent in” throughout the manuscript regarding the spatial clusters to avoid implication of statistical tests.

Changes made: Results “ZFTA-RELA tumors are locally organized into recurrent tumor spatial clusters” section, Figure 4 and legends, Extended Data Figure 6 and legends.

12. Lines 328–330: The authors state: “The distinct MES/Hypoxia-enriched spatial cluster underscores the distinct spatial organization of hypoxic regions found by calculating spatial coherence across cell states/types (Figure 3e).” However, comparing Figures 3e and 4e, similar patterns could also be observed for neuroepithelial-like, endothelial, and neuronal-like cells. As written, the current phrasing may overemphasize the MES/Hypoxia cluster and underrepresent other data patterns. I suggest reframing the sentence to acknowledge these additional spatial patterns while still highlighting the MES/Hypoxia finding as appropriate.

We thank the reviewer for this point, and we have removed this sentence for better clarity and flow in this section of the manuscript.

Changes made: Results “ZFTA-RELA tumors are locally organized into recurrent tumor spatial clusters”

13. Extended Data Figure 6f: The plots and analyses appear to be strongly influenced by single outlier points. I suggest reconsidering their inclusion, as they may not reliably represent broader trends.

We thank the reviewer for highlighting the need for clarification. Our analysis compared cell state/type proportions in spatial clusters enhanced for a given cell state/type to those in corresponding cell state/type-enhanced metaniches. Outlier points indicate the cell state/type with the highest proportion in both the spatial cluster and its matched metaniche, such as endothelial cells in the endothelial-enhanced examples. This supports our finding that both metaniches and spatial clusters are frequently dominated by a single cell state/type. As this corroborates our findings, we have included these outlier points and have revised the plot for clarity (**Figure R6a**). We further assessed the robustness of correlations in cell type proportions between metaniche and spatial clusters by repeating the correlation analysis after excluding cell states/types with proportion greater than 0.25 in either metaniche or spatial clusters (**Figure R6b**). Strong positive correlation still existed in four out of five pairs of metaniche and spatial clusters, suggesting that the cell state/type-enriched spatial clusters even without the most dominating cell state/type correlates strongly in proportion to the matched metaniche. These analyses further demonstrate that the two orthogonal methods for identifying local spatial neighborhoods are mutually reinforcing.

