Investigating the cellular landscape of glioblastoma brain tumours, and how it changes through treatment

Shoaib Ali Ajaib

Supervisor: Dr. Lucy Stead

School of Medicine
University of Leeds

July 2025

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

Publication Statement

I confirm that the work submitted is my own, except where work which has formed part of jointly authored publications has been included. My contribution and the other authors to this work has been explicitly indicated below. I also confirm that appropriate credit has been given within the thesis where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Lead Author Publications

1. **Ajaib, S.** *et al.* GBMdeconvoluteR accurately infers proportions of neoplastic and immune cell populations from bulk glioblastoma transcriptomics data. *Neuro-Oncology* **25**, 1236–1248 (July 2023).

   doi:10.1093/neuonc/noad021.

   This work forms Chapter 2 of this thesis. My contribution included its conception and design; collection and assembly of the data; data analysis and interpretation; and writing of the manuscript.

2. **Ajaib, S.** *et al.* Deconvoluting cellular landscape changes in IDH-wildtype glioblastomas pre- and post-treatment (draft manuscript).

   This work forms Chapter 3 of this thesis. My contribution included its conception and design; collection and assembly of the data; data analysis and interpretation; and writing of the manuscript.

3. **Ajaib, S.** *et al.* Spatial profiling of longitudinal glioblastoma reveals consistent changes in cellular architecture, post-treatment (submitted pre-print).

   doi:10.1101/2025.01.31.635832.

This work forms Chapter 4 of this thesis. My contribution included its conception and design; collection and assembly of the data; data analysis and interpretation; and writing of the manuscript.

Collaborations (not included)

1. Thomas, M. P. H., **Ajaib, S.**, Tanner, G., Bulpitt, A. J. & Stead, L. F. GBMPurity: A Machine Learning Tool for Estimating Glioblastoma Tumour Purity from Bulk RNA-seq Data. *Neuro-Oncology*, noaf026 (Feb. 2025). doi:10.1093/neuonc/noaf026.

2. Tanner, G. Barrow, R., **Ajaib, S.**, *et al.* IDHwt glioblastomas can be stratified by their transcriptional response to standard treatment, with implications for targeted therapy. *Genome Biology* **25**, 45 (Feb. 2024).

   doi:10.1186/s13059-024-03172-3.

3. Marcuccio, F., Chau, C. C., Tanner, G., Elpidorou, M., Finetti, M. A., **Ajaib, S.**, *et al.* Single-cell nanobiopsy enables multigenerational longitudinal transcriptomics of cancer cells. *Science Advances* **10**, eadl0515 (Mar. 2024). doi:10.1126/sciadv.adl0515.

4. Pitaksalee, R., Burska, A. N., **Ajaib, S.**, *et al.* Differential CpG DNA methylation in peripheral naïve CD4+ T-cells in early rheumatoid arthritis patients. *Clinical Epigenetics* **12**, 54 (Apr. 2020). doi:10.1186/s13148-020-00837-1.

Acknowledgements

I would like to thank Dr. Lucy Stead for her guidance, support, and unwavering commitment to my success.

I am also grateful to Dr. Frédérique Ponchel, Prof. Paul Emery, Dr. Richard Hodgett, Dr. Alistair Droop, Dr. Matthew Care, Dr. Arief Gusnanto, Dr. Kave Shams, and Dr. Ai Lyn Tan. Each of you has shaped my journey in your own way, without which I would not have reached this point.

Thank you to all members of the Glioma Genomics group and my previous lab group, for the laughter and healthy competition we shared.

I also extend my appreciation to the dedicated technicians and support staff I encountered at the Wellcome Trust Brenner Building and the Clinical Sciences Building on the St James's Hospital campus.

My sincere thanks to Dr. Henry Wood and Dr. Simon Cockell for taking the time to read this thesis and conduct my viva.

To my family and my friends - thank you for your love and patience through these years. I imagine many of you thought this day might never come!

Finally, I acknowledge the patients affected by the devastating disease that lies at the heart of this work. Their repeated, selfless consent to donate tumour samples made this vital research possible. I hope that one day, their contributions will help bring a cure for those who follow.

Abstract

Glioblastoma (GBM) is the most aggressive primary malignancy of the central nervous system. Despite standard treatment, comprising surgical resection, followed by concomitant radiation and chemotherapy, it is incurable. This devastating prognosis stems from complex multi-layered heterogeneity, which enables GBM tumour cells to resist treatment and reoccur. Advances in genomic technologies have classified GBM tumours at the single-cell resolution, revealing that malignant GBM cells occupy distinct neoplastic cell states which resemble neurodevelopmental hierarchies and wound healing programs. These states are supported by complex interactions which include immune, healthy brain and vasculature cells.

To investigate how the cellular landscape of GBM tumours changes through treatment, I utilised an extensive dataset of paired (pre- and post-treatment) GBM tumour patient samples. These samples were profiled using bulk RNA sequencing (RNA-seq) and thus characterising their cell type composition *in silico*, necessitated the use of cellular deconvolution techniques. Benchmarking of such methods has shown that accuracy and interpretability are highly dependent on the specificity of the cell type reference used. Therefore, I developed a set of GBM-specific cell type markers and used these to validate the optimal deconvolution method, which I also released as a publicly available web application, GBMDeconvoluteR. Using this tool, I characterised 219 paired GBM samples and uncovered consistent cell type changes through treatment. These changes were associated with survival outcomes and aligned with our previously described patient stratification, based on treatment-resistance mechanisms.

To complement these findings, I then applied a novel spatial proteomics method and found that hypoxia drives the layered organisation of the GBM tumour microenvironment (TME) pre-treatment, but post-treatment the GBM TME is less structured, driven instead by reactive astrocytes and infiltrating lymphocytes.

Collectively, this work highlights some key shifts in the cellular landscape of GBM through treatment, which may hold therapeutic potential.

Contents

List of Figures

List of Tables

Abbreviations

1p19q	chromosome 1 short arm and chromosome 19 long arm loss
5-ALA	5-aminolevulinic acid
AC-like	astrocyte-like
ADC	apparent diffusion coefficient image
aKG	α-ketoglutarate
arcsinh	inverse hyperbolic sine function
ATP	adenosine triphosphate
ATRX	alpha thalassemia/mental retardation syndrome X-linked
BBB	blood-brain barrier
BDNF	brain-derived neurotrophic factor
bp	base pairs
BRAIN UK	brain archive information network UK
Breg	regulatory B cell
BSA	body surface area
CAR-T	chimeric antigen receptor T cell
CCNU	Lomustine
CDK2A/B	cyclin-dependent kinase inhibitor 2A/B gene
CE	classification error
Ch +7/-10	chromosome 7 gain and chromosome 10 loss

CI	confidence interval
CL	classical
CN	cellular neighbourhood
CNS	central nervous system
CNS1	WHO CNS classification 1st edition
CNS4	WHO CNS classification 4th edition
CNS5	WHO CNS classification 5th edition
CNV	copy number variation
CODEX	co-detection by indexing
CRISPR	clustered regularly interspaced short palindromic repeats
CSC	cancer stem cell
CSF	cerebrospinal fluid
CT	computed tomography
CTL	cytotoxic T lymphocyte
CyTOF	cytometry by time-of-flight
D2HG	D-2-hydroxyglutarate
DC	dendritic cell
DCV	dendritic cell vaccine
DE	differentially expressed
DNA	deoxyribonucleic acid
DWI	diffusion-weighted image
EGFR	epidermal growth factor receptor
EMT	epithelial-to-mesenchymal transition
EORTC	European organisation for research and treatment of cancer
ER	endoplasmic reticulum
Fc region	fragment crystallizable region

FDR	false discovery rate
FFPE	formalin-fixed paraffin-embedded
FISH	fluorescence *in situ* hybridization
FOXP3	forkhead box P3
FPKM	fragments per kilobase million
GAM	glioma-associated macrophage
GBM	glioblastoma
GLASS	glioma longitudinal analysis consortium
glmnet	generalised linear model with elastic net
GSC	glioma stem cell
GST	glutathione S-transferase
GTR	gross total resection
Gy	gray
H&E	haematoxylin and eosin staining
HIF1α	hypoxia-inducible factor 1-alpha
HPF	high-power fields
HR	hazard ratio
ICC	immunocytochemistry
IDH	isocitrate dehydrogenase
IDHmut	isocitrate dehydrogenase mutated
IDHwt	isocitrate dehydrogenase wild-type
IF	immunofluorescence
IFNγ	interferon-gamma
IgE	Immunoglobulin E
IgG	Immunoglobulin G
IgM	Immunoglobulin M

IHC	immunohistochemistry
IL-10	interleukin 10
IL-11	interleukin 11
IL-6	interleukin 6
IMC	imaging mass cytometry
ISH	*in situ* hybridization
KNN	k-nearest neighbour
KO	knock-out
LDA	linear discriminant analysis
Log-reg	logistic regression
MES	mesenchymal
MES-like	mesenchymal-like
MES1	hypoxia-independent mesenchymal-like
MES2	hypoxia-dependent mesenchymal-like
MES3	Wang *et al.* mesenchymal
MET	mesenchymal epithelial transition
MGMT	O^6-methylguanine DNA methyltransferase
MHC	major histocompatibility complex
MIBI	multiplexed ion beam imaging
MMP	matrix metalloproteinase
MRI	magnetic resonance imaging
MRS	magnetic resonance spectroscopy
MS	multiple sclerosis
MSC	mesenchymal stem cell
multinorm	multinomial loglinear model
MVP	microvascular proliferation

NADP	nicotinamide adenine dinucleotide phosphate
NF1	neurofibromin 1
NGS	next generation sequencing
NHS	National Health Service
NK cell	natural killer cell
NLGN3	neuroligin-3
NMF	non-negative matrix factorisation
nnet	neural network
NOS	not otherwise specified
NPC-like	neural progenitor-like
NPC1	oligodendrocyte-lineage associated neural progenitor-like
NPC2	neuronal-lineage associated neural progenitor-like
NVU	neurovascular unit
OPC	oligodendrocyte progenitor cell
OPC-like	oligodendrocyte progenitor-like
OS	overall survival
OSM	oncostatin M receptor
PaCMAP	pairwise controlled manifold approximation
PAM	PI3K/AKT/mTOR signalling pathway
PBS	phosphate buffered saline
PCA	principal component analysis
PCR	polymerase chain reaction
PD-1	programmed cell death protein 1
PDGFR	platelet derived growth factor receptor
PET	positron emission tomography
PFI	permutation feature importance

PFS	progression-free survival
PH	proportional hazards
pHH3	phosphohistone H3
PN	proneural
PTEN	phosphatase and tensin homolog
PVR	perivascular region
PWI	perfusion-weighted image
QC	quality control
QDA	quadratic discriminant analysis
ranger	random forest
RANO	response assessment in neuro-oncology
Rb	retinoblastoma
RNA	ribonucleic acid
RNA-seq	RNA sequencing
ROC-AUC	receiver operator curve-area under the curve
ROI	regions of interest
ROS	reactive oxygen species
rpart	classification tree
rRNA	ribosomal ribonucleic acid
RTK	receptor tyrosine kinase
SC	spatial context
scRNA-seq	single-cell RNA sequencing
SLM	smart local moving
SMR	supramaximal resection
SNN	shared nearest-neighbour
snRNA-seq	single-nucleus RNA sequencing

SVM	support vector machine
T1-CE	T1-contrast enhanced image
T1-W	T1-weighted image
T2-FLAIR	T2 fluid-attenuated inversion recovery image
T2-W	T2-weighted image
TAM	tumour-associated macrophage
TARA	tumour-associated reactive astrocyte
TCA	tricarboxylic acid cycle
TCGA	the cancer genome atlas
TERT	telomerase reverse transcriptase
TGFβ	transforming growth factor-beta
Th	T helper cell
TI	tumour-intrinsic
TLS	tertiary lymphoid structure
TME	tumour microenvironment
TMZ	temozolomide
TNFα	tumour necrosis factor-alpha
TP53	tumour protein 53
TPM	transcripts per million
Treg	regulatory T cell
TTF	tumour-treating field
UMAP	uniform manifold approximation and projection
VEGF	vascular endothelial growth factor
WCSS	within-cluster sum of squares
WHO	world health organisation
xgboost	extreme gradient boosting

Bismilla Hir Rahma Nir Raheem

Chapter 1

Introduction

Glioblastomas (GBMs) are the most common type of malignant brain tumour, accounting for ∼50% of all primary malignancies of the central nervous system (CNS) and more than half of all gliomas[1,2]. They are characterised by aggressive proliferation, diffuse infiltration and resistance to standard-of-care multimodality treatments, which invariably results in tumour recurrence and mortality on average ∼7 and ∼15 months of diagnosis, respectively[3–5].

1.1 Adult-type diffuse gliomas

The classification scheme for CNS tumours was last revised (fifth edition) in 2021 and is provided by the world health organisation (WHO)[6]. In this framework, termed WHO CNS classification 5th edition (CNS5), tumours are grouped according to the following taxonomy: category, family, type and subtype. GBM tumours are classified within the adult-diffuse glioma family of CNS tumours which is part of the broader glioma "category" of tumours (Figure 1.1 and Figure 1.2).

Gliomas are malignancies of the CNS that originate from glial cells: non-neuronal cells that play a critical role in maintaining homeostasis within the brain and spinal cord[7]. Linking the origin of cancers to normal brain cells was first reported in 1856 by Rudolf Virchow following his rudimentary classification of brain malignancies based on post-mortem samples[8]. Subsequently, in 1926, Percival Bailey and Harvey Cushing, amongst others, introduced the first systematic classification of CNS tumours, which relied on a combination of histological and morphological features and their correlation with patient survival outcomes[9].

The WHO CNS classification 1st edition (CNS1) continued to focus on histopatho-

logical methods to diagnose and classify tumours, comparing tumour cells with normal brain cells: those resembling astrocytes were called astrocytomas, and cells resembling oligodendrocytes were called oligodendrogliomas[10]. In the 2016 revision of the WHO CNS classification 4th edition (CNS4), molecular testing was first added alongside histological techniques, though it was only used in an ancillary capacity to diagnose and classify CNS tumours[11]. Following this in 2021, the most recent classification scheme (CNS5) made molecular diagnostics foundational to the classification of CNS tumours, whilst still utilising histological and immunohistochemical techniques as complementary approaches. As a consequence of this, adult-type diffuse gliomas and pediatric-type gliomas were separated into distinct tumour families for the first time: these tumours are histologically indistinguishable but have very different biology[6]. The other major sub-division within the glioma family of tumours is based on the degree of invasiveness into the surrounding brain tissue. Gliomas which display (based on histology) infiltration into the surrounding normal brain parenchyma are defined as "diffuse" and those with better-demarcated margins are labelled with the term "circumscribed"[12].

The adult-type diffuse glioma family of tumours comprises several tumour types which are graded (from 2-4) according to their respective malignancy under CNS5[6]. Moreover, each tumour type is classified and graded using a combination of histological and molecular features (Figure 1.2). Herein, the term "glioma" will be used to refer to the adult-type, diffuse family of gliomas.

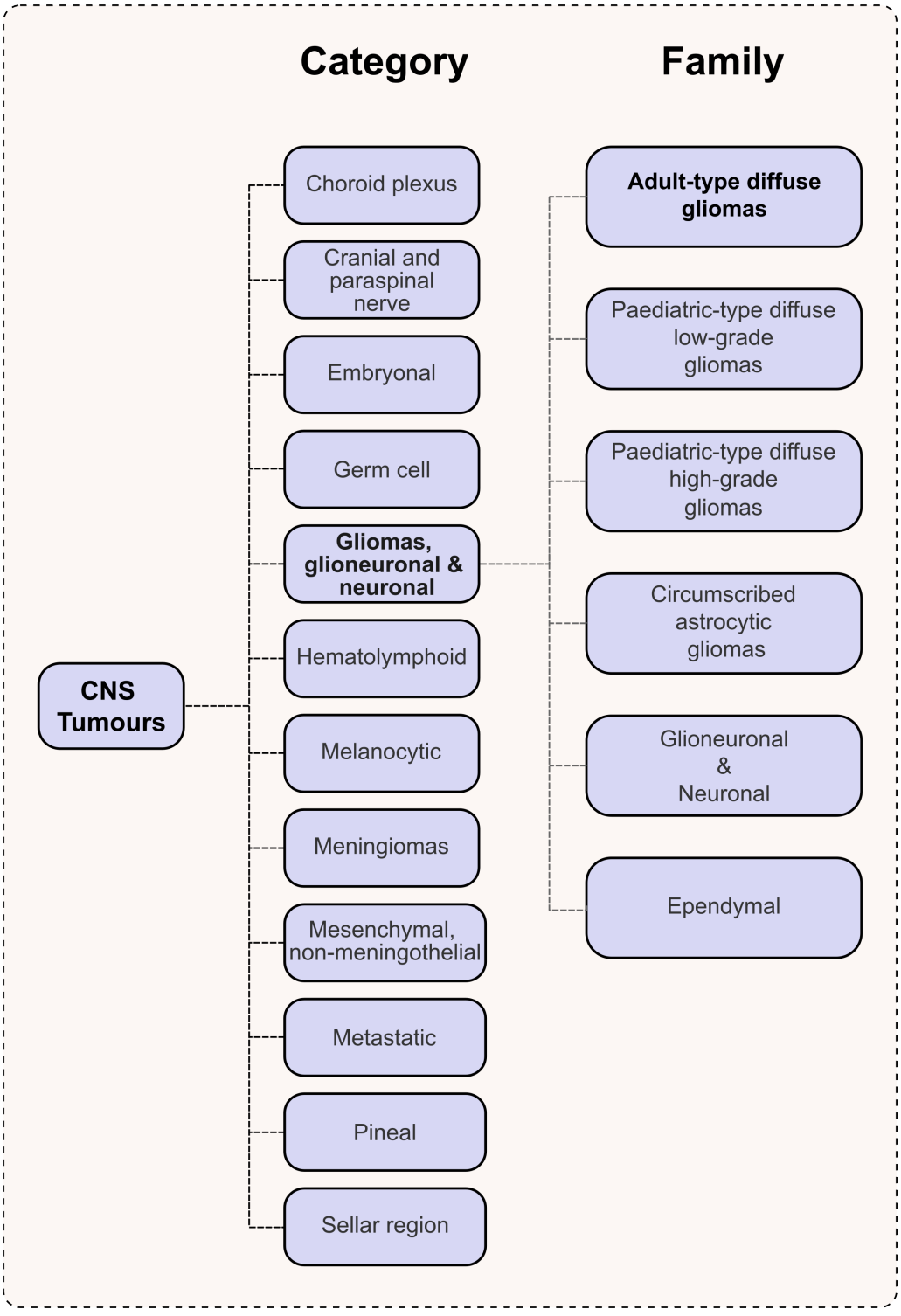


Figure 1.1 | Overview of the 2021 WHO classification of CNS tumours.

Hierarchical tree diagram showing the main CNS tumour categories that make up the 2021 (5th edition) World Health Organization (WHO) classification of CNS Tumours. The full taxonomy of this classification follows the following order: Category; Family; Type and Subtype. All the main CNS tumour categories and only the "Glioma" family of tumours are shown.

