

Supplemental information

Spatially resolved multi-omics

deciphers bidirectional tumor-host

interdependence in glioblastoma

Vidhya M. Ravi, Paulina Will, Jan Kueckelhaus, Na Sun, Kevin Joseph, Henrike Salié, Lea Vollmer, Ugne Kuliesiute, Jasmin von Ehr, Jasim K. Benotmane, Nicolas Neidert, Marie Follo, Florian Scherer, Jonathan M. Goeldner, Simon P. Behringer, Pamela Franco, Mohammed Khiat, Junyi Zhang, Ulrich G. Hofmann, Christian Fung, Franz L. Ricklefs, Katrin Lamszus, Melanie Boerries, Manching Ku, Jürgen Beck, Roman Sankowski, Marius Schwabenland, Marco Prinz, Ulrich Schüller, Saskia Killmer, Bertram Bengsch, Axel K. Walch, Daniel Delev, Oliver Schnell, and Dieter Henrik Heiland

Supplemental information

Spatially resolved multi-omics deciphers bidirectional tumor-host interdependence in glioblastoma

Vidhya M. Ravi, Paulina Will, Jan Kueckelhaus, Na Sun, Kevin Joseph, Henrike Salié, Lea Vollmer, Ugne Kuliesiute, Jasmin von Ehr, Jasim K. Benotmane, Nicolas Neidert, Marie Follo, Florian Scherer, Jonathan M Goeldner, Simon P. Behringer, Pamela Franco, Mohammed Khiat, Junyi Zhang, Ulrich G. Hofmann, Christian Fung, Franz L. Ricklefs, Katrin Lamszus, Melanie Boerries, Manching Ku, Jürgen Beck, Roman Sankowski, Marius Schwabenland, Marco Prinz, Ulrich Schüller, Saskia Killmer, Bertram Bengsch, Axel K. Walch, Daniel Delev, Oliver Schnell, Dieter Henrik Heiland

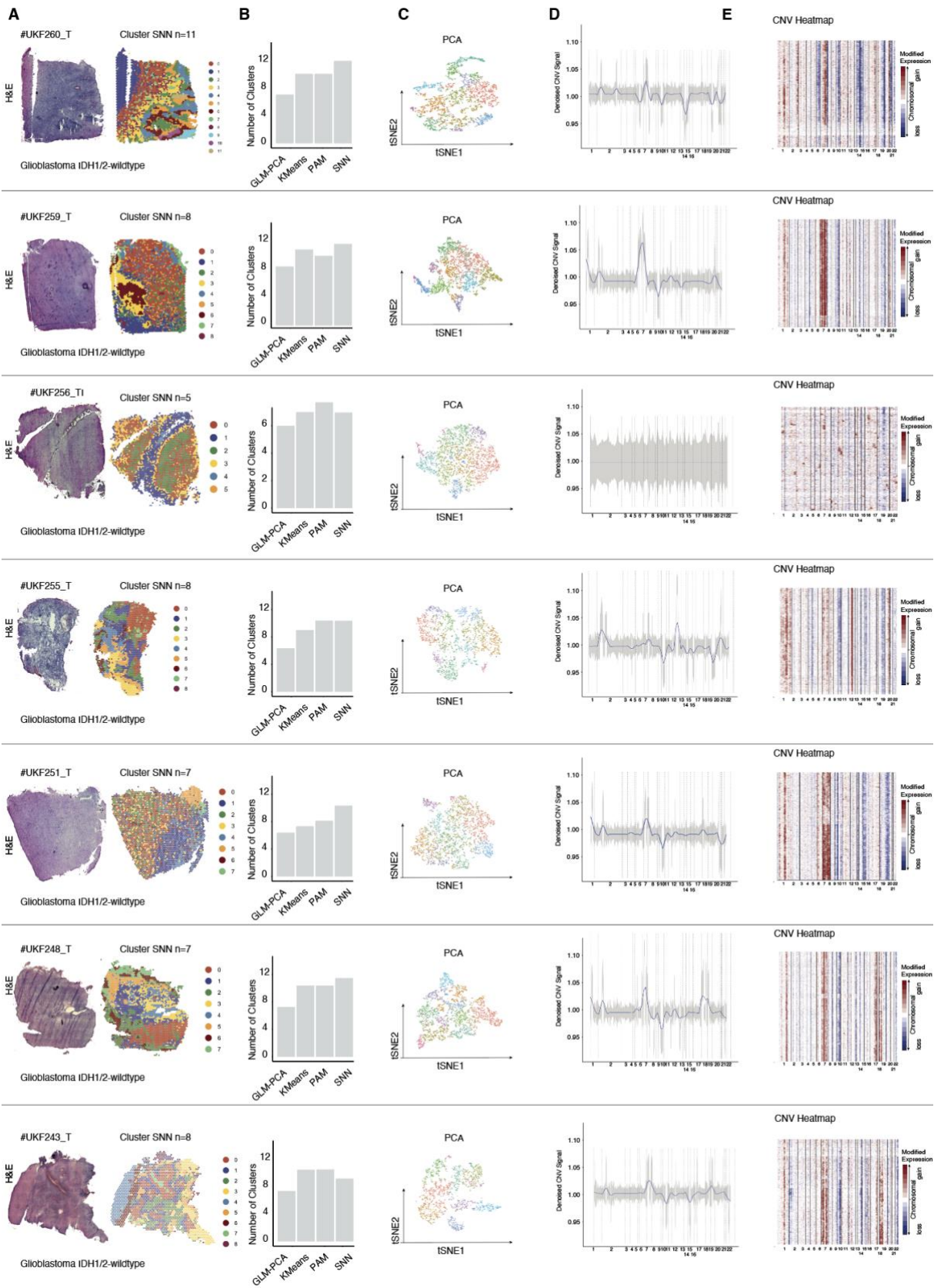
12 **Table S1: Overview of the cohort related to Figure 1**

ID	Tissue Type	Batch	Region	Age	Tumor	Histology	Sex	MALDI	IMC
#275UKF	T	6	frontal	80	IDH-WT	GBM	f	1	1
#270UKF	T	6	frontal	41	IDH-MUT	AA	m		
#269UKF	T	6	occipital	51	IDH-WT	GBM	m		
#268UKF	T	6	frontal	49	IDH-MUT	ODG	f		
#266UKF	T	6	temporal	76	IDH-WT	GBM	f		
#265UKF	C	2	temporal	55	C	Cortex	m		
#265UKF	T	2	temporal	55	IDH-WT	GBM	m		
#262UKF	T	5	temporal	59	IDH-WT	GBM	m	1	1
#260UKF	T	5	temporal	81	IDH-WT	GBM	m	1	1
#259UKF	T	3	temporal	75	IDH-WT	GBM	f	1	1
#256UKF	C	4	temporal	64	C	Cortex	f		
#256UKF	TC	4	temporal	64	IDH-WT	GBM	f		
#256UKF	TI	4	temporal	64	IDH-WT	GBM	f		
#255UKF	T	5	temporal	81	IDH-WT	GBM	f		
#251UKF	T	4	temporal	61	IDH-WT	GBM	m		
#248UKF	C	1	frontal	44	C	Cortex	m	1	1
#248UKF	T	1	frontal	44	IDH-WT	GBM	m		
#243UKF	T	1	frontal	55	IDH-WT	GBM	f		
#242UKF	C	1	frontal	81	C	Cortex	f		
#242UKF	T	1	frontal	81	IDH-WT	GBM	f		
#241UKF	C	1	temporal	19	C	Cortex	m		
#259UKF	T	3	temporal	75	IDH-WT	GBM	f		
#296UKF	T	9	temporal	34	IDH-WT	GBM	m	1	1
#304UKF	T	9	temporal	81	IDH-WT	GBM	m		
#313UKF	C	10	temporal	57	C	Cortex	m		
#313UKF	T	10	temporal	57	IDH-WT	GBM	m		
#334UKF	C	10	frontal	73	C	Cortex	f		
#334UKF	T	10	frontal	73	IDH-WT	GBM	f		

