

A spatiotemporal cancer cell trajectory underlies glioblastoma heterogeneity

Grant de Jong^{*1}, Fani Memi^{*1}, Tannia Gracia^{*1}, Olga Lazareva^{*1,2,3}, Oliver Gould¹, Alexander Aivazidis¹, Manas Dave¹, Qianqian Zhang^{4,5}, Melanie Jensen⁶, Ahmet Sureyya Rifaioglu⁷, Joao D. Barros-Silva⁶, Sabine Eckert¹, Di Zhou¹, Yvette Wood¹, Elizabeth Tuck¹, Sezgin Er^{1,8}, Henry Marshall¹, Kenny Roberts¹, Andrew L. Trinh¹, Shreya Rai¹, Tyler Shaw^{9,10}, Agnes Oszlanczi¹, Hayden Powell¹, Robert Petryszak¹, Zoi Katsirea¹, Irfan Mamun¹, Ilaria Mulas¹, Annelies Quaegebeur^{10,11}, Mayen Briggs⁹, Stanislaw Makarchuk¹, Jessica Cox¹, Jimmy Tszy Hang Lee¹, Laura Rueda^{12,13,14}, Manu Saraswat^{2,3,12,13,14}, Harry Bulstrode⁴, Adam Young⁴, Minal Patel¹, Tarryn Porter¹, Elena Prigmore¹, Moritz Mall^{12,13,14}, Julio Saez-Rodriguez^{7,15}, James Briscoe⁶, David H. Rowitch^{4,5}, Richard Mair^{9,10}, Sam Behjati^{1,5,11#}, Oliver Stegle^{1,2,3,#}, Omer Ali Bayraktar^{1,#}

¹Wellcome Sanger Institute, Cambridge, CB10 1SA, UK.

²Division of Computational Genomics and Systems Genetics, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany.

³European Molecular Biology Laboratory, Genome Biology Unit, Heidelberg 69117, Germany

⁴Cambridge Stem Cell Institute, University of Cambridge, Cambridge, UK.

⁵Department of Paediatrics, University of Cambridge, Cambridge, UK.

⁶The Francis Crick Institute, 1 Midland Road, London NW1 1AT, UK.

⁷Heidelberg University, Faculty of Medicine, and Heidelberg University Hospital, Institute for Computational Biomedicine, Heidelberg, Germany.

⁸International School of Medicine, Istanbul Medipol University, Istanbul, Marmara, Turkey.

⁹Cancer Research UK Cambridge Institute, University of Cambridge, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, UK.

¹⁰Department of Clinical Neurosciences, Box 165, Cambridge Biomedical Campus, Cambridge, CB2 0QQ, UK.

¹¹Cambridge University Hospitals NHS Foundation Trust, Cambridge, CB2 0QQ, UK.

¹²Cell Fate Engineering and Disease Modeling Group, German Cancer Research Center (DKFZ) and DKFZ-ZMBH Alliance; 69120 Heidelberg, Germany.

¹³HITBR Hector Institute for Translational Brain Research gGmbH; 69120 Heidelberg, Germany.

¹⁴Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University; 68159 Mannheim, Germany.

¹⁵European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Hinxton, Cambridgeshire, UK.

*equal contribution

#Correspondence: sb31@sanger.ac.uk (S.B.), oliver.stegle@embl.de (O.S.) and ob5@sanger.ac.uk (O.A.B.)

51 Abstract

52 **Cancer cells display highly heterogeneous and plastic states in glioblastoma, an**
53 **incurable brain tumour. However, how these malignant states arise and whether they**
54 **follow defined cellular trajectories across tumours is poorly understood. Here, we**
55 **generated a deep single cell and spatial multi-omic atlas of human glioblastoma that**
56 **pairs transcriptomic, epigenomic and genomic profiling of 12 tumours across multiple**
57 **regions. We identify that glioblastoma heterogeneity is driven by spatially-patterned**
58 **transitions of cancer cells from developmental-like states towards those defined by a**
59 **glial injury response and hypoxia. This cellular trajectory regionalises tumours into**
60 **distinct tissue niches and manifests in a molecularly conserved manner across**
61 **tumours as well as genetically distinct tumour subclones. Moreover, using a new deep**
62 **learning framework to map cancer cell states jointly with clones *in situ*, we show that**
63 **tumour subclones are finely spatially intermixed through glioblastoma tissue niches.**
64 **Finally, we show that this cancer cell trajectory is intimately linked to myeloid**
65 **heterogeneity and unfolds across regionalised myeloid signalling environments. Our**
66 **findings define a stereotyped trajectory of cancer cells in glioblastoma and unify**
67 **glioblastoma tumour heterogeneity into a tractable cellular and tissue framework.**

68

69 Main

70 The extensive tumour heterogeneity of glioblastoma (GB), an incurable adult brain
71 cancer, presents a major obstacle to treatment. Nearly two decades of molecular profiling has
72 found that malignant cells exist across a spectrum of cellular states within each tumour,
73 resembling developmental neural progenitors as well as mesenchymal-like and injury
74 response states^{1–5}. These states are suggested to vary in abundance across tumours and
75 stratify GBs into subtypes^{1,2}. Furthermore, patient derived xenografts have shown that
76 malignant cells can transition across cellular states, indicating their plasticity^{2,6}. However,
77 despite these extensive profiling and modelling efforts, we still do not understand how cancer
78 cells diversify into distinct states and give rise to GB heterogeneity. In particular, how
79 malignant cells undergo state transitions and whether they follow stereotyped or variable
80 cellular trajectories across GB tumours is unknown. A cancer cell trajectory that is conserved
81 and tractable across GBs would present a highly attractive therapeutic target.

82 Single cell, bulk and spatial transcriptomics have revealed stereotypical aspects of
83 GBs. Multiple malignant cell states can be detected within each tumour^{2,5,7} and their single cell
84 transcriptional profiles can be mapped onto an axis between dev-like and injury response-
85 associated gene expression programs^{4,8}. Furthermore, these states show a recurrent spatial
86 organisation across tumours^{5,9}. Yet, these observations have not been linked together to
87 resolve malignant cellular trajectories in GB, in part due to the lack of multi-modal profiling
88 efforts that can integrate insights from various cell and tissue atlasing technologies.

89 To chart the cellular trajectories of malignant cells in GB, it is necessary to account for
90 their genetic heterogeneity. Whole genome and exome sequencing has identified multiple
91 phylogenetically distinct subclonal lineages within individual GBs^{5,7}. Whereas single cell and
92 spatial transcriptomics suggest diverse malignant cell states can arise within each subclone^{2,9},
93 these efforts inferred subclones from transcriptomic data and do not incorporate “ground truth”
94 somatic cancer mutations identified at the genomic DNA level. Hence, the subclonal origins of

95 diverse malignant cell states and whether GB intratumoural genetic heterogeneity can alter
96 these states is not well understood.

97 Beyond malignant cells, GB heterogeneity extends to the tumour microenvironment
98 (TME). Myeloid cells are the most abundant TME cell type in GB^{2,10}, underlying a highly
99 immunosuppressive environment linked to poor clinical outcomes^{10–13}. While macrophage-
100 derived signals were previously shown to alter GB malignant cell states¹⁴, recent studies have
101 identified highly diverse myeloid subtypes in tumours^{15–17}. How the heterogeneity of malignant
102 and myeloid cells relate to one another in GBs and how distinct myeloid subtypes influence
103 the pathobiology of diverse malignant cell states is not well understood.

104 Here, we mapped the cellular, clonal and tissue architecture of GB heterogeneity by
105 deep multi-region profiling that integrates single cell multi-omics, spatial transcriptomics and
106 spatial whole genome sequencing. We identified a malignant cellular trajectory that is spatially
107 patterned in tumours, that is conserved across tumours and spatially-intermixed subclones
108 within each tumour, and that is accommodated by a regionalised myeloid signalling
109 environment. Hence, we find that a stereotyped trajectory of cancer cells gives rise to GB
110 heterogeneity and we provide a cell and tissue framework that unifies GB tumour
111 heterogeneity.

112

113 Results

114 Deep spatial multi-omic profiling of GB

115 To chart the cellular and tissue organisation of GB, we generated a single cell and
116 spatial multi-omic atlas of 12 isocitrate dehydrogenase wildtype (IDH-wt) primary GB tumours
117 (**Fig. 1A, Table S1**). We combined three major approaches to comprehensively profile GB
118 heterogeneity, extending previous surveys^{2,5,8,9,18}.

119 First, to capture intra-tumour heterogeneity, we profiled each tumour across multiple
120 sites. Our sampling ranged from 4 sites (~1.25 cm³ each) for the majority of cases to 15 sites
121 that effectively covered one whole tumour (**Extended Data Fig. 1A, 1B**), profiling 57 tumour
122 sites in total.

123 Second, to account for inter- and intra-tumour heterogeneity during multi-modal data
124 integration, we generated paired single nuclei and spatial omic data from each tumour and
125 sampling site. We profiled consecutive tissue sections from each site using single nuclei joint
126 transcriptome- and chromatin accessibility-sequencing (“multiome” snRNA+ATAC-seq) and
127 Visium spatial transcriptomics (ST) (10X Genomics) (**Fig. 1A**). Additionally, we assayed select
128 sites with Xenium ST and immunohistochemistry as orthogonal spatial approaches, as well as
129 whole genome sequencing of tumour microbiopsies to interrogate GB clonal architecture
130 (**Extended Data Fig. 1B**).

131 Third, to enable comprehensive analysis at the level of each tumour and tumour
132 subclone, we profiled large numbers of nuclei and multiple spatial tissue sections per site. We
133 generated over 1 million single nuclei multiome profiles and over 0.3 million spatially resolved
134 transcriptomes across 97 Visium sections (**Extended Data Fig. 1, Table S2**). The resulting
135 “GBM-space” dataset, accessible on a user-friendly interactive webportal¹⁹
136 (www.gbm-space.org/), enables comprehensive multi-modal characterisation of GB
137 architecture and presents a unique holistic tumour atlas for cancer research.

138 **Malignant cells vary from developmental-like to gliosis and hypoxia states**

139 To investigate GB heterogeneity in our atlas, we first annotated malignant and TME
140 cell states in our single nuclei multiome data. We performed standard data processing
141 workflows and filtered nuclei based on quality of both RNA and ATAC profiles, resulting in
142 1,025,329 nuclei with over 85,000 nuclei per tumour on average (**Methods, Fig. 1A and**
143 **Extended Data Fig. 2**). We then focused on the snRNA-seq data to annotate transcriptomic
144 cell states, and examined the ATAC data to map GB gene regulatory networks in a companion
145 study²⁰.

146 We integrated the snRNA-seq data using scVI²¹ to account for batch variations such
147 as 10X reaction, tumour and tumour site (**Methods**) (**Extended Data Fig. 2A**). We then
148 distinguished malignant clusters based on inferred copy number (CN) gain of chromosome 7
149 and loss of chromosome 10, as well as enrichment of known GB expression programs
150 compared to TME cells² (**Fig. 1B, Extended Data Fig. 3B,C**). Finally, subclustering yielded
151 28 malignant clusters, including 10 highly common clusters present across 50% of the
152 tumours, as well as 71 TME clusters (**Fig. 1C, Extended Data Fig. 4A,B**).

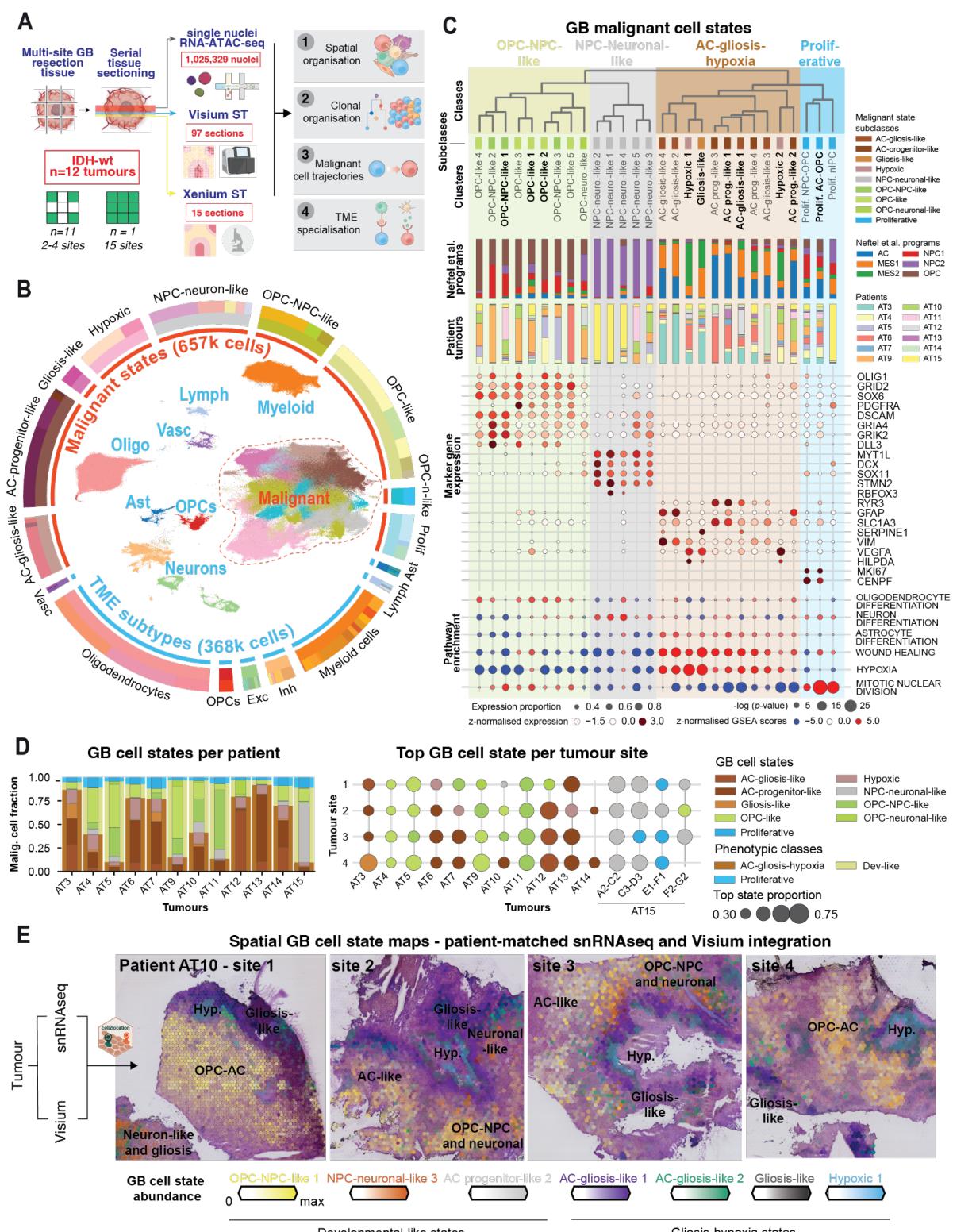
153 We annotated malignant clusters by 1) quantifying their expression of GB cell state
154 gene programmes described by Neftel et al² (**Extended Data Fig. 5A**), 2) relating them to the
155 developing human brain cell types via label transfer from a scRNA-seq atlas²² (**Extended Data**
156 **Fig. 5B**) and 3) gene set (MSigDB) enrichment analysis (**Extended Data Fig. 5C, Table S5**).
157 As a result, we hierarchically grouped malignant clusters into 4 major classes and 9
158 subclasses of GB cell states that were each observed across multiple tumours (**Fig. 1C,D,**
159 **and Extended Data Fig. 4A, Table S4**).

160 Two major classes of GB cell states resembled distinct developmental progenitors.
161 The first developmental-like (dev-like) state, oligodendrocyte precursor- and neural progenitor-
162 like (OPC-NPC-like) cells, showed universally high expression of the “OPC-like” Neftel et al
163 programme and varying degrees of the “NPC1-like” programme (**Fig. 1C**). They were most
164 similar to developing human brain OPCs, expressing OPC markers (*PDGFRA*, *OLIG1*, *SOX6*)
165 and gene sets (**Fig. 1C and Extended Data Fig. 5B**). OPC-like clusters resembled early
166 progenitors (i.e. OPC-like 2) differentiating towards myelinating oligodendrocytes (i.e. OPC-
167 like 4), whereas OPC-NPC- and OPC-neuronal-like clusters additionally expressed neuroblast
168 and neuron-related genes (*GRIK2*, *DLL3*, *GRIA4*) (**Fig. 1C and Extended Data Fig. 5B**).

169 The second dev-like state, neuronal progenitor- and neuronal-like (NPC-neuronal-like)
170 cells, highly expressed the “NPC2-like” Neftel et al programme and were most similar to
171 developmental neuroblasts and neurons (**Fig. 1C, and Extended Data Fig. 5A,B**). They
172 expressed neuronal development genes (*MYT1L*, *STMN2*) and pathways, resembling early
173 neuroblast (e.g. *SOX11*⁺ *DCX*⁺ NPC-neuronal-like 2) differentiation towards neurons (i.e.
174 NPC-neuronal-like 1 with elevated synaptic gene expression) (**Fig. 1C and Extended Data**
175 **Fig. 5B**). These were the least abundant states, consistent with the low frequency of neuronal-
176 like “Neural” GB tumour subtypes¹.

177 The third major GB cell state exhibited a transcriptomic spectrum from astrocyte
178 progenitor (AC-prog)-like to gliosis-like (i.e glial injury response) and hypoxic states. These
179 states broadly expressed core astrocyte markers (*SLC1A3*, *GFAP*) (**Fig. 1C**). The AC-prog-
180 like cells were enriched for the Neftel “AC-like” programme and were most similar to
181 developmental glioblasts (**Fig. 1C and Extended Data Fig. 5A**). In AC-gliosis- and gliosis-like
182 cells, this developmental signature was diminished and replaced by the expression of genes
183 related to wound healing and injury response (*VIM*, *SERPINE1*) and

184



185

186 **Figure 1: Spatially resolved single cell multiomic atlas of GB tumour heterogeneity.**

187 **A)** Schematic illustration of experimental design and primary aims of this study.

188 **B)** Overview of the cellular composition of our GBM-space atlas. UMAP visualisation of the
189 snRNA-seq profiles of malignant and TME cells outlining the major constituent populations,
190 and a ring plot illustrating the relative distribution of cells at three levels of granularity:

191 malignant (657,275 cells) versus TME (368,054 cells) status, coarse-level populations (text
192 labels around the outer ring), and fine grained cell states.

193 **C)** Dendrogram depicting transcriptomic clustering of malignant cell states (top) and their
194 biological characteristics, including their hierarchical annotation, distribution of top Neftel et al.
195 expression signatures, distribution of cell states across patient tumours, marker gene
196 expression, and significantly enriched MSigDB pathways. Malignant clusters present across
197 more than 50% of the tumours are shown in bold.

198 **D)** Intra- and inter-tumour GB cell state heterogeneity across our dataset. The barplot (left)
199 shows cell state distribution summarised across each tumour. The dotplot (right) depicts the
200 top cell state for a given sampling site within a given tumour (colour) and its relative abundance
201 (size).

202 **E)** Visium ST sections from each sampling site of tumour AT10 coloured according to GB cell
203 state abundances derived from cell2location mapping. The top seven most frequently
204 observed states at the tumour-level were visualised, with colour intensity proportional to
205 inferred cell state abundance. Broader trends, particularly areas of cell state co-localisation
206 are highlighted in bold text.

207
208 angiogenesis (VEGFA) (**Fig. 1C**). The remaining cells were characterised by hypoxia
209 (HILPDA) (**Fig. 1C**).

210 The gliosis-like and hypoxic cells were enriched for the Neftel “MES1” and “MES2”
211 programmes, respectively, that were initially described as mesenchymal-like states^{1,2}. Recent
212 studies have interpreted these GB states as similar to reactive astrocytes responding to injury⁴
213 and glial wound healing signatures^{4,8}. Here, we adopted the gliosis and hypoxia nomenclature
214 that better reflected malignant cellular trajectories in our dataset (see sections below).

215 The fourth and final major malignant state represented proliferative cells. While
216 dominated by a proliferation gene expression signature (**Fig. 1C**), these cells also expressed
217 dev-like or gliosis-hypoxia programmes (**Fig. 1C**), as previously observed², suggesting that
218 both dev-like and gliosis-hypoxia states proliferate in GBs.

219 Our snRNA-seq atlas also captured all major cell lineages in the GB TME (**Extended**
220 **Data Fig. 4B and Extended Data Fig. 6**). OPCs, mature oligodendrocytes and myeloid cells
221 were the most abundant, with the latter including diverse subtypes (described in detail in the
222 final section). We finely annotated neurons, including 9 cortical layer-specific excitatory and 8
223 inhibitory subtypes. Astrocytes, while rare, included several homeostatic and reactive
224 populations. Lymphocytes were similarly rare, yet spanned dendritic cells, NK cells, B cells,
225 plasma cells and multiple T cell subtypes including CD8+, CD4+, proliferative and regulatory
226 T cells. Finally, we also captured vasculature spanning pericytes, vascular leptomeningeal
227 cells (VLMCs), and multiple endothelial subtypes, the latter including chr7&10 CN+ cells that
228 may represent tumour derived vasculature²³.