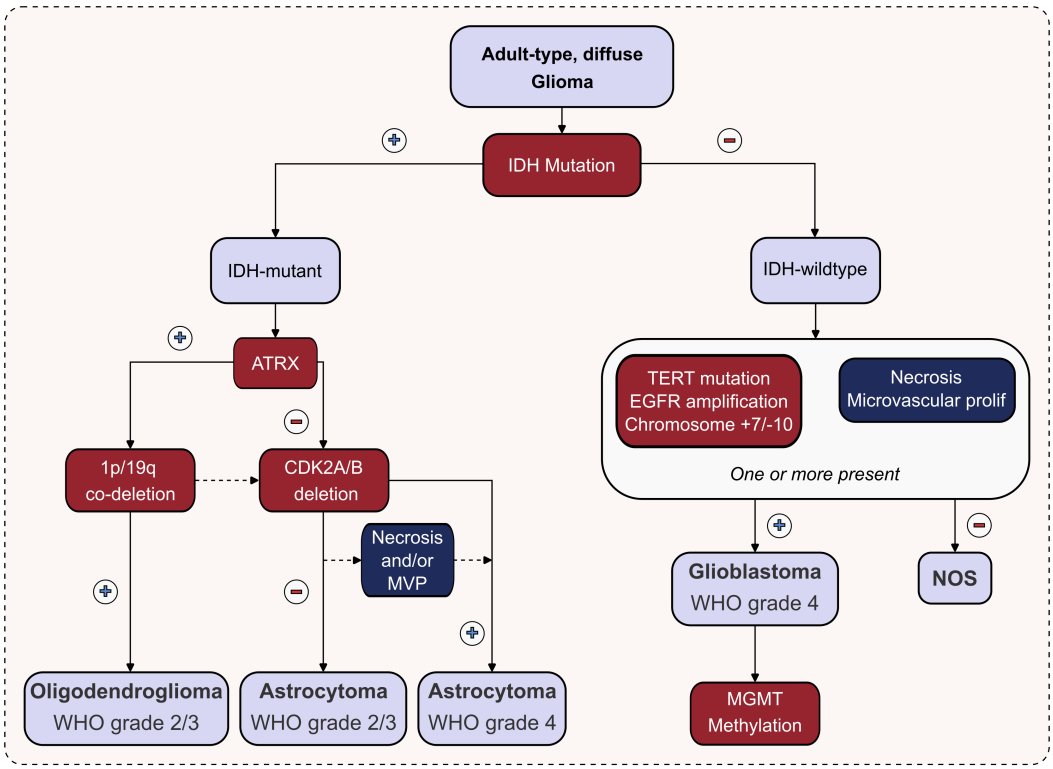


Figure 1.2 | Classification & diagnosis workflow of adult-type diffuse gliomas.

Flow diagram detailing a simplified overview of the classification pathway for adult-type diffuse gliomas based on CNS5. Molecular features are shown shaded in red and histological features are shaded in dark blue. The "+" icon denotes presence/gain, and the "-" icon denotes the loss/maintenance of molecular or histological feature. Isocitrate dehydrogenase (IDH); alpha thalassemia/mental retardation syndrome X-linked (ATRX); cyclin-dependent kinase inhibitor 2A/B gene (CDK2A/B); telomerase reverse transcriptase (TERT); epidermal growth factor receptor (EGFR); microvascular proliferation (MVP); not otherwise specified (NOS); O^6-methylguanine DNA methyltransferase (MGMT).

1.2 IDH mutation status

IDH refers to a family of enzymes that feature prominently in the tricarboxylic acid cycle (TCA), where they catalyse the oxidative decarboxylation of isocitrate, producing α-ketoglutarate (aKG) and the reduced form of nicotinamide adenine dinucleotide phosphate (NADP)[13]. The IDH family comprises of three known isoforms, IDH1, IDH2 and IDH3[14]. Point-mutations within the binding site of IDH1 and IDH2 result in a gain-of-function which enables them to catalyse the reduction of aKG to form the onco-metabolite D-2-hydroxyglutarate (D2HG)[15]. The accumulation of D2HG inhibits aKG–dependent dioxygenases, which include histone demethylases, leading to widespread histone hypermethylation[16]. These epigenetic modifications alter the cell metabolism within IDH mutant cells and make them more susceptible to standard treatment, by reducing their ability to handle reactive oxygen species (ROS) generated during standard therapy[13].

Within adult-diffuse gliomas IDH1 mutations (IDH1:C.395G>A p.R132H) occur in >80% of all tumours[17,18]. These mutations are predominantly found within lower grade, WHO grade 2/3 adult-diffuse gliomas and in secondary GBM tumours that develop from lower grade gliomas[19]. Differences between isocitrate dehydrogenase mutated (IDHmut) and isocitrate dehydrogenase wild-type (IDHwt) tumours are also seen in survival outcomes where IDHmut patients have much better overall survival (OS) and progression-free survival (PFS), irrespective of grade and are also more responsive to certain drugs[20]. Given this and the fact that IDH mutations occur early in in gliomagenesis[21], the current WHO classification of gliomas uses IDH mutation status as the primary molecular biomarker by which to classify and grade between lower and higher-grade gliomas[11].

1.2.1 Oligodendroglioma

Oligodendrogliomas are the least prevalent (∼7% in the US) adult-type diffuse gliomas and have an incidence rate of 0.29 per 100,000 people[1]. In comparison to other adult-type diffuse gliomas, they are associated with a good prognosis and chemotherapeutic responsiveness, with median survival ranging from 11.19 to 20.45 years[22,23]. Molecularly, oligodendrogliomas are classified by the coexistence of IDH1/2 mutations and 1p/19q codeletion[24,25]. They are graded as either grade 2 or 3, with the latter designation being chosen when there is a significant mitotic count (> 6 mitoses per 10 high-power fields (HPF)) and/or MVP[6]; or the presence of CDK2A/B homozygous deletion[26,27].

1.2.2 Astrocytoma

Astrocytomas have an incidence rate of 0.45 per 100,000 people and make up ∼11% of all diffuse adult-type glioma diagnoses[1]. They are rarely diagnosed in adults over the age of 55 as they primarily occur in younger adults, with a median diagnosis age of 37 years and a median survival that ranges from 7.29 to 9.3 years[1]. Molecularly, astrocytomas are classified by the concurrence of mutations in IDH1/2 genes, ATRX gene and the tumour suppressor gene, tumour protein 53 (TP53)[6]. This is in contrast with oligodendrogliomas, where ATRX expression is retained and accumulation of p53 is largely absent[28]. In the WHO CNS4 classification, astrocytomas carrying a IDH1/2 mutation (IDHmut), were assigned to three different tumour types (diffuse astrocytoma, anaplastic astrocytoma, and glioblastoma) based on histological features[11]. However, in the current classification all IDHmut astrocytomas are considered a single type (Astrocytoma, IDHmut) and are graded within CNS WHO grade 2, 3, or 4 6. Moreover, since the publication of the CNS5 guidelines, homozygous deletion of CDK2A/B is now sufficient to classify an astrocytoma as the highest grade (grade 4), even in the absence of high-grade histopathological features (Figure 1.2)[6]. This was done, on account of the poor prognosis and limited OS (∼3 years) in astrocytomas with deletions of CDK2A/B[29,30].

1.3 Grading gliomas

Grading tumours serves an important purpose in that it provides important prognostic information for clinicians based on a tumour's microscopic features and allows them to develop the most appropriate treatment plan.

Gliomas are graded from 2-4, though this process is largely subjective and suffers greatly from inter- and intra-observer variabilities[26,31]. The reason we do not see any grade 1 gliomas is because this grade is reserved for tumours that typically do not infiltrate surrounding tissue[6]. Gliomas are by their very definition, "diffuse" and do infiltrate into the normal surrounding brain which in turn, makes them difficult to completely remove, and challenging to treat as residual tumour facilitates recurrence. Although the process of grading is no longer entirely histological, there are four main histological features that are commonly referred to as the *AMEN* score which are used when determining tumour grade[32].

Nuclear atypia (A) - describes the abnormal size, shape and density (nuclear pleomorphism) of the nucleus within tumour cells and the presence of abnormally

dense chromatin (hyperchromasia). Lower grade (grade 2) gliomas exhibit minimal nuclear atypia, whereas higher grade (grade 3-4) have marked to severe nuclear atypia[33,34].

Mitotic activity (M) – a key indicator of a tumour's proliferative capacity. This is measured using haematoxylin and eosin staining (H&E) which detects the presence of mitotic figures that appear as darkly stained sections and represent cells undergoing chromosomal condensation[35]. Mitoses are counted in ten consecutive HPF and form the mitotic index measure. Lower grade (grade 2) gliomas exhibit minimal mitosis, whereas higher grade (grade 3-4) gliomas show increased mitotic activity that correlates with a poorer prognosis[26]. Immunohistochemistry (IHC) with antibodies for Ki-67 and phosphohistone H3 (pHH3) are also commonly used to assess the degree of malignancy[36–38].

Microvascular proliferation (E) – this was previously known as endothelial proliferation and refers to the aberrant growth of the vasculature within the tumour. The endothelial cells form glomeruloid structures which can be observed using H&E staining. This feature is a hallmark of higher grade (grade 4) gliomas where it is indicative of aggressive tumour behaviours and a worse prognosis[39].

Necrosis (N) - This is detected in varying amounts across multiple different cancer types and refers to areas of dead cells within the tumour[40]. Histologically this is seen using H&E staining with acellular areas that comprise nuclear debris that are surrounded by hypercellular zones of viable tumour cells (pseudopalisading necrosis)[40]. Necrosis is largely absent from lower grade gliomas, though is a key criterion for classifying higher grade (grade 4) gliomas[41].

1.4 IDH-wildtype glioblastoma

IDHwt GBM is the most prevalent type of glioma with an incidence rate of 2.66 per 100,000 people and predominately affects older adults with a median age at diagnosis of ∼65 years old[42]. There is a slight preponderance toward males with a 3:2 (male:female) ratio and white patients are more frequently affected compared with other ethnicities as evidenced by the diverging prevalences between Europeans/North Americans (3-4 per 100,000) and Asians (0.59 per 100,000)[42,43].

The vast majority (∼90%) of IDHwt GBM tumours develop *de novo* and as such are termed *primary* tumours, as opposed to *secondary* GBM tumours which arise from progression of a prior lower-grade glioma[44]. Although *primary* and *secondary* tumours are histologically indistinguishable, they have specific molecular, genetic

and epigenetic profiles. *Secondary* GBM tumours are typically found in younger patients and carry an IDHmut, resulting in a significantly better prognosis compared to *primary* GBM tumours which are highly aggressive and contain larger patterns of necrosis[45].

1.4.1 Clinical presentation & diagnosis

The clinical presentation of GBM is varied and depends on several factors including the tumour size, location and associated edema within the brain. The most widely reported symptoms are: headaches which are triggered by the Valsalva maneuver and further exacerbated by lying down, coughing or bending over; nausea and vomiting due to brainstem compression; cognitive impairment and personality changes such as memory loss, confusion, lethargy and mood swings; focal neurological deficits; and seizures (∼20% of cases)[45,46]. In more acute cases (<2%) some patients experience inter-tumoural hemorrhaging and present with stroke-like symptoms[47]. The pre-operative diagnosis of GBM is initially conducted using the patients clinical history and multimodal imaging techniques such as structural magnetic resonance imaging (MRI) with gadolinium; computed tomography (CT) and positron emission tomography (PET)[48,49]. Historically this process was largely qualitative and relied exclusively on the interpretation of radiologists. However, recent developments in CT, MRI, and artificial intelligence have introduced quantitative components which are used to supplement subjective analysis[49]. More novel techniques such as theranostics have also been explored for use in GBM and remain under active development[49,50]. This technique utilises radioactive nuclides coupled via a linker to an antibody for a cellular target/receptor that can be used both for imaging (with a PET scanner) and delivering therapeutic radiation to cancer cells[50]. Histological and molecular diagnostic techniques are used to confirm an initial diagnosis and also to further refine the tumour classification, after patients have undergone surgical intervention.

1.4.2 Molecular features

The current WHO CNS5 scheme defines IDHwt GBM tumours based on presence of three main genetic aberrations (TERT, EGFR and chromosome 7 gain and chromosome 10 loss (Ch +7/-10)) that are routinely tested for as part of the diagnosis process using a combination of Sanger sequencing; next generation sequencing (NGS); fluorescence *in situ* hybridization (FISH) and polymerase chain reaction (PCR) techniques[6].

MGMT gene promoter methylation and protein expression are both important molecular features in GBM classification. Clinically, however, they are primarily used as prognostic markers for predicting response to alkylating chemotherapy, and therefore, will be discussed in more detail in the following section on treatment.

TERT mutation

TERT refers to the catalytic subunit of telomerase, an enzyme which plays a critical role in maintaining telomere lengths[51]. Telomeres are short, tandem repeats of deoxyribonucleic acid (DNA) sequences (TTAGGG in humans) found at the ends of chromosomes, and are pivotal in maintaining genomic stability and integrity, by preventing DNA degradation and end-to-end fusions[52]. Most somatic cells have limited or absent telomerase activity, leading to gradual telomere attrition over successive cell division cycles; however, in germ cells, stem cells and highly proliferative cells such as cancer cells, telomerase is frequently reactivated[51,52].

In GBM tumours, TERT promoter mutations are found in 70–90% of cases, driving telomerase reactivation, enabling tumour cells to bypass replicative senescence and sustain uncontrolled proliferation[53]. The reactivation of TERT occurs due to mutations located at −124 base pairs (bp) and −146 bp upstream of the TERT translation start site[54]. These mutations create sequences which represent a binding motif for the E26 transformation-specific family transcription factor which increase TERT transcription and telomerase activity[54].

The prognostic significance of TERT mutations depends on IDH status and the presence of chromosome 1 short arm and chromosome 19 long arm loss (1p19q) codeletion. Triple-positive lower grade gliomas (positive for TERT, IDHmut, and 1p19q codeletion), result in a much better prognosis compared with IDHmut GBM tumours, which in turn have an even worse overall prognosis compared with triple-negative gliomas[24].

EGFR amplification

Receptor tyrosine kinases (RTKs) are trans-membrane glycoproteins which mediate cell-cell interactions and are involved in biological functions such as cell motility, differentiation, proliferation and metabolism[55]. Most RTKs are activated via ligand binding to an extracellular domain which induces a conformational change in an intracellular catalytic domain[55]. This activation facilitates the recruitment of additional proteins which trigger signalling cascades that integrate different

signalling pathways[56]. There are 58 known RTKs in humans[57]: several RTK encoding genes such as EGFR, platelet derived growth factor receptor (PDGFR), and the mesenchymal epithelial transition (MET) factor receptor are implicated in GBM development[56].

EGFR is the most frequently amplified RTK in GBM and seen in ∼40-50% of all *primary* GBM tumours and retained through treatment in ∼80% of patients[58,59]. The amplified genes are often located on extrachromosomal circular DNA known as double minutes which are formed during chromothripsis: a chromosomal rearrangement event where one or more chromosomes shatter into pieces, then reassemble in an unnatural order and/or orientation[60]. In addition to amplification, EGFR genes can also harbour point mutations or deletions: the most frequently occurring deletion is found in exons 2-7 within the extracellular domain of EGFR, resulting in a truncated mutant variant (EGFRvIII). This specific deletion is found in ∼50% of all cases and leads to constitutive activation of the EGFR receptor, independent of ligand binding[61,62].

The prevalence of EGFR across multiple cancers has made it a prominent target of therapeutic intervention by inhibition. This approach has been successful in the treatment of colorectal, head, neck and lung cancers[63–65], though to-date has proven unsuccessful in GBM[66]. The reasons for this include differences in tissues, therapeutic accessibility and the molecular heterogeneity amongst EGFR: the EGFR mutations in GBM occur within the extracellular domain whilst in other cancers they occur in the intracellular kinase domain[67].

Chromosome +7/-10 alterations

The co-occurrence Ch +7/-10 is commonly found in IDHwt GBM and distinguishes it from other lower-grade gliomas[60,68]. Although this signature is prevalent and characteristic, its presence alone is not shown to be prognostic in IDHwt GBM, though the partial loss of chromosome 10, compared with the complete loss is associated with a better prognosis[69,70]. Both +7 and -10 are thought to occur early in gliomagenesis, with some studies suggesting that chromosome 10 loss precedes the 7 gains, though the exact ordering remains unclear[71,72]. Three separate regions are most frequently associated with chromosome 10 loss, with one on the shorter arm (10p) and two on the longer chromosomal arms (10q)[73]. The tumour suppressor phosphatase and tensin homolog (PTEN) is located at 10q and is suggested as playing a role in GBM tumourigenesis[74], in addition to other genes such as ANXA7, ADARB2, and KLF6[75,76]. The polysomy of chromosome 7 typically involves the entire chromosome, with EGFR and MET - two key oncogenes

located on the 7p and 7q arms, respectively - believed to contribute significantly to the selective advantage conferred by this alteration[77]. Recently, several studies have suggested that there is no small set of genes which explain the Ch +7/-10 pattern, but rather it is shaped by the accumulation of multiple haploinsufficient and triplosensitive genes[78,79]. What is clear is that this aneuploidy signature is important in identifying IDHwt GBM tumours, which is why it was introduced alongside EGFR and TERT in the most recent WHO classification (CNS5)[6].

1.4.3 Treatment

Surgery

The first treatment stage following diagnosis is to de-bulk the tumour mass and relieve intracranial pressure but also provide tissue samples for further molecular and histopathological testing. The guiding principle for resecting gliomas is termed *maximal safe resection* and attempts to remove as much tumour as possible, whilst preserving healthy neurological function[80]. However, this is an inherently difficult task owing to the diffuse nature of GBM, which extends beyond the visible tumour mass into surrounding brain tissue. Moreover, there is no objective standard for determining the extent of maximal safe resection, leading to variability depending on the surgeon attempting the resection[80]. Recently, the response assessment in neuro-oncology (RANO) classification system was developed, in an effort to standardise the extent of tumour resections, though this isn't routinely used in clinical practice[81]. The classification, grouped resections into four categories based on differing median survival rates; 1.) Supramaximal Resection (∼24 months); 2.) maximal contrast-enhancing resection (∼19 months); submaximal contrast-enhancing resection (∼15 months) and biopsy (∼10 months)[81]. The entire surgical process is guided by MRI, which is the gold standard for diagnosing and monitoring GBM. Surgeons utilise different MRI sequences to obtain complementary information about tumour location, infiltration, and response to treatment (Table 1.1). Traditionally, surgery has focused on gross total resection (GTR), which aims to remove all contrast-enhancing tumour visible on T1-contrast enhanced image (T1-CE) MRI[80,82]. Although, this can extend patient survival to 20-25 months in cases where it is achieved, it does not eliminate non-enhancing infiltrative tumour regions that ultimately contribute to recurrence[82]. To address this limitation, a process known as supramaximal resection (SMR) is also used where possible. This process extends beyond GTR by removing additional tissue, as determined by T2 fluid-attenuated inversion recovery image (T2-FLAIR) regions, accounting for mi-

croscopic tumour infiltration[81–83]. This approach has significant advantages over GTR, with patients who undergo SMR showing improved PFS and OS[84,85].

Awake craniotomies are common-place when resecting GBM tumours, as they allow surgeons to monitor language and motor function during surgery, reducing the risk of postoperative deficits[86]. Additionally, the use of the fluorescent dye 5-aminolevulinic acid (5-ALA) has been shown to improve the extent of tumour resection in GBM patients, as it improves visualisation of tumour margins under a red-light illumination source[87]. The use of artificial intelligence is also starting to be integrated into surgical planning to assist in real-time boundary identification and is expected to be used more widely in future[80].

Table 1.1 | Summary of MRI sequences used in GBM.