13 Tissue Type (Macroscopic): T: Tumor, C: Cortex, TC: Tumor Core: TI: Tumor Infiltrative region

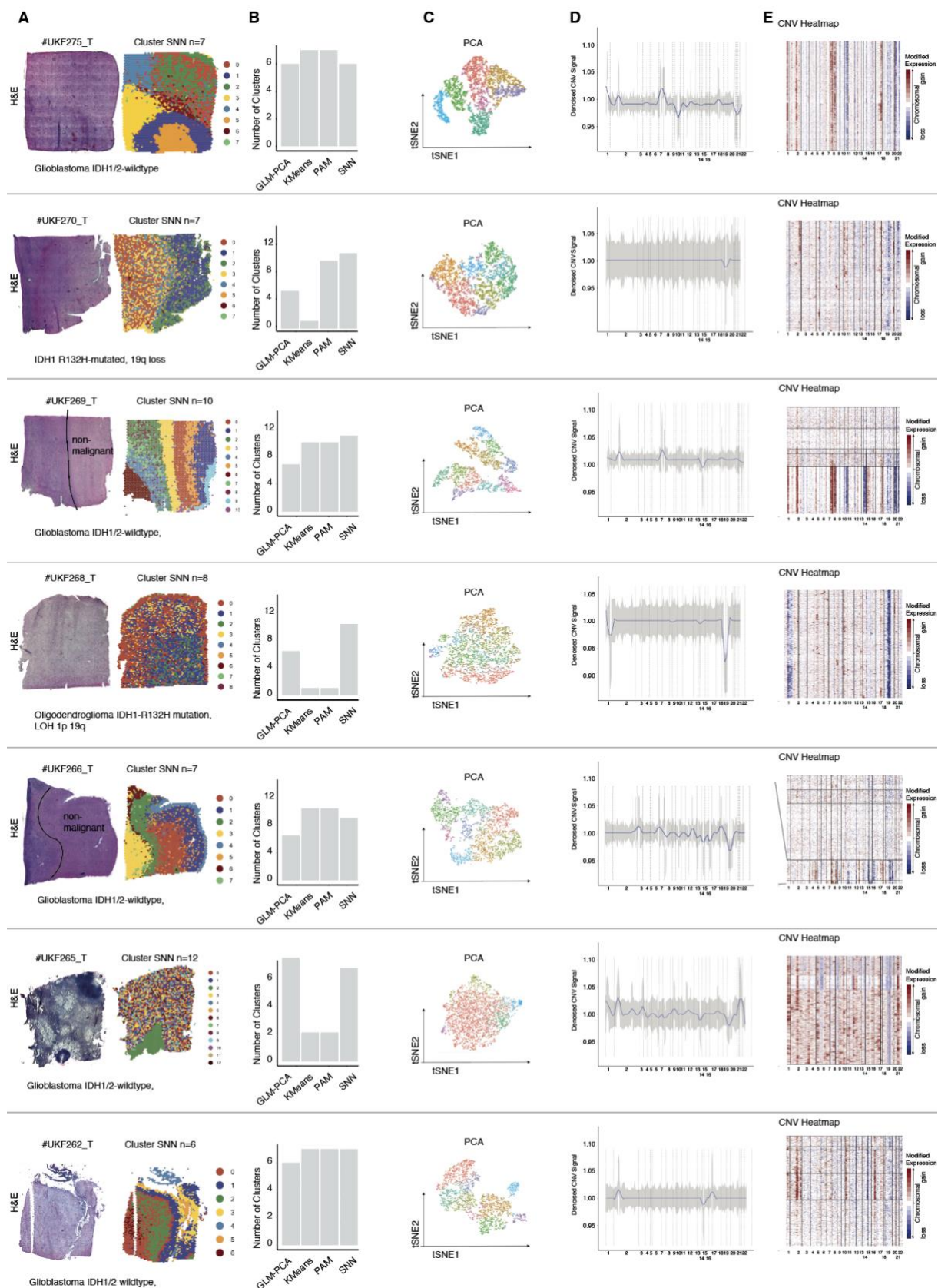
14

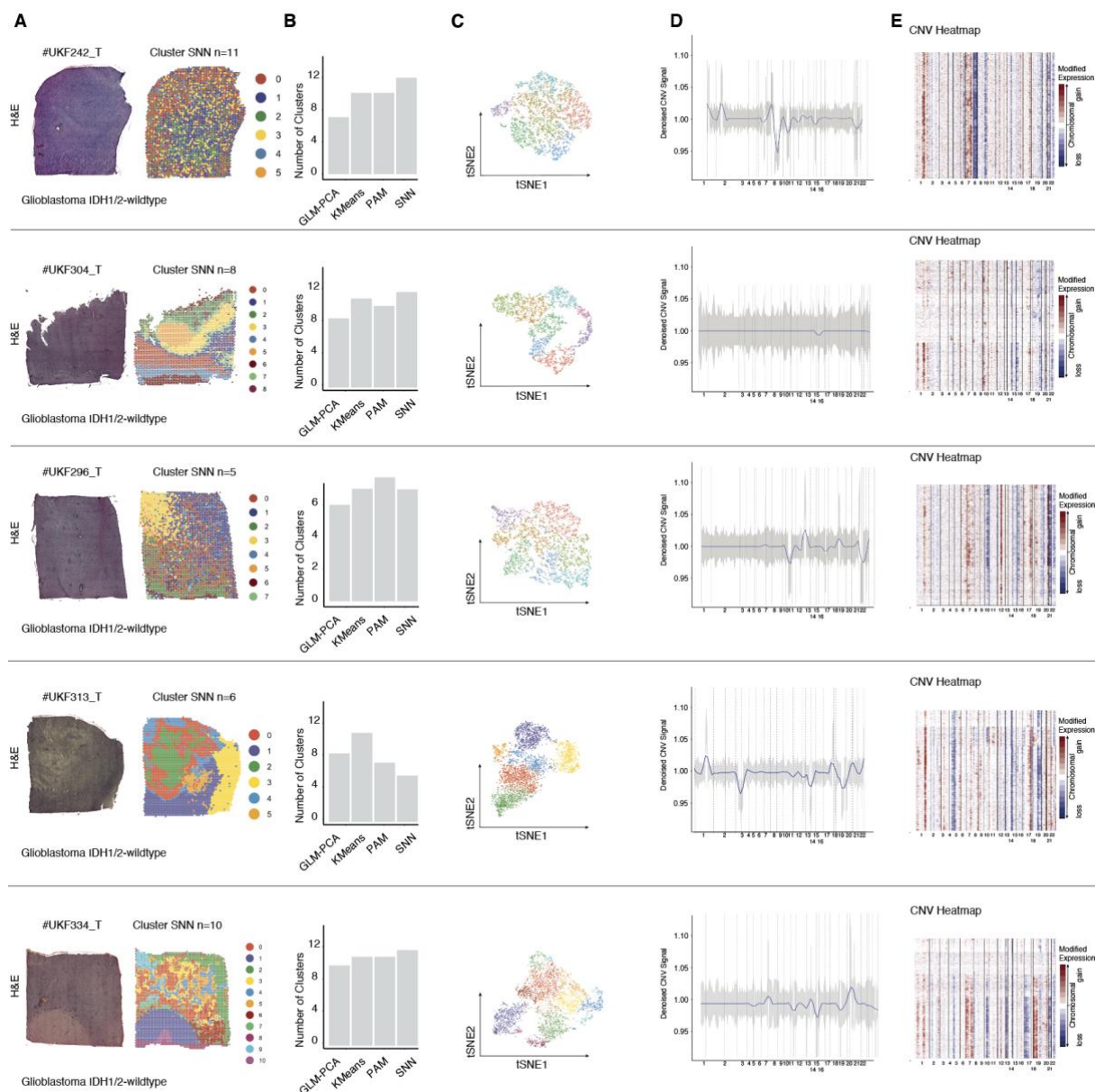
15 **Supplementary Figure 1**



16

17

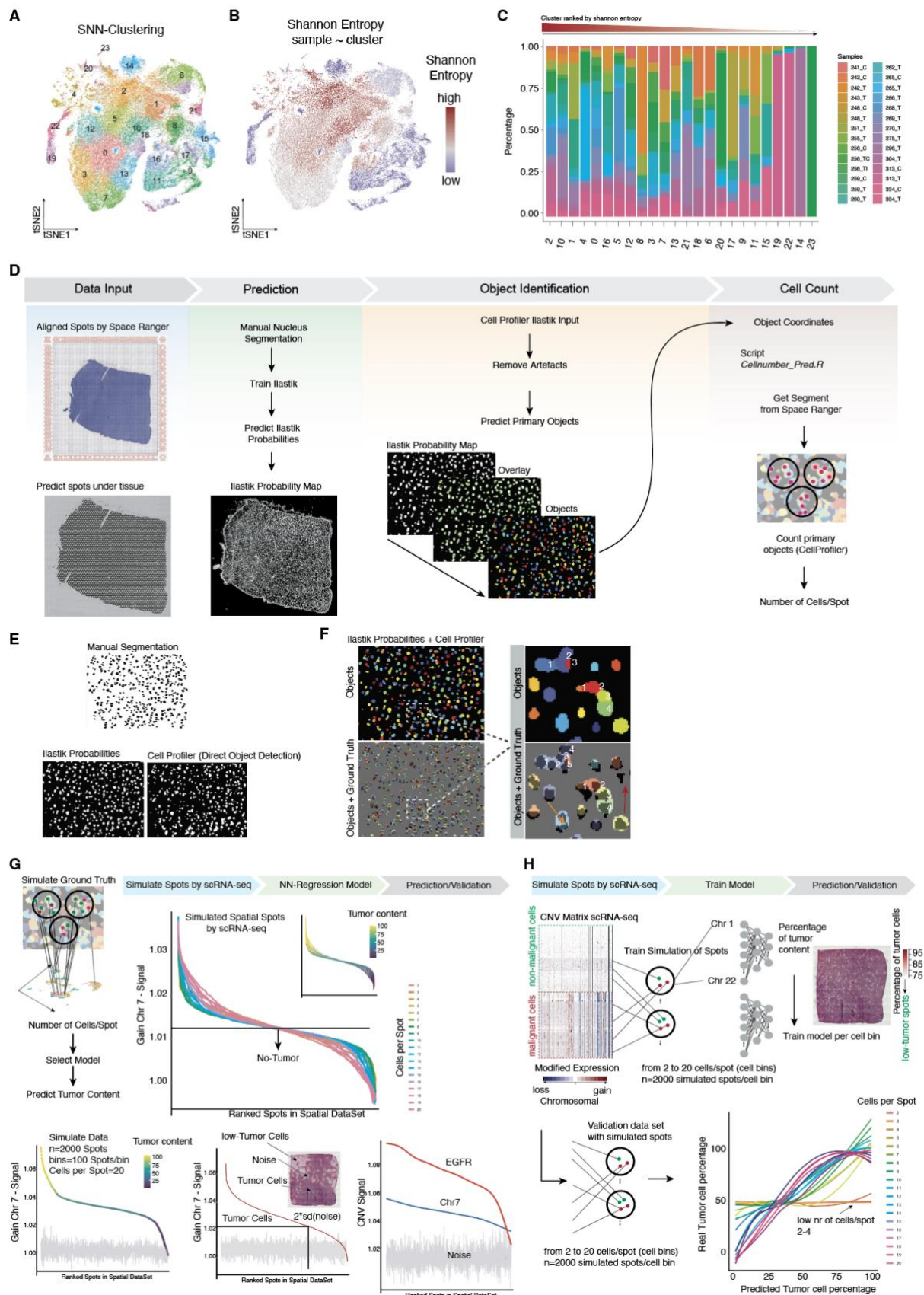




Supplementary Figures 1: Overview of the cohort, related to Figure 1

(A) Illustration of a H&E staining (left) and the optimized SNN cluster analysis at the right side using iterative resolutions. (B) Bar plot of the optimal number of clusters using various algorithms (C) UMAP projection of spots. (D) Illustration of the copy number profile presented as line plot (left) with standard deviations marked in grey. (E) CNA alterations presented as heatmap. Blue indicate a chromosomal loss, red a chromosomal gain.