Figure R6

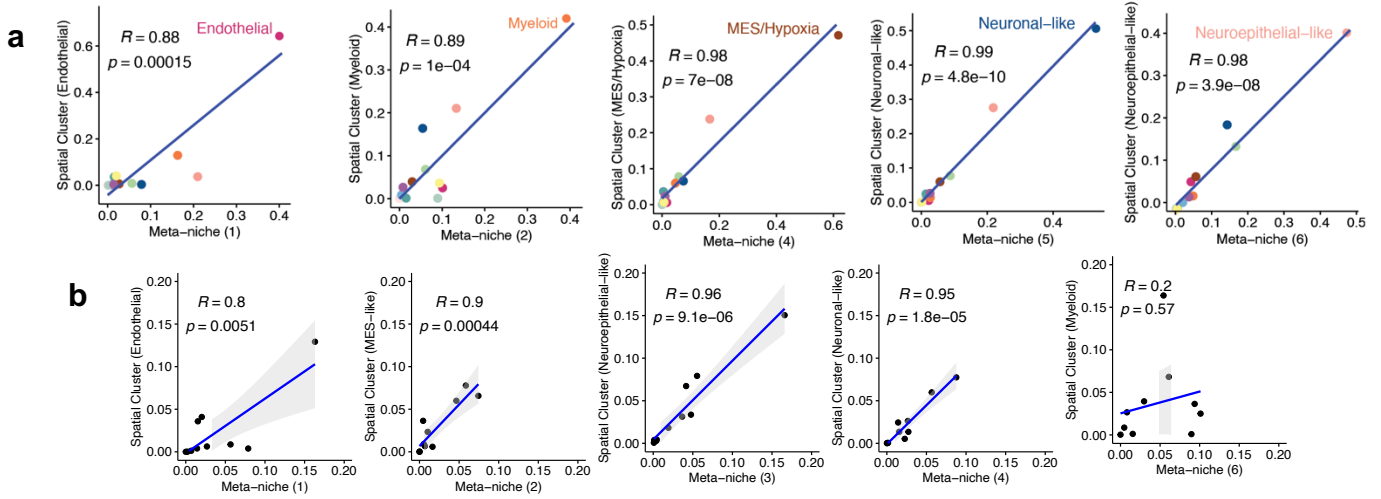
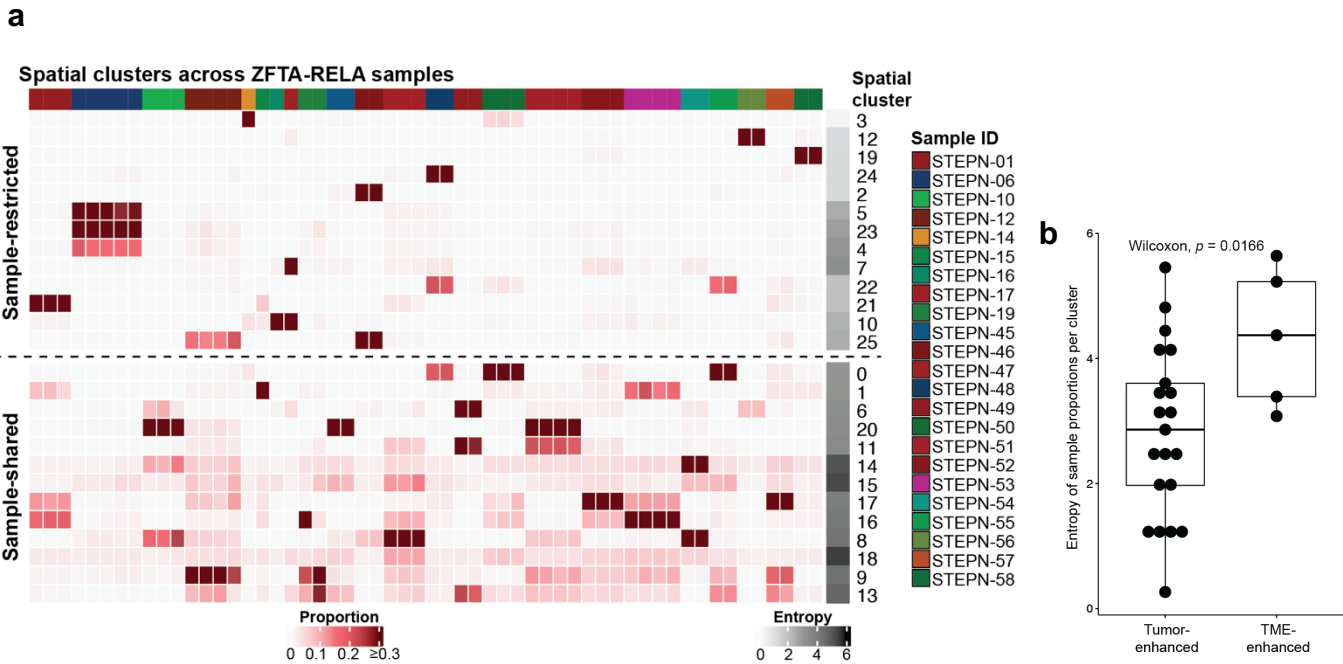


Figure R6. (a) Revised Extended Data Figure 6f. **(b)** Correlation of cell type proportion in CellCharter spatial clusters (y-axis) and in metaniches (x-axis) after cell types with proportion greater than 0.25 in either metaniche or spatial cluster are removed.

Changes made: Extended Data Figure 6f

14. Figure 4d and Extended Data Figure 6g: This is a very interesting analysis! However, in Figure 4d, it appears that spatial clusters 22, 25, and 0 are not sample-specific, whereas cluster 6 may be sample-specific. I suggest revisiting the selection of these clusters, which may require updating the statistical test in Extended Data Figure 6g. Additionally, for consistency, please label the two groups on the x-axis as “enhanced” rather than “enriched.”