229 Taken together, our single nuclei transcriptomic atlas presents a deep, fine-grained
230 cellular census of malignant and TME states in GB.

231 Individual GB tumours span dev-like to gliosis and hypoxia states

232 Next, we examined the inter- and intra-tumour heterogeneity of malignant cell states
233 in snRNA-seq. Most tumours showed a dominant GB cell state across the major classes of
234 OPC-NPC-like (e.g. AT5 and 9), NPC-neuronal-like (e.g. AT15) or AC-gliosis-hypoxia (e.g.
235 AT3 and 6) (**Fig. 1D, left**). Yet, within these groupings, tumours showed highly variable

236 subclass composition (e.g. AC-prog-like states enriched in AT6 versus the full spectrum of
237 AC-gliosis-hypoxia in AT3) (**Fig. 1D**).

238 Importantly, despite their stratification by GB cell states, each tumour contained
239 substantial fractions of both dev-like and AC-gliosis-hypoxia cells. Two tumours (AT4 and 10)
240 showed near equal proportions of both classes, while in others, the minority states represented
241 5 to 21% of all malignant cells (**Fig. 1D, Table S4**). While other studies have also noted the
242 stratification of GBs by their cell state composition^{2,4,24}, our deep profiling is the first to identify
243 the presence of both dev-like and AC-gliosis-hypoxia malignant states within every tumour.

244 At the intra-tumour level, individual sampling sites also contained a mixture of GB cell
245 states that broadly reflected the overall cellular composition of the parent tumour (**Extended**
246 **Data Fig. 7B,C**). Yet, we could often identify at least one site per tumour with a unique cell
247 state composition. In a striking example, tumour AT10 displayed distinct dominant GB cell
248 states at each site (**Fig. 1D, right**). Consistent with other studies^{5,7,18}, these observations
249 highlight regional GB heterogeneity and show that single tumour biopsies can be significantly
250 confounding to stratify GB tumour phenotypes.

251 **Spatial segregation of dev-like and gliosis-hypoxia states**

252 Next, we examined the spatial organisation of GB cell states. While previous studies
253 have applied ST to GB tumours, they could not distinguish fine cellular states as they relied
254 on either the coarse-resolution Visium assay (i.e. 55 µm diameter Visium spots sampling
255 multiple cells per location)⁹ or reference-free deconvolution of Visium into gene expression
256 programs that, for example, could not delineate between neuronal-like cancer cells and non-
257 malignant neurons¹⁸.

258 Here, we leveraged computational integration of our multi-modal atlas and
259 deconvolved fine GB cell states defined by snRNA-seq in Visium ST data (n=97 sections,
260 338,481 spot transcriptomes) (**Fig. 1E**). While integrating paired snRNA-seq and ST from each
261 tumour accounted for GB heterogeneity, it also enabled us to link malignant transcriptional
262 trajectories, subclonal organisation and TME signalling that we subsequently identified in
263 snRNA-seq to tumour spatial organisation.

264 We mapped malignant and TME clusters from each patient's snRNA-seq profile to their
265 matched Visium data using cell2location²⁵ (**Fig. 1E**). This revealed striking spatial segregation
266 of GB cell states across all tumours. In particular, malignant cells associated with gliosis and
267 hypoxia segregated from tumour areas enriched for the dev-like states including OPC-, NPC-
268 neuronal- and AC-prog-like cells (**Fig. 1E, Extended Data Fig. 8A**). Broadly, dev-like GB
269 states tended to spatially intermix, whereas tumour regions dominated by gliosis-hypoxia
270 states had more homogeneous cell composition (**Extended Data Fig. 9**), as previously
271 shown¹⁸. Our spatial mapping suggested the cellular composition of GB tumours are more
272 similar to each other than measured with snRNA-seq, with some tumours showing elevated
273 hypoxic cell states in ST that may be lost during tissue dissociation prior to snRNA-seq
274 (**Extended Data Fig. 7D**). Finally, our ST analysis also captured regional heterogeneity of GB
275 cell states across different tumour sampling sites. For example, tumour AT10 site 1 was
276 enriched for OPC-NPC-like states whereas site 3 was dominated by AC-gliosis-hypoxia (**Fig.**
277 **1E and Extended Data Fig. 7C**).

278 **Stereotyped spatial transitions of malignant cell states**

279 To assess the conservation of GB spatial architecture, we examined whether
280 malignant and TME cell states form tissue niches in a stereotyped manner across tumours

281 (Fig. 2A). To do this, we first identified spatially co-localised cell states (i.e. niches) per tumour
282 via non-negative matrix factorisation of cell2location results²⁵ (**Methods**). This resulted in 192
283 factors summarising individual tissue niches across our cohort (**Table S7**). We then clustered
284 these factors by their cell state similarity and identified 14 major tissue niches that recurrently
285 appear across tumours (Fig. 2B).

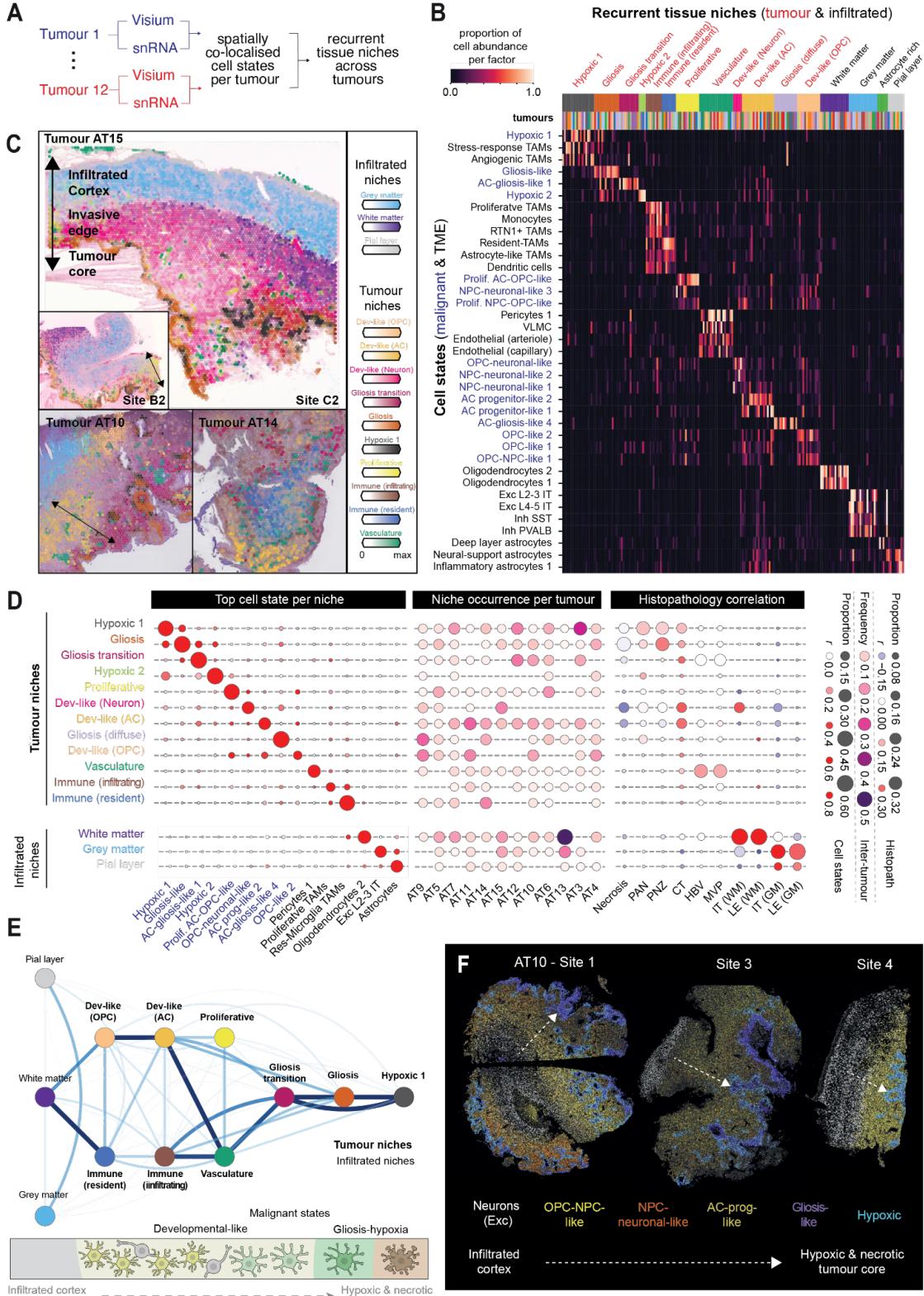
286 We annotated the recurrent GB niches based on their cellular composition and
287 anatomical location (Fig. 2B, 2C). “Tumour” niches represented malignant, immune and
288 vasculature cell states located in the tumour core, in contrast to “infiltrated” niches of non-
289 malignant neurons and glia in the surrounding infiltrated regions (Fig. 2B,C). Most tumour
290 niches were defined by their distinct malignant cell state composition (Fig. 2D). *Dev-like* niches
291 were enriched for either OPC-, NPC-neuronal- or AC-prog-like GB states, often containing
292 multiple spatially co-localised subclusters (e.g. OPC-like 1 and 2) (Fig. 2B, 2D). In contrast,
293 *Gliosis transition*, *Gliosis* and *Hypoxia* niches were dominated by a single respective malignant
294 cluster (Fig. 2B, 2D, and Extended Data Fig. 10A).

295 The majority of tumour niches were observed across multiple donors (Fig. 2D and
296 Extended Data Fig. 10C). Core malignant niches such as *Dev-like* (OPC), *Dev-like* (AC),
297 *Proliferative*, *Gliosis transition*, *Gliosis*, and *Hypoxic 1* were detected in over 50% of tumours
298 (Fig. 2D), while others such as *Dev-like* (Neuron) and *Hypoxic 2* were rare. Crucially, 8 out of
299 12 tumours (e.g. AT3, AT10) showed nearly the full complement of core malignant niches (Fig.
300 2D), suggesting these features are fundamental to GB organisation.

301 Previous ST studies have noted stereotyped anatomical organisation of GB cell states,
302 locating hypoxic mesenchymal-like states at the necrotic core and dev-like states at the
303 infiltrating edge of tumours^{9,18}. To correlate our tissue niches with GB anatomy, we annotated
304 histologically distinct anatomical features on the Visium H&E images according to the Ivy
305 Glioblastoma Atlas (IvyGAP) criteria²⁶ (n=39 Visium sections) (**Methods**, Extended Data Fig.
306 10B). *Dev-like* niches occurred in the cellular tumour (CT), the major tumour compartment
307 bordering the infiltrating tumour (IT) and leading edge (LE) (Fig. 2C, 2D). In contrast, niches
308 associated with gliosis and hypoxia were localised further from the infiltrating regions and
309 formed distinct zones around necrotic areas. *Gliosis transition* (i.e. enriched for AC-gliosis
310 states) and *gliosis* niches were found in the perinecrotic zone (PNZ) at the CT border (Fig.
311 2D, mid panel), forming spatially distinct layers in some tumours (Fig 2C). *Hypoxia* niches
312 were located even deeper in tumours, extending into areas marked by pseudopalisading cells
313 and necrosis (PAN) and necrosis (Fig. 2C,D, and Extended Data Fig. 10B).

314 Our observations thus show that GB is organised into layered tissue niches, consistent
315 with recent findings¹⁸. By incorporating snRNA-seq definitions of malignant cell identities into
316 spatial mapping, our approach further suggested that malignant cell states show stepwise
317 spatial transitions across niches, most evident in tissue areas with gliosis and hypoxia. To
318 quantify their precise spatial arrangement, we calculated the pairwise spatial proximities of GB
319 tissue niches across all tumours (n=97 Visium sections) and summarised them in a network
320 graph (**Methods**).

321 We observed that GB tissue niches are ordered along a major anatomical axis
322 between the infiltrating tumour edge and the necrotic tumour core (Fig. 2E). Amongst core
323 malignant niches, *Dev-like* (OPC) was found closest to infiltrated white matter. *Dev-like* (AC)
324 spatially overlapped with the *Dev-like* (OPC), yet extended deeper into the tumour towards



325

Figure 2: Recurrent tissue niches across GBs indicate stereotyped spatial transitions of malignant cell states.

328 A) Workflow for identification of GB spatial tissue niches that recur across tumours.

B) Distinct cellular compositions of recurrent GB tissue niches. The heatmap depicts the proportion of cell state abundance associated with a given tumour-specific NMF factor. Factors

331 (columns) are ordered by their corresponding spatial niche clusters (top colour bar). The most
332 frequently observed cell states for each niche are shown per row, where black and blue font
333 colours represent TME and malignant cells, respectively.

334 **C) Spatial niche abundances for several Visium ST sections showing key niche distribution**
335 **patterns in an intra-tumoural context (AT15) and across different tumours.**

336 **D) Spatial niche summaries.** Left: Most frequently observed cell state (column) found within
337 each niche (row). Dotplot shows the proportion of the top cell state per spatial niche averaged
338 across tumours (size) and the average Pearson's correlation coefficient (r) across tumours.
339 Centre: Relative frequency of each niche across tumours (size and colour). Right: Correlation
340 of spatial niches with histopathological features annotated on Visium H&E images based on
341 IvyGAP annotations (columns). Dotplot shows the proportion of spots associated with a spatial
342 niche which overlap with a given histological feature (size) and the point biserial correlation
343 coefficient (r) between feature annotations and spatial niche abundance.

344 **E) Network diagram illustrating spatial proximity between recurrent spatial niches.** Edge width
345 and colour represent the mean pairwise Euclidean distance between niche spots.

346 **F) Xenium sections depicting spatial maps of major cell states in tumour AT10.**

347 successive layers of *Gliosis transition*, *Gliosis* and *Hypoxia* (**Fig. 2E**). *Dev-like* (Neuron) was
348 similarly positioned near the infiltrated regions when present (**Fig. 2C**). In the tumour that we
349 grid-profiled in whole (AT15), these gradual transitions from dev-like niches to gliosis and
350 hypoxia were observed across each sampling site (**Extended Data Fig. 11A,B**).

351 We validated these observations with the orthogonal imaging-based Xenium ST
352 technology. We applied a 315-plex probe panel, including 62 markers of GB and TME cell
353 states curated from our snRNA-seq data (**Methods, Table S8**), to 17 sections from 8 tumours.
354 We annotated Xenium data at single cell level by integrating them with patient-matched
355 snRNA-seq data via Tangram²⁷. This validated the zonation of malignant cell states, where
356 dev-like states bordered infiltrated regions marked by neurons, contrasting with the gliosis and
357 hypoxia states situated further around necrotic and perinecrotic regions (**Fig. 2F and**
358 **Extended Data Fig. 11C**). Importantly, profiling centimeter-scale tumour sections, we
359 observed this pattern at each sampling site (**Fig. 2F**), indicating stereotyped spatial
360 organisation throughout each tumour.

361 The GB tissue niches also reflected TME organisation. *Vasculature* captured pericytes,
362 VLMCs and endothelial cells (**Fig. 2B**) and anatomically correlated with microvascular
363 proliferation (MVP) and hyperplastic blood vessels (HBV) (**Fig. 2D**). *Dev-like* (AC), *Gliosis*
364 *transition* and *Proliferative* malignant niches were found in close proximity to *Vasculature* (**Fig.**
365 **2E**). Additionally, two *Immune* niches captured distinct myeloid cell compartments (see last
366 section). Amongst infiltrated regions, *White matter* was found in close proximity to tumour
367 leading edge (**Fig. 2B-E**), consistent with GB invasion observed through cortical white
368 matter²⁸⁻³⁰.

369 Taken together, we reveal stereotyped stepwise spatial transitions of malignant cells
370 from dev-like to gliosis and hypoxia states in GB.

371 **A putative malignant cell trajectory towards gliosis and hypoxia**

372 We hypothesised that the stereotyped spatial organisation of malignant cell states
373 relates to their cellular trajectories. Specifically, we hypothesised that as GB tumours expand
374 at their leading edge into normal brain tissue, the dev-like malignant cells progressively

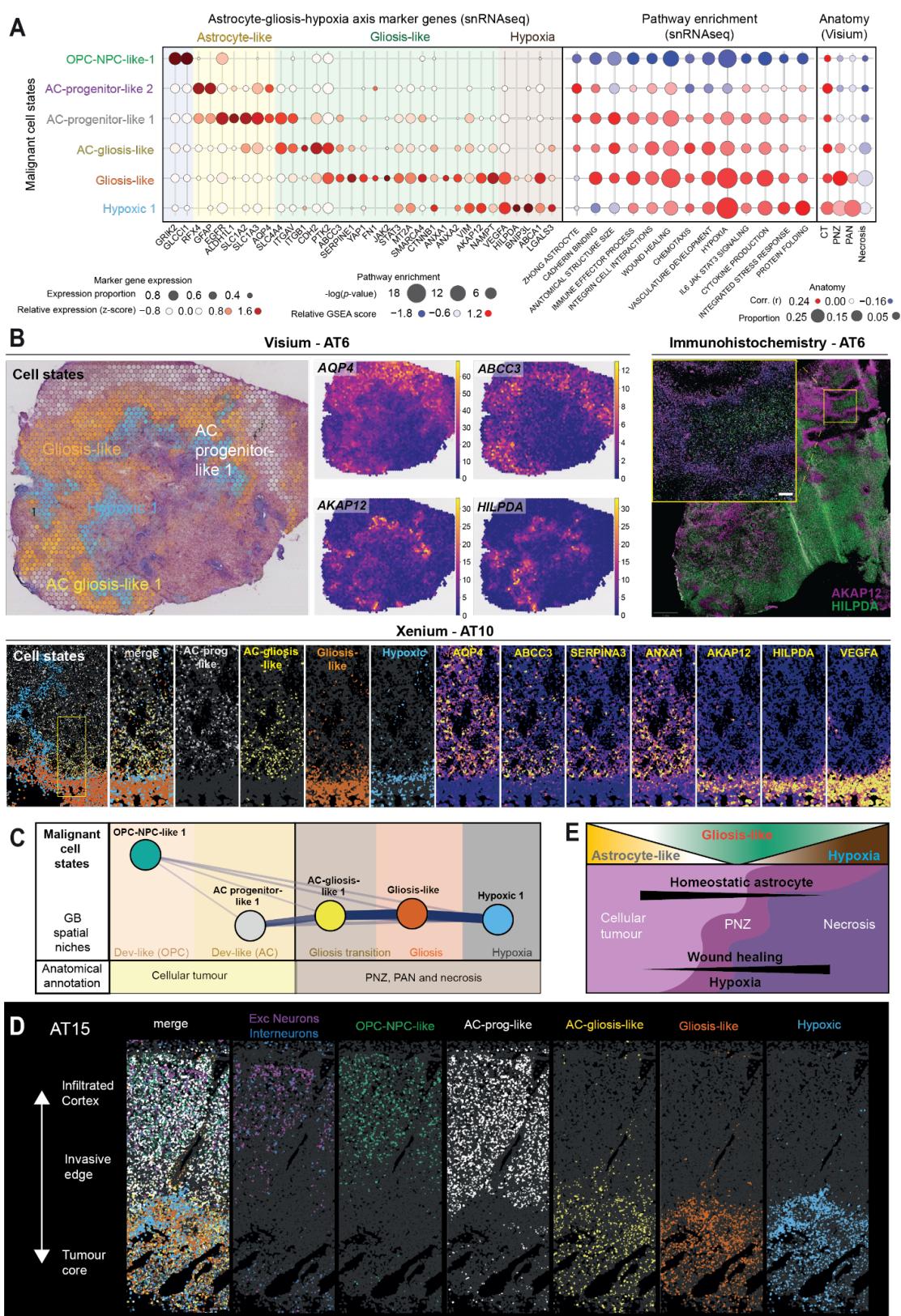
375 transition towards gliosis and hypoxia states in a recurrent manner. To test this, we first
376 examined putative malignant trajectories related to gliosis and hypoxia.

377 In the snRNA-seq data, we observed a gene expression trajectory from AC-prog-like
378 malignant states to gliosis then hypoxia (**Fig. 3A**). The AC-prog-like states showed the highest
379 levels of developmental and homeostatic astrocyte gene expression (*EGFR*, *SLC1A3*, *AQP4*,
380 *ALDH1L1*). These were downregulated in AC-gliosis-like cells and replaced by a
381 transcriptomic program resembling astrogliosis, including genes from the integrin-N-cadherin
382 pathway (*ITGAV*, *ITGB1*, *CDH2*) associated with glial scar formation³¹. This program
383 expanded in gliosis-like cells to genes associated with JAK-STAT signalling (*JAK2*, *STAT3*,
384 *SMARC4*), wound healing (*ANXA2*, *ANXA5*, *YAP1*, *F1*, *AKAP12*), cytokine production (*IFI16*,
385 *IL6R*, *IL1R1*), coagulation (*SERPINE1*) as well as angiogenesis and hypoxia (*VEGFA*)³²⁻³⁴.
386 This trajectory terminated in hypoxia (*HILPDA*, *BNIP3L*, *VEGFA*) and stress response gene
387 expression (*JUN*, *FOS*, *HSPA1B*) (**Fig. 3A and Extended Data Fig. 12A**).