27 **Supplementary Figure 2**



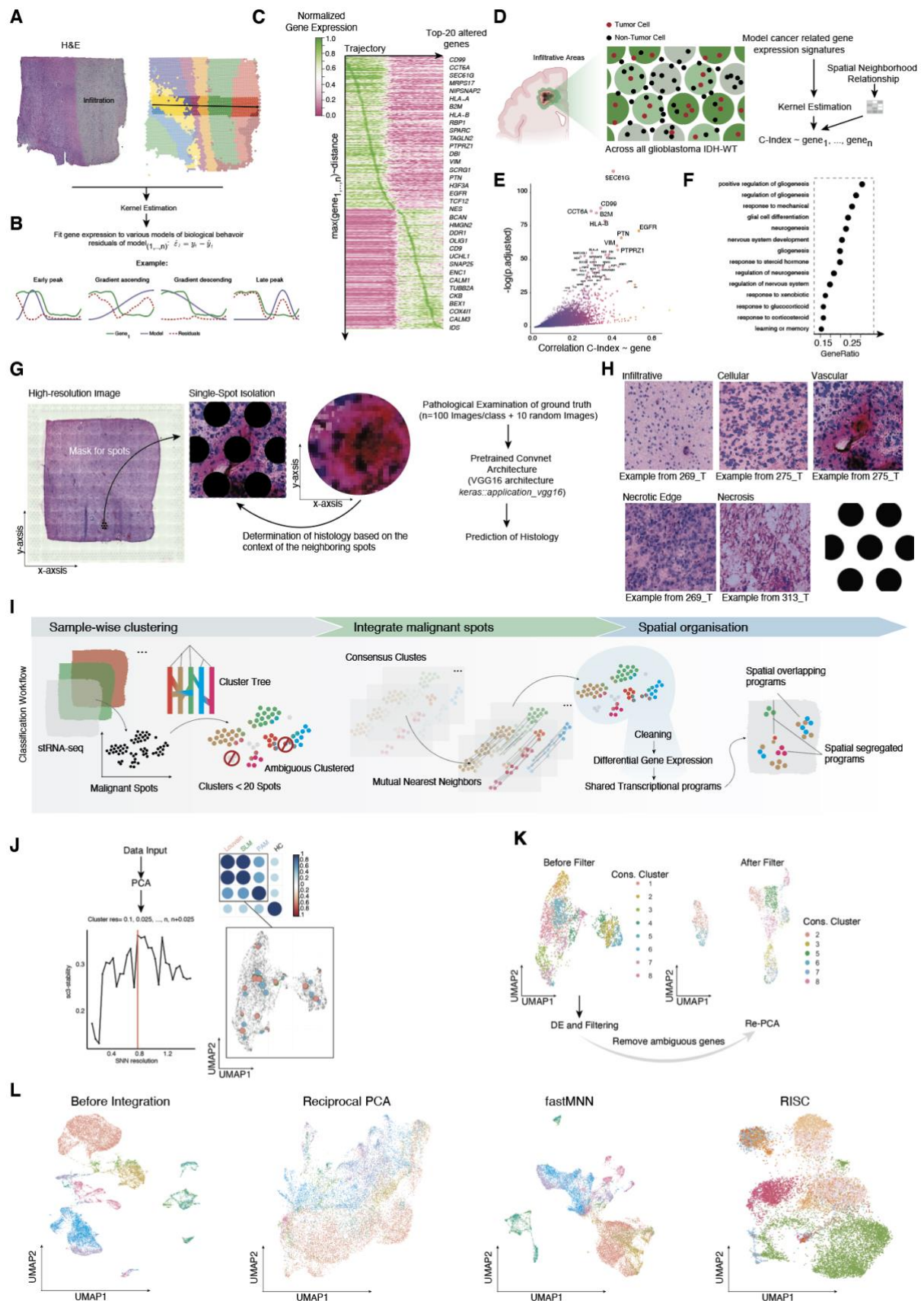
28

29 **Supplementary Figure 2: Data integration stRNA, related to Figure 1**

(A) A dimensional reduction using the tSNE algorithm. Colors indicate clusters from the Shared Nearest Neighbour (SNN) approach. (B) To investigate the extent to which clusters were patient specific or included multiple individuals, we estimated Shannon Entropy between clusters and patients. Clusters are colored by respective entropy in the tSNE representation. (C) Bar plot representation of patient distribution across clusters, sorted by entropy. (D) Illustration of the workflow for nucleus segmentation from stRNA-seq datasets. (E) Validation of segmentation was performed by comparing manual segmentation (ground truth, upper image), ilastik probability maps (bottom left) and the CellProfiler object identification (bottom, right). (F) Example of the combined ilastik+CellProfiler strategy for segmentation. The identified object is illustrated in the upper image and the overlay of the ground truth (bottom). The enlargements on the right show the inaccuracies of automated segmentation. In some cases, clumped objects are not separated correctly. The arrows show that some nuclei are not recognized (red) or are detected incorrectly (orange).

(G) On the upper left, overview of the workflow for spatial data simulation is illustrated. A detailed description of the method can be found within the STAR-Methods. In a nutshell, single-cell sequencing data were used to generate randomly assembled spots with both tumor and non-malignant cells. This dataset was used as ground truth to determine the content of malignant spots. Training and validation datasets were also simulated. Since the spots contain differing number of cells, we simulated spots similarly. We then computed copy number variants at large chromosomal regions and compared the gain of chromosome 7 in simulated spots (plot bottom left) and our stRNA-seq data (middle plot). The gray signal represents the average noise (based on chromosomes without alterations). Chromosome 7 intensity from non-simulated stRNA-seq data exhibits increased noise compared to simulated data. Using a loess regression model, we found that a prediction of the tumor content is highly biased by the signal to noise ratio (bottom right plots). (H) The insufficient accuracy of the simple regression model necessitated a switch to deep learning (artificial neural network, ANN) approach. A detailed description is provided in the STAR-Methods. We trained the ANN to predict the number of malignant cells in our simulated dataset. This was possible with high accuracy in spots composed of 5 cells and more. Low cell numbers per spot lack sufficient prediction (bottom right).

58 **Supplementary Figure 3:**

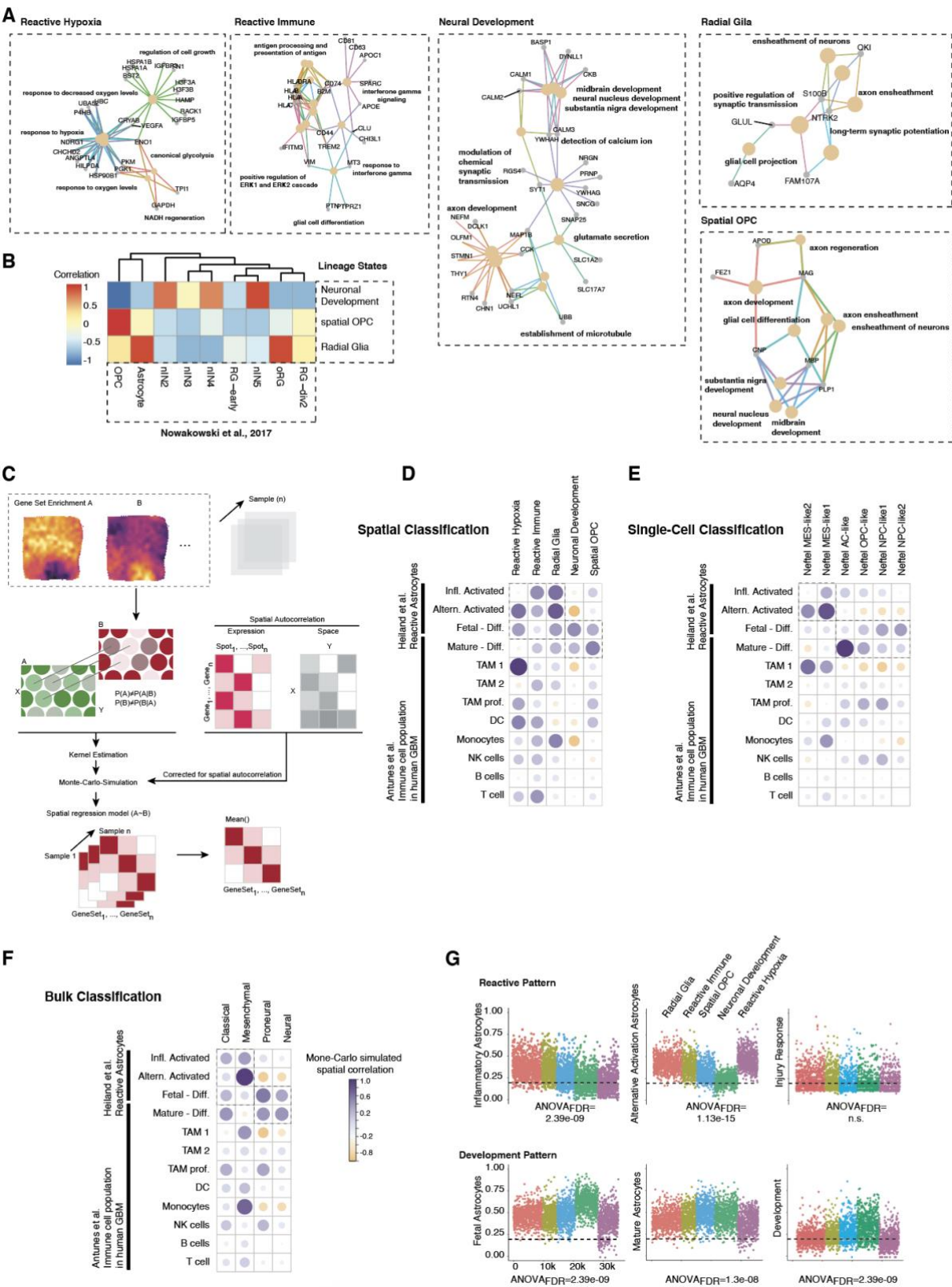


59

60 **Supplementary Figure 3: Analysis of the infiltrative margin and data integration, related to Figure**