We agree with the reviewer that our definition of a “sample-specific” spatial cluster may be misleading, as these are spatial clusters that are only present in a small number (1-2) of samples. We have renamed these clusters as “sample-restricted” to more clearly communicate this point. Spatial cluster 22 and 25 are present in two samples and so are still classified as “sample-restricted”. We also thank the review for pointing out that spatial cluster 0 looks more shared across samples. We have re-checked that this cluster, which is in fact prevalent in three samples, so we have re-allocated this cluster as “sample-shared” (**Figure R7**). As spatial cluster 6 is present in three samples, we have still classified as “sample shared”. In addition, as the statistical test in Extended Data Figure 6g is based on proportions of normal and malignant cell types as well as entropies of sample proportion per cluster, this is **independent** of classifications of spatial clusters. As such, Extended Data Figure 6g is not affected by changes in Figure 4d. Finally, we have revised the x-axis of Extended Data Figure 6g to “enhanced” and adjusted the figure legends accordingly.

Figure R7**Figure R7. (a)** Revised Figure 4d. **(b)** Revised Extended Data Figure 6g

15. #3–19: The authors test whether there is a difference in spatial coherence scores between primary and recurrent tumors. Although the result remains non-significant with the increased number of data points, negative results are still valuable to report. I suggest including a brief statement in the manuscript that this comparison was investigated and that no significant difference was found. This would help provide a complete view of spatial coherence patterns, which is central to one of the main findings of the study.

We have included a brief statement stating that there was no difference in spatial coherence scores between primary and recurrent tumors.

Results: We did not observe a statistically significant correlation between global structural organization and sampling of the tissue (data not shown).

16. #3–28: Upon reviewing the revised manuscript and Figure R46 (Extended Data Figure 7e), it appears that each condition includes only three biological replicates, which is too few to calculate p-values reliably. Treating each single cell as an independent observation would inflate statistical significance. I suggest instead calculating the proportion of Progenitor-like versus Differentiated cells per replicate for each condition (adherent, spheres, coculture, PDX, and patient). The plot could then be presented as a bar plot showing the mean proportion across replicates, with error bars representing variability, and individual replicate points overlaid (n=3). Use descriptive language such as “higher proportion of...” or “predominantly...” rather than statistical terminology. Given the small sample size, avoid using terms like “enrichment” for this analysis.

&

17. #3–29: Given the small number of biological replicates (see #3–28), I suggest softening the conclusion from “Overall, these results highlight how...” to something like “Overall, these results indicate that...”. Additionally, use “enhanced” instead of “enriched” to avoid implying statistical significance.

We thank the reviewer for this suggestion and generated the recommended plot (**Figure R8**). We also removed the chi-square test results in Extended Data Figure 7e, and revised “enriched” to “prevalent in” and “greater abundance of” in **Results** to avoid the term “enrichment” that implies statistical testing.