388 While gliosis and hypoxia-associated malignant states have been described as
389 mesenchymal-like² and wound healing and injury response genes overlap with epithelial-to-
390 mesenchymal transition (EMT) programs^{35,36}, we observed negligible and non-specific
391 expression of key EMT regulators (e.g. *SNAI1/2*, *TWIST1/2*, *ZEB1/2*) in these malignant states
392 (**Extended Data Fig. 12A**). Hence, our nomenclature places these states in the more specific
393 biological context of glial injury response and hypoxia, putatively stemming from
394 developmental astrocyte-like malignant cells.

395 The molecular trajectory was correspondingly patterned in space, manifesting as
396 distinct concentric segregation of AC-gliosis-hypoxia states in Visium and Xenium ST data
397 (**Fig. 3B,C**). AC-prog-like cell states were the most spatially distinct, located in the non-
398 necrotic CT in the *Dev-like* (AC) niche (**Fig. 3B-D**, **Fig. 2D**). AC-gliosis- and gliosis-like states
399 formed *Gliosis transition* and *Gliosis niches*, respectively, that spatially intermixed to varying
400 degrees. In some tumours, there was a clear gradient where AC-gliosis-like was located at
401 the margin of CT and transitioned into gliosis-like within the PNZ (e.g. AT10), whereas in
402 others they broadly overlapped (e.g. AT6) (**Fig. 3B**). Hypoxic cell states were located either
403 within or adjacent to the necrotic core of patient tumours, including the PNZ and PAN (**Fig. 3B**
404 and **Fig. 2D**). Consistently, the marker genes of these states showed spatially zonated
405 expression in ST data including *AQP4* (AC-progenitor), *ABCC3* (AC-gliosis), *AKAP12* (Gliosis)
406 and *HILPDA* (Hypoxia), which we also validated using orthogonal immunohistochemistry (**Fig.**
407 **3B and Extended Data Fig. 12B**).

408 Taken together, we identify putative cellular transitions of malignant cells from
409 developmental astrocyte-like to gliosis then hypoxia dominated states, which unfolds across
410 an anatomical axis from tumour expansion towards hypoxia and necrosis (**Fig. 3E**).



411

412 **Figure 3: Spatial zonation of a putative malignant trajectory from astrocyte progenitor-like states to gliosis and hypoxia.**

413 **A) Transcriptional continuum across AC-gliosis-hypoxia GB cell states. Left: Marker gene expression across states in snRNA-seq, showing relative z-score normalised gene expression**

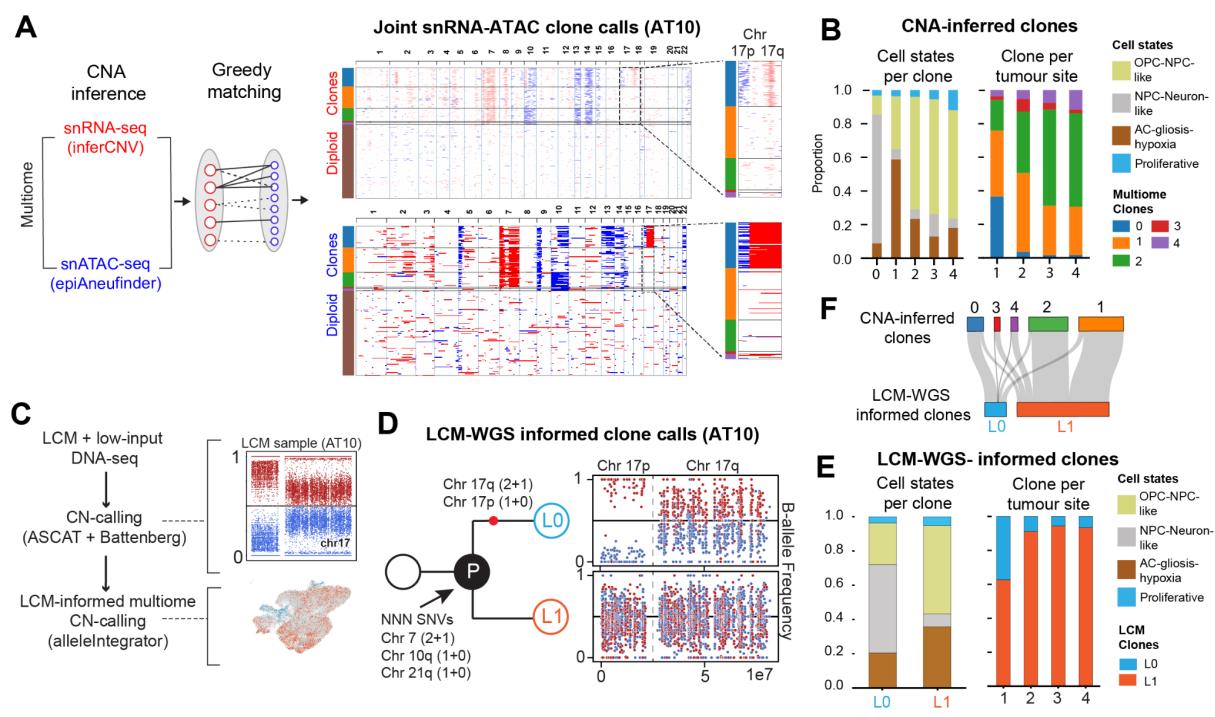
416 (colour), and proportion of cells per state expressing the gene (size). Centre: Pathway
417 enrichment across states in snRNA-seq, showing the z-score normalised enrichment score
418 (colour) and the -log(p-value) of each comparison (size). Right: Spatial correlation of cell states
419 with anatomical features in Visium ST data, showing the Pearson's correlation coefficient (r)
420 (colour) and relative abundance of a cell state within feature-annotated Visium spots (size).
421 B) Spatial zonation of major AC-gliosis-hypoxia GB states and marker gene expression in ST
422 data and immunohistochemistry validation. Top left: Cell state abundance (left panel, colour)
423 and gene expression (right panel, colour) in Visium ST data. Top right: Antibody staining
424 showing distinct zonation of AKAP12 (gliosis-hypoxia) and HILPDA (hypoxia). Bottom: Cell
425 states (left panels, colour) and marker genes (right panels, colour) in Xenium ST data.
426 C) Network diagram summarising the spatial proximity of relevant cell states in Visium data,
427 with coloured banners indicating the primary spatial niche of each cell state, and boxes
428 indicating the associated histological features.
429 D) Xenium section demonstrating zonation of AC-gliosis-hypoxia cell states ranging from
430 infiltrated grey matter to the necrotic tumour core.
431 E) Schematic summarising the AC-gliosis transcriptional gradient in a histological context,
432 including relevant functional axes.

433 Shared clonal origins of dev-like and gliosis-hypoxia malignant cells

434 To characterise GB cell state transitions, we first set out to validate whether dev-like
435 and gliosis-hypoxic states could emerge within the same tumour subclones (i.e. genetically
436 related malignant sublineages). Previous studies have applied copy number aberration (CNA)
437 inference to single cell or spatial RNA-seq data to resolve GB subclones^{2,3,9}. Here, we initially
438 leveraged multiome data to validate inferred CNAs across multiple modalities and performed
439 joint clone calling on snRNA-seq and snATAC-seq profiles of malignant cells (**Fig. 4A**). We
440 inferred clones independently on each modality, then greedily matched them on the basis of
441 shared cells and CNA profile similarity (**Methods**).

442 This resulted in 58 tumour subclones across 12 patient tumours (5 per tumour on
443 average) with chr7 and chr10 CNs distinguishing them from non-malignant diploid cells (**Fig.**
444 **4A and Extended Data Fig. 13A**). The majority of subclones contained both dev-like and
445 gliosis-hypoxic cells (**Fig. 4B, Extended Data Fig. 13B,C, Table S9**), with the minority states
446 making up an average of 15% (ranging from 3-49%) of each clone. The subclones within each
447 tumour showed similar distributions of coarse malignant states (e.g. OPC-NPC vs NPC-
448 neuronal) and tended to mirror the overall cellular composition of tumours, though we
449 observed greater clonal variability in distributions of fine-grained GB states (**Fig. 4B and**
450 **Extended Data Fig. 13B**).

451 Surprisingly, in three tumours, we identified individual subclones with highly distinct
452 cellular compositions. In AT10, clones 0 and 1 were preferentially enriched for NPC-neuronal-
453 like and AC-gliosis-hypoxia states, respectively (**Fig. 4B**). Similarly, AT15 clones 3 and 5 were
454 enriched for AC-gliosis-hypoxia and OPC-like states, respectively, and AT6 clone 0 was
455 enriched for OPC-like states (**Extended Data Fig. 13B**). Distinct CNAs marked these clones,
456 such as chr17p loss and chr17q gain in AT10 clone 0 (**Fig. 4B**). While these clones still
457 contained both dev-like and gliosis-hypoxia cells, these patterns imply that late genomic
458 alterations in GB phylogenies can influence malignant cell states.
459



460

461 **Figure 4: Developmental-like and gliosis-hypoxia states share subclonal origins.**

462 **A)** Multiome CNA-calling workflow. The joint clonal annotation in tumour AT10 is illustrated by
463 heatmaps from paired snRNA- (top) and snATAC-seq (bottom) data showing putative copy
464 gain (red) and copy loss (blue). The harmonised clone annotations are indicated by the left
465 sidebar. Insets highlight a major lineage-defining CN changes to chr17.

466 **B)** Barplots depicting cell state distributions across clones (right) and clone distribution across
467 tumour sampling sites (left) in CNA-inferred clones.

468 **C)** LCM-informed CNA calling workflow to predict snRNA-seq CN events via alleleIntegrator.
469 The scatterplot (top right) represents the B-allele frequency (BAF) for sample
470 PD60966a_lo00005. Colours correspond to SNP allele frequency of the major (red) and minor
471 (blue) alleles. The UMAP (bottom right) shows alleleIntegrator clone calls superimposed on
472 the visualisation such that cells with chr17 CNs (light blue) can be distinguished from cells
473 without chr17 CNs (orange).

474 **D)** LCM genotypes in the snRNA-seq data. Left: phylogeny illustrating lineage defining
475 mutations. Right: BAF plot showing the LCM-based SNP-phasing applied to genotype the
476 single nucleus data for two putative clones.

477 **E)** Barplots depicting cell state distributions across clones (right) and clone distribution across
478 tumour sampling sites (left) in LCM-WGS informed clones.

479 **F)** Sankey plot depicting concordant clone assignment across multiome and LCM-informed
480 CNA calling methods per cell.

481 To identify CNAs and subclones directly at the DNA level rather than inferring them
482 from snRNA/ATAC-seq data, we performed whole genome sequencing (WGS) on tissue
483 collected via laser capture microdissection (LCM)³⁷. Each cut consisted of 100-200 µm² sized
484 regions of interest (ROIs) z-stacked across two adjacent H&E sections. We profiled multiple
485 ROIs per primary tissue section across two sampling sites of two tumours (AT3, AT10),
486 generating 61 high-quality WGS profiles (**Extended Data Fig. 14A**). In general, we found that
487 most detected CN events are truncal, with only a minority of events appearing in specific

488 tumour ROIs (**Extended Data Fig. 14B, Methods**), suggesting that GB subclonal
489 diversification is a relatively late event. Notably, we detected a CN loss in chr17p and gain
490 chr17q in one ROI from AT10 (**Fig. 4C and Extended Data Fig. 14B**), validating our multiome
491 CN calls (**Fig. 4B**).

492 To understand whether the genotypes of subclones can influence their composite GB
493 cell state phenotypes, we defined tumour subclones in the multiome data based on the CN
494 events validated at the DNA level. We used the alleleIntegrator tool³⁸ to leverage allelic
495 imbalances identified in the LCM-WGS data to predict corresponding CNs in matched snRNA-
496 seq profiles. In AT10, this approach validated chr7 gain and chr10 loss in malignant cells and
497 distinguished two subclones based on chr17-alterations (**Fig. 4D**). Both subclones contained
498 a mixture of dev-like and gliosis-hypoxia cells (**Fig. 4E**). Yet, strikingly, clone L0 was enriched
499 for NPC-neuronal-like cells and marked by chr17p loss and chr17q gain (**Fig. 4D,E**). This
500 mirrored the cellular composition and CN profile of clone 0 identified by our orthogonal
501 multiome CNA inference (**Fig. 4A,B**). Correspondingly, the cell assignments to chr17-altered
502 clones were highly consistent across our two CN calling approaches (**Fig. 4F**), validating our
503 joint clone calls from multiome data by the WGS-derived “ground truth”.

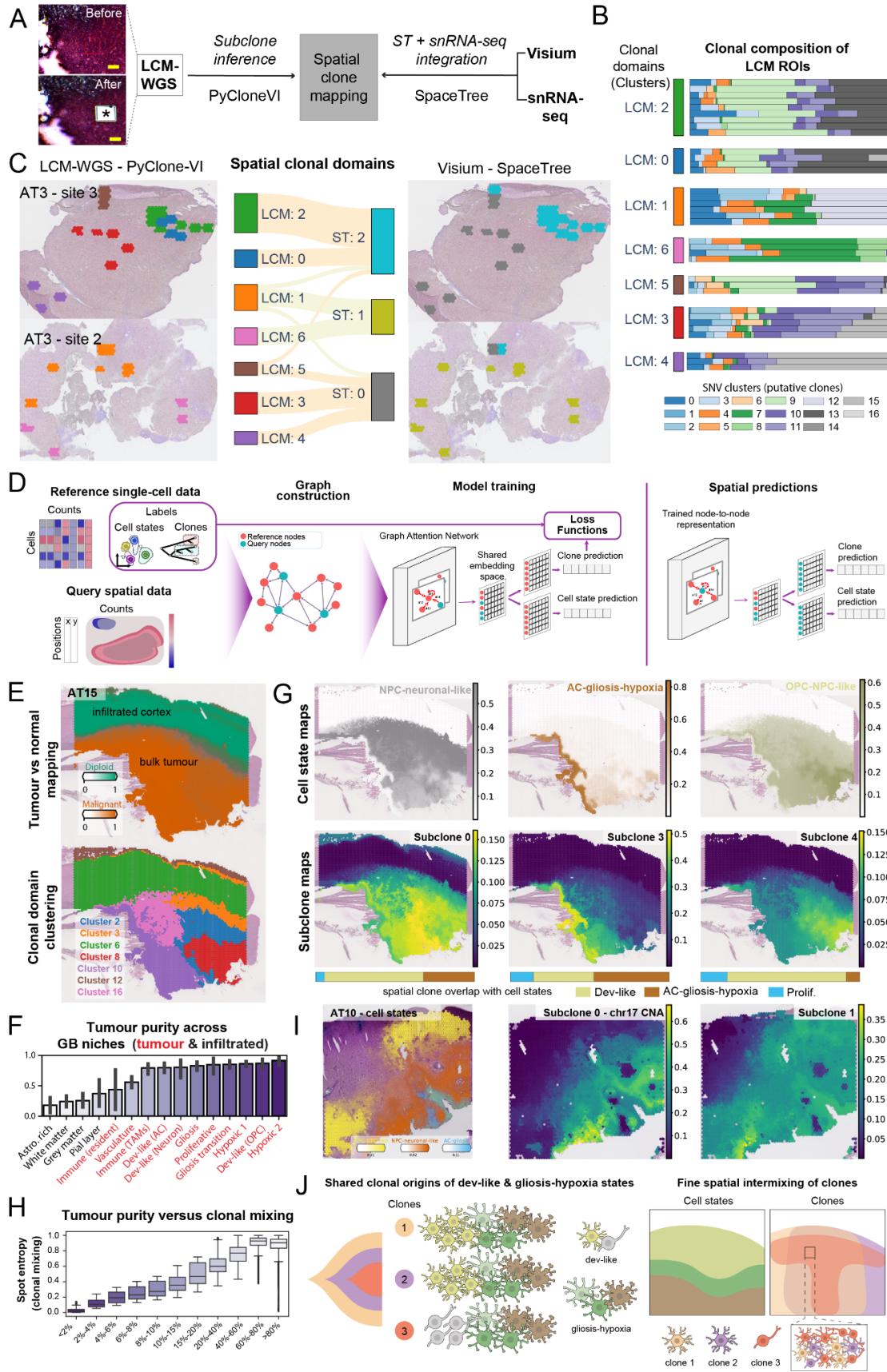
504 Taken together, our paired genomic, epigenomic and transcriptomic data demonstrate
505 shared clonal origins of dev-like and gliosis-hypoxia cell states, consistent with our trajectory
506 hypothesis. Surprisingly, we also identified that late genomic alterations can alter malignant
507 cell states and expand intratumour heterogeneity, the first such demonstration at cellular
508 resolution in GB.

509 **Fine-grained spatial intermixing of GB subclones**

510 Given the spatial segregation of GB cell states, we inquired whether tumour subclones
511 are also spatially organised. Based on the shared clonal origins of dev-like and gliosis-hypoxia
512 cell states (**Fig. 4**), we expected individual subclones to contribute to multiple spatial tissue
513 niches (**Fig. 2**). While GB clonal distribution has been studied at coarse resolution using bulk
514 tumour biopsies^{5,7}, the fine-grained spatial organisation of clones and how they relate to tissue
515 niches is poorly understood.

516 Initially, we investigated clonal organisation at centimeter-scale resolution from tumour
517 sampling sites. Based on multiome CNA inference, we detected multiple subclones within
518 each site (**Fig. 4B and Extended Data Fig. 15A**). The clone distribution was largely consistent
519 across different tumour sites, with notable exceptions including AT10, where the chr17-altered
520 clone 0 is predominantly located in site 1 of 4 (**Fig. 4B and Extended Data Fig. 15A**). Consistently,
521 our orthogonal LCM-WGS approach found the corresponding clone L0 enriched
522 in the same site (**Fig. 4E**). These observations suggest broad spatial intermixing of GB clones,
523 as previously shown^{5,7}, and that late genomic alterations can give rise to spatially restricted
524 clones.

525 To investigate clonal organisation at higher spatial resolution and identify clones at the
526 DNA level, we examined clonal distributions in our LCM-WGS data profiling ~100x100 μm
527 tissue areas. Considering the clonal intermixing described above and that our LCM cuts span
528 multiple cells, we used PyClone-VI³⁹ to infer clonal population structure of LCM ROIs based
529 on single nucleotide variant (SNV) allele frequencies (**Fig. 5A, Methods**). We focused on
530 tumour AT3 that yielded 35 high-quality genomes from LCM ROIs across two sampling sites,
531 and identified 17 SNV clusters (i.e. putative subclones) in total (**Extended Data Fig. 14A,**
532 **right**). Individual SNV clusters were present across multiple LCM ROIs and individual



533

534 **Figure 5: GB spatial tissue niches are polyclonal. (legend on the next page)**
535

536 **A)** LCM-WGS- and SpaceTree-based spatial clone mapping workflows. The H&E images
537 show a tumour ROI before and after LCM cutting, marked by an asterisk. Scale bars: 100 µm.
538 **B)** Spatially intermixed GB subclones in LCM-WGS data. Barplot summarises pyclone-VI
539 clone distributions for each LCM ROI and clonal cluster.
540 **C)** Spatial clonal domains compared across LCM-WGS pyclone deconvolution and Visium ST
541 SpaceTree mapping across matching tissue locations. Domains were defined by Bray-Curtis
542 clustering of clone distributions within each ROI across each modality H&E images depict
543 tissue locations used for LCM colored by SpaceTree clone clusters (left) and pyclone-VI clone
544 clusters (right). The sankey plot demonstrates ROI cluster assignment between modalities.
545 **D)** SpaceTree workflow for joint spatial mapping of cell states and tumour subclones via
546 integration of ST and snRNA-seq data.
547 **E)** SpaceTree distinguishes tumour versus normal regions and resolves spatial clonal domains
548 in tumours. Top: Visium section from tumour AT15 showing SpaceTree mapping of diploid
549 versus malignant subclones to infiltrated and bulk tumour regions, respectively. Bottom:
550 Spatial clonal domains identified by clustering SpaceTree-derived clonal composition of
551 Visium spots.
552 **F)** Tumour purity of infiltrated versus tumour tissue niches estimated by SpaceTree,
553 summarised at the Visium spot level ($n=97$ Visium sections).
554 **G)** GB subclones spatially overlap with both dev-like and gliosis-hypoxia cell states.
555 SpaceTree derived cell state (top panel) and clonal (bottom panel) maps in tumour AT15 are
556 shown. The barplots summarise the distribution of cell states across spots in which a given
557 clone is the most abundant.
558 **H)** Clonal entropy of Visium spots based on SpaceTree mapping binned across tumour purity
559 fractions. High entropy (i.e. clonal intermixing) is observed in tissue regions with high tumour
560 purity.
561 **I)** SpaceTree cell state (left) and clonal (centre and right) maps of tumour AT10 resolves
562 distinct regional distributions of chr17-CNA subclone.
563 **J)** GB clonal organisation summary diagram.

564
565 ROIs contained multiple SNV clusters, indicating clonal intermixing at fine spatial resolution
566 (**Fig. 5B**).