MRI Sequence	Key Features	Clinical Use in GBM
T1-weighted image (T1-W)	Anatomical brain structures.	Preoperative planning/diagnosis.
T1-CE	Areas where the blood-brain barrier is disrupted.	Define the main tumour mass to guide GTR and detect recurrent tumours.
T2-weighted image (T2-W)	Fluid and edema.	Identify tumour-associated swelling and distinguish between tumour and necrotic areas.
T2-FLAIR	Same as T2-W but with normal cerebrospinal fluid (CSF) signals removed.	Define area of tumour spread beyond contrast enhancement to guide SMR and track progression and radiation-related changes.
diffusion-weighted image (DWI)	Water molecule movement in the brain.	Differentiate between high-grade tumours vs cysts and identify ischemic damage during surgery.
apparent diffusion coefficient image (ADC)	Water diffusion within the brain.	Define cell density and determine tumour grade and treatment response.
perfusion-weighted image (PWI)	Blood flow within brain and tumour.	Differentiate high-grade vs. low-grade gliomas and areas of necrosis.
magnetic resonance spectroscopy (MRS)	Metabolic composition of the tumour.	Differentiate GBM from lower grade metastases and assess metabolism to monitor treatment response.

Concomitant radiotherapy and chemotherapy

Following surgical resection, the standard of care treatment for *primary* GBM tumours is known as the *Stupp protocol* and has been used in clinical practice for the past 20 years[88]. In this protocol patients are treated with a combination of radiotherapy and chemotherapy with the oral alkylating agent temozolomide (TMZ). TMZ works by transporting methyl groups at the O^6 and N^7 positions of guanine, and at the N^3 position of adenine during DNA replication, forming cytotoxic O^6-methylguanine, N^7-methylguanine, and N^3-methyl-adenine[89]. These groups form lethal mismatched base pairs lesions, that cause single- and double-stranded DNA breaks, which give rise to cell cycle arrest at G2/M followed by apoptosis[89,90].

The radiotherapy regimen comprises of fractionated focal doses of 2 gray (Gy) per fraction, five days per week, over six weeks, resulting in a total dose of 60 Gy[88]. Concomitantly, patients also receive daily oral TMZ which is dosed according to the patient's body surface area (BSA) at 75mg per square meter of BSA[88,91]. Following the completion of radiotherapy, patients enter the adjuvant phase where they receive six cycles of TMZ, given at a higher dosage (150–200mg per BSA) over five days within 28-day cycles[88,91]. This combined approach has been shown to significantly extend median survival from 12.1 months to 14.6 months, compared with radiotherapy alone and has a two-year survival rate of 26.5% compared to 10.4%[88,92]. A key predictor for the effectiveness of the Stupp protocol is the methylation status of the MGMT gene promoter[93].

MGMT methylation

MGMT is a DNA repair enzyme that protects cells from DNA mismatch damage that, if left unrepaired, leads to apoptosis and cell death[93]. Specifically, it works by repairing the DNA lesions created by TMZ, conferring cells' resistance to TMZ[93]. The expression level of MGMT protein is regulated by epigenetic mechanisms; when the gene promoter region is methylated the gene is silenced and the MGMT protein levels are reduced[93]. Consequently, this reduction causes decreased DNA repair and makes cancer cells more susceptible to TMZ-induced DNA damage, thereby improving patient response to the treatment[93]. This is why MGMT promoter methylation is considered an important prognostic marker and is detailed in the WHO CNS5 scheme as it can help clinicians predict response to standard treatment[6,91].

Other treatments

Patients who have unmethylated MGMT promoters have a poorer response to the standard treatment of care and have a worse OS[92,93]. Although it has reduced effectiveness in this cohort, the recommended chemotherapy strategy for these patients still includes the *Stupp protocol*, as there are currently no better alternatives available[91]. In cases where patients are younger, combining TMZ with another alkylating agent, Lomustine (CCNU) may be an option, though this approach has higher toxicity and its efficacy remains inconclusive[94]. Likewise, at least one study has suggested that TMZ should be omitted altogether in elderly patients (>70 years of age) with unmethylated MGMT promoters as combined therapy results in a worse survival outcome[95]. There are also several novel treatments/technologies that are being developed and are available to patients willing to participate in clinical trials. These can be broadly grouped in to following categories: 1.) Molecular therapies, e.g., EGFR inhibitors; 2.) Immunotherapies, e.g., chimeric antigen receptor T cell (CAR-T) therapy and Dendritic vaccines; 3.) Gene therapies, e.g., clustered regularly interspaced short palindromic repeats (CRISPR) gene editing; 4.) tumour-treating fields (TTFs); 5.) Angiogenesis inhibitors, e.g., Bevacizumab; 6.) Stem cell therapies[96–98]. Although these approaches may show promise in future, to-date none have been superior to either a placebo or the standard treatment of care in randomised clinical trials[96].

1.4.4 Definition of GBM used in this thesis

Throughout the remainder of this thesis, the term GBM will be used to refer to the following tumour samples, unless stated otherwise:

- Grade 4, IDHwt GBM as defined by WHO CNS5.
- Methylated/unmethylated or unknown MGMT promoter status.
- Obtained from patients who received the standard treatment of care: surgical resection followed by the *Stupp protocol*.
- The first *recurrent* resected tumours.
- Only locally recurrent, i.e., *primary* and *recurrent* tumours from the same brain location.

1.5 GBM molecular subtypes

In 2008, the cancer genome atlas (TCGA) project began a large-scale, multi-dimensional analysis of over 20,000 primary cancers and matched normal samples, spanning 33 cancer types[99]. GBM was the first tumour type to be profiled: the genetic profiles of 206 GBM tumours were analysed using RNA-seq[100]. Subsequently, TCGA expanded its scope to include additional layers of cancer biology through integrative multi-omics approaches comprising DNA methylation, copy number variation (CNV), chromatin accessibility (e.g., ATAC-seq), proteomics, and immune profiling[101–104]. More recently, whole genome sequencing was also incorporated to capture non-coding alterations and identify novel cancer driver mutations[105].

In the context of GBM, this initiative identified somatic alterations that were present across 74% of tumours which affected three main pathways[100,101]: 1.) the RTK pathway; 2.) the p53 tumour suppressor pathway; 3.) The retinoblastoma (Rb) cell-cycle regulation pathway. Moreover, they highlighted the heterogeneity GBM tumours display and hinted towards the existence of distinct molecular subgroups.

Building on this work, Phillips *et al.* transcriptionally profiled high-grade gliomas and identified three distinct gene signatures that associated with differing patient survival outcomes: proneural, proliferative and mesenchymal[106]. Verhaak *et al.* later expanded this framework by identifying four molecular subtypes based on transcriptional profiling: proneural, neural, classical and mesenchymal[107]. Each subtype was linked with specific genetic alterations and distinct clinical outcomes, that helped explain differences in treatment response among GBM patients.

Subsequent studies further refined the four Verhaak *et al.* subtypes[108]. The proneural group was found to include many IDHmut GBM tumours, accounting for the younger age and had a better prognosis observed in this subtype[107,109]. Excluding these cases aligned survival outcomes with those reported for IDHwt GBM[43]. Additionally, the neural subtype was removed after it was shown to result from contamination by non-malignant neural cells, as revealed through tumour-intrinsic gene clustering[108]. Thenceforth, IDHwt GBM tumours have been classified into three molecular subtypes: proneural (PN); classical (CL) and mesenchymal (MES) (Figure 1.3A). These subtypes represent distinct biological programs, each associated with specific genetic aberrations, signalling pathways, clinical outcomes and tumour microenvironmental profiles as discussed below.

1.5.1 Proneural

PN subtype tumours feature aberrations in PDGFR and are enriched in genes associated more with neural and oligodendrocytic progenitor cell phenotypes. They have high tumour purity, low levels of immune infiltration and reduced stromal cell signals. Although, they have a more favourable survival outcome compared with MES, they can be less responsive to aggressive treatments and have been shown to switch at tumour recurrence. PN tumours typically have a higher transcriptional simplicity score, i.e., their gene patterns are more homogenous/uniform and have reduced admixture from CL or MES subtype signatures.

1.5.2 Classical

CL tumour subtypes display the Ch +7/-10 chromosomal aneuploidy signature and resemble proliferating astroglial cell phenotypes. Most CL subtype tumours amplify genes involved in RTK signalling, such as EGFR which has 4-fold higher expression. They also commonly (>70% of cases) display deletions in the CDK2A/B gene, which is a well-established tumour suppressor present across multiple cancers. These tumours have low levels of immune cell infiltration and are the most sensitized to intensive standard treatments, particularly at recurrence, resulting in a better overall patient survival.

1.5.3 Mesenchymal

MES subtype tumours are most prevalent at recurrence and have the worst prognosis, particularly in cases where tumours are more transcriptionally homogenous. They are enriched for genes associated with inflammation, wound-healing and extracellular remodelling, expressing markers such as MET and CHI3L1/YKL40. MES subtype tumours consistently exhibit low tumour purity and high immune cell infiltration, mainly from the myeloid lineage immune cells such as M2 macrophages and brain-resident microglia. Moreover, 53% of MES tumours show frequent inactivation of the neurofibromin 1 (NF1) gene. NF1 loss through inactivation results in altered cytokine production and recruitment of immune cells which promotes an immunosuppressive and pro-tumourigenic environment.

Although GBM tumours can be classified into the three discrete molecular subtypes detailed above, these labels only represent a dominant phenotype and do not capture their full biological complexity. In reality, GBM tumours are highly het-

erogeneous, comprising distinct cell populations and regions that exhibit different subtype signatures within the same tumour mass[110]. This heterogeneity partially explains the observed subtype switching that can occur during GBM disease progression, such as the transition from PN to MES at recurrence - a shift commonly associated with treatment resistance.

1.6 Glioma stem cells

Prior to the studies which molecularly profiled GBM tumours, Singh *et al.* in 2004 presented evidence for a rare population of stem-like cell, which were termed glioma stem cells (GSCs)[111]. These cells were identified *in vivo* using immunodeficient mice that were transplanted with cells expressing stemness associated cell membrane markers such as CD133, CD15, CD44, A2B5 or intracellular markers such as SOX2 and Nestin[111–113]. Singh *et al.* and others, reported that GSCs could self-renew, proliferate continuously and recapitulate a GBM tumour's original histology, suggesting that such properties enabled GBM tumours to resist therapy, repopulate and recur. It is thought that cellular heterogeneity in GBM is directly correlated with the size of the GSC population in the GBM tumour microenvironment (TME): GBM tumours with higher proportions of GSCs have a worse survival outcome[114].

Initially, the prevailing view was that GSCs behaved much like other cancer stem cells (CSCs); in that they followed a unidirectional hierarchy (Figure 1.3B) where stem cells create progenitors which then leads to more differentiated cells[115,116]. In this model reversibility is restricted between closely related progenitors and CSCs. However, Bhat *et al.*, demonstrated how PN GSCs could switch to a mesenchymal signature when exposed to the cytokine tumour necrosis factor-alpha (TNFα) that is released by macrophages[117]. Moreover, this study also mechanistically linked GSC plasticity with TME cell signals, showing how they both facilitate treatment resistance. More recently, a growing body of work supports the hypothesis that GBM cells including GSCs and non-stem tumour cells can stochastically alter their phenotypes and also revert to CSC phenotypes (Figure 1.3A)[118–120]. This plasticity has significant implications, as removing the GSCs population alone is not sufficient, and treatments must also account for the ability of differentiated cells to revert to more treatment resistant stem-like cells.

The specific mechanisms GSCs employ to resist treatment include increasing DNA damage repair, blocking apoptotic signalling, altering DNA checkpoints, expressing of adenosine triphosphate (ATP)-binding cassette transporters to increase drug

efflux and inactivating drugs by upregulating glutathione S-transferases (GSTs) detoxication enzymes[121].

1.7 GBM malignant cell states

Bulk transcriptional analyses such as those used to define the molecular subtypes above, provide an aggregated view of millions of cells within a GBM tumour, but do not adequately capture the variability between individual cells. In order to achieve this much higher resolution is required and the development of single-cell RNA sequencing (scRNA-seq) in the 2010s provided us with such a technology. Patel *et al.* was the first to profile GBM at the single-cell level, sequencing 430 cells across five patients[122]. Cells for the same patient tumours varied greatly in their expression of markers for different molecular subtypes and oncogenic signalling (EGFR, PDGFR). Moreover, the key finding from Patel *et al.* was that GBM cells reside on a continuum of stemness, hypoxia, proliferation and quiescence. This intra-tumoural transcriptional heterogeneity may explain why novel single-target treatments such as molecular therapies show such low efficacy, as subpopulations of GBM cells not reliant on that target, persist and drive tumour recurrence.

1.7.1 Neftel *et al.* (2019)

More recently, as the cost of scRNA-seq technologies has reduced and become more widely accessible, larger studies have confirmed and further expanded upon Patel *et al.*'s seminal findings. In 2019, Neftel *et al.* used an integrative approach with scRNA-seq, cell lineage tracing and combined functional assays to comprehensively profile 28 GBM tumours[123]. They identified that malignant GBM cell heterogeneity is driven by four neoplastic cellular states (Figure 1.3A): astrocyte-like (AC-like), oligodendrocyte progenitor-like (OPC-like), neural progenitor-like (NPC-like), and a mesenchymal-like (MES-like) state[123]. The NPC-like, OPC-like, and AC-like states express signatures corresponding to neurodevelopmental programs, whilst MES-like cells do not mirror any normal brain cells. Each tumour comprised a mixture of multiple neoplastic cell states, present in varying proportions, though the dominant state in a given tumour matched its bulk molecular subtype. They also confirmed Patel *et al.*'s finding of plasticity between the cell states using single-cell lineage tracing experiments in mice, further supporting the hypothesis that GSCs do not follow a hierarchical differentiation organisation. In parallel to Neftel *et al.* numerous other studies found similar gradients of GBM cell

plasticity, with some incorporating additional cellular properties such as cellular specialization, metabolism, TME and injury responses[124–132].

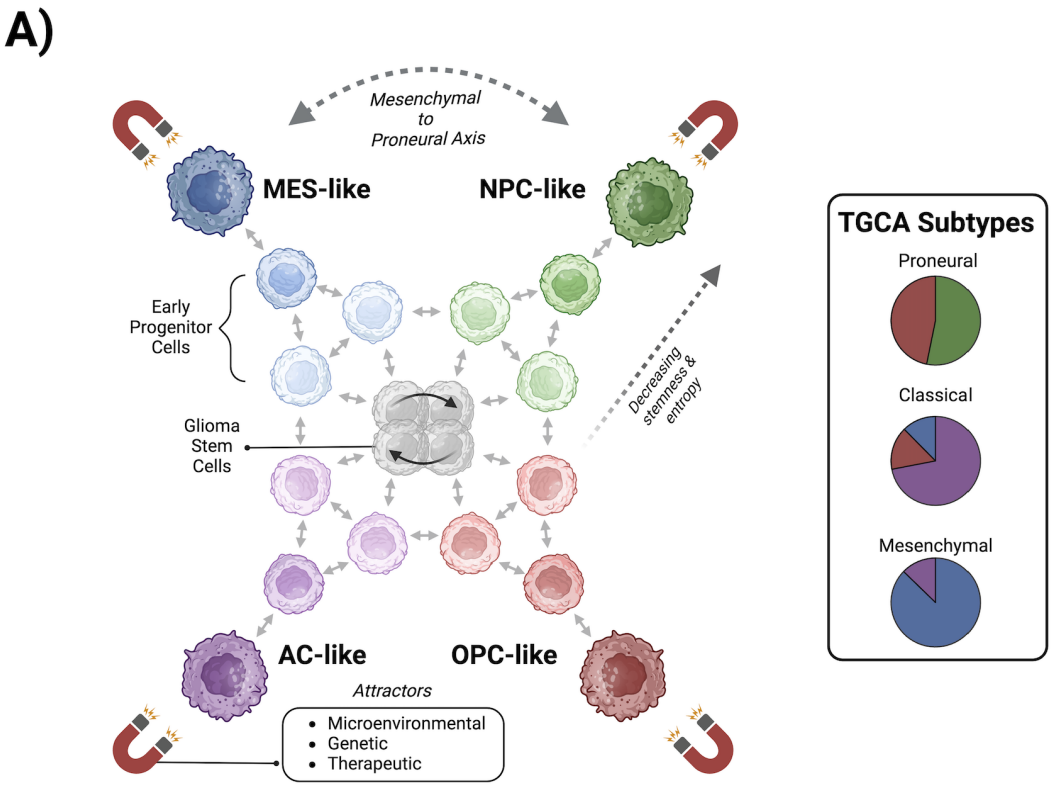
1.7.2 Wang *et al.* (2019)

While Neftel *et al.*'s four-state model is a prominent framework for understanding malignant GBM cell heterogeneity, other important and overlapping models have also been proposed. Notably, Wang *et al.* used scRNA-seq to describe GBM heterogeneity as a bi-modal continuum along a single axis of variation (Figure 1.3A), spanning from PN to MES phenotypes[124]. The PN end was marked by high expression of oligodendrocyte and neuronal genes and increased cell proliferation. Conversely, the MES end was characterised by mesenchymal markers, reduced proliferation and markers related to hypoxia and proinflammatory cytokines. A key focus of Wang *et al.*'s work was to examine the classical GBM subtype, which they found comprised of a mix of PN and MES phenotypes, enriched with AC-like cells, mesenchymal GSCs, and variable immune cell infiltration[124].

Crucially, the bi-directional PN to MES model aligns with Neftel *et al.*'s four-cell-state model: the two MES states correspond directly, and the combined NPC-like and OPC-like states align with PN. The AC-like state doesn't perfectly align with either PN or MES and is considered to be a transitional state, leaning more towards the MES.

A recent meta-analysis of over 100 single-cell and bulk transcriptomic datasets supports the view that GBM phenotypic variation occurs primarily along a PN to MES axis[133]. The authors argued that other previously defined states (AC-like and OPC-like) and subtypes (classical and neural) are better interpreted as intermediate or transitional stages along this continuum, rather than discrete entities. Notwithstanding, this evolving view reflects the dynamic and continuous nature of GBM malignant cell identity. What remains clear is that malignant GBM cell states mirror neurodevelopmental cell hierarchies and/or adopt programmes associated with wound-healing (MES-like) that promote GBM malignant cellular heterogeneity. This heterogeneity poses major challenges for treating GBM: the presence of multiple distinct malignant cell populations means that resistance and disease progression only requires the survival of one such subpopulation.

The following section will briefly discuss what is currently known about the processes and mechanisms that facilitate treatment resistance among malignant GBM cell subpopulations.