61 **1-2**

(A) In order to analyze infiltrated areas, we performed spatial trajectory analysis, described (Kueckelhaus et al., 2020) in detail recently and in the STAR methods. An illustration helps visualize the concept of spatial trajectory analysis. (B) Feature behavior (gene expression/ gene set enrichments or other phenotype features) is fitted to a library of predefined biological behavior's and validated by its residuals (integral of the residuals: area under the curve, AUC). Features with low AUC were filtered to screen for defined transcriptional adaptations along the trajectory from tumor towards the infiltrated cortex. (C) Heatmap of filtered genes with a low AUC (< 2) with the biological behavior: "early peak", "one peak", "late peak", smoothed by a gaussian kernel and ranked by maximum expression along the trajectory. In the transition between tumor and infiltrative cortex, an increased expression of Radial Glia marker (*VIM*, *PTPRZ*) is demonstrated. (D) Illustration of the second approach to investigate the infiltrated areas. Only a minority of cells per spot are tumor cells, therefore CN-Analysis is used to identify low tumor content based on our simulated single cell data. Using the tumor content probability (Cancer (C)-index), a spatially weighted correlation (bi-weighted 2D Kernel estimator) is displayed, to identify genes that correlate with the probability of tumor localization. (E) Genes that were also identified in the transition area between cellular tumor and infiltrated cortex (F) The dot plot shows that gene set enrichment indicates increased gliogenesis related to migration and differentiation. (G) Illustration of the workflow to determine the histology at spot-level. A detailed description of the method is given in the STAR-Methods. (H) Example of the different histology's in relation to the spot size. (I) Illustration of the workflow for cluster analysis. A detailed step-by-step explanation is given in the STAR-Methods. (J) Illustration of an example in which multiple cluster algorithms were integrated to identify consensus clusters, based on a shared cluster tree structure. On the left, the iterative approach to determine the optimal cluster resolution is illustrated. On the x-axis, the resolution is shown, on the y-axis the determined cluster stability (based in the clustertree package). (K) After consensus cluster determination, the spots were filtered by its cluster affiliation. The illustrated UMAPs indicate the spot distribution before (left) and after filtering (right). (L) An example of horizontal data integration (n=6 spatial datasets) with features as anchors. From left to right, baseline status without integration, reciprocal PCA (Andreatta and Carmona, 2020), mutual nearest neighbors (Haghverdi et al., 2018) and RISC (Liu et al., 2021)



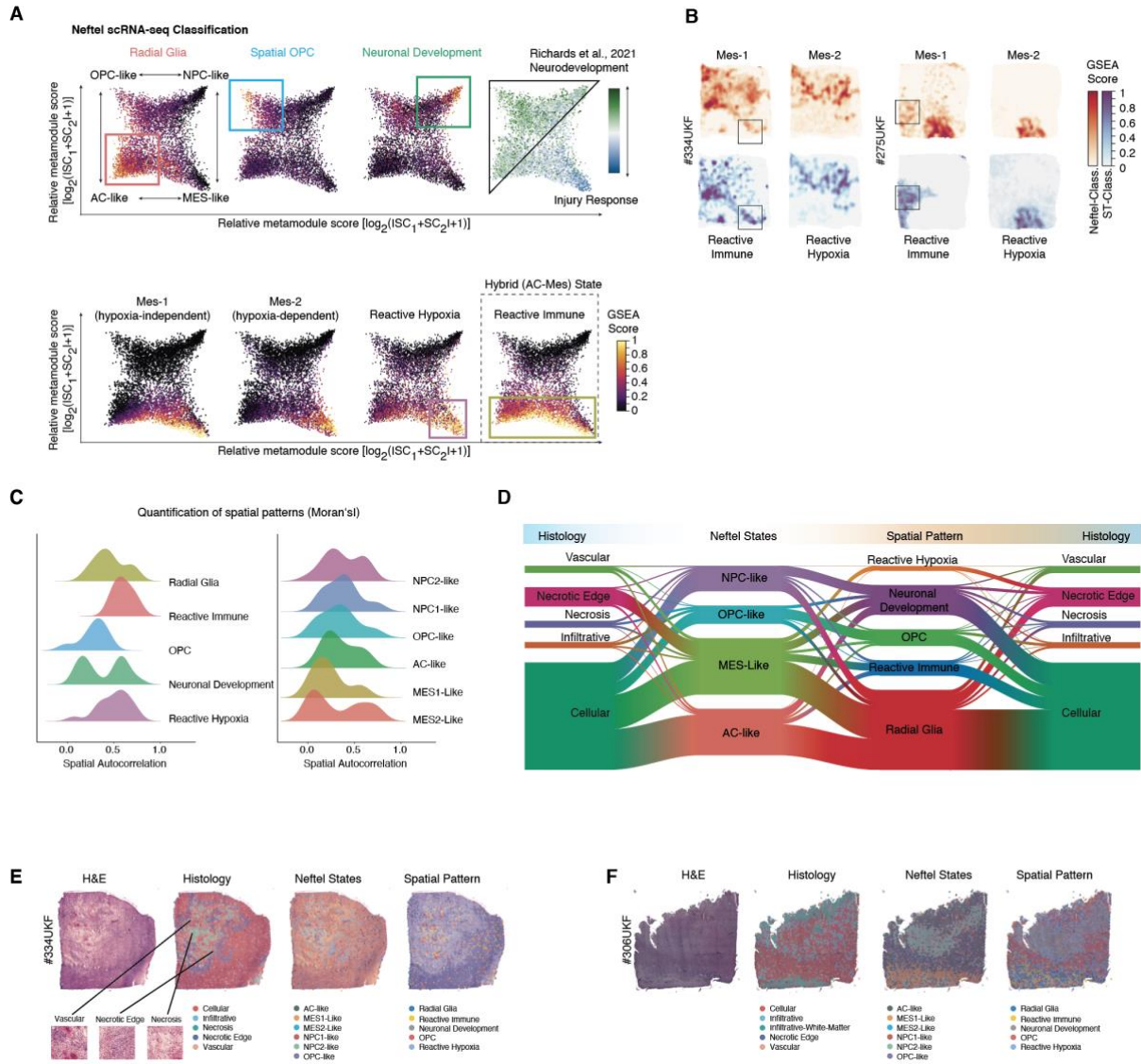
92

93 **Supplementary Figure 4: Analysis of the spatially distinct transcriptional programs, related to**