Figure R8

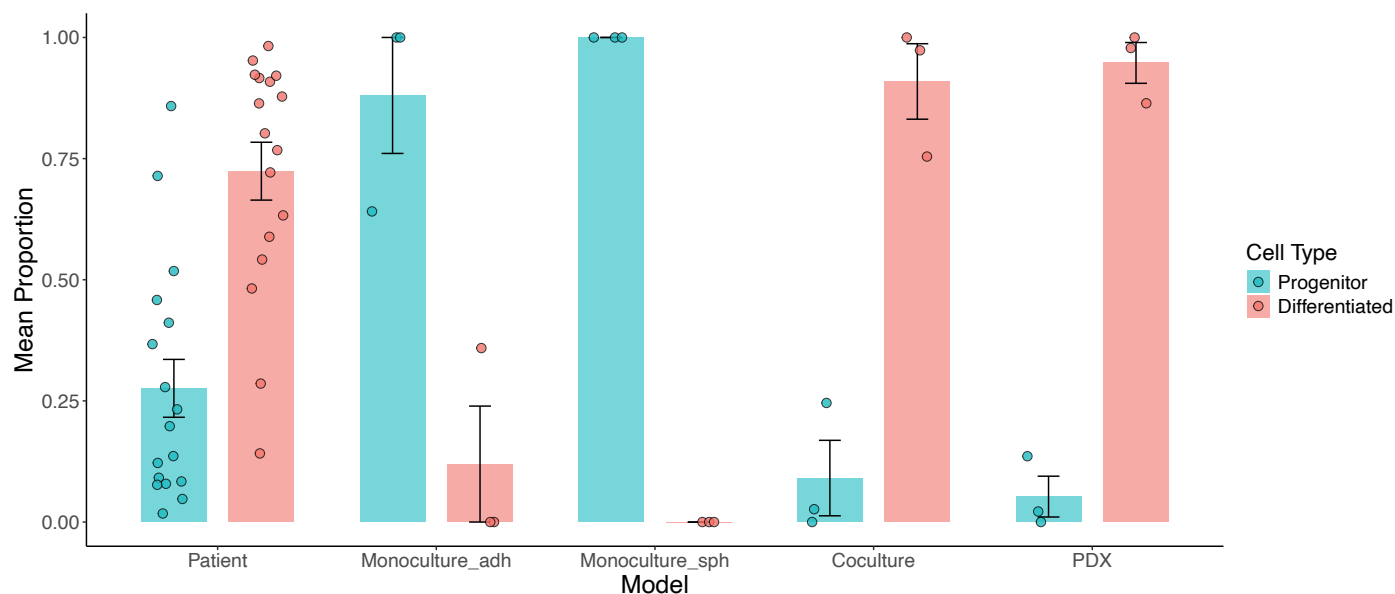


Figure R8. Mean proportion of progenitor (embryonic-like and neuroepithelial-like) and differentiated (neuronal-like and ependymal-like) populations of each cell line (EP1NS, BT165, VBT242) in each condition (monoculture adherent, monoculture spheres, coculture with rat E19 cortical cells, PDX), compared with mean proportion of progenitor-like and differentiated cells in patient samples. Error bars depict SE.

Results: ZFTA-RELA cells grown as monoculture (adherent and spheroid) were prevalent in neuroepithelial-like cells (Extended Data Figure 7b-d), while ZFTA-RELA cells grown in coculture exhibited a greater abundance of neuronal-like cells (Extended Data Figure 7a-d) and neuronal function-related GO terms (Extended Data Figure 7f-g), as well as some early ependymal-like signatures (Extended Data Figure 7b-d).

Methods

18. For reproducibility, please add package version for the different analysis packages used. E.g., Seurat (v.x.x.x). It looks like it is missing for some packages in the methods section.

We have amended the Methods section to include the missing package versions, including Seurat (v5.0.2), SingleCellExperiment (v1.28.1), edgeR (v0.27), clusterProfiler (v4.6.2), UCell (v2.8.0), CellCharter (v0.3.4).

Changes made: Methods

19. Lines 886–888: The authors state: “ST-EPN cells/nuclei were annotated based on the maximum expression score for the respective MP signatures after excluding the one related to cell cycle (‘cycling’).” However, in Figure 2, cycling cells appear in the UMAP in 2c but not in 2d. Could the authors clarify: are these cells excluded from all downstream analyses and only shown in 2c, or are they only excluded in 2d? If they are only excluded in 2d, I suggest including them for consistency across metaprograms. If there is a reason for their exclusion from the main figure, an additional supplementary figure including the cycling cells would be helpful. Also, please clarify whether cycling cells are included or excluded in Extended Data Figures 2e–f.

We thank the reviewer for this comment. We exclude the cycling metaprogram from the scoring during downstream analysis after Figure 2b, as “cycling” is a gene expression signature that likely reflects activity during a physiological process, not necessarily a distinctly stable cell state⁴. To clarify, we do not remove the cells belonging to the cycling program, but rather re-annotate the cells without the cycling gene signature. For consistency, we have amended Figure 2c UMAP (Figure R9) to not include the cycling program.

Figure R9

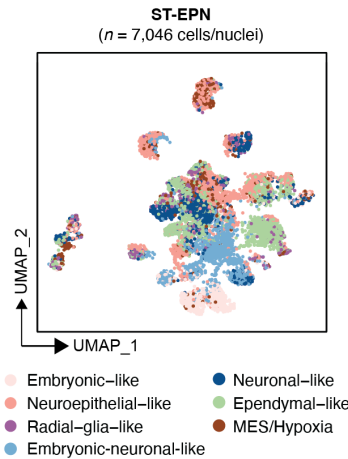


Figure R9. Revised Figure 2c

Changes made: Figure 2d, Extended Data Figure 2 legends

20. Line 1049: I think this could be a typo for the name of the R-package. Should it not be ‘stats’ instead of ‘stast’? I think it is present at more places in the manuscript.

We have corrected the typo in the Methods section.

Methods: We then clustered the resulting scores across all NMF programs by hierarchical clustering as explained in section: Program-wise hierarchical clustering, and used the function *cutree* from the ‘stats’ R-package to automatically extract $n = 15-20$ highly correlated MPs....Finally, we used the function *cutree* from the ‘stats’ R-package to automatically extract $n = 5-6$ highly correlated MNs.