A)
Mesenchymal
to
Proneural Axis
MES-like
NPC-like
Early
Progenitor
Cells
Glioma
Stem
Cells
Decreasing
stemness &
entropy
AC-like
OPC-like
Attractors
Microenvironmental
Genetic
Therapeutic
TGCA Subtypes
Proneural
Classical
Mesenchymal

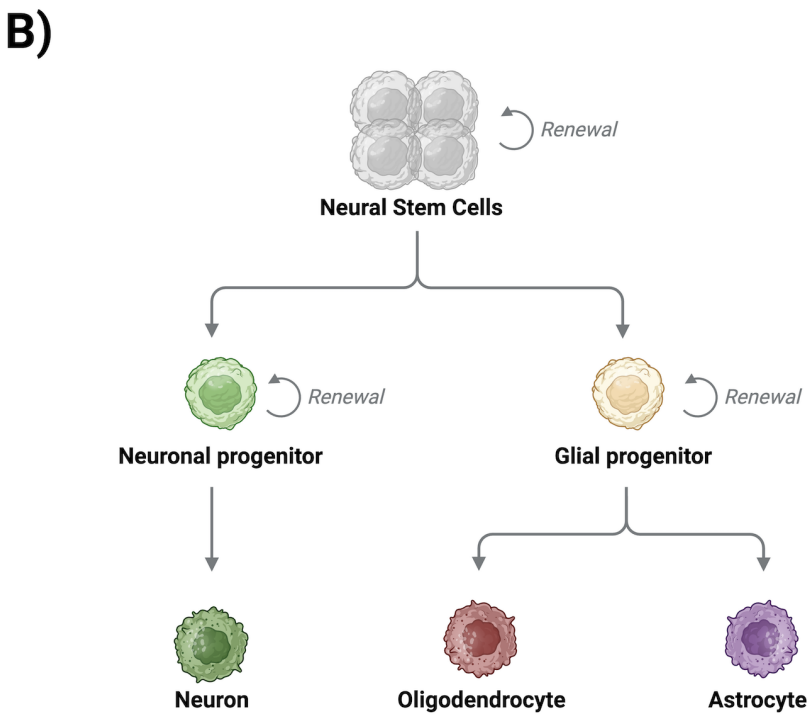

B)
Renewal
Neural Stem Cells
Renewal
Renewal
Neuronal progenitor
Glial progenitor
Neuron
Oligodendrocyte
Astrocyte

Figure 1.3 | Glioma stem cells, cell states and molecular subtypes.

A) Malignant GBM cells exist across several transcriptionally distinct states that resemble neurodevelopmental progenitors: OPC-like, NPC-like, AC-like, and MES-like cells. These neoplastic populations are highly dynamic and exhibit plasticity, transitioning between states in a non-hierarchical manner. GSCs sit at the apex of the GBM stemness/entropy hierarchy, facilitating these transitions. Each progenitor-like state is influenced by specific genetic, microenvironmental, and therapeutic cues, which act as attractors driving transitions between states. GBM tumours comprise varying proportions of neoplastic cell states, which collectively underpin TCGA molecular subtypes: Proneural are enriched in NPC-like and OPC-like; classical are dominated by AC-like; and MES are characterised by the MES-like cell state. **B)** Illustration of normal, unidirectional, neuronal cell differentiation. Progenitor cells can self-renew; however, they do not switch once committed to a lineage. Created using BioRender.

1.8 GBM treatment resistance

Treatment resistance of malignant GBM cells following standard therapy is a major therapeutic challenge against limiting GBM disease progression. It is well-established that malignant GBM cells survive standard-of-care treatment, by activating mechanisms that protect against and/or reverse the effects of standard therapy[121]. However, it is unclear whether therapeutic intervention is the driver of resistance in GBM, i.e. does it exert a selective pressure that facilitates the survival of resistant subclones in a Darwinian-like manner?

To answer this, Barthel *et al.* and Körber *et al.* independently profiled the evolutionary trajectories of paired GBM tumour samples[71,72]. Both studies found that genetic alterations present in *primary* GBM tumours, persisted and remained unchanged through treatment. These findings suggest that standard treatment exerts minimal selective pressure on GBM cell populations and that resistant cells arise from enrichment of pre-existing genetic diversity rather than newly acquired mutations. Building on this, Varn *et al.* and Hoogstrate *et al.* considered how transcriptional, rather than genetic changes, between paired GBM tumour samples contribute to treatment resistance[134,135].

Varn *et al.* found enrichment of profiles associated with neuronal signalling and mesenchymal transition within recurrent GBM cells[135]. This shift was largely driven by specific interactions within the TME, most notably, increases in myeloid cells associated with the MES-like GBM cell state. Similarly, Hoogstrate *et al.*

found that GBM tumour progression through treatment was characterised by decreased tumour purity, marked by a reduction in malignant GBM cells and an increase in normal brain cells, tumour-associated macrophages and vascular-mediated remodelling of the extracellular matrix[134]. Wang *et al.* further validated these findings using a multi-omics approach that integrated single-cell and spatial transcriptomics with proteomics to profile 86 matched GBM tumour samples[136]. Building on their previous work which identified the PN to MES axis of variation in GBM cells, they found that most GBM patients undergo a shift (shaped by treatment) from PN to MES at recurrence. This transition toward a more resistant phenotype is driven by the upregulation of transcriptional programmes associated with inflammation/invasiveness and sustained by specific immune cell interactions with GBM cells, such as tumour-associated macrophages[136].

We recently profiled 107 paired GBM tumour samples and stratified patients into Up and Down responders based on the differential dysregulation of cancer cell-intrinsic genes regulated by a chromatin remodelling complex[137]. Up responders were more quiescent, expressing proneural phenotype markers, whereas Down responders were more proliferative, expressing mesenchymal phenotype markers. These differences were demonstrated *in vitro* and *ex vivo* using GBM models and were also found to be influenced by distinct non-malignant cell populations within the GBM TME.

Collectively, all these studies indicate that GBM evolution is driven by reorganisation and adaptation of the TME with significant involvement of non-malignant cells, including immune, stromal and vascular components.

1.9 GBM tumour microenvironment

GBM tumours do not solely comprise of a homogenous mass of malignant cells, instead malignant GBM cells are heavily influenced by their surrounding TME[138,139]. The GBM TME is a highly dynamic niche which through complex interactions facilitates tumour growth, immune invasion, treatment resistance and also directly influences the frequency of cells in each neoplastic cell state[123]. In addition to malignant cells, the GBM TME is comprised of non-cancerous, normal brain cells (astrocytes, oligodendrocytes, neurons and microglia); tumour-resident and infiltrating immune cells (of myeloid and lymphoid lineages); vasculature components (pericytes and endothelial cells); and a rich milieu of extracellular components such as proteins, cytokines, and metabolic factors[138,140].

1.9.1 Non-cancerous brain cells

Astrocytes

Astrocytes are the most abundant cells within the brain (∼50% of all cells) and are critical for maintaining neuronal homeostasis[141]. They achieve this by performing varying functions depending on their phenotype and location within the brain e.g., they regulate the blood-brain barrier (BBB) through interactions with vasculature[141]. They also modulate the release of neurotransmitters between neurons and are involved in repairing damaged tissue following injury. With respect to the latter, astrocytes undergo a process known as astrogliosis where they switch to a reactive astrocyte phenotype that promotes wound-healing by upregulating growth factors and adhesion proteins[141]. In GBM, malignant cells can hijack this response and spawn tumour-associated reactive astrocytes (TARAs)[141,142]. These abnormal astrocytes facilitate GBM tumourigenesis by secreting chemokines such as transforming growth factor-beta (TGFβ) and interleukin 6 (IL-6), which is associated with a more inflammatory, MES-like cell state. TARAs are also implicated in being able to assist GBM cells boost their cell metabolism, by horizontal transfer of mitochondria via microtubules[143]. They have also been shown to able to form gap junctions between GBM cells in order to exchange toxins and small molecules to avoid programmed cell death[138,144].

Oligodendrocytes

Oligodendrocytes are another important glial cell within the brain responsible for maintaining homeostasis by mediating neural plasticity and coordinate metabolic support to neurons through axon myelination[145,146]. Myelin is a fatty-acid layer which ensheathes neurons fibres, increasing the speed and efficiency of neural electrical signals, that are essential for cognitive function[147]. Although the turnover of myelin in the brain is high, oligodendrocyte turnover is extremely low (1 in 3000 cells per annum), suggesting that this process is facilitated by mature oligodendrocytes[148]. In GBM, they are most commonly found at the tumour margins suggesting they are implicated more with GBM migration/infiltration at the resection border, as opposed to proliferation[145]. This has been determined experimentally where inhibiting cytokines released by oligodendrocytes such angiopoietin-2 resulted in reduced infiltration and increased patient survival[149]. One way they support GBM cell infiltration into the brain parenchyma is by promoting angiogenesis[150].

Neurons

Parenchymal neurons are a crucial constituent of the GBM TME as they directly interact with malignant GBM cells, exerting mitogenic effects that promote tumour growth. They do this by secreting the synaptic protein neuroligin-3 (NLGN3), and to a lesser extent the brain-derived neurotrophic factor (BDNF) which induces the PI3K/AKT/mTOR signalling pathway (PAM)[151]. The PAM pathway is highly conserved in eukaryotes and regulates cell survival, growth and proliferation by tightly modulating many additional pathways: its dysregulation is observed across multiple cancer types[152]. GBM cells have also been shown to be able to integrate into neural circuits by expressing synaptic genes and forming glutamate receptor-dependent synapses with neurons[153,154]. These synapses facilitate the transfer of electrical signals between normal neurons and GBM cells. Moreover, because GBM cells are often connected to each other via gap junctions this results in the formation of electrically coupled networks, where electrical impulses can easily propagate throughout the tumour[138,154]. This phenomenon could also explain why some GBM patients experience seizures that are typically only seen in cases with the worst prognosis[45]. The degree of connectivity between GBM cells and neurons in the GBM TME correlates with a worse patient survival and increased impairment of cognition, as demonstrated in patients when performing lexical retrieval language tasks during awake surgery[153].

1.9.2 Immune cells

The CNS (and by extension the GBM TME) has historically been considered to be immune-privileged, owing to the presence of the BBB. Consequently, brain tissue defence heavily relies on innate mechanisms mediated by normal brain cells such as neurons, astrocytes and also brain-resident immune cells such as microglia[155]. This view has now shifted with the discovery of lymphatic drainage systems such as the meningeal lymphatic compartment located under the skull and drainage to the cervical lymph node compartment[156]. These systems have been demonstrated within mouse models[157], but the relevance to GBM remains unknown: the brain parenchyma is not directly connected to these compartments, and primary GBM metastases are rare[158]. More recently, tertiary lymphoid structures (TLSs) have been found in GBM which may serve as sites for T cell priming and immune infiltration[159].

Myeloid cells

Macrophages and microglia (collectively known as glioma-associated macrophages (GAMs)) are the most abundant immune cells found within the GBM TME, comprising up to 50% of all cells[160]. Microglia are brain-resident macrophages which derive from yolk sac precursors during embryonic development and self-renew in the CNS[160]. In contrast, macrophages are derived from bone-marrow progenitors that migrate to the CNS as monocytes and then differentiate in response to local stimuli[160]. Microglia are found located within the brain parenchyma, whereas macrophages are found near blood vessels and in the brain meninges – the outer layer between the brain and the inner skull bone[161]. Further, in GBM, microglial GAMs are found in the tumour margins and infiltrating zones, whilst macrophage GAMs are more prominent in tumour recurrence and areas of necrosis where there is high hypoxia[162]. Phenotypically, GAMs are heterogeneous, sitting on continuum between two functional extremes: a pro-inflammatory (M1) state and an immunosuppressive, tumour-promoting (M2) state[162,163]. They also exhibit a large degree of plasticity being able to move between M1 and M2 phenotypes in response to external stimuli.

M1 GAMs directly kill tumour cells by phagocytosis and induce the adaptive immune response by presenting antigens to T cells and also releasing cytokines such as TNFα, and IL-6[164,165]. M2 GAMs promote GBM tumourigenesis by releasing cytokines such as TGFβ, vascular endothelial growth factor (VEGF) which promotes angiogenesis[166]. They also release matrix metalloproteinases (MMPs) such as MMP2 and MMP9 which degrade and remodel the GBM extracellular matrix, hampering cell-cell communication and creating favourable conditions for tumour infiltration[167]. Hypoxia induced M2 GAMs release VEGF which can promote epithelial-to-mesenchymal transition (EMT) programs in GBM cells[168]. Similarly, M2 GAMs can also directly induce the MES-like cell state in GBM cells by binding to the oncostatin M receptor (OSM)[132]. Immunosuppression by M2 GAMs is achieved by releasing interleukin 11 (IL-11) to recruit additional GAMs which can then inhibit the activation of T cells, natural killer cells (NK cells) and dendritic cells (DCs)[169]. The paracrine mechanisms employed by M2 GAMs that create an immunosuppressive TME can equally be initiated by GBM cells[160].

DCs are another important myeloid lineage cell within the GBM TME. DCs are potent antigen-presenting cells which capture GBM tumour antigens and co-ordinate the adaptative immune response by activating T cells and B cells[170]. Although they comprise a very small proportion of all cells within the GBM TME, they act as an important bridge between the innate and adaptive immune response[171].

Given this fact, a group of emerging treatments have attempted utilise DCs to promote a tumour-specific immune response through effective tumour antigen presentation in the form of dendritic cell vaccines (DCVs)[172]. The purported benefit being that DCV generate cytotoxic anti-tumour responses against multiple GBM tumour antigens compared with other therapies which are more antigen specific such as CAR-T therapies[171,172].

Lymphoid cells

NK cells are lymphoid cells that form part of the innate immune response and can recognise GBM tumour antigens without any prior sensitisation, triggering a cytotoxic cascade which kills tumour cells[173]. They are located within the perivascular niches and tumour margins in the GBM TME, though can migrate into the brain parenchyma when the BBB is compromised and more permeable[173]. *In vitro* studies have shown that healthy NK cells can directly kill GBM tumour cells by releasing granzymes, perforins and also indirectly, by releasing pro-inflammatory cytokines such as interferon-gamma (IFNγ)[174]. To counteract this, GBM cells within the TME release immunosuppressive cytokines such as TGFβ which directly inhibits NK cell receptors, resulting in them displaying an "exhausted" phenotype[175]. Most, nucleated, healthy cells express major histocompatibility complex (MHC) Class I proteins on their surface which NK cells use as an inhibitory ligand to identify "self" cells[176]. GBM tumour cells, which either lose the ability to, or which lowly express MHC class I, exploit this process by expressing a different inhibitory ligand HLA-E[176,177]. This results in NK cells being present in the GBM TME but often in non-functional or inhibited states, which may present a therapeutic opportunity[175].

Other lymphocytes within the GBM TME include B cells and T cells, which form part of the adaptive immune response[178]. Although they have not received as much attention as myeloid lineage cells or T cells, B cells play an important role within the GBM TME[179]. Their primary role involves differentiating into mature Plasma B cells, which secrete antibodies in response to GBM antigens and also antigen-presenting to T cells, via MHC class II surface proteins[180]. However, they can also undermine immunity by secreting immunosuppressive cytokines such as interleukin 10 (IL-10) and TGFβ by acquiring a regulatory B cell (Breg) phenotype through interactions with GAMs[181]. Given these dual roles, their effects on GBM tumourigenesis are highly location- and cell interaction-dependent.

T cells are the main effector cells involved in the adaptative immune response within GBM tumours. They can both, promote and suppress tumour progres-

sion, which is partly attributed to presence of different T cell subpopulations that collectively comprise ∼1-5% of all cells within the GBM TME[182,183]. The main subsets of T cells that are widely discussed in relation to GBM tumours include $CD8^+$ cytotoxic T lymphocytes (CTLs); $CD4^+$ T helper cells (Ths); and $CD4^+$ regulatory T cells (Tregs)[182].

$CD8^+$ CTLs are effector immune cells which recognise and lyse MHC class I, tumour antigens that are expressed on antigen-presenting cells such as DC or GAMs[184]. $CD4^+$ Ths are much less prevalent than CTLs and are found located within the perivascular niches[183]. These cells are called "helper" cells as they can induce $CD8^+$ CTLs activity by releasing cytokines and also are involved in promoting B cell activation and proliferation[183]. $CD4^+$ Tregs are defined by the additional expression of forkhead box P3 $(FOXP3)^+$ and are tumour-promoting cells and associated with immunosuppression and a worse patient prognosis[185]. Moreover, they are also specifically associated with hypoxic conditions and an ability to suppress CTLs[186]. T cells within the GBM TME exist in a number of dysfunctional cell states: tolerance, ignorance, anergy and exhaustion[183,187]:

Tolerance - T cells fail to respond to antigen-presentation signals either due to a lack of a positive co-stimulatory signal or the presence of inhibitory signals[188].

Ignorance - fully functional T cells fail to elicit an immune response either due to GBM tumour antigens being are anatomically separated and therefore, not easily accessible, or because GBM cells have downregulated their tumour antigens[189].

Anergy – T cell activation requires both an antigen-specific and non-antigen specific signal (co-stimulation); when the latter is missing, T cells become unresponsive[190]. GBM cells contribute to this by reducing the expression of co-stimulatory molecules on antigen-presenting cells[189].

Exhaustion – a process where naïve T cells are exposed to a chronically high antigen loads such as in GBM tumours, causing to them enter into a reversible, evolutionarily conserved (to avoid causing damage to normal tissue) hypo-functional state[190,191]. When in this state T cells upregulate the expression of inhibitory immune checkpoint molecules, such as programmed cell death protein 1 (PD-1)[187]. GBM cells also express these same inhibitory immune checkpoint molecules, inducing T cell exhaustion and dampening the cytotoxic effect of $CD8^+$ CTLs[191].

1.9.3 Vasculature components

The BBB is a distinguishing feature of brain tissue. In healthy individuals it provides a highly selective barrier that regulates cerebral blood circulation into the brain parenchyma, protecting it from infectious agents and toxic substances, including many therapeutic agents[192,193]. The BBB is made up of a complex network of arteries, arterioles, capillaries, venules, and veins that are organised into neurovascular units (NVUs)[193]. The specific cells which comprise the NVUs include endothelial cells, mural cells and astrocytic foot processes. Additionally, microglia are also thought to interact with the NVUs in order to regulate and repair the BBB in response to injury[192,193].

Endothelial cells make up the vessel lumens that form the BBB and are functionally and morphologically distinct compared with other endothelial cells in the body. Specifically, they are sealed (with endothelial cells) via tight, adheren and gap junctions, and also express low levels of leukocyte antigens that limit immune cells from entering the brain parenchyma[194,195].

Mural cells include vascular smooth muscle cells and pericytes that are both located on the abluminal surface of endothelial cells[192]. The smooth muscle cells are found within larger arteries and veins and are separated from the endothelium by an extracellular matrix layer[192]. In contrast, pericytes are directly embedded between endothelial cells and astrocyte end feet projections[196].

1.9.4 Microenvironmental heterogeneity

The cells within the GBM TME are compartmentalised into three anatomically distinct regions or niches (Figure 1.4): perivascular region (PVR), hypoxic tumour core and the invasive edge[178,197]. Each of these regions facilitate GBM tumourigenesis in different ways and favour specific cell interactions between GBM, stromal and immune cells, with vasculature being a central player across all three regions.

Perivascular region (PVR)

GBM tumours feature abnormal angiogenesis, mainly caused by the increased expression of VEGF[198]. Neoplastic GBM cells ($CD133^+$) actively express VEGF and other angiogenic factors which cause the detachment of pericytes from the basement membrane of the BBB, by destroying tight junction proteins[199]. This results in enlarged, "leaky" blood vessels that present as glomeruloid structures

known as MVP; which is a hallmark of GBM[39]. The compromised BBB allows for circulating immune cells such as monocytes to enter into the brain parenchyma and differentiate into more M2-like GAMs[200]. These activated GAMs further promote tumour progression by releasing cytokines that increase vascular permeability, inducing cerebral edema[201]; hamper the T cell response by causing an exhausted state[202]; facilitating neoplastic GBM cell proliferation and resistance to standard treatment[203]; and by recruiting additional myeloid lineage cells[204].