94 **Figure 2**

(A) Go-term enrichment and pathway enrichment analysis of the spatially distinct transcriptional programs using the clusterProfiler with the parameters *ont* = "ALL", *OrgDb* = "org.Hs.eg.db". The plots were built using the *cnetplot()* function. Pathways are indicated in bold; genes and pathways are presented as nodes. Connections across pathways are indicated by edges and colored in accordance to the respective pathway. (B) Heatmap of the correlation matrix between the spatially distinct transcriptional programs (only non-reactive programs) to the reference dataset of developmental stages of the human brain recently published by (Nowakowski et al., 2017). The analysis demonstrate that the identified lineage stages of tumor cells reassemble developmental stages of the human brain such as: *Neural Development*: nN2, nN4 and nN5; *Radial Glia*: Astrocytes and RG and RG-div2; *Spatial OPC*: OPC and RG-div2. (C) Overview of spatially weighted correlation analysis across multiple samples. A detailed description on the methods is given in the STAR-Methods. (D-F) Analysis of spatially weighted correlation between defined cell types of the glioma-associated environment. Cell signatures of glioblastoma associated glial cells (Henrik Heiland et al., 2019)(activated astrocytes), various myeloid cell types and T cells (Pombo Antunes et al., 2020). The correlation was illustrated by a heat-dot-plot in which correlation is indicated by size and color. (G) The illustrated scatterplot illustrates that glioblastoma reflect activation patterns and cellular states seen in non-malignant astrocytes. Signature genes of inflammatory-transformed astrocytes were enriched in the Radial-Glia and Reactive Immune programs (ANOVA and TukeyHSD, FDR=2.39e-09). Reactive Hypoxia is marked by similarity to the alternative activation signature (ANOVA and TukeyHSD, FDR=1.13e-15) (Clarke et al., 2018; Henrik Heiland et al., 2019; Liddel et al., 2017; Yun et al., 2018).

Supplementary Figure 5

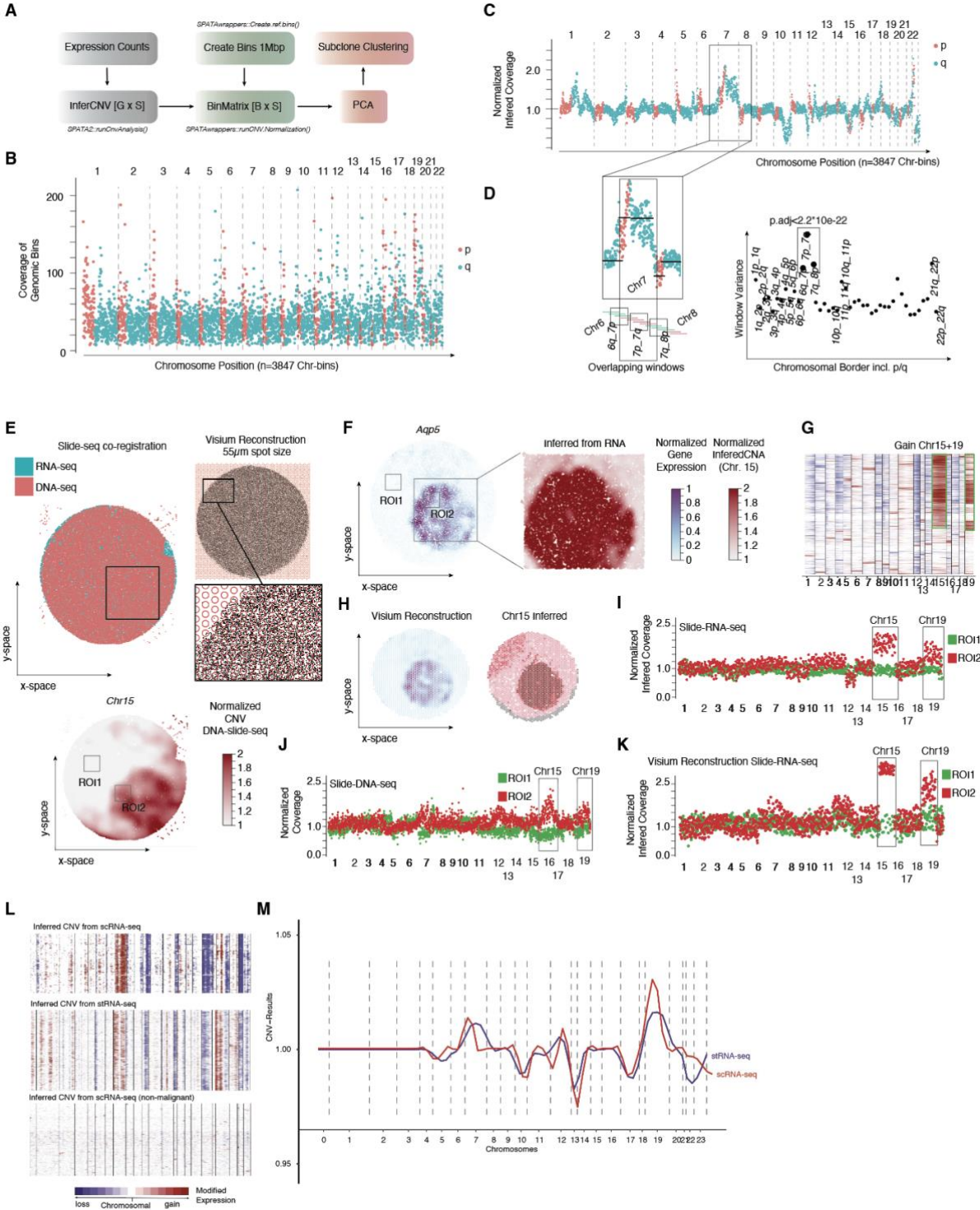


Supplementary Figure 5: Comparison of spatial and single-cell transcriptomics, related to Figure 2

(A) Scatter plot of the 2D representation of the four major cellular states in Glioblastoma (Neftel-classification). The color code represents geneset enrichment scores for the spatially distinct transcriptional programs or the two-state classification from Richards (Richards et al., 2021) (upper right). At the bottom, the enrichment of the MES1 and MES2 state (Neftel) is presented (left) and the reactive states from our spatially distinct transcriptional programs (right) are illustrated. (B) Gene expression surface plots of the sample #334UKF (right) and #275UKF (left). The plots illustrate the spatial overlap of MES-1 gene signatures (Neftel et al., 2019) to Reactive Immune and MES-2 gene signatures (Neftel et al., 2019) to Reactive Hypoxia. The highlighted regions indicate areas containing differences between both subgroups. (C) Statistical evaluation of the spatial exclusivity based on

130 Morans's I statistics. A detailed description is given in the STAR-Methods. The density of
131 autocorrelations of all samples is presented.
132 (D) River plot of the subgroup annotations across all patients between histological annotations, cellular
133 states and spatially distinct transcriptional programs. (E-F) Examples of two patients with annotated
134 histological pattern, cell states and spatially distinct transcriptional programs.
135