21. Lines 1158–1161: The statistical strategy should be revised. From the methods, it appears that the Kruskal–Wallis test may not have been applied correctly (e.g., Figure R55). The Kruskal–Wallis test is a global test that only indicates whether at least one group differs; it does not specify which groups are different. To identify specific group differences, a significant Kruskal–Wallis result should be followed by post hoc pairwise Wilcoxon rank-sum tests with appropriate multiple-comparison correction. A recommended framing would be: “If one or more groups did not follow a normal distribution (Shapiro–Wilk p-value < 0.05), we used the non-parametric Kruskal–Wallis test for comparisons across more than two groups. If the Kruskal–Wallis test was significant, we performed post hoc pairwise comparisons using the Wilcoxon rank-sum test with appropriate (e.g., Bonferroni) multiple testing correction.”

We thank the reviewer for highlighting the lack of clarity in our approach. To clarify, we performed Kruskal–Wallis test first to check whether there was at least one group that differed from the rest (**Figure R10** top), and after finding that the test was significant, we performed Dunn’s multiple comparisons test as a post hoc test to determine which specific groups differed from each other (**Figure R10** bottom). We have revised the Methods section to include this approach, and we thank the reviewer for their constructive feedback.

Figure R10

Kruskal-Wallis test	
P value	<0.0001
P value summary	****
Do the medians vary significantly? (P < 0.05)	Yes
Number of groups	4
Kruskal-Wallis statistics	114.5

Comparison	Mean rank diff.	Significant?	Summary	Adjusted P Value
Neuroepithelial-like-1 vs. Neuroepithelial-like-2	-59.76	Yes	**	0.0046
Neuroepithelial-like-1 vs. Immature ependymal-like	33.77	No	ns	0.2864
Neuroepithelial-like-1 vs. Neuronal-like	-70.52	Yes	***	0.0006
Neuroepithelial-like-2 vs. Immature ependymal-like	93.52	Yes	****	<0.0001
Neuroepithelial-like-2 vs. Neuronal-like	-10.77	No	ns	>0.9999
Immature ependymal-like vs. Neuronal-like	-104.3	Yes	****	<0.0001

Figure R10. (top) Kruskal-Wallis test results across all cell states (Neuroepithelial-like-1, neuroepithelial-like2, immature ependymal-like, neuronal-like) **(bottom)** Dunn’s multiple comparison test results for comparisons between cell migration speed values of different cell states

Methods “Statistical Analysis”: If one or more groups did not follow a normal distribution (Shapiro-Wilk P-value < 0.05), we used the non-parametric Kruskal-Wallis test for comparisons across more than two groups. If the Kruskal-Wallis test was significant, we performed post hoc pairwise comparisons using the Wilcoxon rank-sum test with appropriate (e.g. Bonferroni test) multiple testing correction, or the Dunn’s multiple comparison test.

Changes made: Methods “Statistical Analysis” section

22. Extended Data Figure 1d (Figure R26): Please define the abbreviations CNA and NM in the figure legend. If CNA and CNV refer to the same concept, consider using only one term for consistency and easier interpretation.

We have revised “CNA” to “CNV” in the Extended Data Figure 1d for consistency (**Figure R11**) and have defined “NM” as “non-malignant” in Extended Data Figure 1d legends.