Hypoxic (necrotic) tumour core

A compromised BBB leads to the inconsistent flow of oxygen around the GBM tumour. Initially, this causes regions of hypoxia to develop, but later this also results in areas of cell necrosis[197]. In response to this, GBM cells elongate their nuclei and align into rows around the centres of necrosis, termed pseudopalisading necrosis[178,197]. The exact reason as to why GBM cells organise into palisade-like structures is not entirely clear and may include edema-enhanced vessel collapse; vascular regression, or microscopic intravascular thrombosis driven by overexpression of pro-coagulants[205,206]. These regions are regulated by the increased expression of hypoxia-inducible factor 1-alpha (HIF1α) which, promotes GBM cell stemness; causes the release of immunosuppressive cytokines; increases resistance to therapy; and overall, is associated with more aggressive GBM tumours[206,207].

High hypoxia, necrotic regions are also typically associated with immunosuppressed immune cells such as macrophages and T cells[178]. It is suggested that these cells are initially attracted by proinflammatory signals caused by necrotic cell death and then later sequestered by a combination of being trapped (surrounded by pseudopalisading cells) and/or via upregulation of HIF1α[208]. More recently, several studies have found that the MES cell state comprises increased hypoxic regions, and is associated with macrophages which induce the neoplastic state in GBM[132,208].

Invasive (infiltrating) edge

Diffuse tumour infiltration into the healthy brain parenchyma is a distinguishing feature of GBM and makes maximal surgical resection an impractical task[6]. GBM cells do not metastasise via lymphatic systems or through the vasculature, instead they spread by displacing astrocytes that are associated with blood vessels[209]. The invasive edge denotes regions located at the tumour periphery or margins. As GBM tumours progress, this niche becomes incorporated into the main tumour

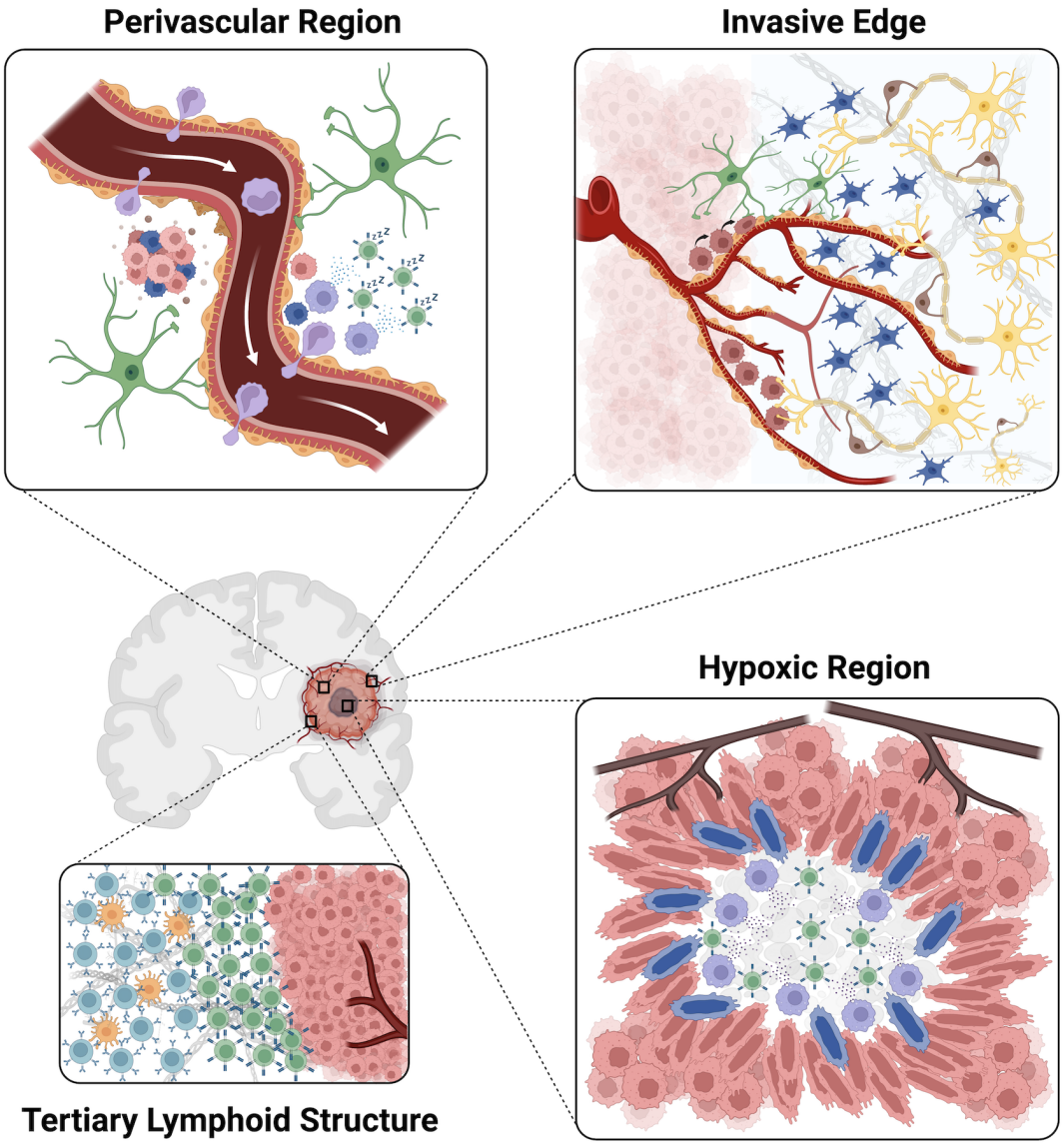
mass and subsequently transitions, first into the PVR and then finally into hypoxia tumour core niches[210].

GBM cells at the invasive edge integrate into neural circuits by enhancing synapse-associated gene expression and directly forming synapses with healthy neurons[151,153,154]. Astrocytes are also important stromal cells located at the invasive edge as GBM cells link with them via tumour microtubules to transfer molecules and organelles to facilitate increased metabolic processes[143,144,154]. The invasive edge also comprises increased numbers of reactive astrocytes which produce growth factors, metabolites and cytokines as part of a wound healing response that increases gliomagenesis[131,211]. Compared with the PVR and tumour core regions, the invasive edge is also enriched in microglia which promote invasion by releasing matrix metalloproteinases that degrade the extracellular matrix within the healthy brain parenchyma[212].

1.9.5 Spatial heterogeneity

Recently, Greenwald *et al.* applied spatial transcriptomics and proteomics to further characterise the known anatomical regions of GBM at the single-cell resolution[213]. The main focus of this work was the *in situ* spatial mapping of the malignant GBM cellular sates that were previously identified by Neftel *et al.*[123]. Their key finding was that GBM tumours are characterised by a combination of spatially structured and disorganised regions, with the former being particularly prominent in high-grade gliomas such as IDHwt GBM. Within these structured regions they identified a higher-order organisation comprising of five repeating and adjacent layers, extending from hypoxic cores outward to areas resembling normal brain parenchyma. The layers are maintained by recurring pairwise interactions between specific malignant cell states and non-malignant cell types and primarily driven by hypoxia, with less hypoxic areas showing reduced structural organisation.

The five layers proposed by Greenwald *et al.* each aligned with the classical anatomical regions found in GBM (Figure 1.4), but extended this understanding by adding a molecular and cellular framework to explain how these histological features arise. Layers 1 and 2 correspond to the hypoxic, necrotic tumour core and the surrounding hypoxia-adjacent zone, respectively. Layer 3 aligns with the PVR and is adjacent to Layer 4 which represents the neurodevelopmental malignant cell states (OPC-like, AC-like and NPC-like). Layer 5 aligns with the invasive (infiltrating) edge and is adjacent to the normal brain parenchyma.


Perivascular Region
Invasive Edge
Hypoxic Region
Tertiary Lymphoid Structure

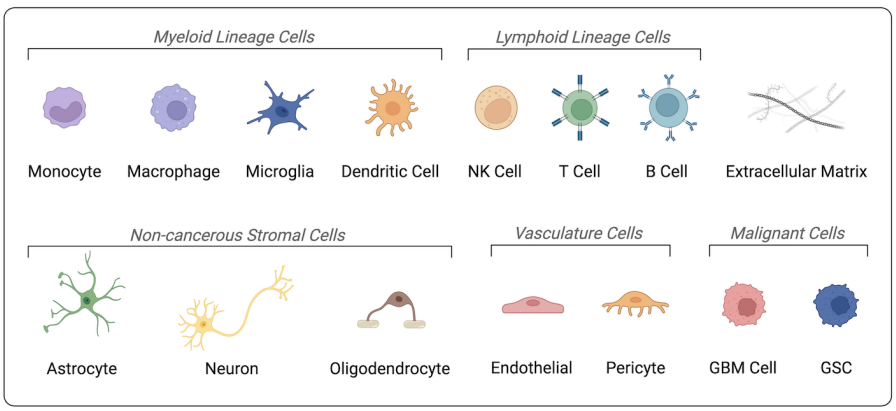

Myeloid Lineage Cells
Lymphoid Lineage Cells
Monocyte
Macrophage
Microglia
Dendritic Cell
NK Cell
T Cell
B Cell
Extracellular Matrix
Non-cancerous Stromal Cells
Vasculature Cells
Malignant Cells
Astrocyte
Neuron
Oligodendrocyte
Endothelial
Pericyte
GBM Cell
GSC

Figure 1.4 | GBM TME cellular heterogeneity.

Illustration depicting the three main anatomical regions found within GBM tumours: PVR, hypoxic tumour core and the invasive edge. The PVR features a compromised BBB and immune cell infiltration of monocytes which differentiate into macrophages. The activated macrophages release cytokines which suppress T cells in the PVR. GSCs within the PVR release cytokines, breaking down vasculature cells to promote tumour growth and therapy resistance. The hypoxic tumour core comprises a pseudopalisading rim and the necrotic core. The rim is associated with suppressed macrophages and T cells and an increased presence of MES-like neoplastic GBM cells. The invasive edge is the interface between the tumour margin and the brain parenchyma. The GBM cells migrate along the blood vessels displacing astrocytes, degrading the healthy brain extracellular matrix components and forming direct connections with neurons. This region is also associated with a large proportion of brain-resident-microglia and astrocytes which exhibit a more reactive phenotype. Also depicted are TLSs, which are thought to serve as localized sites of immune cells which facilitate T cell priming and immune cell infiltration into the GBM TME.

1.10 Computational investigation of the GBM TME

Several genomic techniques and computational tools exist to allow researchers to study the transcriptome of GBM tumours[214]. The most widely used approach in this space, involves quantifying ribonucleic acid (RNA) abundances (RNA sequencing (RNA-seq)) for differential gene expression analysis across different tissues and conditions (e.g., normal vs diseased tissue)[215]. More recent advancements have also integrated higher-resolution, single-cell information (including spatio-temporal information), enabling us to gain much deeper insights into the different layers of heterogeneity present within the GBM TME[216].

The work presented in this thesis is focussed on RNA-seq data as this represented the primary form of molecular profiling available for our paired GBM tumour samples. Notwithstanding, it is also important to acknowledge that other genomic sequencing technologies, including whole-exome and whole-genome sequencing which consider DNA rather than RNA, offer complementary insights into the mutational and structural alterations underpinning GBM tumour biology.

1.10.1 RNA sequencing

Bulk RNA sequencing

Bulk RNA-seq combines the gene expression profiles of all cells within a tumour sample. This approach has a number of advantages over other RNA-seq techniques: 1.) It is well-established, making it cost-effective across large samples sizes; 2.) There are many, well-supported methods/tools available for downstream analyses such as read alignment and abundance quantification; 3.) It has good read depth range between 5 million – 200 million reads per sample, allowing better detection of low-abundance transcripts[214,215,217]. However, the main limitation of this method is that averaging signals within entire tissue samples makes it difficult to differentiate signals in heterogenous tissues such as GBM[218]. This averaging can also obscure biological variability between samples by masking rare cell population signals such as neoplastic GBM cells or infiltrating immune cells, in favour of more dominant cell types[218].

Single-cell RNA sequencing

scRNA-seq overcomes the main limitation of bulk RNA-seq by quantifying transcripts across individual cells within a sample. This enables better identification of rare cell subpopulations and gene patterns associated with specific cell functions and/or phenotypes. Moreover, this technique also enables lineage tracing of individual cells providing information about cell development and differentiation processes[218]. In order to profile individual cells, they need to be dissociated from the tissue which is a significant technical limitation of this approach as it induces stress and can result in the loss of more fragile cells such as neurons[214]. The technique is also expensive and labour-intensive, making it currently impractical for large-scale studies[219]. There is also the increased prevalence of "dropout" events with scRNA-seq caused by low amounts of RNA in individual cells, inefficient RNA capture, and the stochasticity of RNA expression[219].

Single-nucleus RNA sequencing

Single-nucleus RNA sequencing (snRNA-seq) is a variant of scRNA-seq which attempts to overcome the limitation of cell dissociation by only isolating and sequencing the nucleus of a cell[220]. This is particularly useful in GBM samples where enzymatic dissociation is challenging due to the presence of large fragile

cells that are often tightly connected[221]. There is also no requirement for live cells making it ideal for analysing archived formalin-fixed paraffin-embedded (FFPE) and frozen GBM samples. This reduces the dissociation-inducing cell stresses, resulting in a more in situ representation of gene expression[222]. The limitation of this technique is that it only captures nascent RNA transcripts; while this makes it useful for understanding gene regulation processes, it does neglect cytoplasmic RNA signals[220]. Consequently, it captures fewer overall transcripts compared to scRNA-seq and also results in signals that comprise more intronic regions[223].

1.10.2 Cell deconvolution

Deconvolution refers to a computational approach used to separate signals corresponding to different sources of variations within RNA-seq data[224]. Whilst this approach can be used in the removal of unwanted sources of variation, known as batch-effect correction[225], we will only focus on its use for delineating biological signals corresponding to cell phenotypes. In the context of the GBM TME, deconvolution can tell us what fraction of a tumour sample is composed of immune cells, stromal cells and neoplastic GBM cells using only the bulk gene expression data. Being able to obtain extra resolution from bulk RNA-seq, *in silico*, is an attractive proposition, not least because the alternative single-cell based technologies are prohibitive towards this end, owing to their labour-intensiveness and high cost[224]. Although, there are a plethora of different deconvolution tools available to users (∼50 in the past 15 years), these broadly fall into two distinct categories: supervised (reference-based) and unsupervised (reference-free) tool[226].

Supervised (reference-based) approaches

Supervised cellular deconvolution methods rely on *a priori* knowledge, supplied in the form of a non-negative expression matrix that contains the gene expression profiles of the cell-types of interest[224,226,227]. These reference profiles can be derived from purified cells, sorted populations, or scRNA-seq data. Most supervised deconvolution methods use either linear regression or machine learning algorithms (e.g., support vector regression) to determine the combination of reference profiles that best "fit" the bulk RNA-seq data[227]. CIBERSORT and CIBERSORTx are two popular examples of supervised approaches, though many other techniques also exist[228,229]. There are also a smaller number of supervised methods such as MuSiC which employ probabilistic models to estimate cell type proportions by assessing the similarity between bulk and reference profiles[230]. The accuracy of

all supervised methods depends greatly on the quality of the reference data supplied[231]. Poor references that contain cell types which are missing in the bulk sample, can lead to miss-assignment or omission of cell populations[231].

Unsupervised (reference-free) approaches

Unsupervised cellular deconvolution methods do not require a pre-defined reference; instead, they treat deconvolution as a blind source separation problem and attempt to infer cell type proportions directly from bulk RNA-seq data[224,226,231]. These methods rely on non-negative matrix factorisation (NMF) to decompose bulk RNA-seq matrices into two smaller matrices that represent cell type expression profiles and their proportions[224]. A popular example of an unsupervised deconvolution method is DECODER[232]. Although these methods offer the convenience of not needing a reference they tend to underperform in benchmarking[226]. This may be because they rely entirely on mathematical decomposition which can produce solutions that lack clear biological interpretation. Moreover, some unsupervised methods assume that cell types are not correlated, which is not often the case with some cell types (e.g. immune cells) that share common expression profiles[226,227].

Semi-supervised approaches

Semi-supervised cell deconvolution methods offer a blended approach where they use partial a priori knowledge in the form of cell-type specific marker to estimate proportions from bulk RNA-seq data[231,233]. These methods also utilise both NMF and linear regression algorithms, but they constrain the decomposition solution space using the supplied cell-type markers[233]. A popular example of an unsupervised deconvolution method is MCPCounter[234]. Although this approach offers the advantage of being able to detect cell populations that may not be as well characterised, its accuracy relies on marker completeness and specificity to the bulk RNA-seq data[224,226,231,233].

1.10.3 Spatial transcriptomics

Spatial transcriptomics is an emerging field which combines sequencing and imaging in order profile RNA, whilst preserving spatial information[235]. This is particularly useful in the context of GBM, as it can provide a multidimensional view of the TME by combining cell phenotypes with neighbouring interactions[216]. Fur-

ther, this information can be used to better understand how cell-cell interactions change throughout the different anatomical regions within the GBM TME and also through treatment. This is a rapidly changing field, with new technologies being developed all the time. The current methods can be grouped according to how they capture spatial information: array based (solid-phase) methods and *in situ* (comprising RNA-seq and fluorescence-based imaging) methods[216].

Array-based methods

These methods use a solid surface such as a glass slide comprising millions of barcoded oligonucleotides[236]. Tissue sections are placed directly onto these slides and permeabilised to remove the RNA which is then reverse-transcribed, sequenced and then aligned back to original slide using the positional barcodes. The 10X Genomics Visium platform is a popular example of this method, though it differs in that glass slides are organised into arrays made up of 5000 spots (each $\sim$55μm in diameter and spaced 100μm apart) to capture the transcripts from multiple cells per spot[216,236]. These methods are high-throughput and can be used with both FFPE and frozen samples. One major limitation with these methods is that older approaches do not offer single-cell resolution: it is limited to the size of each spot which itself often comprises multiple cells[216]. Notwithstanding, some newer technologies have addressed this by using more densely packed spots to capture near single-cell resolution[237,238].

In situ methods

GeoMX and CosMX digital spatial profiling, developed by Nanostring, are both examples of selective barcoding *in situ* methods, offering whole transcriptome and single-cell resolution, respectively[216]. They differ from solid-phase methods by using fluorescently-labelled RNA probes that are cleaved, in situ (using ultraviolet light) and then sequenced over multiple cycles[239]. One limitation of this technique is that currently, CosMx uses a pre-defined 6000-gene marker panel which may not accurately capture specific cell phenotypes and states.

In situ hybridization (ISH) is an established technique that is routinely used in clinical practice for diagnosing and classifying GBM tumours[6,216]. It involves the use of radioactive-, fluorescent- or colorimetric-labelled nucleic acid probes to visualise DNA and RNA[240]. The advantage of this method is that it offers high spatial resolution (sub-cellular, nanometer scale), but with much lower throughput and an higher technical demand due to it being probe dependent[216].

1.10.4 Spatial proteomics

IHC remains a cornerstone of clinical pathology for disease diagnosis and monitoring[241], but its limited capacity for multiplexing restricts its utility in complex environments like the GBM TME[242]. Spatial proteomics overcomes this limitation by enabling the simultaneous mapping of dozens of proteins across whole tissue sections[243]. Unlike spatial transcriptomics, it offers a more stable view of cell phenotypes, as protein expression is less stochastic than RNA expression and can capture features such as post-translational modifications[242–244]. Spatial proteomics is a rapidly evolving field, with current technologies broadly divided into mass spectrometry–based and multiplexed fluorescence–based approaches[216].