136 **Supplementary Figure 6**



137

138 **Supplementary Figure 6: Validation of CNA Analysis, related to Figure 2**

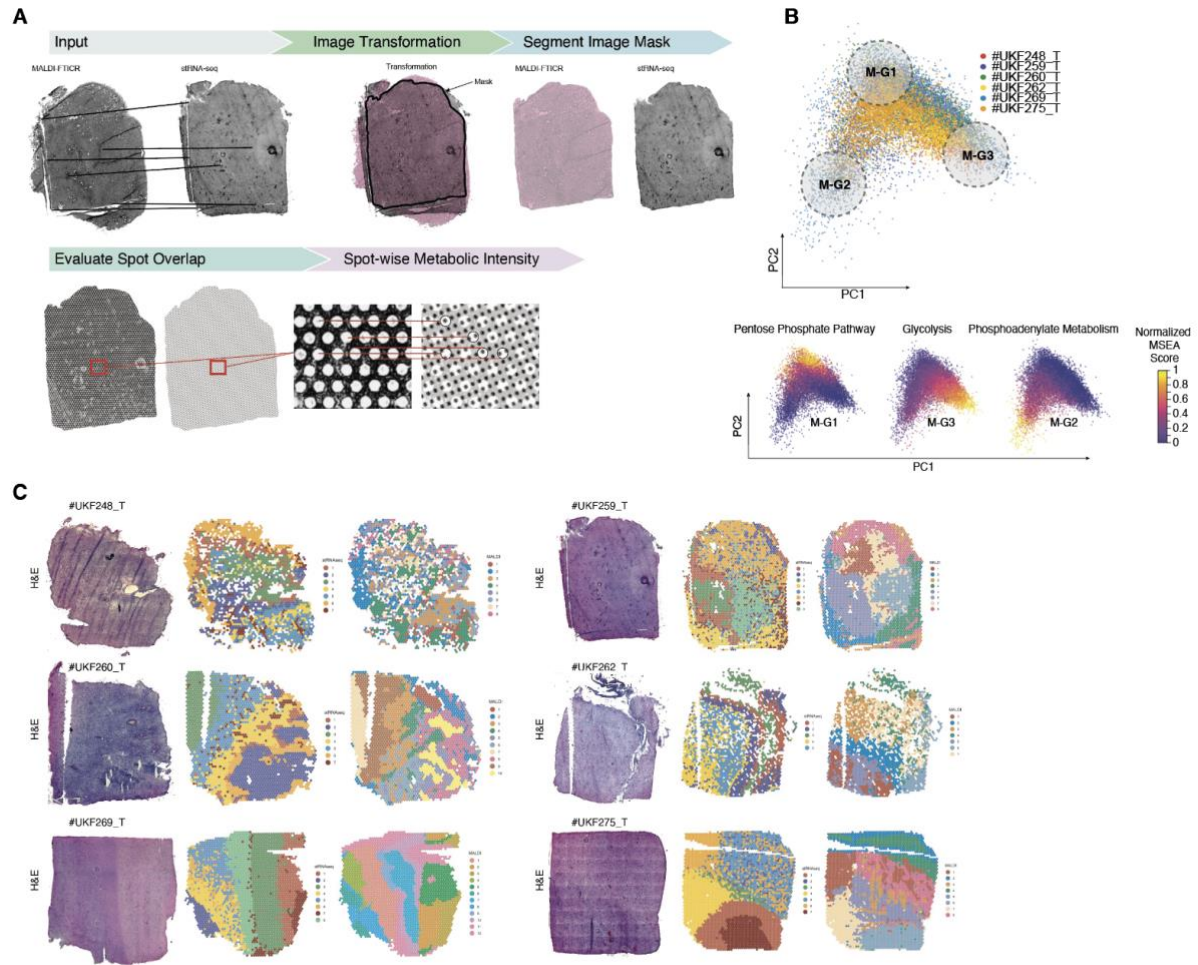
139 (A) Illustration of the workflow for copy number analysis. The input from the inferCNV algorithm was
140 used to create defined 1mbp bins of the chromosome. The SPATAwapper package was used to
141 normalize the copy-number signal to the bins and estimate real copy-number values. Further subclonal
142 analysis was performed by PCA and kMeans clustering. (B) At 1Mbp bins the coverage was 25.45 at
143 10kbp 5.5 genes per bin. The coverage plot illustrates the variance of coverage across chromosomal

regions. In the smaller chromosomes (13-22) the p-arm was underrepresented causing a potential bias.

(C) A representative example of the copy number alterations seen in #UKF275T. (D). Illustration to what extent chromosomal gains/losses respect chromosomal borders. The significance of a chromosomal loss/gain was calculated using ANOVA and the resulting p-value was corrected using FDR. On the right, the variance of overlapping windows is illustrated (variance on the y-axis). (E) Validation of inferred CNVs with measured CNAs from DNA sequencing. In order to validate our algorithm, we used the recent published combined DNA-, RNA-slide seq dataset (Zhao et al., 2022). An overview of the alignment of both DNA- and RNA-slide-seq datasets is given in (E). Colors indicate the dataset origin. Since the slide-seq technology has a spatial resolution of 10µm which is 30.23 times lower compared to Visium spot (55µm diameter). In order to correct for these differences, we created a Visium spot mask (55µm diameter, 100µm center-to-center) and averaged the gene expression within each Visium spot (a mean number of 30.32 slide-seq spots). (F) Gene expression surface plot of *Aqp5* indicate the right bottom liver metastasis. The gains of chromosome 15 from the inferred CNA analysis is indicated at the right side. (G) CNA heatmap of the samples shows areas of tumor (Chr 15 & 19 gain).