Figure R11

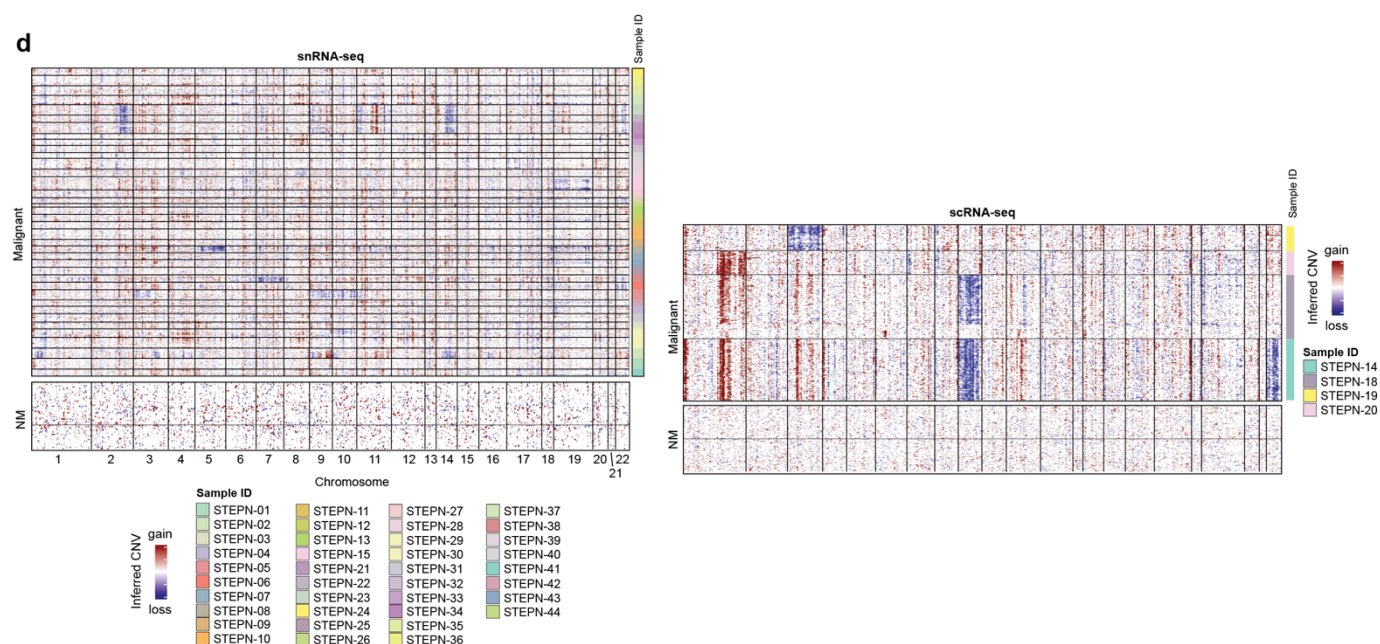


Figure R11. Revised Extended Data Figure 1d

Changes made: Extended Data Figure 1d and legends

23. I was unable to access the deposited datasets. For publication, just a reminder to please make sure to include the annotation metadata of all the metaprograms and meta-niches the spatial and sc/snRNAseq data was annotated with in the final manuscript version.

We thank the reviewer for this reminder, and we have amended the GEO deposited datasets to also include the metaniche and *CellCharter* spatial cluster annotations. The GEO accession number is **GSE260452** that will be made public if the manuscript is accepted.

References

1. Phipson, B. *et al.* propeller: testing for differences in cell type proportions in single cell data. *Bioinformatics* 38, 4720–4726 (2022).
2. Cebrian-Silla, A. *et al.* Neural stem cell relay from B1 to B2 cells in the adult mouse ventricular-subventricular zone. *Cell Rep* 44, 115264 (2025).
3. Leca, I., Phillips, A. W., Ushakova, L., Cushion, T. D. & Keays, D. A. Codon modification of Tuba1a alters mRNA levels and causes a severe neurodevelopmental phenotype in mice. *Sci Rep* 13, 1215 (2023).
4. Tirosh, I. & Suva, M. L. Cancer cell states: Lessons from ten years of single-cell RNA-sequencing of human tumors. *Cancer Cell* 42, 1497–1506 (2024).

RE: Nature manuscript 2024-08-16431

“Multidimensional profiling of heterogeneity in supratentorial ependymomas”

Point-by-point response to reviewer comments

We thank the reviewers for their thoughtful feedback to the manuscript. Please see below for response to Reviewer #3’s remaining comment.

Referee #3 (Remarks to the Author):

Dear Authors,

Thank you for the revised version of the manuscript, and the response to my feedback. For me, only one small concern remains before the manuscript is suitable for publication.

Line 260. The authors added the sentence “We did not observe a statistically significant correlation between global structural organization and sampling of the tissue (data not shown).”. However, as it reads now, it is not clear what this text associates to, and in the worst case could be interpreted as evidence against some of the authors main conclusions. It should be clarified that this refers to differences in spatial coherence scores between primary and recurrent tumors, (which is an important negative result to report).

We thank the reviewer for this feedback. We have changed the language for clarity.

Results: We did not observe a statistically significant difference in the global structural organization between primary and recurrent tumors (data not shown).