Mass spectrometry-based methods

These methods extend cytometry by time-of-flight (CyTOF) technology by quantifying antibodies (up to 40) for proteins of interest that have been tagged with heavy metal isotopes[245]. The antibodies are used to stain tissue sections, which are ablated pixel by pixel ($\sim 1\mu$m) and the vaporized tissue is analysed by CyTOF. The resulting multiplexed images feature pixels where each one has multiple value corresponding to abundances of the proteins used. The benefits of this approach are that it can be applied retrospectively, using archived tissues in FFPE and also there is very little signal contamination compared with fluorescence-based techniques[245]. Imaging mass cytometry (IMC) and multiplexed ion beam imaging (MIBI) are prominent examples of mass-spectrometry-based spatial proteomics[242].

Multiplexed fluorescence-based methods

These methods extend immunofluorescence microscopy to detect up to 50 proteins by iteratively staining tissue sections in cycles[242].
Co-detection by indexing (CODEX) is one such popular example that utilises fluorescently labelled DNA-barcoded antibodies[246]. These methods offer high-sensitivity and high-resolution (sub-cellular) whilst also being non-destructive, unlike mass-cytometry based methods, allowing additional assays and H&E staining to be performed[242]. However, the fluorophores used to identify proteins of interest can generate background signals (autofluorescence) – an effect that can be compounded over multiple imaging cycles[242].

1.11 Hypothesis

Malignant GBM cells exhibit high plasticity, transitioning between distinct transcriptional states in response to standard therapy. These neoplastic cell states align with neurodevelopmental hierarchies and wound-healing programs, and are shaped by the non-malignant components of the GBM TME, comprising immune, stromal and vasculature cells. Furthermore, the GBM TME is organised into sub-anatomical regions and structured layers, each with unique cellular compositions and interactions that sustain and promote gliomagenesis. To understand and ultimately target treatment resistance in GBM, it is essential to characterise how cell type proportions, interactions, and dynamics within the TME change through treatment.

1.12 Aims & Objectives

In order to effectively investigate treatment-associated changes in the GBM cellular landscape requires longitudinal samples across multiple patients at scale. Currently, the largest dataset we hold that meets this criterion is at bulk RNA-seq level, generated from a large cohort of paired (primary and recurrent) GBM patient samples. However, these data lack the resolution provided by single-cell approaches, making it difficult to accurately resolve the cellular landscape of GBM.

To overcome this, I will apply cellular deconvolution techniques to our bulk RNA-seq data in order to estimate the cellular composition of each tumour sample. I will first develop a GBM-specific deconvolution tool tailored to this purpose (Chapter 2). Following this, I will deconvolute our paired, bulk RNA-seq GBM patient samples and examine how the cellular landscape changes through treatment and how these changes are associated with clinical outcomes or treatment response (Chapter 3). Finally, I will validate the findings using spatial proteomics to gain higher-resolution insights into GBM TME cell-cell interactions (Chapter 4).

The chapter-specific aims & objectives are detailed below and in Figure 1.5. In addition, details of my individual contributions, including the extent of any inherited work, are also provided.

1.12.1 Chapter 2 - aims & objectives

Aim

Derive a GBM tissue-specific reference signature for cell deconvolution.

Objectives

1. Collate publicly available scRNA-seq data from brain, GBM and other high-grade glioma tissues.
2. Integrate datasets, annotate distinct immune cell types.
3. Derive a reference signature matrix and marker list comprising markers that delineate each immune cell type.
4. Benchmark cell deconvolution methods and select the best-performing method for inferring cell types using bulk RNA-seq data.

Contribution & prior work

The data used in this chapter were obtained from publicly available sources prior to the start of my work. My contributions began with the quality control and standardisation of the datasets, followed by the identification of tissue-specific cell type markers. This included the incorporation of neoplastic cell type markers previously defined by Neftel *et al.*[123]. The cell deconvolution methods employed were established and published approaches, with their core methodologies left unchanged. However, the input components - specifically, the marker gene panel and single-cell expression reference matrix - were curated and expanded to better reflect GBM tissue-specific profiles. The optimal combination of method and input was then implemented as a web application (GBMDeconvoluteR), which I fully developed and deployed.

Additionally, I performed all analyses presented in the published manuscript associated with this chapter, including orthogonal validation using IMC. Further details of my contribution to the IMC experimental work can be found in subsection 1.12.3.

1.12.2 Chapter 3 - aims & objectives

Aim

Deconvolute bulk RNA-seq data from paired GBM patient samples.

Objectives

1. Obtain cell-type proportion estimates for all paired GBM tumour samples using the best-performing cell deconvolution method (from chapter 2).
2. Investigate how cell-type proportions change through treatment and which specific cell types are correlated with changes through treatment.
3. Associate cell type changes with patient survival outcomes.
4. Identify if there any prognostic and predictive features based on changes in cell-type scores through treatment using supervised and unsupervised approaches.

Contribution & prior work

The paired GBM patient tumour samples analysed in this study were either generated in-house or obtained as raw data through international data-sharing agreements, with contributions from institutions in the USA, Europe, and Korea. Prior to my involvement, the raw read counts had already been pre-processed, aligned, and normalised for sequencing depth.

My contributions began with the downstream analyses, which included quality control, cell type deconvolution using GBMDeconvoluteR, and all additional analyses presented in the accompanying draft manuscript. I also analysed publicly available, pre-processed spatial proteomics data (CODEX) to validate the key findings related to oligodendrocyte and B cell populations. In addition, I also performed the orthogonal experimental validation using IMC; further details of this can be found in subsection 1.12.3.

Experimental validation and analysis using the 10X Visium spatial transcriptomics platform were carried out by Sophie Williams, who is listed as a named author on the draft manuscript associated with this chapter.

1.12.3 Chapter 4 - aims & objectives

Aim

Characterise the GBM TME cell-cell interactions at the spatial resolution.

Objectives

1. Identify protein targets that delineate key cell types and cell states within the GBM TME.
2. Analyse spatial proteomics data from IMC applied to paired GBM tumour patient samples.
3. Identify and compare cell-types interactions, pre- and post-treatment.
4. Identify spatial organisational patterns such as cell neighbourhoods and look at how these change through treatment.

Contribution & prior work

The selection of tumour sample regions of interest, along with all initial experimental procedures including antibody validation, tissue preparation, and execution of the IMC workflow was carried out by the named authors listed in the manuscript. Similarly, all histopathological assessments of the patient tumour samples were also performed by the named authors.

My contributions began upon receipt of the raw IMC data, that were provided as multi-stack TIFF image files. From that point onward, I conducted all subsequent analyses described in the accompanying manuscript, including cell segmentation, phenotyping, and downstream analyses such as examining cell–cell interactions and spatial neighbourhood structures.

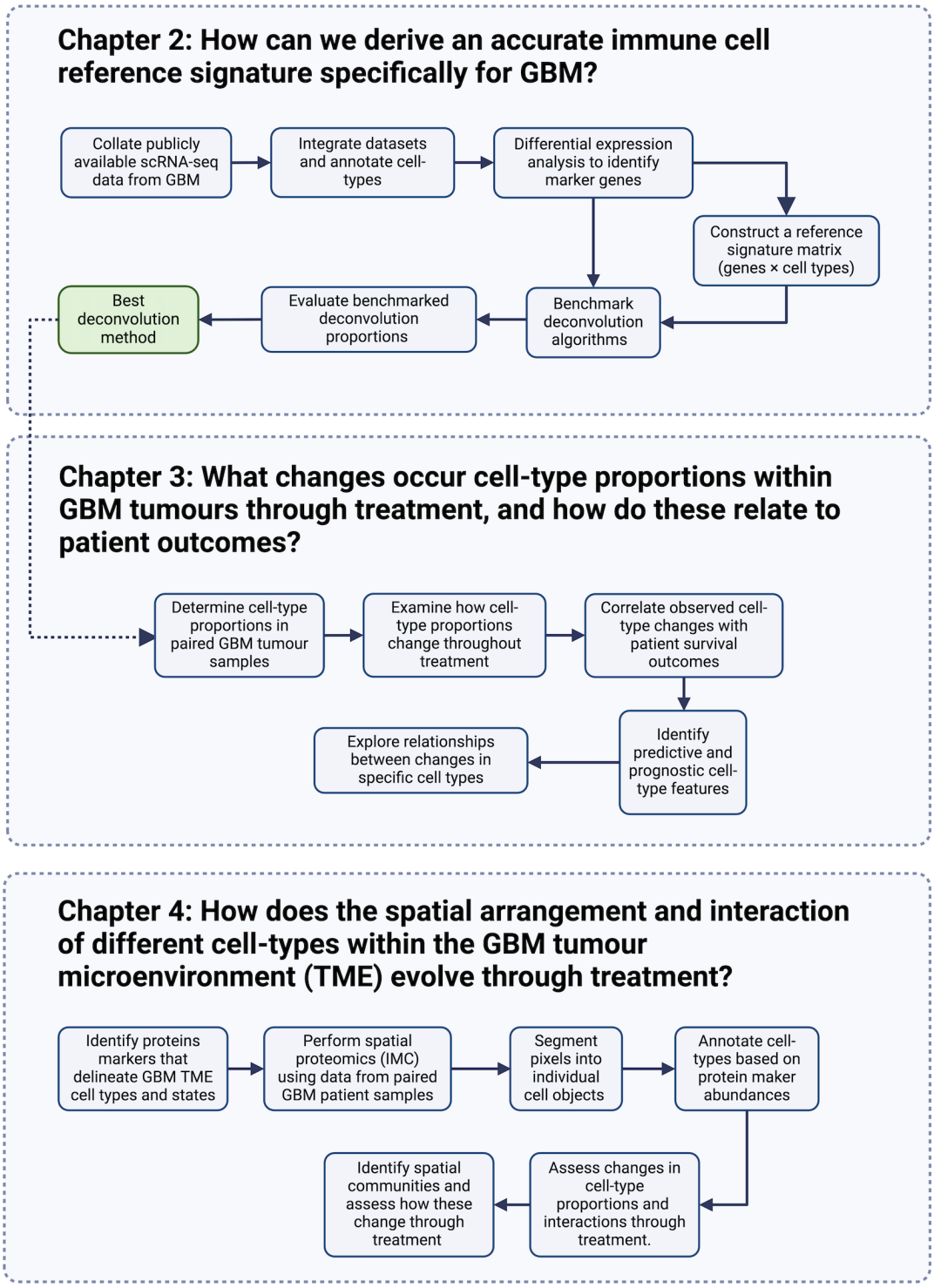


Figure 1.5 | Flowchart of chapter-specific aims & objectives.

References

1. Price, M. *et al.* CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2017–2021. *Neuro-Oncology* **26**, vi1–vi85 (Oct. 2024). doi:10.1093/neuonc/noae145.
2. Xiao, D. *et al.* National Brain Tumour Registry of China (NBTRC) statistical report of primary brain tumours diagnosed in China in years 2019–2020. *The Lancet Regional Health: Western Pacific* **34**, 100715 (Feb. 2023). doi:10.1016/j.lanwpc.2023.100715.
3. Wen, P. Y. *et al.* Glioblastoma in adults: a Society for Neuro-Oncology (SNO) and European Society of Neuro-Oncology (EANO) consensus review on current management and future directions. *Neuro-Oncology* **22**, 1073–1113 (Aug. 2020). doi:10.1093/neuonc/noaa106.
4. Oronsky, B., Reid, T. R., Oronsky, A., Sandhu, N. & Knox, S. J. A Review of Newly Diagnosed Glioblastoma. *Frontiers in Oncology* **10**, 574012 (Feb. 2021). doi:10.3389/fonc.2020.574012.
5. Tan, A. C. *et al.* Management of glioblastoma: State of the art and future directions. *CA: A Cancer Journal for Clinicians* **70**, 299–312 (2020). doi:10.3322/caac.21613.
6. Louis, D. N. *et al.* The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Neuro-Oncology* **23**, 1231 (June 2021). doi:10.1093/neuonc/noab106.
7. He, F. & Sun, Y. E. Glial cells more than support cells?. *The International Journal of Biochemistry & Cell Biology* **39**, 661–665 (2007). doi:10.1016/j.biocel.2006.10.022.
8. Virchow, R. *The Pathological Tumors: Thirty Lectures Held During the Winter Semester 1862-1863 at the University of Berlin: Third Volume, First Half* (BoD–Books on Demand, 2021).
9. Bailey, P. & Cushing, H. *A classification of the tumors of the glioma group on a histogenetic basis with a correlated study of prognosis* (JB Lippincott Company, 1926).
10. Kleihues, P., Burger, P. C. & Scheithauer, B. W. The New WHO Classification of Brain Tumours. *Brain Pathology* **3**, 255–268 (July 1993). doi:10.1111/j.1750-3639.1993.tb00752.x.

11. Louis, D. N. *et al.* The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathologica* **131**, 803–820 (June 2016). doi:10.1007/s00401-016-1545-1.

12. Lan, Z., Li, X. & Zhang, X. Glioblastoma: An Update in Pathology, Molecular Mechanisms and Biomarkers. *International Journal of Molecular Sciences* **25**, 3040 (Mar. 2024). doi:10.3390/ijms25053040.

13. Waitkus, M. S., Diplas, B. H. & Yan, H. Isocitrate dehydrogenase mutations in gliomas. *Neuro-Oncology* **18**, 16–26 (Jan. 2016). doi:10.1093/neuonc/nov136.

14. Dalziel, K. Isocitrate dehydrogenase and related oxidative decarboxylases. *FEBS Letters* **117**, K45–K55 (1980). doi:10.1016/0014-5793(80)80569-2.

15. Dang, L. *et al.* Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* **462**, 739 (Dec. 2009). doi:10.1038/nature08617.

16. Xu, W. *et al.* Oncometabolite 2-Hydroxyglutarate Is a Competitive Inhibitor of a-Ketoglutarate-Dependent Dioxygenases. *Cancer cell* **19**, 17–30 (Jan. 2011). doi:10.1016/j.ccr.2010.12.014.

17. Yan, H. *et al.* IDH1 and IDH2 Mutations in Gliomas. *The New England journal of medicine* **360**, 765–773 (Feb. 2009). doi:10.1056/NEJMoa0808710.

18. Balss, J. *et al.* Analysis of the IDH1 codon 132 mutation in brain tumors. *Acta Neuropathologica* **116**, 597–602 (Dec. 2008). doi:10.1007/s00401-008-0455-2.

19. Watanabe, T., Nobusawa, S., Kleihues, P. & Ohgaki, H. IDH1 Mutations Are Early Events in the Development of Astrocytomas and Oligodendrogliomas. *The American Journal of Pathology* **174**, 1149–1153 (Apr. 2009). doi:10.2353/ajpath.2009.080958.

20. Mellinghoff, I. K. *et al.* Vorasidenib in IDH1- or IDH2-mutant Low-Grade Glioma. *The New England journal of medicine* **389**, 589–601 (Aug. 2023). doi:10.1056/NEJMoa2304194.

21. Figueroa, M. E. *et al.* Leukemic IDH1 and IDH2 Mutations Result in a Hypermethylation Phenotype, Disrupt TET2 Function, and Impair Hematopoietic Differentiation. *Cancer Cell* **18**, 553–567 (Dec. 2010). doi:10.1016/j.ccr.2010.11.015.

22. Pekmezci, M. *et al.* Adult infiltrating gliomas with WHO 2016 integrated diagnosis: additional prognostic roles of ATRX and TERT. *Acta Neuropathologica* **133**, 1001–1016 (June 2017). doi:10.1007/s00401-017-1690-1.

23. Aoki, K. *et al.* Prognostic relevance of genetic alterations in diffuse lower-grade gliomas. *Neuro-Oncology* **20**, 66–77 (Jan. 2018). doi:10.1093/neuonc/nox132.

24. Eckel-Passow, J. E. *et al.* Glioma Groups Based on 1p/19q, IDH, and TERT Promoter Mutations in Tumors. *New England Journal of Medicine* **372**, 2499–2508 (June 2015). doi:10.1056/NEJMoa1407279.

25. Rincon-Torroella, J. *et al.* Current and Future Frontiers of Molecularly Defined Oligodendrogliomas. *Frontiers in Oncology* **12**, (July 2022). doi:10.3389/fonc.2022.934426.

26. Komori, T. Grading of adult diffuse gliomas according to the 2021 WHO Classification of Tumors of the Central Nervous System. *Laboratory Investigation* **102**, 126–133 (Feb. 2022). doi:10.1038/s41374-021-00667-6.

27. Appay, R. *et al.* PL1.1 CDKN2A homozygous deletion is a strong adverse prognosis factor in diffuse malignant IDHmutant gliomas. *Neuro-Oncology* **21**, iii1 (Sept. 2019). doi:10.1093/neuonc/noz126.000.

28. Network, T. C. G. A. R. Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. *New England Journal of Medicine* **372**, 2481–2498 (June 2015). doi:10.1056/NEJMoa1402121.

29. Appay, R. *et al.* CDKN2A homozygous deletion is a strong adverse prognosis factor in diffuse malignant IDH-mutant gliomas. *Neuro-Oncology* **21**, 1519–1528 (Dec. 2019). doi:10.1093/neuonc/noz124.

30. Reis, G. F. *et al.* CDKN2A Loss Is Associated with Shortened Overall Survival in Lower Grade (World Health Organization II-III) Astrocytomas. *Journal of neuropathology and experimental neurology* **74**, 442–452 (May 2015). doi:10.1097/NEN.0000000000000188.

31. Van den Bent, M. J. Interobserver variation of the histopathological diagnosis in clinical trials on glioma: a clinician's perspective. *Acta Neuropathologica* **120**, 297–304 (2010). doi:10.1007/s00401-010-0725-7.

32. Perry, A. & Brat, D. J. in *Practical Surgical Neuropathology: A Diagnostic Approach (Second Edition)* (eds Perry, A. & Brat, D. J.) Second Edition, 1–17 (Elsevier, 2018). doi:https://doi.org/10.1016/B978-0-323-44941-0.00001-1.

33. Lee, S. C. Diffuse Gliomas for Nonneuropathologists: The New Integrated Molecular Diagnostics. *Archives of Pathology & Laboratory Medicine* **142**, 804–814 (May 2018). doi:10.5858/arpa.2017-0449-RA.

34. Collins, J. M. & Christoforidis, G. A. in *Handbook of Neuro-Oncology Neuroimaging (Second Edition)* (ed Newton, H. B.) 439–460 (Academic Press, San Diego, Jan. 2016). doi:10.1016/B978-0-12-800945-1.00039-2.

35. Coons, S. W. & Pearl, D. K. Mitosis identification in diffuse gliomas. *Cancer* **82**, 1550–1555 (1998). doi:10.1002/(SICI)1097-0142(19980415)82:8<1550::AID-CNCR17>3.0.CO;2-3.

36. Takei, H., Bhattacharjee, M. B., Rivera, A., Dancer, Y. & Powell, S. Z. New Immunohistochemical Markers in the Evaluation of Central Nervous System Tumors: A Review of 7 Selected Adult and Pediatric Brain Tumors. *Archives of Pathology & Laboratory Medicine* **131**, 234–241 (Feb. 2007). doi:10.5858/2007-131-234-NIMITE.

37. Habberstad, A. H., Gulati, S. & Torp, S. H. Evaluation of the proliferation markers Ki-67/MIB-1, mitosin, survivin, pHH3, and DNA topoisomerase IIa in human anaplastic astrocytomas - an immunohistochemical study. *Diagnostic Pathology* **6**, 43 (May 2011). doi:10.1186/1746-1596-6-43.

38. Duregon, E. *et al.* Ki-67 proliferation index but not mitotic thresholds integrates the molecular prognostic stratification of lower grade gliomas. *Oncotarget* **7**, 21190–21198 (Mar. 2016). doi:10.18632/oncotarget.8498.