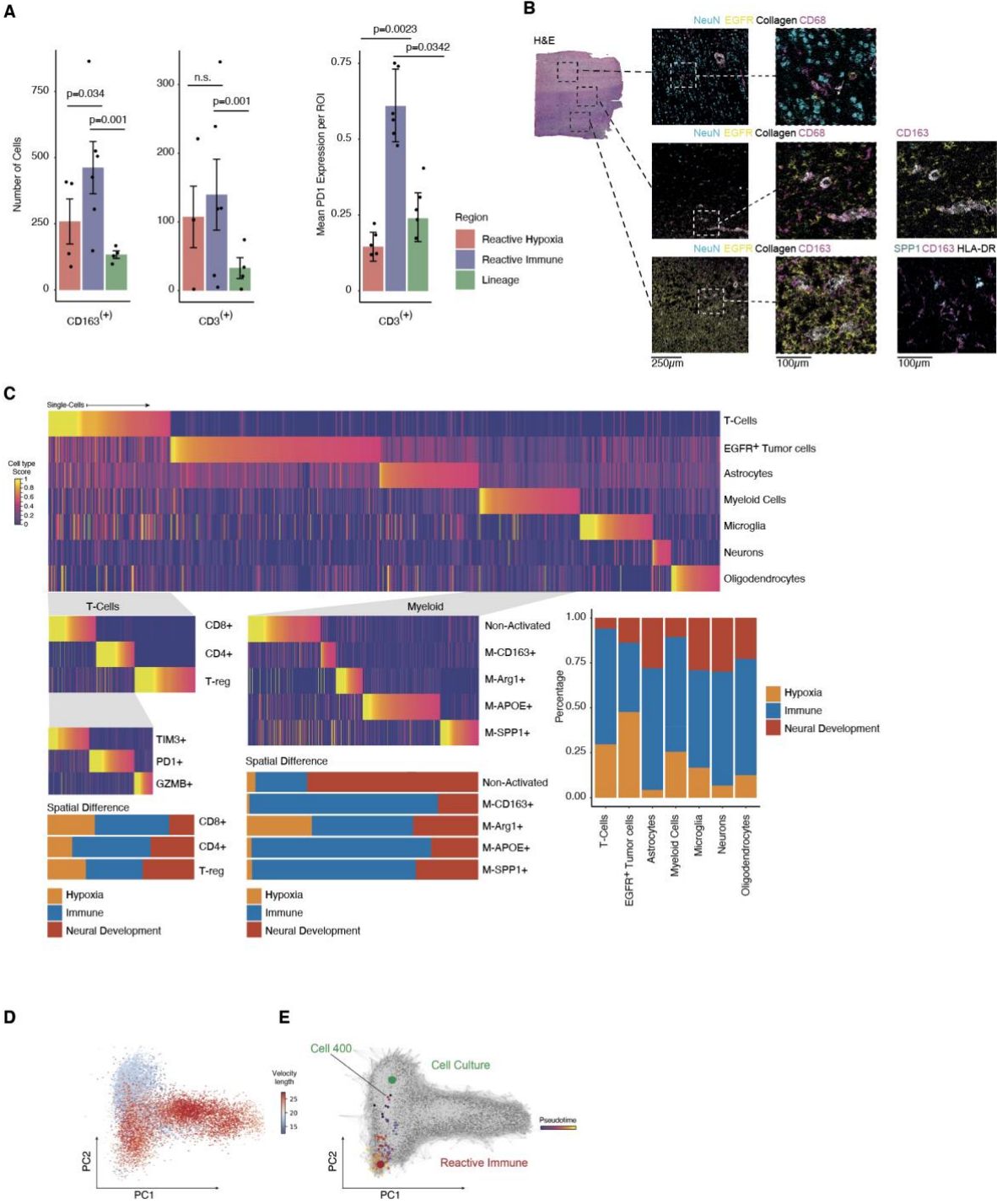
(H) A surface plot of *Aqp5* and the inferred gain of chromosome 15 was illustrated. (I-K) Direct comparison of CNAs across chromosome 1-19 from inferred CNA slide-seq(I) DNA data (J) and from the Visium reconstruction (K). Colors indicate the ROIs. Significance is tested by two-sided t-test demonstrated at the bottom. I) Heatmap of CNAs from scRNA-seq (upper panel), stRNA-seq (middle panel) and non-malignant reference. M) Direct comparison of mean CNA obtained from sc- or stRNA-seq data. Both data types were from the same tissue source.

Supplementary Figure 7



Supplementary Figure 7: Integration of MADi and stRNA-seq related to Figure 4

A) Workflow of the image processing and alignment of metabolic and stRNA-seq data. A detailed explanation is given in the STAR Methods B) PCA based dimensional reduction of metabolic spots, colors indicate the patients. (bottom) Metabolic enrichment of Pentose Phosphate Pathway (PPP) (left), Glycolysis (middle) and the Phosphoadenylate Metabolism (right). Colors indicate the enrichment score. C) Overview of aligned metabolic and transcriptional data.



176

177 **Supplementary Figure 8: Analysis of IMC Data, related to Figure 7 - 8**

178 A) Bar plots of cell counts for tumor-associated macrophages (CD163⁺) and T cells (CD3⁺) (on the left)

179 in different ROIs. On the right, the expression of PD1 in CD3⁺ cells within different areas was quantified.

180 Error bars illustrate the standard error and significance levels were determined using ANOVA. B)

181 Sample with tumor and infiltration areas (#UKF_269). ROI's used for IMC are marked in the H&E image.

182 IMC images (right) from all regions illustrate the different distribution of tumor cells (EGFR), neurons

(NeuN) and myeloid cells (CD68 and CD163). Bottom-Right: immunostaining (IMC) of SPP1, HLA-DR and CD163 illustrate the typical tumor-associated activated myeloid cells. C) Illustration of the cell type identification in IMC data. A heatmap of cellular distribution in different spots is illustrated. At the top, all regions are summarized (~82,000 cells). Middle panel: T cells and subgroups (bottom right). In the middle panel, the myeloid cells are further characterized. D) PCA dimensional reduction. Colors indicate the velocity length. E) A velocity graph embedded in the PCA dimensions reduction is illustrated. After deriving a distribution over root cells from the graph, the average number of steps needed to reach a cell after walking along the graph starting from the root cells is measured. Here, a random cell of the cell culture is selected.

Table S4: IMC Antibodies, related to Figure 7

Metal	Target	Clone	Company	Cat#	LOT	Dilution
89Y	CD45	D9M8I	CST	13917BF	8	200
111Cd	CD4	RPA-T4	BioLegend	300502	B273528	50
113In	VEGFA	G153-694	BD	555036	9344458	100
115In	HLA-DR	TAL-1B5	Abcam	ab176408	GR3314722-2	400
141Pr	Iba1	EPR16588	Abcam	ab178846	GR3290113-2	100
142Nd	EGFR	A19002A	BioLegend	933902	B297820	400
143Nd	GFAP	EPR1034Y	Abcam	834501	B305731	100
144Nd	CHI3L1	EPR19078-157	Abcam	ab255864	GR3281325-2	200
145Nd	CD3	Poly	Fluidigm	85061BF	4	200
146Nd	Ki-67	B56	Fluidigm	556003	8116751	200
147Sm	CD163	EDhu-1	Fluidigm	3147021D	242001	200
148Nd	NeuN	1B7	BioLegend	834501	B305731	400
149Sm	SPP1	A15059B	BioLegend	691302	B246000	100
151Eu	CCL2	Poly5360	BioLegend	536002	B264582	100
152Sm	CCL4	W15138A	BioLegend	686202	B217782	100
153Eu	HOPX	Poly	Abcam	ab230544	GR3283075-3	100
154Sm	Tim-3	D5D5R	Fluidigm	3154024D	3251902	200
155Gd	FoxP3	236A/E7	Thermo	14-4777-82	2129676	100
157Gd	OLIG1	EPR6790	Abcam	ab248054	GR3316742-1	400
158Gd	PD-1	EH12.2H7	BioLegend	329902	B292191	100
159Tb	CD68	KP1	Fluidigm	3159035D	2421901	5000
160Gd	CALM2	EPR5028	Abcam	ab247990	GR3321108-1	100
161Dy	THY1	5E10	Fluidigm	3161009B	0022002	3200
162Dy	CD8a	C8/144B	Fluidigm	372902	B298974	800
163Dy	SNAP25	O91D3/SNAP-25	BioLegend	850302	B240762	100
164Dy	Arginase-1	D4E3M	Fluidigm	93668BF	5	100
165Ho	NTRK2	A15019C	BioLegend	695102	B231674	5000
166Er	VIM	O91D3	BioLegend	677802	B267131	1600

167Er	GranzymeB	EPR20129-217	BioLegend	3167021D	2771902	400
168Er	CD24	ML5	BioLegend	311127	B240406	200
169Tm	Collagen	Poly	Fluidigm	3169023D	281902	800
170Er	CD44	BJ18	BioLegend	338811	B229015	3200
171Yb	APOE	A17065A	BioLegend	852802	B257864	200
172Yb	CD11b	ICRF44	BioLegend	301337	B291844	400
173Yb	CD45RO	UCHL1	Fluidigm	304202	B265383	400
174Yb	C3	K13/16	BioLegend	518101	B202439	1600
175Lu	CCR4	L291H4	BioLegend	359402	B280308	100
176Yb	cCasp3	5A1E	CST	9664S	23	50
191/193Ir	DNA		Fluidigm	201192A	P19H2305	2000

194

195

196