567 Clustering ROIs based on their subclonal composition (**Fig. 5B**), we identified broad
568 spatial clonal domains (**Fig. 5C**). Neighboring tissue regions had similar polyclonal
569 composition, giving rise to several millimeter-wide clonal domains (e.g. LCM ROI domains 0
570 & 2 on site 3). Broad spatial gradients of subclones demarcated neighbouring domains, such
571 as SNV cluster 9 spanning domains 0, 2 and 5 and SNV cluster 15 appearing in domains 3
572 and 4 (**Fig. 5C**). Different tumour sampling sites showed largely distinct polyclonal composition
573 (**Fig. 5C**), displaying more extensive differences than we observed with multiome CNA
574 inference (**Extended Data Fig. 15C,D**), likely owing to the better sensitivity of SNV-based
575 clone calls from WGS data. We inferred phylogenetic trees of AT3 subclones and observed
576 early SNV clusters to be widely spatially spread while late clusters showed variable spatial
577 distribution, including those restricted to few sites (**Extended Data Fig. 15C**).

578 **Joint spatial modeling reveals GB tissue niches are polyclonal**

579 Next, given their broad spatial distribution, we sought to validate that GB subclones
580 contribute to multiple tissue niches spanning the putative trajectory from dev-like to gliosis-
581 hypoxia states. Consistently, in the LCM-WGS data, we observed that individual clones (e.g.

582 SNV cluster 9) and larger clonal domains (e.g. LCM cluster 0, 2 & 5) spanned tissue areas
583 with diverse malignant cell states and tissue niches based on the adjacent Visium sections
584 (**Extended Data Fig. 15D**). However, our LCM-WGS findings were limited to select ROIs from
585 two tumours due to the high cost and low throughput of this methodology. Hence, we aimed
586 to resolve spatial clonal patterns in our Visium ST dataset that comprehensively profiles 12
587 tumours (n=97 sections) at high resolution. The ST profiles in Visium data also enabled us to
588 directly relate clonal domains to transcriptomic cell states and tissue niches (**Fig. 2**) on the
589 same tissue section.

590 Previous studies have used CNA inference on Visium data to chart subclones in GB⁹
591 and other tumours⁴⁰. However, these approaches assumed monoclonal composition of Visium
592 spots, which could not resolve the finely intermixed GB subclones. To address this challenge,
593 we developed SpaceTree, a computational framework to jointly deconvolve tumour subclones
594 and cell states in ST data by leveraging a reference scRNA-seq dataset (**Fig. 5D**). SpaceTree
595 utilises a multi-task graph neural network architecture coupled with label propagation to
596 integrate the reference scRNA-seq and query ST datasets, enabling joint spatial mapping of
597 clones and cell states annotated in the single cell data (**Fig. 5D**, described in detail in **Supp.**
598 **Comp. Note**).

599 We applied SpaceTree to integrate our GB snRNA-seq and Visium ST data per tumour
600 site, leveraging our granular cell annotation (**Fig. 1B,C**) and joint multiome clone calls (**Fig.**
601 **4B**). We optimised SpaceTree hyperparameters to match cell2location cell type mapping
602 (**Supp. Comp. Note**). Initially, we validated that Spacetree can distinguish between infiltrated
603 (i.e. largely normal cells) versus tumour regions. Consistently, SpaceTree mapped diploid
604 clones containing TME cells to sparsely infiltrated regions, while malignant subclones were
605 enriched in the tumour core (**Fig. 5E, top panel**). We then quantified the fraction of malignant
606 cells per Visium spot (i.e. tumour purity) from SpaceTree estimates and summarised it by GB
607 tissue niches across all tumours (**Fig. 2C**). Consistently, this showed predominant enrichment
608 of malignant cells to malignant GB niches (e.g. *Dev-like*, *Hypoxic*) over infiltrated (e.g. *Grey*
609 *matter*) or TME (e.g. *Vasculature*) niches (**Fig. 5F**).

610 SpaceTree revealed widespread spatial intermixing of subclones across all tumours
611 (**Fig. 5G and Extended Data Fig. 16A**). To quantify the extent of clonal mixing, we calculated
612 the clonal entropy per Visium spot across all tumours. Spots with low fraction of malignant
613 cells (i.e. infiltrated and TME tissue niches with low tumour purity) showed low clonal entropy,
614 suggesting they are dominated by diploid clones (**Fig. 5H**). In contrast, spots with high fraction
615 of malignant cells (i.e. malignant tissue niches with high tumour purity) showed high clonal
616 entropy, indicating fine-grained spatial mixing of GB subclones and the polyclonal nature of
617 GB tissue niches (**Fig. 5H**).

618 SpaceTree also resolved broad spatial domains of subclones across GBs (**Fig. 5E and**
619 **Extended Data Fig. 16B**). Individual clonal domains and their constituent subclones
620 overlapped with diverse malignant cell states (**Extended Data Fig. 16D**). For example, AT15
621 subclones 0, 3 and 4 each overlapped with both dev-like and gliosis-hypoxia cell states (**Fig.**
622 **5G**). Yet, we also observed regional clonal enrichment, such as clone 3 contributing more
623 heavily to gliosis-hypoxia (**Fig. 5G**). Notably, AT10 clone 0, distinguished by chr17 alterations
624 (**Fig. 4**), was spatially enriched in tissue regions with NPC-neuronal-like cells (**Fig. 5I**).

625 Finally, we also validated SpaceTree by orthogonal LCM-WGS mapping of adjacent
626 tissue sections (**Methods**). Across matching tissue regions, we observed concordance
627 between clonal domains mapped by SpaceTree deconvolution of Visium ST data and those
628 identified by SNV-based deconvolution of LCM-WGS data (**Fig. 5B, Methods**). While SNV-

629 based mapping provided higher granularity, the spatial clonal domains broadly matched
630 across the two modalities (**Fig. 5B**).

631 In summary, we present a new framework for the cellular and spatial organisation of
632 GB subclones (**Fig. 5I**). We demonstrate that individual GB subclones give rise to both dev-
633 like and gliosis-hypoxia states, and that they finely spatially intermix in tumours, organising
634 into broad clonal domains cutting across GB tissue niches. We also show that late genomic
635 alterations (i.e. chr17 CNA in AT10) can alter the cellular composition and localisation of GB
636 subclones to expand intra-tumour heterogeneity.

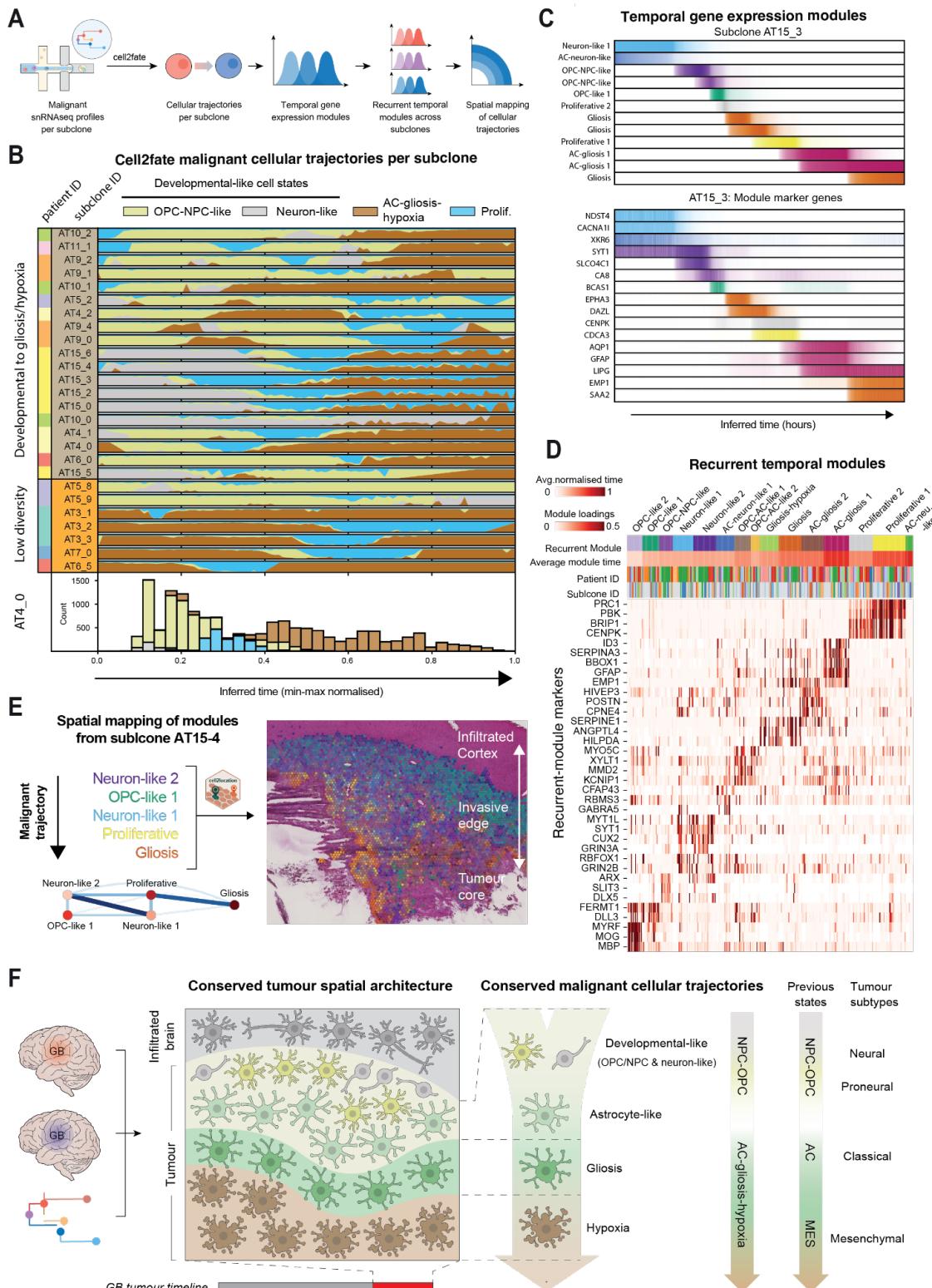
637 **A recurrent trajectory from dev-like to gliosis and hypoxia states dominates GBs**

638 Having validated the shared clonal origins of dev-like and gliosis-hypoxia states in
639 single cell (**Fig. 4**) and spatial data (**Fig. 5**), we investigated whether they form a cellular
640 trajectory tractable across GBs. Specifically, we examined whether the transition from dev-like
641 to gliosis-hypoxia states represents the major transcriptomic trajectory across distinct GB
642 subclones and tumours, and whether this involves conserved gene expression programs.

643 Given their spatial zonation observed throughout each tumour (**Fig. 2E,F**), we
644 hypothesised dev-like malignant cells continuously transition into gliosis and hypoxia states
645 during tumour expansion. The rapid expansion rate of GBs, which double in size every 50
646 days⁴¹ supported the likelihood that our “snapshot” snRNA-seq profiling can capture the
647 different stages of these putative transitions. To infer cellular trajectories in an unbiased
648 manner from snRNA-seq data, we used cell2fate, a probabilistic RNA velocity model⁴² (**Fig.**
649 **6A**).

650 We applied cell2fate to the snRNA-seq profiles of individual tumour subclones,
651 avoiding confounding genetic differences across clones and tumours. Out of 58 subclones, 26
652 subclones from 9 tumours showed robust cellular trajectories based on their inferred
653 duration⁴², where GB cell states segregated across distinct latent time windows (**Fig. 6B and**
654 **Extended Data Fig. 17A, Methods**). The remaining subclones showed poorly discernible
655 trajectories based on their inferred duration, likely due to either small numbers of cells or low
656 state diversity captured per clone (**Extended Data Fig. 17B**). Across the reconstructed
657 trajectories, the majority of subclones (19/26, >73%) displayed temporal transitions from dev-
658 like to AC-gliosis-hypoxia (**Fig. 6B**). We observed this trajectory across 7 out of the 9 tumours
659 (e.g. AT4, AT10, AT15) and quantified temporal ordering of GB states from dev-like to AC-
660 gliosis-hypoxia (**Fig. 6B, bottom panel, Extended Data Fig. 17D**). The minority of clones that
661 did not show this pattern (7/26) were dominated by an individual state and displayed relatively
662 lower cell state transition quality scores (**Fig. 6B, Extended Data Fig. 13B, Extended Data**
663 **Fig. 17A, Methods**).

664 At a finer-grained cellular level, both NPC-neuronal-like (AT15_0) or OPC-NPC-like
665 (AT6_0) states led to AC-gliosis-hypoxia (**Fig. 6B**). Proliferating cells distributed across this
666 trajectory and were most frequently observed mid-latent time (**Extended Data Fig. 17D**).
667 Within dev-like or AC-gliosis-hypoxia classes, fine-grained malignant states did not show
668 highly distinct temporal ordering and consistently displayed low cell state transition scores
669 (**Extended Data Fig. 17B**).



670

671 **Figure 6: A recurrent malignant cell state trajectory across GB tumours.**

672 **A) Cellular trajectory analysis workflow.**

673 **B) Cell2fate analysis resolving a cellular trajectory from dev-like to AC-gliosis-hypoxia states**
674 **across GB subclones and tumours. Top: Streamgraph plots show cell state occurrence across**
675 **cell2fate-inferred trajectory time across 26 tumour subclones (rows). Clones are clustered**

676 based on the ranked temporal order of coarse-grained cell states, distinguishing those
677 spanning diverse GB cell states (top panels) versus low diversity clones dominated by one of
678 the aforementioned cell states (bottom panels). Bottom: Barplot showing the abundance of
679 cell states in clone AT4_0 across inferred time.

680 **C)** Activation patterns of genes (top) and meta-modules (bottom) for clone AT15_3
681 summarising the inferred cellular trajectory. Cell2fate derived activation plots of temporal gene
682 expression modules (top) and top marker genes of each module (bottom) with matched
683 colours.

684 **D)** Recurrent temporal gene expression modules are conserved across the trajectories of 26
685 tumour subclones. Heatmap displays the module loadings for top module markers across each
686 individual subclone. Columns are clustered by Jaccard similarity index of top genes ($n=200$).
687 Colour bars correspond to recurrent temporal module annotations, average latent time across
688 each recurrent module, tumour and clone labels.

689 **E)** Spatial mapping of key temporal modules (colours) from subclone AT15_4 onto its
690 matching Visium section (right). Network diagram (bottom) shows the mean pairwise
691 euclidean distance between temporal modules for the associated Visium section.

692 **F)** Summary diagram of the spatiotemporal malignant cell trajectory identified across GBs.
693 The diagram relates different stages of this trajectory to GB cell states and tumour subtypes
694 described by previous studies.

695 Importantly, we observed the transition from dev-like to AC-gliosis-hypoxia states
696 commonly across tumours of all presumed GB subtypes, including the OPC-NPC-like enriched
697 “proneural” (e.g. AT5), the NPC-neuronal-like enriched “neural” (i.e. AT15) and the AC-gliosis-
698 hypoxia enriched “classical-mesenchymal” (e.g. AT6) tumours (**Fig. 6B and 1D**). Within each
699 tumour, multiple subclones also shared this trajectory (e.g. AT15_0 and AT15_3) (**Fig. 6B**).
700 This included tumour AT10, where both NPC-neuronal- and OPC-NPC-enriched clones
701 distinguished by chr17 alterations (**Fig. 4A,B**), converged on AC-gliosis-hypoxia (**Fig. 6B**).
702 Taken together, these findings identify a dominant malignant cellular trajectory across GBs,
703 manifesting across tumour subtypes as well as genetically distinct clonal lineages.

704 Next, we examined whether this cellular trajectory is molecularly conserved across
705 subclones and tumours, leveraging temporal gene expression modules (i.e. RNA velocity
706 modules) identified by cell2fate⁴². Malignant cells in each subclone progressed through
707 successive gene expression modules that were sequentially activated and temporally
708 overlapped (**Fig. 6C**). Clustering modules based on their marker gene similarity (**Methods**),
709 we identified 15 meta-modules representing recurrent temporal transcriptional programs
710 across subclones and tumours (**Fig. 6D, Table S10**). Based on their GB cell state (**Extended**
711 **Data Fig. 18A**) and MSigDB enrichment (**Table S10**), we annotated these meta-modules
712 using a nomenclature similar to GB cell states (**Fig. 6D**). Developmental-like modules occurred
713 early in the GB cellular trajectory across subclones and tumours, and shared marker genes
714 with OPC-like (e.g. *MBP*, *MOG*) and NPC-neuronal-like (e.g. *MYT1L*, *SYT1*) cell states. In
715 contrast, AC-gliosis-hypoxia modules occurred late and displayed corresponding marker
716 genes (e.g. *GFAP*, *SERPINE1*, *ANGPTL4*, *HILPA*) (**Fig. 6D**). Proliferative modules frequently
717 preceded or followed AC-gliosis-hypoxia. Hence, these modules demonstrate a shared
718 malignant gene expression trajectory across GBs.

719 While the recurrent temporal gene expression modules broadly reflected GB cell state
720 phenotypes and included known regulators of dev-like (e.g. *ASCL1*, *SOX10*)^{43,44} and gliosis-
721 hypoxia (e.g. *FOS*)⁴⁵ states (**Extended Data Fig. 18C, Table S10**), they also pointed at novel

722 candidate regulators and effectors of GB trajectories. For example, *MYRF*, which
723 transcriptionally co-regulates oligodendrocyte maturation and myelination with *SOX10*^{46,47},
724 was enriched in the earliest developmental modules (**Fig. 6D**). GBM cells upon white matter
725 invasion have been shown to acquire mature oligodendrocyte-like states induced by *SOX10*⁴⁸.
726 In contrast, *HIVEP3*, which regulates transcription mediated by nuclear factor kappa-B (NF-
727 κB)⁴⁹ and c-Jun⁵⁰, was enriched in late gliosis-hypoxia modules (**Fig. 6D**). Accordingly, c-Jun
728 is implicated in regulation of mesenchymal-like states⁵¹ and *HIVEP3* is also implicated in
729 Interleukin signalling in glioma⁵².

730 To relate temporal malignant gene expression trajectories to clinical outcomes, we
731 performed survival analysis on bulk RNA-seq data from TCGA and CPTAC cohorts (n=235
732 GBs)⁵³⁻⁵⁵ using gene sets from the recurrent gene expression modules. We found that patients
733 with high relative expression of late-occurring Gliosis-hypoxia and Proliferative module
734 markers showed worse survival than those enriched for early-occurring OPC-like 1 module
735 ($p<0.007$) (**Extended Data Fig. 18D**), consistent with poor prognosis of “mesenchymal-like”
736 GBs^{1,56}.

737 To spatially validate the GB cell state trajectory identified above, we mapped clone-
738 specific temporal gene expression modules to our Visium ST data. As expected, we observed
739 Dev-like modules adjacent to the infiltrated cortex, followed by proliferative and gliosis-like
740 modules near to the necrotic and hypoxic core (**Fig. 6E**). Pairwise distance computation also
741 suggested a trajectory in which dev-like modules transition into proliferative states and
742 terminate in gliosis-like gene expression programs (**Fig. 6E**).