39. Soni *et al.* The spectrum of microvascular patterns in adult diffuse glioma and their correlation with tumor grade. *Journal of Pathology and Translational Medicine* **58**, 127–133 (May 2024). doi:10.4132/jptm.2024.03.11.

40. Rong, Y., Durden, D. L., Van Meir, E. G. & Brat, D. J. 'Pseudopalisading' Necrosis in Glioblastoma: A Familiar Morphologic Feature That Links Vascular Pathology, Hypoxia, and Angiogenesis. *Journal of Neuropathology & Experimental Neurology* **65**, 529–539 (June 2006). doi:10.1097/00005072-200606000-00001.

41. Figarella-Branger, D. *et al.* Prognostic Relevance of Histomolecular Classification of Diffuse Adult High-Grade Gliomas with Necrosis. *Brain Pathology* **25**, 418–428 (Dec. 2014). doi:10.1111/bpa.12227.

42. Ostrom, Q. T. *et al.* CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2016—2020. *Neuro-Oncology* **25**, iv1–iv99 (Oct. 2023). doi:10.1093/neuonc/noad149.

43. Louis, D. N. *et al.* The 2007 WHO Classification of Tumours of the Central Nervous System. *Acta Neuropathologica* **114**, 97–109 (Aug. 2007). doi:10.1007/s00401-007-0243-4.

44. Delgado-López, P. D. & Corrales-García, E. M. Survival in glioblastoma: a review on the impact of treatment modalities. *Clinical and Translational Oncology* **18**, 1062–1071 (Nov. 2016). doi:10.1007/s12094-016-1497-x.

45. McKinnon, C., Nandhabalan, M., Murray, S. A. & Plaha, P. Glioblastoma: clinical presentation, diagnosis, and management. *BMJ* **374**, n1560 (July 2021). doi:10.1136/bmj.n1560.

46. Chang, S. M. *et al.* Patterns of Care for Adults With Newly Diagnosed Malignant Glioma. *JAMA* **293**, 557–564 (Feb. 2005). doi:10.1001/jama.293.5.557.

47. Richard, S. A., Ye, Y., Li, H., Ma, L. & You, C. Glioblastoma multiforme subterfuge as acute cerebral hemorrhage: A case report and literature review. *Neurology International* **10**, 7558 (Apr. 2018). doi:10.4081/ni.2018.7558.

48. Melhem, J. M., Detsky, J., Lim-Fat, M. J. & Perry, J. R. Updates in IDH-Wildtype Glioblastoma. *Neurotherapeutics* **19**, 1705–1723 (Oct. 2022). doi:10.1007/s13311-022-01251-6.

49. Hooper, G. W., Ansari, S., Johnson, J. M. & Ginat, D. T. Advances in the Radiological Evaluation of and Theranostics for Glioblastoma. *Cancers* **15**, 4162 (Aug. 2023). doi:10.3390/cancers15164162.

50. Burkett, B. J. *et al.* A Review of Theranostics: Perspectives on Emerging Approaches and Clinical Advancements. *Radiology: Imaging Cancer* **5**, e220157 (July 2023). doi:10.1148/rycan.220157.

51. Heidenreich, B. & Kumar, R. *TERT* promoter mutations in telomere biology. *Mutation Research/Reviews in Mutation Research* **771**, 15–31 (Jan. 2017). doi:10.1016/j.mrrev.2016.11.002.

52. Chow, T. T., Zhao, Y., Mak, S. S., Shay, J. W. & Wright, W. E. Early and late steps in telomere overhang processing in normal human cells: the position of the final RNA primer drives telomere shortening. *Genes & Development* **26**, 1167–1178 (June 2012). doi:10.1101/gad.187211.112.

53. Olympios, N. *et al.* TERT Promoter Alterations in Glioblastoma: A Systematic Review. *Cancers* **13**, 1147 (Jan. 2021). doi:10.3390/cancers13051147.

54. Bell, R. J. *et al.* The transcription factor GABP selectively binds and activates the mutant TERT promoter in cancer. *Science (New York, N.Y.)* **348**, 1036–1039 (May 2015). doi:10.1126/science.aab0015.

55. Schlessinger, J. Cell Signaling by Receptor Tyrosine Kinases. *Cell* **103**, 211–225 (Oct. 2000). doi:10.1016/S0092-8674(00)00114-8.

56. Snuderl, M. *et al.* Mosaic Amplification of Multiple Receptor Tyrosine Kinase Genes in Glioblastoma. *Cancer Cell* **20**, 810–817 (Dec. 2011). doi:10.1016/j.ccr.2011.11.005.

57. Robinson, D. R., Wu, Y.-M. & Lin, S.-F. The protein tyrosine kinase family of the human genome. *Oncogene* **19**, 5548–5557 (Nov. 2000). doi:10.1038/sj.onc.1203957.

58. Eskilsson, E. *et al.* EGFR heterogeneity and implications for therapeutic intervention in glioblastoma. *Neuro-Oncology* **20**, 743–752 (May 2018). doi:10.1093/neuonc/nox191.

59. Lassman, A. B. *et al.* Epidermal Growth Factor Receptor (EGFR) Amplification Rates Observed in Screening Patients for Randomized Trials in Glioblastoma. *Journal of neuro-oncology* **144**, 205–210 (Aug. 2019). doi:10.1007/s11060-019-03222-y.

60. Bigner, S. H. *et al.* Relationship between gene amplification and chromosomal deviations in malignant human gliomas. *Cancer Genetics and Cytogenetics* **29**, 165–170 (Nov. 1987). doi:10.1016/0165-4608(87)90045-8.

61. Gan, H. K., Cvrljevic, A. N. & Johns, T. G. The epidermal growth factor receptor variant III (EGFRvIII): where wild things are altered. *The FEBS Journal* **280**, 5350–5370 (2013). doi:10.1111/febs.12393.

62. Frederick, L., Wang, X. Y., Eley, G. & James, C. D. Diversity and frequency of epidermal growth factor receptor mutations in human glioblastomas. *Cancer Research* **60**, 1383–1387 (Mar. 2000).

63. Shepherd, F. A. *et al.* Erlotinib in Previously Treated Non–Small-Cell Lung Cancer. *New England Journal of Medicine* **353**, 123–132 (July 2005). doi:10.1056/NEJMoa050753.

64. Lynch, T. J. *et al.* Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of Non–Small-Cell Lung Cancer to Gefitinib. *New England Journal of Medicine* **350**, 2129–2139 (May 2004). doi:10.1056/NEJMoa040938.

65. Früh, M. & Pless, M. EGFR IHC score for selection of cetuximab treatment: Ready for clinical practice?. *Translational Lung Cancer Research* **1**, 145–146 (June 2012). doi:10.3978/j.issn.2218-6751.2012.03.01.

66. Westphal, M., Maire, C. L. & Lamszus, K. EGFR as a Target for Glioblastoma Treatment: An Unfulfilled Promise. *CNS Drugs* **31**, 723–735 (2017). doi:10.1007/s40263-017-0456-6.

67. Vivanco, I. *et al.* Differential Sensitivity of Glioma- versus Lung Cancer-specific EGFR mutations to EGFR Kinase Inhibitors. *Cancer Discovery* **2**, 458–471 (May 2012). doi:10.1158/2159-8290.CD-11-0284.

68. Kim, T.-M. *et al.* Functional genomic analysis of chromosomal aberrations in a compendium of 8000 cancer genomes. *Genome Research* **23**, 217–227 (Feb. 2013). doi:10.1101/gr.140301.112.

69. Galbraith, K. *et al.* Molecular Correlates of Long Survival in IDH-Wildtype Glioblastoma Cohorts. *Journal of Neuropathology & Experimental Neurology* **79**, 843–854 (Aug. 2020). doi:10.1093/jnen/nlaa059.

70. Stichel, D. *et al.* Distribution of EGFR amplification, combined chromosome 7 gain and chromosome 10 loss, and TERT promoter mutation in brain tumors and their potential for the reclassification of IDHwt astrocytoma to glioblastoma. *Acta Neuropathologica* **136**, 793–803 (Nov. 2018). doi:10.1007/s00401-018-1905-0.

71. Körber, V. *et al.* Evolutionary Trajectories of IDHWT Glioblastomas Reveal a Common Path of Early Tumorigenesis Instigated Years ahead of Initial Diagnosis. *Cancer Cell* **35**, 692–704.e12 (Apr. 2019). doi:10.1016/j.ccell.2019.02.007.

72. Barthel, F. P. *et al.* Longitudinal molecular trajectories of diffuse glioma in adults. *Nature* **576**, 112–120 (Dec. 2019). doi:10.1038/s41586-019-1775-1.

73. Ichimura, K., Schmidt, E. E., Miyakawa, A., Goike, H. M. & Collins, V. P. Distinct patterns of deletion on 10p and 10q suggest involvement of multiple tumor suppressor genes in the development of astrocytic gliomas of different malignancy grades. *Genes, Chromosomes and Cancer* **22**, 9–15 (1998). doi:10.1002/(SICI)1098-2264(199805)22:1<9::AID-GCC2>3.0.CO;2-1.

74. Li, D.-M. & Sun, H. TEP1, Encoded by a Candidate Tumor Suppressor Locus, Is a Novel Protein Tyrosine Phosphatase Regulated by Transforming Growth Factor beta1. *Cancer Research* **57**, 2124–2129 (June 1997).

75. Yadav, A. K. *et al.* Monosomy of Chromosome 10 Associated With Dysregulation of Epidermal Growth Factor Signaling in Glioblastomas. *JAMA : the journal of the American Medical Association* **302**, 276–289 (July 2009). doi:10.1001/jama.2009.1022.

76. Masilamani, A. P. *et al.* KLF6 depletion promotes NF-kappaB signaling in glioblastoma. *Oncogene* **36**, 3562–3575 (June 2017). doi:10.1038/onc.2016.507.

77. Wullich, B., Sattler, H. P., Fischer, U. & Meese, E. Two independent amplification events on chromosome 7 in glioma: amplification of the epidermal growth factor receptor gene and amplification of the oncogene MET. *Anticancer Research* **14**, 577–579 (1994).

78. Davoli, T. *et al.* Cumulative Haploinsufficiency and Triplosensitivity Drive Aneuploidy Patterns to Shape the Cancer Genome. *Cell* **155**, 948–962 (Nov. 2013). doi:10.1016/j.cell.2013.10.011.

79. Sack, L. M. *et al.* Profound Tissue Specificity in Proliferation Control Underlies Cancer Drivers and Aneuploidy Patterns. *Cell* **173**, 499–514.e23 (Apr. 2018). doi:10.1016/j.cell.2018.02.037.

80. Roh, T. H. & Kim, S.-H. Supramaximal Resection for Glioblastoma: Redefining the Extent of Resection Criteria and Its Impact on Survival. *Brain Tumor Research and Treatment* **11**, 166–172 (July 2023). doi:10.14791/btrt.2023.0012.

81. Karschnia, P. *et al.* Prognostic validation of a new classification system for extent of resection in glioblastoma: A report of the RANO resect group. *Neuro-Oncology* **25**, 940–954 (Aug. 2022). doi:10.1093/neuonc/noac193.

82. Roh, T. H. *et al.* Long-term outcomes of concomitant chemoradiotherapy with temozolomide for newly diagnosed glioblastoma patients. *Medicine* **96**, e7422 (July 2017). doi:10.1097/MD.0000000000007422.

83. Yordanova, Y. N., Moritz-Gasser, S. & Duffau, H. Awake surgery for WHO Grade II gliomas within "noneloquent" areas in the left dominant hemisphere: toward a "supratotal" resection. (Aug. 2011). doi:10.3171/2011.3.JNS101333.

84. Pessina, F. *et al.* Maximize surgical resection beyond contrast-enhancing boundaries in newly diagnosed glioblastoma multiforme: is it useful and safe? A single institution retrospective experience. *Journal of Neuro-Oncology* **135**, 129–139 (Oct. 2017). doi:10.1007/s11060-017-2559-9.

85. Li, Y. M., Suki, D., Hess, K. & Sawaya, R. The influence of maximum safe resection of glioblastoma on survival in 1229 patients: Can we do better than gross-total resection?. *Journal of neurosurgery* **124**, 977–988 (Apr. 2016). doi:10.3171/2015.5.JNS142087.

86. Gerritsen, J. K. W. *et al.* Effect of awake craniotomy in glioblastoma in eloquent areas (GLIOMAP): a propensity score-matched analysis of an international, multicentre, cohort study. *The Lancet Oncology* **23**, 802–817 (June 2022). doi:10.1016/S1470-2045(22)00213-3.

87. Stummer, W. *et al.* Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma: a randomised controlled multicentre phase III trial. *The Lancet Oncology* **7**, 392–401 (May 2006). doi:10.1016/S1470-2045(06)70665-9.

88. Stupp, R. *et al.* Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma. *New England Journal of Medicine* **352**, 987–996 (Mar. 2005). doi:10.1056/NEJMoa043330.

89. Lee, S. Y. Temozolomide resistance in glioblastoma multiforme. *Genes & Diseases* **3**, 198–210 (May 2016). doi:10.1016/j.gendis.2016.04.007.

90. JIAPAER, S., FURUTA, T., TANAKA, S., KITABAYASHI, T. & NAKADA, M. Potential Strategies Overcoming the Temozolomide Resistance for Glioblastoma. *Neurologia medico-chirurgica* **58**, 405–421 (Oct. 2018). doi:10.2176/nmc.ra.2018-0141.

91. Bao, J., Sun, R., Pan, Z. & Wei, S. Current chemotherapy strategies for adults with IDH-wildtype glioblastoma. *Frontiers in Oncology* **14**, 1438905 (July 2024). doi:10.3389/fonc.2024.1438905.

92. Stupp, R. *et al.* Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *The Lancet Oncology* **10**, 459–466 (May 2009). doi:10.1016/S1470-2045(09)70025-7.

93. Hegi, M. E. *et al.* MGMT Gene Silencing and Benefit from Temozolomide in Glioblastoma. *New England Journal of Medicine* **352**, 997–1003 (Mar. 2005). doi:10.1056/NEJMoa043331.

94. Herrlinger, U. *et al.* Lomustine-temozolomide combination therapy versus standard temozolomide therapy in patients with newly diagnosed glioblastoma with methylated MGMT promoter (CeTeG/NOA–09): a randomised, open-label, phase 3 trial. *The Lancet* **393**, 678–688 (Feb. 2019). doi:10.1016/S0140-6736(18)31791-4.

95. Cao, J. Q. *et al.* Hypofractionated radiotherapy with or without concurrent temozolomide in elderly patients with glioblastoma multiforme: a review of ten-year single institutional experience. *Journal of Neuro-Oncology* **107**, 395–405 (Apr. 2012). doi:10.1007/s11060-011-0766-3.

96. Angom, R. S., Nakka, N. M. R. & Bhattacharya, S. Advances in Glioblastoma Therapy: An Update on Current Approaches. *Brain Sciences* **13**, 1536 (Oct. 2023). doi:10.3390/brainsci13111536.

97. Le Rhun, E. *et al.* Molecular targeted therapy of glioblastoma. *Cancer Treatment Reviews* **80**, 101896 (Sept. 2019). doi:10.1016/j.ctrv.2019.101896.

98. Brown, C. E. *et al.* Locoregional delivery of IL-13Ralpha2-targeting CAR-T cells in recurrent high-grade glioma: a phase 1 trial. *Nature Medicine* **30**, 1001–1012 (2024). doi:10.1038/s41591-024-02875-1.

99. *National Cancer Institute (NCI).* The Cancer Genome Atlas (TCGA) Program. https://www.cancer.gov/ccg/research/genome-sequencing/tcga (accessed 13 Mar 2025).

100. McLendon, R. *et al.* Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455**, 1061–1068 (Oct. 2008). doi:10.1038/nature07385.

101. Brennan, C. W. *et al.* The somatic genomic landscape of glioblastoma. *Cell* **155**, 462–477 (Oct. 10, 2013). doi:10.1016/j.cell.2013.09.034.

102. Thorsson, V. *et al.* The immune landscape of cancer. *Immunity* **48**, 812–830.e14 (Apr. 17, 2018). doi:10.1016/j.immuni.2018.03.023.

103. Corces, M. R. *et al.* The chromatin accessibility landscape of primary human cancers. *Science* **362**, eaav1898 (Oct. 26, 2018). doi:10.1126/science.aav1898.

104. Bailey, M. H. *et al.* Comprehensive characterization of cancer driver genes and mutations. *Cell* **173**, 371–385.e18 (Apr. 5, 2018). doi:10.1016/j.cell.2018.02.060.

105. Carrot-Zhang, J. *et al.* Whole-genome characterization of lung adenocarcinomas lacking alterations in the RTK/RAS/RAF pathway. *Cell Reports* **34**, (Feb. 2, 2021). doi:10.1016/j.celrep.2021.108707.

106. Phillips, H. S. *et al.* Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. *Cancer Cell* **9**, 157–173 (Mar. 2006). doi:10.1016/j.ccr.2006.02.019.

107. Verhaak, R. G. W. *et al.* Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in *PDGFRA*, *IDH1*, *EGFR*, and *NF1*. *Cancer Cell* **17**, 98–110 (Jan. 2010). doi:10.1016/j.ccr.2009.12.020.

108. Wang, Q. *et al.* Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. *Cancer Cell* **33**, 152 (Jan. 2018). doi:10.1016/j.ccell.2017.12.012.

109. Noushmehr, H. *et al.* Identification of a CpG Island Methylator Phenotype that Defines a Distinct Subgroup of Glioma. *Cancer cell* **17**, 510–522 (May 2010). doi:10.1016/j.ccr.2010.03.017.

110. Sottoriva, A. *et al.* Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proceedings of the National Academy of Sciences* **110**, 4009–4014 (Mar. 2013). doi:10.1073/pnas.1219747110.

111. Singh, S. K. *et al.* Identification of human brain tumour initiating cells. *Nature* **432**, 396–401 (Nov. 2004). doi:10.1038/nature03128.

112. Galli, R. *et al.* Isolation and characterization of tumorigenic, stem-like neural precursors from human glioblastoma. (Oct. 2004).

113. Bao, S. *et al.* Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature* **444**, 756–760 (Dec. 2006). doi:10.1038/nature05236.

114. Pallini, R. *et al.* Cancer Stem Cell Analysis and Clinical Outcome in Patients with Glioblastoma Multiforme. *Clinical Cancer Research* **14**, 8205–8212 (Dec. 2008). doi:10.1158/1078-0432.CCR-08-0644.

115. Lan, X. *et al.* Fate mapping of human glioblastoma reveals an invariant stem cell hierarchy. *Nature* **549**, 227 (Aug. 2017). doi:10.1038/nature23666.

116. Chen, J. *et al.* A restricted cell population propagates glioblastoma growth after chemotherapy. *Nature* **488**, 522–526 (Aug. 2012). doi:10.1038/nature11287.

117. Bhat, K. P. L. *et al.* Mesenchymal Differentiation Mediated by NF-kappaB Promotes Radiation Resistance in Glioblastoma. *Cancer Cell* **24**, 331–346 (Sept. 2013). doi:10.1016/j.ccr.2013.08.001.

118. Chen, R. *et al.* A Hierarchy of Self-Renewing Tumor-Initiating Cell Types in Glioblastoma. *Cancer Cell* **17**, 362–375 (Apr. 2010). doi:10.1016/j.ccr.2009.12.049.

119. Wang, J. *et al.* CD133 negative glioma cells form tumors in nude rats and give rise to CD133 positive cells. *International Journal of Cancer* **122**, 761–768 (2008). doi:10.1002/ijc.23130.