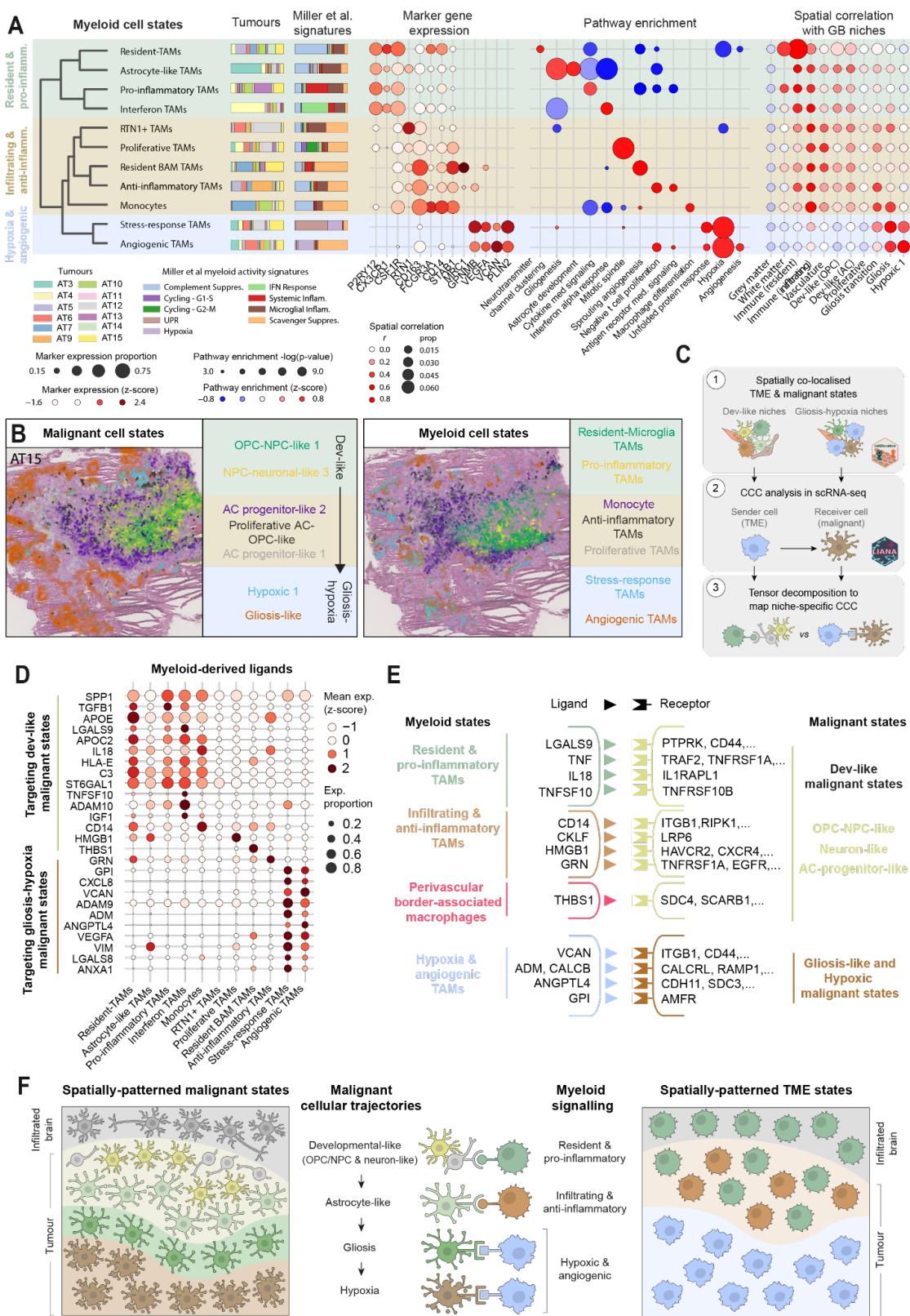
743 Taken together, we identify a spatially patterned and molecularly conserved malignant
744 cell trajectory across GBs. Combined with our fine-grained mapping of AC-gliosis-hypoxia
745 trajectories (**Fig. 3E**), these data indicate stepwise transitions of malignant cells from dev-like
746 towards gliosis and hypoxia-defined states throughout tumour expansion (**Fig. 6F**).

747 **The malignant cell trajectory is coupled to a regionalised myeloid microenvironment**

748 Finally, we investigated whether the GB TME is specialised to accommodate the
749 trajectory of malignant cells from dev-like towards gliosis and hypoxia states. We focused on
750 myeloid cells that were the most abundant TME cell type residing in tumour niches (**Table S3**,
751 **Fig. 2B**). We annotated myeloid subclusters in our snRNA-seq atlas by their expression of 1)
752 canonical marker genes, 2) MSigDB gene sets and 3) gene programs distinguishing functional
753 myeloid subtypes across gliomas from Miller et al.¹⁶. We classified 11 myeloid subtypes across
754 3 major classes, with tumour associated macrophages (TAMs) making up the majority (**Fig.**
755 **7A**).

756 The first major class consisted of resident and pro-inflammatory myeloid cells,
757 distinguished by expression of microglial marker genes (*P2RY12*, *CX3CR1*), and related gene
758 sets and programs (**Fig. 7A**). Resident TAMs appeared closest to homeostatic microglia,
759 whereas pro-inflammatory and interferon TAM populations showed elevated cytokine-
760 mediated signalling and IFN-α signalling, respectively, associated with a strong anti-tumour
761 response^{57,58}.

762 The second class consisted of infiltrating and anti-inflammatory myeloid cells,
763 distinguished by expression of monocyte-derived macrophage (MDM) marker genes (*CD163*,
764 *STAB1*, *CD14*) (**Fig. 7A**). Monocytes expressed classical (*CD14*) and intermediate markers
765 (*FCGR3A/CD16*), and macrophage differentiation gene sets. The remaining populations
766 showed a scavenger immunosuppressive transcriptional phenotype as



767
768

769 **Figure 7: The malignant cell state trajectory is coupled to a regionalised myeloid TME**
770 **and signalling.**

771 **A) Myeloid heterogeneity in our multimodal atlas. Dendrogram depicts transcriptomic**

772 clustering of myeloid subcluster snRNA-seq profiles annotated to three major phenotypic

773 classes (background colors). Stacked bar plots depict the relative distribution of myeloid states
774 across tumours and the proportion of top myeloid activity signatures from Miller et al. per state.
775 Dotplots show key marker gene expression and GOBP term enrichment across states. The
776 spatial correlation dotplot depicts the sum cell2location abundance of a given myeloid cell
777 state relative to total cell2location abundance across spots belonging to a given
778 histopathological feature (prop) and the point biserial correlation between cell state
779 abundances and histopathological feature (r).
780 **B)** Spatial co-localisation of distinct malignant and myeloid states. Visium sections (left) show
781 cell2location abundances of top cell states in tumour AT15 with colours mapping to
782 corresponding labels highlighted in the boxes to the right. Malignant cell states are grouped
783 according to their trajectory and myeloid cell states are grouped according to their phenotypic
784 classes and their spatial organisation.
785 **C)** Schematic detailing the spatially resolved cell-cell communication analysis strategy to
786 detect signalling from distinct myeloid subtypes to their spatially co-localised malignant states
787 that recur across multiple tumours.
788 **D)** Myeloid ligands predicted to signal to distinct malignant cell states by spatially-resolved
789 cell-cell communication analysis. Genes correspond to top ligands identified in myeloid
790 signalling factors identified by tensor decomposition analysis. Dotplot represents the
791 proportion of myeloid cells expressing the ligand (size) and relative expression across cell
792 states (colour).
793 **E)** Overview of major cell-cell interactions identified between regional myeloid (sender) and
794 distinct malignant (receiver) cell states.
795 **F)** Summary diagram illustrating spatially organised myeloid signalling to malignant cell states
796 across their cellular trajectory.

797 described by Miller et al¹⁶, including anti-inflammatory TAMs marked by negative regulation of
798 T cells. Perivascular-border associated macrophages (LYVE1+ Resident BAM-TAMs) were
799 included here based on their elevated expression of scavenger receptors (MRC1) (**Fig. 7A**)

800 The final class consisted of stress-response and angiogenic TAMs, marked by hypoxia
801 and angiogenesis related gene expression (VEGFA, VCAN). These populations also
802 expressed markers of “lipid-laden” macrophages (GPNMB, PLIN2) that were recently shown
803 to reside in hypoxic compartments of GB tumours¹⁵ and promote proliferation of hypoxic
804 malignant cells^{15,59}.

805 Previous studies have identified spatial segregation of myeloid subtypes in GBs^{16,17}.
806 To precisely relate the spatial patterning of myeloid and GB cell states, we compared their
807 distributions across GB tissue niches in our integrated snRNA-seq and Visium ST data (**Fig.**
808 **2B**). The three classes of myeloid cells established spatially distinct immune compartments
809 across tumours (**Fig. 7B**). Resident and pro-inflammatory populations were enriched in the
810 *Immune (resident)* niche (**Fig. 2B,C, Fig. 7A**), forming immune hotspots in the CT that spatially
811 overlapped with dev-like malignant niches such as *Dev-like* (OPC) (**Fig. 2D,E, and Extended**
812 **Data Fig. 19B-D**). These cells also extended into neighbouring white matter (**Fig. 7A**),
813 consistent with their microglial phenotype. Infiltrating and anti-inflammatory populations
814 formed the distinct *Immune (infiltrated)* niche in the CT (**Fig. 7A, Fig. 2B-D**) and were located
815 in closer proximity to *Dev-like* (AC) and *Gliosis transition* niches (**Fig. 2E and Extended Data**
816 **Fig. 19B-D**). Consistent with their infiltrating phenotype, these cells also occurred near
817 *Vasculature*. Finally, stress-response and angiogenic TAMs were highly enriched in the *Gliosis*

818 and *Hypoxia* niches in the PNZ, PAN and necrotic compartments (**Fig. 7A, Fig. 2B, and**
819 **Extended Data Fig. 19B-D**), as previously described^{15,59}.

820 These observations indicate distinct myeloid niches for dev-like and gliosis-hypoxia
821 GB cell states. To determine whether these malignant states are exposed to distinct signals
822 from their myeloid neighbours, we performed cell-cell communication (CCC) analysis in
823 snRNA-seq data amongst spatially co-localised TME (including vasculature and lymphoid
824 cells) and malignant states, assigning cells to either dev-like or gliosis-hypoxia niches (**Fig.**
825 **7C, Tables S11 and Extended Data Fig. 12**). We then used tensor decomposition to identify
826 myeloid ligands that target receptors on either dev-like or gliosis-hypoxia malignant cells in a
827 conserved manner across tumours (**Fig. 7C, Extended Data Fig. 20A-B, Methods**).

828 Our spatially refined CCC analysis identified distinct myeloid signalling environments
829 across the GB cellular trajectory (**Fig. 7D,E, Table S13**). In dev-like niches, we predicted
830 diverse myeloid signals targeting dev-like malignant states. TRAIL (TNFSF10), expressed
831 predominantly by interferon TAMs, showed a putative interaction with TRAIL receptor 2
832 (TNFRSF10B) (**Fig. 7E and Extended Data Fig. 21A**) that can induce apoptosis in glioma
833 cells⁶⁰. Conversely, progranulin (GRN), highly expressed in anti-inflammatory TAMs, has been
834 shown to promote glioma growth and therapeutic resistance^{61,62}. Several other predicted TAM
835 ligands such as Interleukin-18 (IL18), Galectin-9 (LGALS9) and high mobility group box 1
836 (HMGB1) are well known regulators of tumour immune cells^{63,64,65}, but how they influence
837 malignant cells is not well understood.

838 In contrast, in gliosis-hypoxia niches, angiogenic and stress-response TAMs
839 expressed ligands such as VEGFA and VIM that are known to be associated with hypoxia and
840 angiogenesis in GBs^{14,66}. These TAMs were also predicted to target gliosis-hypoxia malignant
841 cells via expression of extracellular matrix proteins such as versican (VCAN) that has multiple
842 roles in glioma⁶⁷ including regulation of cell adhesion via β1 integrin (ITGB1)⁶⁸. Additional
843 TAM-expressed genes in these niches include wound healing regulators such as annexin-1
844 (ANXA1) and angiopoietin-like 4 (ANGPTL4), with the latter linked to wound healing through
845 cadherin-11 (CDH11)⁶⁹, and peptide hormones such as adrenomedullin (ADM) and calcitonin
846 B (CALCB) that promote leukemia relapse through the CALCRL receptor⁷⁰ (**Fig. 7D,E, and**
847 **Extended Data Fig. 21A**).

848 In summary, signals from diverse TAM populations are linked to pro and anti-tumour
849 functions across dev-like niches, whereas TAM signals are linked to wound healing, hypoxia
850 and angiogenesis in gliosis-hypoxia niches. Taken together, we find that the myeloid TME is
851 regionalised across the malignant cells trajectory and exposes malignant cells to distinct
852 signals (**Fig. 7F**).

853 Discussion

854 The extensive tumour heterogeneity and plasticity of GB has undermined efforts to
855 pinpoint the cellular trajectories of cancer cells and develop targeted therapies. Here, we
856 applied a holistic multi-modal mapping approach to jointly characterise the genetic, cellular
857 and tissue architecture of GB tumours and dissect cancer cell trajectories. We created the
858 most comprehensive GB atlas to date, combining large-scale single cell transcriptomics and
859 epigenomics with spatial transcriptomics of each tumour as well as spatial whole genome
860 sequencing.

861 Our findings show remarkable intratumour heterogeneity of GBs, extending previous
862 surveys. Yet, our multimodal atlas and integrated analysis approach enabled us to trace GB
863 heterogeneity to a spatiotemporally-patterned cancer cell trajectory from developmental-like
864 towards gliosis and hypoxia states. This cellular trajectory is molecularly conserved across

865 tumours, including multiple presumed GB subtypes, as well different subclonal lineages within
866 each tumour. Furthermore, it is coupled to spatially patterned TME signalling from distinct
867 myeloid populations. Hence, this trajectory represents a fundamental feature of cancer cell
868 and TME organization across GB tumours.

869 Our multimodal atlasing approach facilitated the biological reinterpretation of
870 spatiotemporal cancer cell states in GB. For example, using new transcriptomic and spatial
871 characterisation, we found that the previously reported mesenchymal cell state^{1,2} was in fact
872 a malignant cell gliosis and hypoxia response. Our integrative analysis demonstrated the
873 stepwise transitions of malignant cells from developmental-like states resembling neuronal
874 and oligodendrocyte progenitors into those resembling astrocytes, followed by the gliosis and
875 hypoxia response. This trajectory is compatible with observations in mouse studies using GB
876 patient-derived xenografts (PDX)^{2,6}, yet we are the first human study to map this trajectory in
877 patient tumours considering both intrinsic cancer cell and TME characteristics. Moreover,
878 wound healing-like states and hypoxia are present across multiple cancers including
879 pancreatic⁷¹ and colon⁷² cancer, suggesting the transitions of developmental-like cancer cells
880 into these states may be a conserved feature across multiple solid tumours⁷³. While hypoxia
881 is a prominent feature of the GB cancer cell trajectory, whether other cell intrinsic and extrinsic
882 cues push developmental-like cancer states into gliosis remains to be identified.

883 The spatiotemporal trajectory of cancer cells challenges the stratification of GB
884 tumours into distinct subtypes. We captured remarkable regional heterogeneity of cancer cell
885 states within each tumour. This suggests that GB tumour subtypes, commonly classified from
886 single tumour biopsies, rather represent region-specific cancer cell states and tissue niches.
887 We profiled one tumour in whole (AT15), ruling out such sampling confounders, and found it
888 highly enriched for neuronal progenitor-like states. While this could represent a genuine
889 tumour subtype as recently proposed⁷⁴, we speculate that the temporal trajectory of cancer
890 cells across tumour progression poses another challenge. For instance, tumours enriched for
891 developmental-like or gliosis-hypoxia states may represent early or advanced stages of GB,
892 respectively, possibly accounting for the worse survival outcomes associated with
893 “mesenchymal” GBs¹.

894 Previous studies have shown that genetic alterations can promote distinct cancer cell
895 states across GB tumours². While our results suggest that the GB malignant trajectory is
896 shared across the GB genetic hierarchy, we also observed that cell states can be modulated
897 across genetically distinct subclonal lineages, exemplified by the association of chr17 CN
898 events with the neuronal progenitor-like states in one tumour. These results show that
899 intratumour genetic heterogeneity can expand cell state diversity in GB. Furthermore, they
900 imply that mutations can bias the GB malignant cell trajectory towards specific states,
901 presumably through perturbations of key cancer cell state regulators.

902 We identify multi-scale spatial architecture of GB cancer cell subclones using
903 complementary spatial whole genome sequencing and spatial transcriptomics approaches.
904 We developed SpaceTree, a new computational approach to jointly map cancer cell states
905 and clonal lineages in spatial transcriptomics data. These identified that, while GB subclones
906 parcellate tumours into broad spatial domains, they finely spatially intermix within each domain
907 in an intimate manner. This pattern contrasts with the spatial segregation of clones observed
908 in breast⁷⁵ and prostate⁷⁶ cancer. We speculate this reflects the lack of physical barriers in the
909 brain, such as ductal structures in the breast⁷⁵, that can restrict clonal outgrowth, as well as
910 the migratory nature of GB cancer cells.

911 Here, we defined GB cancer cell trajectories at a transcriptomic level. In a companion
912 study, we analysed the joint single nuclei RNA- and ATAC-seq data generated in our atlas to

913 define GB cancer cell states and their cellular transitions at the level of gene regulatory
914 networks²⁰. Beyond identifying transcriptional regulators and plasticity potential of distinct GB
915 cell states, this orthogonal approach suggested highly specific transition routes between
916 cancer states and validated the GB trajectory from developmental-like states towards gliosis
917 and hypoxia.

918 We found that the spatial tumour niches and signalling of myeloid cells correlates with
919 the malignant cell trajectory. These results present TME heterogeneity as a distinct axis of
920 regulation for the broad spectrum of cancer cell states in GB. Whether these myeloid states,
921 given their plasticity¹⁶, are induced by cues from cancer cells or represent independent drivers
922 of tumour biology remains to be resolved. Nevertheless, our results predict highly regionalised
923 myeloid functions tightly linked to cancer cell phenotypes as well as processes such as
924 angiogenesis and immune privilege in GB.

925 Our study has several limitations. First, we do not validate the GB cancer cell
926 trajectories using lineage tracing or experimental perturbations and instead focus on
927 identifying conserved cellular patterns across human GB. Second, while our study examines
928 the dominant cellular trajectories of cancer cells in primary GB, we do not examine their
929 plasticity potential. Specifically, we do not demonstrate whether GB cancer cells can revert
930 from gliosis and hypoxia into developmental-like states. Yet, we note that published lineage
931 tracing studies using PDX models recapitulate the trajectory demonstrated here^{2,6} and our
932 companion study predicts limited cell state transitions of gliosis-hypoxia cells based on gene
933 regulatory networks²⁰. Together, these observations suggest the possibility of limited cancer
934 cell plasticity in GB beyond the trajectory illustrated here.

935 The significant intratumoural heterogeneity of GB underscores the challenges to
936 therapeutic intervention. Our findings propose targeting the spatiotemporal trajectory of cancer
937 cells and destabilising the myeloid TME associated with this trajectory, two conserved features
938 of GB tumours identified here, as potential avenues for exploration.

939 **Data and code availability**

940 The processed single cell and spatial cell atlasing datasets can be explored and downloaded
941 from our interactive webportal (www.gbm space.org/). The raw datasets will be deposited on
942 EGA and the ST datasets will be deposited on the BioImage Archive. SpaceTree code is
943 available at <https://github.com/PMBio/spaceTree> and tutorials are provided at
944 <https://pmbio.github.io/spaceTree/>. ATAC datasets will be made accessible on the data portal
945 upon peer review. All code necessary for re-analysis of the data presented in this paper is
946 available upon request.

947 **Acknowledgements**

948 We thank Jason Swedlow and Josh Titlow for their unwavering support and advice, Dave
949 Adams, Muzz Haniffa and Aidan Maartens for comments on the manuscript, Koen Rademaker
950 for assistance with Xenium data analysis, and Martin Prete for assistance with the GBM-space
951 data portal. This work was supported by Wellcome Leap as part of the Delta Tissue
952 programme (to J.B., D.H.R., S.B., O.S. and O.A.B.) and funded by the Wellcome Trust Grant
953 institutional grants to the Wellcome Sanger Institute (206194 and 220540/Z/20/A), the Chan
954 Zuckerberg Initiative DAF, an advised fund of the Silicon Valley Community Foundation (2023-
955 323357, to O.S. and O.A.B.), a Wellcome Trust personal fellowship (223135/Z/21/Z, to S.B.)
956 and investigator award (108139/Z/15/Z, to D.H.R), an ERC Advanced grant (789054, to
957 D.H.R), the NIHR Cambridge Biomedical Research Centre (NIHR203312), the Medical Data
958 Scientist Program of Heidelberg Faculty of Medicine, Heidelberg University (to A.S.R.), and

959 the National Institute for Health and Care Research (NIHR) Academic Clinical Fellowship (to
960 M. J.). J.S., M.J. J.B. were in part supported by the Francis Crick Institute, which receives its
961 core funding from Cancer Research UK, the UK Medical Research Council and the Wellcome
962 Trust (all under CC001051). The views expressed are those of the authors and not necessarily
963 those of the NHS, the NIHR or the Department of Health.

964 **Author contributions**

965 *Conception:* G.D.J., F.M., T.G., O.L., S.B., O.S., and O.A.B. conceived the study. *Data*
966 *generation:* R.M. provided GB tissue samples, aided by T.S., A.Y. and H.B.. F.M. led
967 experimental design, histology, Visium ST data generation and tissue processing for multiomic
968 profiling. E.T, S.R., I.Mu. and M.P. contributed to histology and Visium ST data generation.
969 T.G. coordinated multiomic profiling, led single nuclei multiome data generation, sample
970 logistics and management. S.Ec., D.Z, A.O., H.P., I.Ma. Z.K. and E.P. contributed to single
971 nuclei multiome data generation. O.G. led sequence library generation for multiome and
972 Visium ST data, aided by Y.W., S.Ec. and T.P. M.D. generated LCM-WGS data. K.R. and
973 A.L.T. generated Xenium data. J.D.BS. generated immunohistochemistry data.
974 *Analysis/interpretation:* G.D.J. led multiome, ST and LCM-WGS data processing and analysis,
975 and interpreted results. M.J. performed histopathological annotations of Visium ST data,
976 validated by A.Q. and M.B. O.L. developed the SpaceTree model, applied the model to Visium
977 ST data and interpreted results. A.A. and S.Er. performed cell2fate trajectory analysis and
978 interpreted results. Q.Z. and A.S.R. contributed to myeloid cell state annotation and CCC
979 analysis. H.M. analysed Xenium data. J.T.H.L. and S.M. assisted with raw Visium ST data
980 processing. R.P. developed the GBM-space data portal for multiomic data with contributions
981 from J.C.. O.G. coordinated raw data submission to data repositories. L.R., M.S. and M.M.
982 contributed to interpretation of GBM cell states. *Supervision:* J.S.R. supervised CCC analysis.
983 J.B. supervised histopathological annotation and immunohistochemistry validation. D.H.R.
984 supervised myeloid data interpretation. S.B. supervised LCM-WGS data generation, analysis
985 and interpretation. O.S. supervised SpaceTree model development, application and data
986 interpretation. O.A.B. supervised cell atlasing data generation, analysis and interpretation.
987 *Manuscript preparation:* G.D., F.M. and O.A.B. prepared figures. G.D. and O.A.B. wrote the
988 manuscript with feedback from all authors.

989 **Conflicts of Interest**

990 J.S.R. reports funding from GSK, Pfizer and Sanofi & fees/honoraria from Traver
991 Therapeutics, Stadapharm, Astex, Owkin, Pfizer, Grunenthal, Moderna and Tempus. O.S. is
992 a paid advisor of Insitro. The other authors declare no competing interests.

993 **Extended data figures and tables**

994 **Extended data figures**

995 Extended Data figures 1-21 are provided in an accompanying file.

996 **Supplementary information**

997 **Supplementary Computational Note**

998 Supplementary note is provided in an accompanying file.

999 **Supplementary tables**

1000 Supplementary tables are provided in an accompanying file.

Table S1	Tumour sample and patient cohort summary
Table S2	Data modality summary
Table S3	Annotation hierarchy and associated cell counts
Table S4	Malignant cell state distribution per tumour
Table S5	Significantly enriched pathways and significantly up-regulated genes per granular cell state
Table S6	Spapros marker gene identification of malignant and TME clusters
Table S7	Recurrent spatial niche factor annotation
Table S8	Curated Xenium probe panel addition
Table S9	Cell state counts per tumour subclone
Table S10	Cell2fate recurrent module annotation
Table S11	CCC malignant niche overview
Table S12	CCC malignant niche cell state identities
Table S13	Myeloid ligands identified across dev-like and gliosis-hypoxia niches
Table S14	Serial sectioning and multiomics plan of tumour tissue samples

1003

1004 **References**

- 1005 1. Verhaak, R. G. W. *et al.* Integrated genomic analysis identifies clinically relevant
1006 subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and
1007 NF1. *Cancer Cell* **17**, 98–110 (2010).
- 1008 2. Neftel, C. *et al.* An Integrative Model of Cellular States, Plasticity, and Genetics for
1009 Glioblastoma. *Cell* **178**, 835–849.e21 (2019).
- 1010 3. Bhaduri, A. *et al.* Outer Radial Glia-like Cancer Stem Cells Contribute to Heterogeneity
1011 of Glioblastoma. *Cell Stem Cell* **26**, (2020).
- 1012 4. Richards, L. M. *et al.* Gradient of Developmental and Injury Response transcriptional
1013 states defines functional vulnerabilities underpinning glioblastoma heterogeneity. *Nature
1014 Cancer* **2**, 157–173 (2021).
- 1015 5. Mathur, R. *et al.* Glioblastoma evolution and heterogeneity from a 3D whole-tumor
1016 perspective. *Cell* **187**, 446–463.e16 (2024).
- 1017 6. Venkataramani, V. *et al.* Glioblastoma hijacks neuronal mechanisms for brain invasion.

- 1018 *Cell* **185**, 2899–2917.e31 (2022).
- 1019 7. Sottoriva, A. *et al.* Intratumor heterogeneity in human glioblastoma reflects cancer
1020 evolutionary dynamics. *Proceedings of the National Academy of Sciences of the United
1021 States of America* **110**, (2013).
- 1022 8. Albiach, A. M. *et al.* Glioblastoma is spatially organized by neurodevelopmental
1023 programs and a glial-like wound healing response. *bioRxiv* 2023.09.01.555882 (2023)
1024 doi:10.1101/2023.09.01.555882.
- 1025 9. Ravi, V. M. *et al.* Spatially resolved multi-omics deciphers bidirectional tumor-host
1026 interdependence in glioblastoma. *Cancer Cell* **40**, 639–655.e13 (2022).
- 1027 10. Klemm, F. *et al.* Interrogation of the microenvironmental landscape in brain tumors
1028 reveals disease-specific alterations of immune cells. *Cell* **181**, 1643–1660.e17 (2020).
- 1029 11. Abdelfattah, N. *et al.* Single-cell analysis of human glioma and immune cells identifies
1030 S100A4 as an immunotherapy target. *Nat. Commun.* **13**, 767 (2022).
- 1031 12. Jackson, C. M., Choi, J. & Lim, M. Mechanisms of immunotherapy resistance: lessons
1032 from glioblastoma. *Nat. Immunol.* **20**, 1100–1109 (2019).
- 1033 13. Pang, L., Khan, F., Heimberger, A. B. & Chen, P. Mechanism and therapeutic potential
1034 of tumor-immune symbiosis in glioblastoma. *Trends Cancer* **8**, 839–854 (2022).
- 1035 14. Hara, T. *et al.* Interactions between cancer cells and immune cells drive transitions to
1036 mesenchymal-like states in glioblastoma. *Cancer Cell* **39**, 779–792.e11 (2021).
- 1037 15. Kloosterman, D. J. *et al.* Macrophage-mediated myelin recycling fuels brain cancer
1038 malignancy. *Cell* **187**, 5336–5356.e30 (2024).
- 1039 16. Miller, T. E. *et al.* Programs, origins and immunomodulatory functions of myeloid cells in
1040 glioma. *Nature* 1–11 (2025).
- 1041 17. Haley, M. J. *et al.* Hypoxia coordinates the spatial landscape of myeloid cells within
1042 glioblastoma to affect survival. *Sci. Adv.* **10**, eadj3301 (2024).
- 1043 18. Greenwald, A. C. *et al.* Integrative spatial analysis reveals a multi-layered organization
1044 of glioblastoma. *Cell* **187**, 2485–2501.e26 (2024).
- 1045 19. Li, T. *et al.* WebAtlas pipeline for integrated single-cell and spatial transcriptomic data.

- 1046 *Nat. Methods* **22**, 3–5 (2025).
- 1047 20. Saraswat, M. *et al.* Decoding Plasticity Regulators and Transition Trajectories in
1048 Glioblastoma with Single-cell Multiomics. Preprint at *bioRxiv*. (2025).
- 1049 21. Lopez, R., Regier, J., Cole, M. B., Jordan, M. I. & Yosef, N. Deep generative modeling
1050 for single-cell transcriptomics. *Nat Methods* **15**, 1053–1058 (2018).
- 1051 22. Braun, E. *et al.* Comprehensive cell atlas of the first-trimester developing human brain.
1052 *Science* **382**, eadf1226 (2023).
- 1053 23. Ricci-Vitiani, L. *et al.* Tumour vascularization via endothelial differentiation of
1054 glioblastoma stem-like cells. *Nature* **468**, 824–828 (2010).
- 1055 24. Varn, F. S. *et al.* Glioma progression is shaped by genetic evolution and
1056 microenvironment interactions. *Cell* **185**, 2184–2199.e16 (2022).
- 1057 25. Kleshchevnikov, V. *et al.* Cell2location maps fine-grained cell types in spatial
1058 transcriptomics. *Nat. Biotechnol.* **40**, 661–671 (2022).
- 1059 26. Puchalski, R. B. *et al.* An anatomic transcriptional atlas of human glioblastoma. *Science*
1060 (*New York, N.Y.*) **360**, 660 (2018).
- 1061 27. Biancalani, T. *et al.* Deep learning and alignment of spatially resolved single-cell
1062 transcriptomes with Tangram. *Nat Methods* **18**, 1352–1362 (2021).
- 1063 28. Scherer, H. Structural Development in Gliomas. *American Journal of Cancer* **34**, 333–
1064 351 (1938).
- 1065 29. Wang, J. *et al.* Invasion of white matter tracts by glioma stem cells is regulated by a
1066 NOTCH1-SOX2 positive-feedback loop. *Nat. Neurosci.* **22**, 91–105 (2019).
- 1067 30. Gieseks, A. & Westphal, M. Glioma invasion in the central nervous system.
1068 *Neurosurgery* **39**, 235–50;discussion250–2 (1996).
- 1069 31. Hara, M. *et al.* Interaction of reactive astrocytes with type I collagen induces astrocytic
1070 scar formation through the integrin-N-cadherin pathway after spinal cord injury. *Nat Med*
1071 **23**, 818–828 (2017).
- 1072 32. Escartin, C. *et al.* Reactive astrocyte nomenclature, definitions, and future directions.
1073 *Nat. Neurosci.* **24**, 312–325 (2021).

- 1074 33. You, T. *et al.* IL-17 induces reactive astrocytes and up-regulation of vascular endothelial
1075 growth factor (VEGF) through JAK/STAT signaling. *Sci. Rep.* **7**, 41779 (2017).
- 1076 34. Huynh, J., Chand, A., Gough, D. & Ernst, M. Therapeutically exploiting STAT3 activity in
1077 cancer - using tissue repair as a road map. *Nat. Rev. Cancer* **19**, 82–96 (2019).
- 1078 35. Katsuno, Y. & Deryck, R. Epithelial plasticity, epithelial-mesenchymal transition, and
1079 the TGF- β family. *Dev. Cell* **56**, 726–746 (2021).
- 1080 36. Yang, J. *et al.* Guidelines and definitions for research on epithelial-mesenchymal
1081 transition. *Nat. Rev. Mol. Cell Biol.* **21**, 341–352 (2020).
- 1082 37. Oliver, T. R. W. *et al.* Clonal diversification and histogenesis of malignant germ cell
1083 tumours. *Nat Commun* **13**, 4272 (2022).
- 1084 38. Trinh, M. K. *et al.* Precise identification of cancer cells from allelic imbalances in single
1085 cell transcriptomes. *Communications Biology* **5**, 1–8 (2022).
- 1086 39. Gillis, S. & Roth, A. PyClone-VI: scalable inference of clonal population structures using
1087 whole genome data. *BMC Bioinformatics* **21**, 1–16 (2020).
- 1088 40. Ma, C. *et al.* Inferring allele-specific copy number aberrations and tumor
1089 phylogeography from spatially resolved transcriptomics. *Nature Methods* **21**, 2239–2247
1090 (2024).
- 1091 41. Stensjøen, A. L. *et al.* Growth dynamics of untreated glioblastomas in vivo. *Neuro.*
1092 *Oncol.* **17**, 1402–1411 (2015).
- 1093 42. Aivazidis, A. *et al.* Cell2fate infers RNA velocity modules to improve cell fate prediction.
1094 *Nat. Methods* 1–10 (2025).
- 1095 43. Myers, B. L. *et al.* Transcription factors ASCL1 and OLIG2 drive glioblastoma initiation
1096 and co-regulate tumor cell types and migration. *Nat. Commun.* **15**, 10363 (2024).
- 1097 44. Wu, Y. *et al.* Glioblastoma epigenome profiling identifies SOX10 as a master regulator
1098 of molecular tumour subtype. *Nat. Commun.* **11**, 6434 (2020).
- 1099 45. Wang, L. *et al.* A single-cell atlas of glioblastoma evolution under therapy reveals cell-
1100 intrinsic and cell-extrinsic therapeutic targets. *Nat. Cancer* **3**, 1534–1552 (2022).
- 1101 46. Aprato, J. *et al.* Myrf guides target gene selection of transcription factor Sox10 during

- 1102 oligodendroglial development. *Nucleic Acids Res.* **48**, 1254–1270 (2020).
- 1103 47. Hornig, J. *et al.* The transcription factors Sox10 and Myrf define an essential regulatory
1104 network module in differentiating oligodendrocytes. *PLoS Genet.* **9**, e1003907 (2013).
- 1105 48. Brooks, L. J. *et al.* The white matter is a pro-differentiative niche for glioblastoma. *Nat.*
1106 *Commun.* **12**, 2184 (2021).
- 1107 49. Allen, C. E., Mak, C.-H. & Wu, L.-C. The kappa B transcriptional enhancer motif and
1108 signal sequences of V(D)J recombination are targets for the zinc finger protein
1109 HIVEP3/KRC: a site selection amplification binding study. *BMC Immunol* **3**, 10 (2002).
- 1110 50. Oukka, M., Wein, M. N. & Glimcher, L. H. Schnurri-3 (KRC) interacts with c-Jun to
1111 regulate the IL-2 gene in T cells. *J. Exp. Med.* **199**, 15–24 (2004).
- 1112 51. Luo, D. *et al.* The PDK1/c-Jun pathway activated by TGF-β induces EMT and promotes
1113 proliferation and invasion in human glioblastoma. *Int. J. Oncol.* **53**, 2067–2080 (2018).
- 1114 52. Bartolomé, R. A. *et al.* Schnurri-3 drives tumor growth and invasion in cancer cells
1115 expressing interleukin-13 receptor alpha 2. *Cell Death Dis.* **14**, 742 (2023).
- 1116 53. Wang, L.-B. *et al.* Proteogenomic and metabolomic characterization of human
1117 glioblastoma. *Cancer Cell* **39**, 509–528.e20 (2021).
- 1118 54. Liu, J. *et al.* An integrated TCGA Pan-Cancer Clinical Data Resource to drive high-
1119 quality survival outcome analytics. *Cell* **173**, 400–416.e11 (2018).
- 1120 55. Ceccarelli, M. *et al.* Molecular profiling reveals biologically discrete subsets and
1121 pathways of progression in diffuse glioma. *Cell* **164**, 550–563 (2016).
- 1122 56. Phillips, H. S. *et al.* Molecular subclasses of high-grade glioma predict prognosis,
1123 delineate a pattern of disease progression, and resemble stages in neurogenesis.
1124 *Cancer Cell* **9**, 157–173 (2006).
- 1125 57. Zitvogel, L., Galluzzi, L., Kepp, O., Smyth, M. J. & Kroemer, G. Type I interferons in
1126 anticancer immunity. *Nat. Rev. Immunol.* **15**, 405–414 (2015).
- 1127 58. Borden, E. C. Interferons α and β in cancer: therapeutic opportunities from new insights.
1128 *Nat. Rev. Drug Discov.* **18**, 219–234 (2019).
- 1129 59. Sankowski, R. *et al.* Multiomic spatial landscape of innate immune cells at human

- 1130 central nervous system borders. *Nat Med* **30**, 186–198 (2024).
- 1131 60. Shah, K., Tang, Y., Breakefield, X. & Weissleder, R. Real-time imaging of TRAIL-
1132 induced apoptosis of glioma tumors *in vivo*. *Oncogene* **22**, 6865–6872 (2003).
- 1133 61. Liau, L. M. *et al.* Identification of a human glioma-associated growth factor gene,
1134 granulin, using differential immuno-absorption. *Cancer Res* **60**, 1353–1360 (2000).
- 1135 62. Bandey, I., Chiou, S.-H., Huang, A.-P., Tsai, J.-C. & Tu, P.-H. Progranulin promotes
1136 Temozolomide resistance of glioblastoma by orchestrating DNA repair and tumor
1137 stemness. *Oncogene* **34**, 1853–1864 (2015).
- 1138 63. Ihim, S. A. *et al.* Interleukin-18 cytokine in immunity, inflammation, and autoimmunity:
1139 Biological role in induction, regulation, and treatment. *Front Immunol* **13**, 919973 (2022).
- 1140 64. Yang, R. *et al.* Galectin-9 interacts with PD-1 and TIM-3 to regulate T cell death and is a
1141 target for cancer immunotherapy. *Nat Commun* **12**, 832 (2021).
- 1142 65. Kang, R., Zhang, Q., Zeh, H. J., 3rd, Lotze, M. T. & Tang, D. HMGB1 in cancer: good,
1143 bad, or both? *Clin Cancer Res* **19**, 4046–4057 (2013).
- 1144 66. Murdoch, C., Muthana, M., Coffelt, S. B. & Lewis, C. E. The role of myeloid cells in the
1145 promotion of tumour angiogenesis. *Nat. Rev. Cancer* **8**, 618–631 (2008).
- 1146 67. Hu, F. *et al.* Glioma-derived versican promotes tumor expansion via glioma-associated
1147 microglial/macrophages Toll-like receptor 2 signaling. *Neuro Oncol* **17**, 200–210 (2015).
- 1148 68. Wu, Y., Chen, L., Zheng, P.-S. & Yang, B. B. beta 1-Integrin-mediated glioma cell
1149 adhesion and free radical-induced apoptosis are regulated by binding to a C-terminal
1150 domain of PG-M/versican. *J Biol Chem* **277**, 12294–12301 (2002).
- 1151 69. Teo, Z. *et al.* Angiopoietin-like 4 induces a β-catenin-mediated upregulation of ID3 in
1152 fibroblasts to reduce scar collagen expression. *Scientific Reports* **7**, 6303 (2017).
- 1153 70. Larrue, C. *et al.* Adrenomedullin-CALCRL axis controls relapse-initiating drug tolerant
1154 acute myeloid leukemia cells. *Nat Commun* **12**, 422 (2021).
- 1155 71. Cui Zhou, D. *et al.* Spatially restricted drivers and transitional cell populations cooperate
1156 with the microenvironment in untreated and chemo-resistant pancreatic cancer. *Nat
1157 Genet* **54**, 1390–1405 (2022).

- 1158 72. Qi, L., Chen, J., Yang, Y. & Hu, W. Hypoxia Correlates With Poor Survival and M2
1159 Macrophage Infiltration in Colorectal Cancer. *Front Oncol* **10**, 566430 (2020).
- 1160 73. MacCarthy-Morrogh, L. & Martin, P. The hallmarks of cancer are also the hallmarks of
1161 wound healing. *Sci Signal* **13**, (2020).
- 1162 74. Drexler, R. *et al.* A prognostic neural epigenetic signature in high-grade glioma. *Nat Med*
1163 **30**, 1622–1635 (2024).
- 1164 75. Lomakin, A. *et al.* Spatial genomics maps the structure, nature and evolution of cancer
1165 clones. *Nature* **611**, 594–602 (2022).
- 1166 76. Erickson, A. *et al.* Spatially resolved clonal copy number alterations in benign and
1167 malignant tissue. *Nature* **608**, 360–367 (2022).

1168 **Methods**

1169 **Human Subjects**

1170 Patients with suspected GB were identified pre-operatively and consented for entry into the
1171 study. Surgery was performed at Cambridge University Hospitals NHS Foundation Trust.
1172 Written and informed consent was obtained in accordance with the guidelines in The
1173 Declaration of Helsinki 2000. Ethical approval for the use of these tissues was obtained from
1174 the Cambridge Local Research Ethics Committee (REC 18/EE/0172). All patients underwent
1175 5-ALA guided tumour resection as per local protocols. During tumour debulking, regions of
1176 high fluorescence were identified, their spatial location recorded and the tissue samples were
1177 collected for this study. Matched whole blood was taken during surgery for germline
1178 characterisation. The clinical characteristics of the patient cohort and samples are detailed in
1179 **Table S1**.

1180 **Serial tissue sectioning for paired multi-omic analysis**

1181 Tissue was sampled from multiple sites of each GB tumour, targeting superior, anterior,
1182 posterior and middle regions where possible. Each tissue sample was immediately washed in
1183 saline buffer and embedded in OCT medium (Scigen OCT Compound, #4586) using a dry ice-
1184 cooled bath of isopentane at -75 °C. OCT-embedded samples were sectioned using a cryostat
1185 (Leica CX3050S). The fresh frozen tissue blocks were trimmed until the tissue surface was
1186 fully exposed. Two to three 10 µm thick sections were collected to check RNA integrity and
1187 another section was processed for Haematoxylin and Eosin (H&E) staining to assess tissue
1188 morphology. The RNA quality of each sample was evaluated by Tapestation (Tapestation RNA
1189 ScreenTape, Agilent) after isolating RNA (Qiagen EZ2 Automated RNA Extraction using EZ2
1190 RNA/miRNA Tissue Kit Cat. No.959035). Only samples with RNA integrity number (RIN)
1191 values >7 were used for omic profiling. Blocks with good morphology and RIN were
1192 subsequently cryosectioned serially for different profiling applications as listed below (see
1193 **Table S14** for the sectioning plan of each tumour tissue block).

1194 1) *Sectioning for single nuclei isolation*: A series of 50 µm thick sections, totalling 500
1195 to 900 µm thick volume depending on the size of each tumour block, were collected in pre-
1196 chilled homogenization glass tubes and kept on dry ice until processing.

1197 2) *Sectioning for Visium and Visium TO*: Sectioning for Visium ST was done according
1198 to the manufacturer's guidelines. The Visium Spatial Gene Expression (GE) (catalog no.
1199 2000233, 10x Genomics) slide was pre-chilled in a cryostat and four 10 µm sections were
1200 mounted onto the 0.42 cm² capture areas. 2 sections were processed for Visium from each
1201 tumour block. For tumour samples larger than the capture areas, the blocks were scored and
1202 each half was placed on an individual capture area. Sections from two different blocks were
1203 collected on one Visium GE slide. Sectioning for Visium TO was done similarly, with four
1204 sections collected per block and two blocks per slide (catalog no. 3000394, 10x Genomics)
1205 were collected. Slides were then placed in a slide mailer and stored in a bag with desiccant in
1206 -80°C. Adjacent tissue sections were assessed with H&E staining.

1207 3) *Sectioning for Visium CytAssist*: 10 µm cryosections were mounted on a
1208 Superfrost™ Plus Microscope Slide (Fisherbrand™) and stored in -80°C. Adjacent tissue
1209 sections were assessed with H&E staining.

1210 4) *Sectioning for Xenium ST*: 10 µm cryosections were mounted onto Xenium slides
1211 (10x Genomics 3000941) and stored at -80°C for up to 4 weeks.

1212 5) *Sectioning for LCM-WGS*: For LCM, sections were collected on a PEN
1213 MembranSlide (cat 11505158, Leica) and the slides were kept in the cryostat chamber until
1214 moved to -80°C. Adjacent tissue sections were assessed with H&E staining.

1215 6) *Sectioning for Immunohistochemistry*: 10 µm cryosections were mounted onto
1216 Superfrost™ Plus Microscope Slides (Fisherbrand™) and stored at -80°C.

1217 **H&E staining and imaging**

1218 Tissue sections were removed from -80°C and air dried before fixation in 70% ethanol for
1219 3 minutes (min). After rinsing with deionised water, slides were stained in Gill II haematoxylin
1220 solution (Leica) for 7 seconds (s). Slides were completely rinsed in 2 washes of tap water for
1221 20 and 25 s. Slides were then stained with aqueous eosin 1% (Leica) for 10 s and rinsed
1222 twice with deionised water for 20s and 25s, followed by dehydration through an ethanol
1223 series (70%, 70%, 100%, 100%; 20 s each) and cleared twice in 100% xylene for 10 s.
1224 Slides were coverslipped and allowed to air dry before being imaged on a Hamamatsu
1225 Nanozoomer 2.0HT digital slide scanner.

1226 **Single-nuclei extraction**

1227 Nuclei were extracted from fresh frozen tissue sections that were homogenised using a glass
1228 Dounce homogenizer (Sigma) in nuclei isolation buffer (3mM MgCl₂, 10mM NaCl, 10mM Tris
1229 (buffer pH7.4), 1 mM DTT, 0.1% Tween-20, 0.1% Nonidet P40, 1% BSA and 0.01% Digitonin)
1230 in the presence of Protector RNase Inhibitor (Roche) at 0.2 U/µl. Tissue was homogenised
1231 using 10 strokes with pestle A and then 10 strokes with pestle B. Nuclei were then filtered
1232 through a 40 µM filter, collected at 500 RCF and resuspended in 0.25 ml of storage buffer
1233 (PBS containing 1% BSA and Protector RNase Inhibitor (Roche) 1 U/µl). An aliquot of the
1234 nuclei suspension was incubated with Trypan Blue (Gibco 15250061) for counting and purified
1235 from debris using a Percoll gradient. The cleaned nuclei suspension was stained with Trypan
1236 blue and counted.

1237 **10x Genomics Chromium GEX and ATAC library preparation and sequencing**
1238 For the snRNA-seq experiments, two to three 10x reactions were prepared per tumour site
1239 and loaded onto the 10X chromium controller according to the manufacturer's protocol for the
1240 Chromium Next GEM Single Cell Multiome ATAC + Gene Expression assay. Post-GEM-RT
1241 cleanup, cDNA amplification and 3' gene expression library construction were carried out as
1242 per the Chromium Single Cell 3' Reagent Kits v3 User Guide, to obtain between 5000-10,000
1243 nuclei per reaction. Libraries were paired end-sequenced on a NovaSeq 6000 System
1244 (Illumina) using the Novaseq S4 Flowcell, targeting a minimum coverage of 100,000 read pairs
1245 per nuclei per modality. The following sequencing formats were employed for GEX and ATAC
1246 respectively:

- 1247 - GEX: Read 1: 28 cycles i7 Index: 10 cycles i5 Index: 10 cycles Read 2: 90 cycles
1248 - ATAC: Read 1N: 50 cycles i7 Index: 8 cycles i5 Index: 24 cycles Read 2N: 49 cycles

1249 **Visium tissue optimisation**
1250 Tissue sections on TO slides were fixed in chilled methanol and H&E-stained according to the
1251 Visium Spatial Tissue Optimization User Guide (catalog no. CG000238 Rev A, 10x
1252 Genomics). Initial optimization experiments assessed 7 different permeabilization times at 6
1253 min intervals and consistently indicated 12 to 30 min as optimal pre-treatment. Hence,
1254 subsequent TO experiments assessed 4 different permeabilization times (12, 18, 24 and 30
1255 min) per tissue block, testing 2 blocks per TO slide. Brightfield histology images were taken
1256 using a 40 \times objective on a Hamamatsu Nanozoomer 2.0HT digital slide scanner and
1257 Fluorescent images were taken with a Cy3 filter using a 20 \times air objective and 200-ms exposure
1258 time on an Opera Phenix™ Plus.

1259 **10x Genomics Visium library preparation and sequencing**
1260 Libraries were generated according to the manufacturer's instructions (Visium Spatial Gene
1261 Expression User Guide, CG000239 Rev A, 10x Genomics). Briefly, sections were fixed with
1262 cold methanol, stained with haematoxylin and eosin and imaged on a Hamamatsu
1263 NanoZoomer 2.0HT before permeabilization (most blocks processed for 12-24 mins), reverse
1264 transcription and cDNA synthesis using a template-switching protocol. Second-strand cDNA
1265 was liberated from the slide and single-indexed libraries were prepared using a 10x Genomics
1266 PCR-based protocol. Sequencing was performed on the NovaSeq 6000, aiming for a minimum
1267 of 50000 read-pairs per spot, with the following sequencing format; read 1: 28 cycles, i7 index:
1268 10 cycles, i5 index: 10 cycles and read 2: 90 cycles.

1269 **Visium CytAssist**
1270 Visium CytAssist slides were dried at RT for 5 min and fixed in cold methanol at -20°C for 30
1271 min, then H&E-stained and imaged on Hamamatsu NanoZoomer 2.0HT. After destaining,
1272 human whole transcriptome probe pairs were hybridised and ligated to the tissue RNA. The
1273 CytAssist instrument was used to facilitate the transfer of transcriptomic probes from the
1274 standard glass slide to the Visium CytAssist Spatial Gene Expression Slide (v2, 11 mm
1275 capture area). Probe hybridization, probe ligation, release, extension, pre-amplification, and
1276 library preparation followed the Visium CytAssist Spatial Gene Expression Reagent Kits User
1277 Guide (CG000495).

1278 Libraries were sequenced with paired-end dual-indexing (28 cycles Read 1, 10 cycles
1279 i7, 10 cycles i5, 90 cycles Read 2) on a Illumina NovaSeq 6000, aiming for a minimum 70,000

1280 read pairs per spot. The Space Ranger v2.1.0 (10x Genomics) and the GRCh38-2020-A
1281 reference were used to process FASTQ files.

1282

1283 **Xenium Spatial Transcriptomics**

1284 Samples were processed using the Xenium Sample Preparation Kit (10x Genomics 1000460)
1285 as per the manufacturer's protocols (CG000581 for pre-treatment including fixation with
1286 formaldehyde and permeabilization with SDS and methanol, followed by CG000582 for probe
1287 hybridisation and rolling circle product generation). Following autofluorescence quenching and
1288 nuclei counterstaining as per the manufacturer's instructions (CG000582), slides were
1289 transferred onto a Xenium Analyzer instrument alongside Decoding Reagents (10x Genomics
1290 1000461) and Decoding Consumables (1000487) all prepared according to the
1291 manufacturer's protocol (CG000584). The DAPI-stained nuclei were exported from the
1292 instrument in the format of ome.tif images.

1293 **Laser capture microdissection and DNA library preparation**

1294 Fresh-frozen tissue samples were fixed using PAXgene fixative (Qiagen) and embedded in
1295 paraffin. Sections were cut at 10 µm thickness, with additional reference slides constructed at
1296 4 µm thickness to ensure the correct region of the tumour was identified. All slides were stained
1297 with H&E (Leica). Regions of interest were isolated by laser capture microdissection (LCM)
1298 and lysed using the Arcturus PicoPure Kit (Applied Biosystems) in accordance with the
1299 manufacturer's instructions. Low-input, whole-genome sequencing libraries were created from
1300 each LCM lysate^{77,78}. Sequencing was performed on the Illumina NovaSeq platform to
1301 generate 150 bp paired-end reads to a target coverage of 30X.

1302 **Immunohistochemistry and microscopy**

1303 Fresh frozen GB tissue sections were fixed in 10% neutral buffered formalin (NBF) for 30 min.
1304 Slides were washed and incubated at 4°C for 10 min in solution of 2% H₂O₂ in methanol to
1305 block endogenous peroxidase. Slides were then washed and rinsed in 100% IMS and air dried
1306 before loading onto an automated staining platform (Leica Bond Rx). Antigen
1307 retrieval/stripping steps between each antibody were performed with Epitope Retrieval
1308 Solution 1 (pH9) (Leica, AR9961) for 20 min at 95°C. Slides were then incubated with 0.1%
1309 BSA-PBST solution for 5 min at RT for protein blocking.

1310 For multiplexed immunofluorescence, primary antibodies (**Table 1**) were labelled with
1311 Opal reagents (Akoya Biosciences). Anti-rabbit Polymer HRP (Leica, RE7260-CE, RTU) was
1312 used as a conjugated secondary antibody. The antibody order, dilution and incubation time of
1313 primary antibody, secondary antibody and Opal dyes are detailed on **Table 2**. Slides were
1314 counterstained with DAPI (Thermo Scientific, 62248; 1:2500), mounted with Prolong Gold
1315 Antifade reagent (Invitrogen, P36934), and scanned following the Phenolmager whole slide
1316 workflow, Akoya Phenolmager HT (formerly Vectra Polaris) using Phenolmager HT 2.0
1317 software.

1318

1319 Table 1: Primary antibodies.

1320

Antibodies	Host/Clonality	Source	Identifier
anti-AKAP12	– C	Rabbit polyclonal terminal	Abcam ab204559

1321 anti-HIG2 (HILPDA) Rabbit polyclonal Abcam ab252976
1322

1323 Table 2: Multiplex immunofluorescence antibody parameters

Position	Ab	AR	Primary dilution	Primary incubation (min)	Secondary	Secondary incubation (min)	Opal	Opal dilution	Opal incubation (min)
1	AKAP12	ER1 20	1 in 750	15	Rabbit Polymer	8	690	1 in 150	10
2	HILPDA	ER1 20	1 in 400	15	Rabbit Polymer	8	520	1 in 300	10

1324 Diluents (Akoya)

1325 OPAL diluent – 1xPlus Automation Amplification Diluent cat number FP1609

1326 Antibody diluent - 1X Antibody Diluent / Block, 100ml cat number ARD1001EA

1328 Pathological annotation

1329 A pathologist annotated digitised images of the H&E stained tissue sections with histologically-
1330 distinct anatomic features. Histological annotations were defined according to the IVY
1331 Glioblastoma Atlas Project reference; an existing taxonomy system with established inter-
1332 neuropathologist agreement and clinical relevance²⁶. Histological annotations included:

- 1333 1) *Leading edge (LE)*: the outermost boundary of the tumour where the laminar
1334 architecture of the cortical layers is frequently evident.
- 1335 2) *Infiltrating tumour (IT)*: the intermediate zone between the leading edge and cellular
1336 tumour, frequently marked by perineuronal satellitosis.
- 1337 3) *Cellular tumour (CT)*: the major part of the tumour core.
- 1338 4) *Necrosis*: dead or dying tissue, marked by the presence of karyorrhectic or cellular
1339 debris and the absence of crisp cytological architecture.
- 1340 5) *Pseudopalisading cells around necrosis (PAN)*: the narrow boundary of cells arranged
1341 like pseudopalisades along the perimeter of necrosis in the tumour core.
- 1342 6) *Perinecrotic zone (PNZ)*: a boundary of tumour cells in the tumour core along the edge
1343 of necrosis that lacks a clear demarcation of pseudopalisading cells around necrosis.
- 1344 7) *Microvascular proliferation (MVP)*: two or more blood vessels sharing a common
1345 vessel wall of endothelial and smooth muscle cells typically in the tumour core, and
1346 arranged in the shape of a glomerulus or garland of multiple interconnected blood
1347 vessels.
- 1348 8) *Hyperplastic blood vessels (HPV)*: blood vessels with thickened walls found anywhere
1349 in the tumour.

1350 Two additional pathologists independently validated the histological annotations.

1351 Single nuclei multiome data processing and quality control

1352 We aligned reads from each snRNA-seq and ATAC-seq library to a custom-made genome
1353 consisting of 10X Genomics' GRCh38 3.0.0 pre-mRNA reference genome and 10X Genomics
1354 Cell Ranger ARC 2.0.1 ATAC genome. To perform quantification and initial quality control, we
1355 used the default parameters in the Cell Ranger ARC software (v2.0.1; 10X Genomics)^{79,80}.
1356 This was followed by CellBender⁸¹ (v0.2.0), which was applied to the Cell Ranger output to
1357 correct for background noise and identify empty droplets. The parameters --expected-droplets
1358 and --total-droplets-included were identified based on sample UMI curves. Quality control of

1359 the RNA data was performed using Cell Ranger ARC filtered count matrices with Scanpy⁸²
1360 (v1.9.3). This involved removing nuclei with total gene counts <500, total counts UMI <1000,
1361 and nuclei with >10% of reads mapping to mitochondrial content. We then applied Scrublet⁸³
1362 (v0.2.3) on each library individually and filtered our data based on a two-step method adapted
1363 from previously described median absolute deviation (MAD) thresholding. Individual cells were
1364 first filtered based on methods described in Pijuan-Sala et al.⁸⁴ (FDR < 0.05) and then clusters
1365 were subsequently filtered using similar methods^{84,85} to ensure the removal of probable
1366 subthreshold doublets (FDR < 0.1). Following doublet detection and removal, any remaining
1367 cell with a total UMI count >75,000 was removed from the dataset. Finally, we calculated cell
1368 cycling scores using Scanpy based on the gene list described by Tirosh et al.⁸⁶. ATAC data
1369 were processed according to the ArchR⁸⁷ workflow (minimum fragments >1000 and >4
1370 transcriptional start sites). These barcodes were then filtered by the list of barcodes passing
1371 RNA quality control filters, effectively removing the majority of low-quality ATAC barcodes.
1372 Finally, the RNA data was subset by barcodes passing ATAC filters to ensure symmetry
1373 between modalities for downstream applications.

1374 **snRNA-seq integration**

1375 snRNA-seq libraries were concatenated into a single dataset and log-normalised for highly
1376 variable gene (HVG) selection. HVGs were identified with Scanpy using tumour ID as a batch
1377 key and dispersion-based feature selection with following non-default parameters:
1378 (min_mean=0.0075, max_mean=4, min_disp=0.1). The resulting list of 10,701 HVGs was then
1379 used for batch integration with scVI²¹ (v0.19.0), for which we used each 10x reaction as the
1380 batch and included tumour ID, tumour site, 10x reaction date, and cell cycle phase as model
1381 covariates. Model hyperparameters included 50 latent space dimensions and 2 hidden layers,
1382 with 1024 nodes per layer. After training the model, leiden clustering was performed using the
1383 nearest neighbours graph constructed from the resulting scVI embeddings. Leiden clusters
1384 corresponding to putative and broadly defined TME populations were then annotated (e.g.
1385 oligodendrocytes, myeloid cells, neurons, lymphocytes) based on marker gene expression.

1386 **snRNA-seq CNA inference**

1387 The detection of cell-specific copy number aberrations (CNAs) was performed on each tumour
1388 individually using inferCNVpy (<https://github.com/icbi-lab/infercnvpy>), a python
1389 implementation of the inferCNV⁸⁸ method, with parameters consistent with previous GB
1390 studies⁸⁹ and a window size of 250 genes. InferCNV also requires a set of non-malignant
1391 reference cells, for which we used the least ambiguous TME populations identified based on
1392 marker gene expression: macrophages, oligodendrocytes, neurons, and lymphocytes.

1393 **Clustering and cell state annotation**

1394 Following integration and coarse-grained clustering, the first step in atlas annotation was
1395 accurately delineating between malignant and non-malignant cells. This followed an approach
1396 adapted from previous GB studies², in which two separate CNA metrics were used to score
1397 cells: CNA signal and CNA correlation. We calculated the absolute CNA signal by averaging
1398 absolute CNA scores originating from the diagnostically relevant chromosome 7 gain and
1399 chromosome 10 loss events. CNA correlation was computed by calculating the correlation
1400 between the CN profile of a given cell and mean CN profile of cells identified as non-malignant
1401 through marker gene expression. Malignant cell signal was defined as a CNA signal > 0.02
1402 and a CNA correlation > 0.3. As static malignant cell thresholds may not identify all malignant

1403 cells and are dependent on the coverage in expression-based CN inference methods, we then
1404 annotated clusters as putatively malignant on the basis of the percentage of malignant cells.
1405 The majority of malignant clusters exhibited a malignant percentage > 20% but malignant
1406 clusters with a percentage $\geq 3\%$ were flagged as putatively malignant.

1407 We then reintegrated and sub-clustered each major TME population and all malignant
1408 cells, individually. This was followed by a subsequent CNA evaluation to ensure there were
1409 no misannotated subclusters. TME cell types were then identified manually using lists of well-
1410 established human brain marker genes (**Extended Data Fig. 6A**). TME subclusters with a
1411 malignant cell percentage $\geq 3\%$ were also flagged. Malignant cells subclusters were annotated
1412 using composite of four different approaches: (1) We calculated a series of gene module
1413 scores based on the workflow outlined by Neftel et al.² in conjunction with the *score_genes*
1414 function in Scanpy, which is analogous to their methodology. (2) We ran mapped our GB data
1415 onto a developing brain atlas²² using scPoli⁹⁰ with default parameters to identify developing
1416 brain analogs to malignant cell state identities. (3) Markers specific to each cluster were
1417 identified using *scanpy.tl.rank_genes_groups* based on Wilcoxon rank-sum. (4) Gene set
1418 enrichment analysis was performed on differential gene expression results using decoupler⁹¹
1419 with MGSigDB signatures^{92,93}.

1420 **Visium spatial cell type mapping and tissue niche identification**

1421 Visium spatial RNA-sequencing data were mapped to the official Cell Ranger GRCh38 using
1422 10x Genomics Space Ranger Software Suite (v2.0 for standard Visium, v2.1 for Visium
1423 Cytassist). To spatially map malignant and TME cell states in Visium data, we used the
1424 cell2location model²⁵ to deconvolve the Visium spot transcriptomes to cell state abundances
1425 based on paired snRNA-seq reference signatures. We generated unique reference signatures
1426 for each tumour to better reflect malignant cell gene expression heterogeneity. This was done
1427 using the negative binomial regression model packaged with cell2location and included the
1428 following changes to default training parameters: *max_epochs*=400, *batch_size*=10000, and
1429 *lr*=0.002. The cell2location model was run on each individual tumour (i.e. batching multiple
1430 Visium sections) with the following modifications to default parameters: model parameters
1431 were adjusted to *N_cells_per_location*=30 and *detection_alpha*=200; training parameters
1432 adjustments included *max_epochs*=6000, and a *batch_size* set to approximately 25% of the
1433 total number of spots profiled for a given tumour.

1434 To identify tissue niches, we applied the scikit-learn NMF model packaged with
1435 cell2location to our cell state abundance results²⁵. This allowed us to decompose matrices of
1436 inferred cell state abundances into components representing patterns of cell state co-
1437 localisation, analogous to microenvironmental niches. The model was trained using 16 factors
1438 for each individual tumour. Of the resulting matrices, we normalised the matrices composed
1439 of factor loadings per cell state and performed a cross-tumour annotation of associated
1440 factors. Factors were grouped into meta-factors, described as niches, using agglomerative
1441 hierarchical clustering with cosine distance. These grouped factors were annotated according
1442 to the most frequently observed cell states in each factor, in addition to the correlation between
1443 cell state and factor abundances.

1444 To define a measure of spatial intermixing we calculated Shannon entropy based on
1445 cell2location abundance frequencies for each Visium spot. To ensure spots were highly
1446 representative of a given cell state, only spots in the 98th percentile of abundance for each cell
1447 state were considered.

1448 To quantify the relative spatial organisation of cell states, niche factors, and cell2fate
1449 modules, we calculated the minimum pairwise spot distance between all component pairs for
1450 each of these features independently using k-d trees. As the inferred abundance of these
1451 features are continuous variables, we imposed thresholds based on feature abundance to
1452 quantify spots associated with high feature abundance. This included two filters: abundance
1453 in a spot must be greater or equal to the median abundance of a feature, and, to remove
1454 abundances with poor signal, the abundance must also be greater than a base threshold. This
1455 allowed us to focus on comparing regions of high density while also accounting for biologically
1456 relevant differences in abundance. The base threshold varied depending on the application
1457 and was based on the distribution of abundances (cell2location = 4, niche factor = 10, and
1458 cell2fate modules = 2). As some spatial features have inherently different spatial distributions
1459 – e.g. the compactness and density of vasculature compared to malignant niches – we
1460 summarised these distances by calculating the bottom 25th percentile of nearest neighbour
1461 distances.

1462 **Xenium panel design and data processing**

1463 Our Xenium panel included an additional 62 probes alongside the standard 266-gene Xenium
1464 human brain panel (<https://www.10xgenomics.com/products/xenium-panels>). The additional
1465 custom probes were largely selected by applying a probe selection method, Spapros⁹⁴, to our
1466 snRNA-seq data. Spapros was run without selecting a feature number cutoff (n=None) and
1467 default parameters. Malignant and TME cells were run inclusively to ensure useful cell state
1468 delineation; however, the atlas was downsampled for each tumour and cell state to reduce the
1469 computational burden. We then selected high-quality markers based on the spatial expression
1470 according to our Visium data. For each cell state associated with a given marker, we calculated
1471 Pearson correlations between the marker voxel expression and cell state abundance inferred
1472 via cell2location. Markers with the highest median correlation and strongest inter-tumoural
1473 concordance were selected. Markers that were already present in the human brain panel were
1474 omitted. This resulted in 44 Spapros-derived cell state markers which were further
1475 supplemented by 17 known GB markers which were expressed in our Visium data.

1476 Default on-instrument nucleus segmentation and expansion was performed on Xenium
1477 datasets (Xenium Onboard Analyser versions xenium-1.5.0.3 and 1.6.0.8). Nuclei were
1478 segmented from DAPI images, followed by the definition of cell boundary as either a 15 µm
1479 expansion or the presence of another cell boundary. Single cell Xenium gene expression data
1480 were pre-processed using standard Scanpy workflow⁹⁵. Cell-by-gene matrices were filtered to
1481 exclude cells with low numbers of transcripts (total gene count < 10) and genes that expressed
1482 in fewer than 10 cells across the dataset. Gene counts were normalised and log-transformed
1483 prior to downstream analysis. Tangram²⁷ (version 1.0.4) was used to transfer coarse cell state
1484 labels from our snRNA-seq atlas onto Xenium sections using default parameters. Patient
1485 matched snRNA-seq profiles were used as reference for each Xenium dataset. Intersecting
1486 and non-zero expressed genes from both modalities were used as training sets (315 genes).
1487 Tangram was trained with default learning rate (0.1) for 500 epochs, in ‘clusters’ mode with a
1488 uniform density prior, and most probable cell type labels were transferred to Xenium cells from
1489 output mapping.

1490 **Joint snRNA- and snATAC-seq subclone calling**

1491 We leveraged multiome data to jointly infer subclones from multiome data. For snRNA-seq,
1492 we used the inferCNV results described above. For snATAC-seq data, we detected CNs

1493 directly on the Cell Ranger ARC ATAC output using epiAneufinder⁹⁶ with a 1MB window size
1494 and blacklisted regions of hg38 provided by the authors. Cells with total fragment counts
1495 <20,000 were filtered to ensure sufficient chromosomal coverage.

1496 Clones were called by applying leiden clustering on a neighbourhood graph generated
1497 from genome-wide CN calls for each modality. We observed that scATAC-seq-based CNAs
1498 exhibited greater variability than RNA-based CNAs, likely due to the higher sensitivity of
1499 ATAC-seq in detecting open chromatin regions, resulting in a larger number of events. To be
1500 more conservative and reduce false positives, we matched clones from scATAC-seq to
1501 scRNA-seq clones. To reduce clone redundancy, we assessed the similarity between
1502 scATAC-seq clones using correlation distance. Clones with pairwise distances below a
1503 predefined threshold were merged to eliminate highly similar clones.

1504 To mitigate differences in resolution between modalities, we applied interpolation to
1505 the RNA-based CNV profiles, aligning them with the larger genomic windows of scATAC-seq
1506 data. For matching clones across modalities, we computed two metrics: cell overlap and
1507 cosine similarity between cluster centroids. Cell overlap was defined as the proportion of cells
1508 shared between scRNA-seq and scATAC-seq clones, while cosine similarity was calculated
1509 based on the mean CNV profile for each clone. These metrics were averaged to create a
1510 composite similarity score.

1511 Next, we retained the top 20% of clone matches based on this composite similarity
1512 score. For each RNA clone, we greedily selected the most similar scATAC-seq clones,
1513 assigning cells shared between matched RNA and ATAC clones to a new joint clone identifier.
1514 Finally, we applied label spreading to propagate clone labels to classified RNA cells based on
1515 their CNV profiles. Each cell was assigned a probability of belonging to a clone, allowing
1516 flexibility in downstream analyses by using different confidence thresholds. For most analyses,
1517 we applied a confidence threshold of 0.9 to ensure high-confidence clone assignments.
1518 Overall, the methodological concordance with LCM-based WGS orthogonally validated the
1519 efficacy of counts-based methods at identifying core clone populations. Furthermore, while
1520 there may be minor clones with small artifactual signals introduced by inferential nature of this
1521 method, we reasoned that redundant granularity would allow us to conservatively control for
1522 genetic heterogeneity when applied to downstream trajectory analyses.

1523 **LCM WGS data variant calling, filtering and alleleIntegrator analysis**

1524 Following sequencing, all low-input WGS data was mapped to the hg38 reference genome
1525 using BWA-MEM⁹⁷. All mutations were called against matched normal blood samples using
1526 the Wellcome Sanger Institute variant calling pipeline.

1527 Copy number variants were called using ASCAT⁹⁸ and Battenberg⁹⁹. Battenberg copy
1528 number results were used for subsequent analyses. For each patient, LCM libraries with
1529 sufficient DNA concentrations (10 µg/mL) and copy number diversity were used as the basis
1530 for alleleIntegrator³⁸. CNAs were called in the snRNA-seq data using alleleIntegrator. The
1531 standard workflow was applied using Battenberg CN segments alongside the standard
1532 alleleIntegrator workflow. Clones were identified by first subsetting cells with support for a chr7
1533 gain, which is diagnostically relevant to GB, followed by support for putative lineage-specific
1534 CNAs (maxPostProb > 0.7).

1535 CaVEMan¹⁰⁰ was used to call substitutions. Variants were filtered to account for
1536 artefacts consistent with other low-input WGS studies^{37,78}, which involved applying a filter
1537 designed to remove false positive variants such as erroneously processed cruciform DNA
1538 (<https://github.com/MathijsSanders/SangerLCMFiltering>). Subsequently, variants were

1539 retained if they met a minimum median alignment score (ASMD) of ≥ 140 , fewer than half reads
1540 were clipped ($CLPM = 0$), were supported by at least 4 variant reads, and had a total
1541 sequencing depth of 10x or greater. The minimum thresholds for mapping quality and base
1542 quality were set to 30 and 25, respectively. Finally, we used a beta-binomial model adapted
1543 from the Shearwater variant caller by Coorens et al.¹⁰¹ designed to remove both germline
1544 variants and consistently observed low-quality variants. Site-specific error rates were
1545 calculated for each variant and only retained if the FDR-correct p -value was less than 0.001.
1546 Sample phylogenies were generated using Sequoia¹⁰² with default parameters and no
1547 application of the binomial mixture model. To ensure Sequoia compatibility with our tumour
1548 data, substitutions in CN variable regions were omitted.

1549 Following filtering, PyClone-VI³⁹ was used to cluster variants outside of patient-
1550 associated CNAs into putative subclones for each patient, separately. Each patient was run
1551 with the following non-default parameters: number of grid points = 150, number of restarts =
1552 100, number of maximum expected clusters = 30. Relationships between inferred clone
1553 populations, including phylogenetic analysis, was performed using PairTree^{103,104} with, apart
1554 from default parameters, a burn-in of 0.4 and the RPROP method for fitting subclones to trees.
1555 LCM samples were grouped into clusters based on clone population frequencies with
1556 agglomerative clustering using Bray-Curtis dissimilarity and complete linkage.

1557 **Spatial clone deconvolution with SpaceTree**

1558 To assess the spatial distribution of multiome-inferred clones (i.e. called from joint RNA and
1559 ATAC) in Visium data, we developed the SpaceTree model for joint clone and cell state
1560 deconvolution (**Supp. Comp. Methods**). This method follows a multi-step process, beginning
1561 with graph construction and followed by the use of a multi-task GNN to predict labels for each
1562 Visium spot. We used cell2location cell state mapping results to further evaluate the
1563 performance of SpaceTree, showing a high degree of concordance (see **Supp. Comp.**
1564 **Methods**, section 2.5.3). Additional orthogonal validation of this method was performed
1565 through comparison with LCM-derived analyses, including spatial clone cluster concordance
1566 (**Fig. 5C**). Cluster concordance evaluation involved summarising SpaceTree voxel
1567 probabilities that were histologically consistent with LCM ROIs into equivalent ROIs. These
1568 were hierarchically clustered using the same methodology as LCM samples (agglomerative
1569 clustering with Bray-Curtis dissimilarity and complete linkage). Tumour purity was calculated
1570 according to SpaceTree clone mapping results and represents the sum probability of all
1571 malignant clones in a given Visium spot. Shannon entropy was calculated on clone
1572 probabilities per spot and defined more generally as clonal entropy.

1573 **Cell2fate cell trajectory analysis**

1574 Cell fate trajectories were inferred using a Bayesian generative model, cell2fate⁴² from snRNA-
1575 seq data with default parameters. This method infers temporal expression modules resulting
1576 from the factorisation of RNA velocity data and then applies associated models of gene
1577 activation to infer cell-specific time along a trajectory. A key advantage of the Bayesian
1578 approach is that this method also returns posterior uncertainty estimates for every parameter,
1579 allowing for the evaluation of the confidence of cell state transitions. As input, we generated
1580 spliced and unspliced counts using STARsolo¹⁰⁵ v2.7.10a. Cell2fate analysis was run for each
1581 tumour clone separately using 10x reaction as a batch covariate, with 20 minimum shared
1582 genes, and 3000 HVGs in total. Individual cell2fate runs were evaluated first based on
1583 maximum cell-specific time ($\text{maximum}(T_c) > 20$), and then flagged as high-quality if average

1584 cell state transition scores exceeded 0.25. To quantify cell state transition scores, posterior
1585 distributions of cell-specific times were compared between two cell states, with the score
1586 equalling to the percentage of cells in state B that have a greater inferred time than the 90th
1587 percentile of cells in state A (**Extended Data Fig. 17A,C**) . For applications involving inter-
1588 clonal comparison across tumours, inferred times were min-max normalised.

1589 RNA velocity modules were subsequently hierarchically clustered using Jaccard
1590 distance of the top 200 genes in each module for the purposes of inter-clone module
1591 annotation. The “meta”-modules were then manually annotated according to (1) the identity of
1592 shared top genes, (2) the results of over-representation analysis of MGSigDB signatures for
1593 this gene list, and (3) enrichment of a given meta-module across malignant cell state
1594 annotations. Meta-module markers were compiled by taking the top 50 genes associated with
1595 each module, and sorting them by the degree of overlap across all modules in a given meta-
1596 module.

1597 Using the standard cell2location workflow and the methods outlined in Aivazidis et al⁴²,
1598 we also spatially mapped RNA velocity modules from cell2fate for a subset of tumours.
1599 Sections were selected if the multiome clone associated with the cell2fate run had a frequency
1600 of at least 30% in the associated site of sampling.

1601 **Survivorship analysis of bulk RNA-seq samples**

1602 Count matrices from CPTAC⁵³ and TCGA^{54,55} bulk RNA-seq samples were downloaded from
1603 the Genomic Data Commons (GDC). Samples were aggregated and log-normalised using
1604 Scanpy⁸². Relative expression scores were calculated for each meta-module based on
1605 commonly observed top genes (n=200) and samples were annotated as “enriched” for a meta-
1606 module if the expression scores were greater than the 75th percentile. Kaplan-Meier analysis
1607 followed by the log rank test were performed using the python package lifelines¹⁰⁶ between
1608 meta-modules with early and late inferred times.

1609 **Cell-cell communication analysis**

1610 We performed two sets of cell-cell communication (CCC) analysis to predict TME-derived
1611 ligands targeting receptors expressed on malignant cell states across different dev-like versus
1612 gliosis-hypoxia niches, focusing on myeloid signalling.

1613 In step 1, we identified TME cell states spatially co-localised with dev-like versus
1614 gliosis-hypoxia malignant cell states from Visium ST data (**Fig. 2**). These included the entire
1615 GB TME including myeloid, vasculature and T cells, and was summarised across all tumours
1616 (**Table S11**). We then selected malignant cell states present within dev-like versus gliosis-
1617 hypoxia niches per each individual tumour (**Table S12**).

1618 In step 2, we performed CCC analysis using the LIANA+ framework (version 1.2.0)¹⁰⁷,
1619 which integrates multiple computational methods and prior knowledge resources into a
1620 consensus-based approach to identify ligand-receptor interactions. We employed the
1621 “consensus” approach, combining five different methods implemented in LIANA+:
1622 CellPhoneDB, Connectome, log2FC, NATMI, SingleCellSignalR, and CellChat. Consensus
1623 scores across these methods were calculated using the robust rank aggregation method¹⁰⁸,
1624 with an 'expr_prop' (filtering criteria based on the proportion of cells per cell type) value of 0.05
1625 and 'n_perms' (number of permutations) set to 1000. Here, TME cell states were set as the
1626 sources (ligands) and the malignant cell states as the targets (receptors). An expression
1627 threshold of 0.05 was applied to filter out ligands and receptors below this threshold. We
1628 applied the LIANA analysis separately to the dev-like and gliosis-hypoxia niches.

1629 In step 3, the resulting ligand-receptor pairs were analyzed using the Tensor-cell2cell
1630 method¹⁰⁹ (version 0.7.4), which utilised the LIANA+ results as input to identify shared CCC
1631 events across tumours as factors. This method automatically determined the number of factors
1632 using elbow analysis and employed robust tensor factorization to ensure consistent results
1633 across donors. The main parameters and their values used in this analysis were as follows:
1634 'score_key = magnitude_rank' (the LIANA+ score used for factor analysis), 'tf_optimization =
1635 robust' (the robustness setting for tensor factorization), 'upper_rank = 15' (the upper limit for
1636 the number of factors), and 'tf_svd = numpy_svd' (the singular value decomposition method
1637 used). We then identified factors shared across tumours that involve myeloid cells, specifically
1638 Factors 1 and 2 from the dev-like analysis and Factor 1 from the gliosis-hypoxia analysis
1639 (**Extended Data Fig. 20**).

1640 In step 4, we sorted cell interactions associated with each factor based on their
1641 loadings and selected interactions in the top 50th percentile. We then focused on ligands from
1642 the myeloid populations that are predicted to target receptors on their respective malignant
1643 cell states. We filtered the ligand-receptor pair list to ensure at least one relevant cell state
1644 showed mean gene expression ≥ 0.5 and had a total expression fraction ≥ 0.1 for each
1645 associated gene pair. This resulted in the identification of 80 unique interactions in the dev-
1646 like niches (involving 37 ligands), 35 unique interactions in the gliosis analysis (involving 18
1647 ligands), and a total of 63 shared interactions shared across both niches (involving 28 ligands).

1648 **References (methods)**

- 1649 77. Lee-Six, H. *et al.* The landscape of somatic mutation in normal colorectal
1650 epithelial cells. *Nature* **574**, 532–537 (2019).
- 1651 78. Ellis, P. *et al.* Reliable detection of somatic mutations in solid tissues by laser-
1652 capture microdissection and low-input DNA sequencing. *Nat. Protoc.* **16**, 841–871
1653 (2021).
- 1654 79. Zheng, G. X. Y. *et al.* Massively parallel digital transcriptional profiling of
1655 single cells. *Nat Commun* **8**, 14049 (2017).
- 1656 80. Satpathy, A. T. *et al.* Massively parallel single-cell chromatin landscapes of
1657 human immune cell development and intratumoral T cell exhaustion. *Nat Biotechnol* **37**,
1658 925–936 (2019).
- 1659 81. Fleming, S. J. *et al.* Unsupervised removal of systematic background noise
1660 from droplet-based single-cell experiments using CellBender. *Nat. Methods* **20**, 1323–
1661 1335 (2023).
- 1662 82. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene
1663 expression data analysis. *Genome Biol* **19**, 15 (2018).
- 1664 83. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: Computational Identification

- 1665 of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst* **8**, 281–291.e9 (2019).
- 1666 84. Pijuan-Sala, B. *et al.* A single-cell molecular map of mouse gastrulation and
1667 early organogenesis. *Nature* **566**, 490–495 (2019).
- 1668 85. Popescu, D.-M. *et al.* Decoding human fetal liver haematopoiesis. *Nature*
1669 **574**, 365–371 (2019).
- 1670 86. Tirosh, I. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma
1671 by single-cell RNA-seq. *Science* **352**, 189–196 (2016).
- 1672 87. Granja, J. M. *et al.* ArchR is a scalable software package for integrative
1673 single-cell chromatin accessibility analysis. *Nat Genet* **53**, 403–411 (2021).
- 1674 88. Tickle, T., Tirosh, I., Georgescu, C., Brown, M. & Haas, B. *inferCNV of the*
1675 *Trinity CTAT Project*. Preprint at <https://github.com/broadinstitute/inferCNV> (2019).
- 1676 89. Venteicher, A. S. *et al.* Decoupling genetics, lineages, and microenvironment
1677 in IDH-mutant gliomas by single-cell RNA-seq. *Science (New York, N.Y.)* **355**, eaai8478
1678 (2017).
- 1679 90. De Donno, C. *et al.* Population-level integration of single-cell datasets
1680 enables multi-scale analysis across samples. *Nat. Methods* **20**, 1683–1692 (2023).
- 1681 91. Badia-I-Mompel, P. *et al.* decoupleR: ensemble of computational methods to
1682 infer biological activities from omics data. *Bioinform. Adv.* **2**, vbac016 (2022).
- 1683 92. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based
1684 approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S.*
1685 *A.* **102**, 15545–15550 (2005).
- 1686 93. Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark
1687 gene set collection. *Cell Syst.* **1**, 417–425 (2015).
- 1688 94. Kuemmerle, L. B. *et al.* Probe set selection for targeted spatial
1689 transcriptomics. *Nat. Methods* **21**, 2260–2270 (2024).
- 1690 95. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene
1691 expression data analysis. *Genome Biol.* **19**, 15 (2018).
- 1692 96. Ramakrishnan, A. *et al.* epiAneufinder identifies copy number alterations from

- 1693 single-cell ATAC-seq data. *Nat Commun* **14**, 5846 (2023).
- 1694 97. Li, H. Aligning sequence reads, clone sequences and assembly contigs with
1695 BWA-MEM. *arXiv [q-bio.GN]* (2013).
- 1696 98. Van Loo, P. et al. Allele-specific copy number analysis of tumors. *Proc. Natl.*
1697 *Acad. Sci. U. S. A.* **107**, 16910–16915 (2010).
- 1698 99. Nik-Zainal, S. et al. The life history of 21 breast cancers. *Cell* **149**, 994–1007
1699 (2012).
- 1700 100. Jones, D. et al. CgpCaVEManWrapper: Simple execution of CaVEMan in
1701 order to detect somatic single nucleotide variants in NGS data. *Curr. Protoc.*
1702 *Bioinformatics* **56**, 15.10.1–15.10.18 (2016).
- 1703 101. Coorens, T. H. H. et al. Embryonal precursors of Wilms tumor. *Science* **366**,
1704 1247–1251 (2019).
- 1705 102. Coorens, T. H. H. et al. Reconstructing phylogenetic trees from genome-wide
1706 somatic mutations in clonal samples. *Nat. Protoc.* **19**, 1866–1886 (2024).
- 1707 103. Kulman, E., Wintersinger, J. & Morris, Q. Reconstructing cancer phylogenies
1708 using Pairtree, a clone tree reconstruction algorithm. *STAR Protoc.* **3**, 101706 (2022).
- 1709 104. Wintersinger, J. A. et al. Reconstructing complex cancer evolutionary
1710 histories from multiple bulk DNA samples using Pairtree. *Blood Cancer Discov.* **3**, 208–
1711 219 (2022).
- 1712 105. Kaminow, B., Yunusov, D. & Dobin, A. STARsolo: accurate, fast and versatile
1713 mapping/quantification of single-cell and single-nucleus RNA-seq data. *Bioinformatics*
1714 (2021).
- 1715 106. Davidson-Pilon, C. lifelines: survival analysis in Python. *J. Open Source*
1716 *Softw.* **4**, 1317 (2019).
- 1717 107. Dimitrov, D. et al. LIANA+ provides an all-in-one framework for cell-cell
1718 communication inference. *Nat Cell Biol* **26**, 1613–1622 (2024).
- 1719 108. Kolde, R., Laur, S., Adler, P. & Vilo, J. Robust rank aggregation for gene list
1720 integration and meta-analysis. *Bioinformatics* **28**, 573–580 (2012).

- 1721 109. Baghdassarian, H. M., Dimitrov, D., Armingol, E., Saez-Rodriguez, J. &
1722 Lewis, N. E. Combining LIANA and Tensor-cell2cell to decipher cell-cell communication
1723 across multiple samples. *Cell Rep Methods* **4**, 100758 (2024).