120. Ogden, A. T. *et al.* IDENTIFICATION OF A2B5+CD133– TUMOR-INITIATING CELLS IN ADULT HUMAN GLIOMAS. *Neurosurgery* **62**, 505 (Feb. 2008). doi:10.1227/01.neu.0000316019.28421.95.

121. Fares, J., Kanojia, D., Cordero, A., Ulasov, I. & Lesniak, M. S. in *Glioblastoma Resistance to Chemotherapy: Molecular Mechanisms and Innovative Reversal Strategies* (eds Paulmurugan, R. & Massoud, T. F.) 587–634 (Academic Press, Jan. 2021). doi:10.1016/B978-0-12-821567-8.00014-2.

122. Patel, A. P. *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396–1401 (June 2014). doi:10.1126/science.1254257.

123. Neftel, C. *et al.* An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. *Cell* **178**, 835–849.e21 (Aug. 2019). doi:10.1016/j.cell.2019.06.024.

124. Wang, L. *et al.* The Phenotypes of Proliferating Glioblastoma Cells Reside on a Single Axis of Variation. *Cancer Discovery* **9**, 1708–1719 (Dec. 2019). doi:10.1158/2159-8290.CD-19-0329.

125. Eyler, C. E. *et al.* Single-cell lineage analysis reveals genetic and epigenetic interplay in glioblastoma drug resistance. *Genome Biology* **21**, 174 (July 2020). doi:10.1186/s13059-020-02085-1.

126. Liau, B. B. *et al.* Adaptive Chromatin Remodeling Drives Glioblastoma Stem Cell Plasticity and Drug Tolerance. *Cell Stem Cell* **20**, 233–246.e7 (Feb. 2017). doi:10.1016/j.stem.2016.11.003.

127. Bhaduri, A. *et al.* Outer Radial Glia-Like Cancer Stem Cells Contribute to Heterogeneity of Glioblastoma. *Cell stem cell* **26**, 48–63.e6 (Jan. 2020). doi:10.1016/j.stem.2019.11.015.

128. Couturier, C. P. *et al.* Single-cell RNA-seq reveals that glioblastoma recapitulates a normal neurodevelopmental hierarchy. *Nature Communications* **11**, 3406 (July 2020). doi:10.1038/s41467-020-17186-5.

129. Castellan, M. *et al.* Single-cell analyses reveal YAP/TAZ as regulators of stemness and cell plasticity in Glioblastoma. *Nature cancer* **2**, 174–188 (Feb. 2021). doi:10.1038/s43018-020-00150-z.

130. Garofano, L. *et al.* Pathway-based classification of glioblastoma uncovers a mitochondrial subtype with therapeutic vulnerabilities. *Nature cancer* **2**, 141–156 (Feb. 2021). doi:10.1038/s43018-020-00159-4.

131. Richards, L. M. *et al.* Gradient of Developmental and Injury Response transcriptional states defines functional vulnerabilities underpinning glioblastoma heterogeneity. *Nature Cancer* **2**, 157–173 (Feb. 2021). doi:10.1038/s43018-020-00154-9.

132. Hara, T. *et al.* Interactions between cancer cells and immune cells drive transitions to mesenchymal-like states in glioblastoma. *Cancer Cell* **39**, 779–792.e11 (June 2021). doi:10.1016/j.ccell.2021.05.002.

133. Bv, H. & Jolly, M. K. Proneural-mesenchymal antagonism dominates the patterns of phenotypic heterogeneity in glioblastoma. *iScience* **27**, (Mar. 2024). doi:10.1016/j.isci.2024.109184.

134. Hoogstrate, Y. *et al.* Transcriptome analysis reveals tumor microenvironment changes in glioblastoma. *Cancer Cell* **41**, 678–692.e7 (Apr. 2023). doi:10.1016/j.ccell.2023.02.019.

135. Varn, F. S. *et al.* Glioma progression is shaped by genetic evolution and microenvironment interactions. *Cell* **185**, 2184–2199.e16 (June 2022). doi:10.1016/j.cell.2022.04.038.

136. Wang, L. *et al.* A single-cell atlas of glioblastoma evolution under therapy reveals cell-intrinsic and cell-extrinsic therapeutic targets. *Nature Cancer* **3**, 1534–1552 (Dec. 2022). doi:10.1038/s43018-022-00475-x.

137. Tanner, G. *et al.* IDHwt glioblastomas can be stratified by their transcriptional response to standard treatment, with implications for targeted therapy. *Genome Biology* **25**, 45 (Feb. 2024). doi:10.1186/s13059-024-03172-3.

138. Sharma, P., Aaroe, A., Liang, J. & Puduvalli, V. K. Tumor microenvironment in glioblastoma: Current and emerging concepts. *Neuro-Oncology Advances* **5**, vdad009 (Feb. 2023). doi:10.1093/noajnl/vdad009.

139. Bikfalvi, A. *et al.* Challenges in glioblastoma research: focus on the tumor microenvironment. *Trends in Cancer* **9**, 9–27 (Jan. 2023). doi:10.1016/j.trecan.2022.09.005.

140. Tamai, S. *et al.* Tumor Microenvironment in Glioma Invasion. *Brain Sciences* **12**, 505 (Apr. 2022). doi:10.3390/brainsci12040505.

141. Brandao, M., Simon, T., Critchley, G. & Giamas, G. Astrocytes, the rising stars of the glioblastoma microenvironment. *Glia* **67**, 779–790 (2019). doi:10.1002/glia.23520.

142. Henrik Heiland, D. *et al.* Tumor-associated reactive astrocytes aid the evolution of immunosuppressive environment in glioblastoma. *Nature Communications* **10**, 2541 (June 2019). doi:10.1038/s41467-019-10493-6.

143. Watson, D. C. *et al.* GAP43-dependent mitochondria transfer from astrocytes enhances glioblastoma tumorigenicity. *Nature Cancer* **4**, 648–664 (May 2023). doi:10.1038/s43018-023-00556-5.

144. Osswald, M. *et al.* Brain tumour cells interconnect to a functional and resistant network. *Nature* **528**, 93–98 (Dec. 2015). doi:10.1038/nature16071.

145. Hide, T. *et al.* Oligodendrocyte Progenitor Cells and Macrophages/Microglia Produce Glioma Stem Cell Niches at the Tumor Border. *eBioMedicine* **30**, 94–104 (Apr. 2018). doi:10.1016/j.ebiom.2018.02.024.

146. White, J., White, M. P. J., Wickremesekera, A., Peng, L. & Gray, C. The tumour microenvironment, treatment resistance and recurrence in glioblastoma. *Journal of Translational Medicine* **22**, 540 (June 2024). doi:10.1186/s12967-024-05301-9.

147. Bercury, K. K. & Macklin, W. B. Dynamics and Mechanisms of CNS Myelination. *Developmental Cell* **32**, 447–458 (Feb. 2015). doi:10.1016/j.devcel.2015.01.016.

148. Yeung, M. S. Y. *et al.* Dynamics of Oligodendrocyte Generation and Myelination in the Human Brain. *Cell* **159**, 766–774 (Nov. 2014). doi:10.1016/j.cell.2014.10.011.

149. Kloepper, J. *et al.* Ang-2/VEGF bispecific antibody reprograms macrophages and resident microglia to anti-tumor phenotype and prolongs glioblastoma survival. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 4476–4481 (Apr. 2016). doi:10.1073/pnas.1525360113.

150. Huang, Y. *et al.* Oligodendrocyte Progenitor Cells Promote Neovascularization in Glioma by Disrupting the Blood–Brain Barrier. *Cancer Research* **74**, 1011–1021 (Feb. 2014). doi:10.1158/0008-5472.CAN-13-1072.

151. Venkatesh, H. S. *et al.* Neuronal Activity Promotes Glioma Growth through Neuroligin-3 Secretion. *Cell* **161**, 803–816 (May 2015). doi:10.1016/j.cell.2015.04.012.

152. Glaviano, A. *et al.* PI3K/AKT/mTOR signaling transduction pathway and targeted therapies in cancer. *Molecular Cancer* **22**, 138 (Aug. 2023). doi:10.1186/s12943-023-01827-6.

153. Krishna, S. *et al.* Glioblastoma remodelling of human neural circuits decreases survival. *Nature* **617**, 599–607 (May 2023). doi:10.1038/s41586-023-06036-1.

154. Venkataramani, V. *et al.* Glutamatergic synaptic input to glioma cells drives brain tumour progression. *Nature* **573**, 532–538 (Sept. 2019). doi:10.1038/s41586-019-1564-x.

155. Klemm, F. *et al.* Interrogation of the Microenvironmental Landscape in Brain Tumors Reveals Disease-Specific Alterations of Immune Cells. *Cell* **181**, 1643–1660.e17 (June 2020). doi:10.1016/j.cell.2020.05.007.

156. Louveau, A. *et al.* Structural and functional features of central nervous system lymphatics. *Nature* **523**, 337–341 (July 2015). doi:10.1038/nature14432.

157. Song, E. *et al.* VEGF-C-driven lymphatic drainage enables immunosurveillance of brain tumours. *Nature* **577**, 689–694 (Jan. 2020). doi:10.1038/s41586-019-1912-x.

158. Noch, E. K., Sait, S. F., Farooq, S., Trippett, T. M. & Miller, A. M. A case series of extraneural metastatic glioblastoma at Memorial Sloan Kettering Cancer Center. *Neuro-Oncology Practice* **8**, 325–336 (Feb. 2021). doi:10.1093/nop/npaa083.

159. Van Hooren, L. *et al.* Agonistic CD40 therapy induces tertiary lymphoid structures but impairs responses to checkpoint blockade in glioma. *Nature Communications* **12**, 4127 (July 2021). doi:10.1038/s41467-021-24347-7.

160. Wang, J., Li, S., Lan, Y., Liu, X. & Li, W. Glioma-associated macrophages: unraveling their dual role in the microenvironment and therapeutic implications. *Current Medicine* **3**, 4 (Apr. 2024). doi:10.1007/s44194-024-00031-y.

161. Andersen, B. M. *et al.* Glial and myeloid heterogeneity in the brain tumour microenvironment. *Nature Reviews Cancer* **21**, 786–802 (Dec. 2021). doi:10.1038/s41568-021-00397-3.

162. Pombo Antunes, A. R. *et al.* Single-cell profiling of myeloid cells in glioblastoma across species and disease stage reveals macrophage competition and specialization. *Nature Neuroscience* **24**, 595–610 (Apr. 2021). doi:10.1038/s41593-020-00789-y.

163. Zhang, S.-Y. *et al.* Tumor-associated macrophages: A promising target for a cancer immunotherapeutic strategy. *Pharmacological Research* **161**, 105111 (Nov. 2020). doi:10.1016/j.phrs.2020.105111.

164. Mantovani, A., Allavena, P., Marchesi, F. & Garlanda, C. Macrophages as tools and targets in cancer therapy. *Nature Reviews Drug Discovery* **21**, 799–820 (Nov. 2022). doi:10.1038/s41573-022-00520-5.

165. Peng, Y. *et al.* Tumor-associated macrophages as treatment targets in glioma. *Brain Science Advances* **6**, 306–323 (Dec. 2020). doi:10.26599/BSA.2020.9050015.

166. Cui, X. *et al.* Hacking macrophage-associated immunosuppression for regulating glioblastoma angiogenesis. *Biomaterials* **161**, 164–178 (Apr. 2018). doi:10.1016/j.biomaterials.2018.01.053.

167. Quintero-Fabián, S. *et al.* Role of Matrix Metalloproteinases in Angiogenesis and Cancer. *Frontiers in Oncology* **9**, 1370 (Dec. 2019). doi:10.3389/fonc.2019.01370.

168. Zhang, G., Tao, X., Ji, B. & Gong, J. Hypoxia-Driven M2-Polarized Macrophages Facilitate Cancer Aggressiveness and Temozolomide Resistance in Glioblastoma. *Oxidative Medicine and Cellular Longevity* **2022**, 1614336 (Aug. 2022). doi:10.1155/2022/1614336.

169. Tan, Y. *et al.* Tumor-Associated Macrophages: A Potential Target for Cancer Therapy. *Frontiers in Oncology* **11**, 693517 (June 2021). doi:10.3389/fonc.2021.693517.

170. Macri, C., Pang, E. S., Patton, T. & O'Keeffe, M. Dendritic cell subsets. *Seminars in Cell & Developmental Biology* **84**, 11–21 (Dec. 2018). doi:10.1016/j.semcdb.2017.12.009.

171. Gardam, B., Gargett, T., Brown, M. P. & Ebert, L. M. Targeting the dendritic cell-T cell axis to develop effective immunotherapies for glioblastoma. *Frontiers in Immunology* **14**, (Oct. 2023). doi:10.3389/fimmu.2023.1261257.

172. Datsi, A. & Sorg, R. V. Dendritic Cell Vaccination of Glioblastoma: Road to Success or Dead End. *Frontiers in Immunology* **12**, 770390 (Nov. 2021). doi:10.3389/fimmu.2021.770390.

173. Sedgwick, A. J., Ghazanfari, N., Constantinescu, P., Mantamadiotis, T. & Barrow, A. D. The Role of NK Cells and Innate Lymphoid Cells in Brain Cancer. *Frontiers in Immunology* **11**, (July 2020). doi:10.3389/fimmu.2020.01549.

174. Castriconi, R. *et al.* NK Cells Recognize and Kill Human Glioblastoma Cells with Stem Cell-Like Properties1. *The Journal of Immunology* **182**, 3530–3539 (Mar. 2009). doi:10.4049/jimmunol.0802845.

175. Morimoto, T. *et al.* Natural Killer Cell-Based Immunotherapy against Glioblastoma. *International Journal of Molecular Sciences* **24**, 2111 (Jan. 2023). doi:10.3390/ijms24032111.

176. Wang, M., Zhou, Z., Wang, X., Zhang, C. & Jiang, X. Natural killer cell awakening: unleash cancer-immunity cycle against glioblastoma. *Cell Death & Disease* **13**, 1–10 (July 2022). doi:10.1038/s41419-022-05041-y.

177. Lee, N. *et al.* HLA-E is a major ligand for the natural killer inhibitory receptor CD94/NKG2A. *Proceedings of the National Academy of Sciences* **95**, 5199–5204 (Apr. 1998). doi:10.1073/pnas.95.9.5199.

178. Read, R. D., Tapp, Z. M., Rajappa, P. & Hambardzumyan, D. Glioblastoma microenvironment—from biology to therapy. *Genes & Development* **38**, 360–379 (May 2024). doi:10.1101/gad.351427.123.

179. Hou, D. *et al.* B-cells Drive Response to PD-1 Blockade in Glioblastoma Upon Neutralization of TGFbeta-mediated Immunosuppression. *Research Square,* rs.3.rs–2399170 (Jan. 2023). doi:10.21203/rs.3.rs-2399170/v1.

180. Akkaya, M., Kwak, K. & Pierce, S. K. B cell memory: building two walls of protection against pathogens. *Nature Reviews Immunology* **20**, 229–238 (Apr. 2020). doi:10.1038/s41577-019-0244-2.

181. Lee-Chang, C. *et al.* Myeloid-Derived Suppressive Cells Promote B cell–Mediated Immunosuppression via Transfer of PD-L1 in Glioblastoma. *Cancer Immunology Research* **7**, 1928–1943 (Dec. 2019). doi:10.1158/2326-6066.CIR-19-0240.

182. González-Tablas Pimenta, M. *et al.* Tumor cell and immune cell profiles in primary human glioblastoma: Impact on patient outcome. *Brain Pathology* **31**, 365–380 (2021). doi:10.1111/bpa.12927.

183. Cordell, E. C., Alghamri, M. S., Castro, M. G. & Gutmann, D. H. T lymphocytes as dynamic regulators of glioma pathobiology. *Neuro-Oncology* **24**, 1647–1657 (Mar. 2022). doi:10.1093/neuonc/noac055.

184. Reiser, J. & Banerjee, A. Effector, Memory, and Dysfunctional CD8+ T Cell Fates in the Antitumor Immune Response. *Journal of Immunology Research* **2016**, 8941260 (2016). doi:10.1155/2016/8941260.

185. Heimberger, A. B. *et al.* Incidence and Prognostic Impact of FoxP3+ Regulatory T Cells in Human Gliomas. *Clinical Cancer Research* **14**, 5166–5172 (Aug. 2008). doi:10.1158/1078-0432.CCR-08-0320.

186. Miska, J. *et al.* HIF-1alpha Is a Metabolic Switch between Glycolytic-Driven Migration and Oxidative Phosphorylation-Driven Immunosuppression of Tregs in Glioblastoma. *Cell Reports* **27**, 226–237.e4 (Apr. 2019). doi:10.1016/j.celrep.2019.03.029.

187. Woroniecka, K. *et al.* T Cell Exhaustion Signatures Vary with Tumor Type and are Severe in Glioblastoma. *Clinical cancer research : an official journal of the American Association for Cancer Research* **24**, 4175–4186 (Sept. 2018). doi:10.1158/1078-0432.CCR-17-1846.

188. Xing, Y. & Hogquist, K. A. T-Cell Tolerance: Central and Peripheral. *Cold Spring Harbor Perspectives in Biology* **4**, a006957 (June 2012). doi:10.1101/cshperspect.a006957.

189. Woroniecka, K. I., Rhodin, K. E., Chongsathidkiet, P., Keith, K. A. & Fecci, P. E. T-cell Dysfunction in Glioblastoma: Applying a New Framework. *Clinical Cancer Research* **24**, 3792–3802 (Aug. 2018). doi:10.1158/1078-0432.CCR-18-0047.

190. Abe, B. T. & Macian, F. Uncovering the mechanisms that regulate tumor-induced T-cell anergy. *Oncoimmunology* **2**, e22679 (Feb. 2013). doi:10.4161/onci.22679.

191. Mirzaei, R., Sarkar, S. & Yong, V. W. T Cell Exhaustion in Glioblastoma: Intricacies of Immune Checkpoints. *Trends in Immunology* **38**, 104–115 (Feb. 2017). doi:10.1016/j.it.2016.11.005.

192. Hovis, G. *et al.* Understanding the Role of Endothelial Cells in Glioblastoma: Mechanisms and Novel Treatments. *International Journal of Molecular Sciences* **25**, 6118 (June 2024). doi:10.3390/ijms25116118.

193. Kane, J. R. The Role of Brain Vasculature in Glioblastoma. *Molecular Neurobiology* **56**, 6645–6653 (Sept. 2019). doi:10.1007/s12035-019-1561-y.

194. Aird, W. C. Phenotypic Heterogeneity of the Endothelium. *Circulation Research* **100**, 158–173 (Feb. 2007). doi:10.1161/01.RES.0000255691.76142.4a.

195. Lochhead, J. J., Yang, J., Ronaldson, P. T. & Davis, T. P. Structure, Function, and Regulation of the Blood-Brain Barrier Tight Junction in Central Nervous System Disorders. *Frontiers in Physiology* **11**, 914 (Aug. 2020). doi:10.3389/fphys.2020.00914.

196. Armulik, A. *et al.* Pericytes regulate the blood–brain barrier. *Nature* **468**, 557–561 (Nov. 2010). doi:10.1038/nature09522.

197. Hambardzumyan, D. & Bergers, G. Glioblastoma: Defining Tumor Niches. *Trends in Cancer* **1**, 252–265 (Dec. 2015). doi:10.1016/j.trecan.2015.10.009.

198. Pore, N., Liu, S., Haas-Kogan, D. A., O'Rourke, D. M. & Maity, A. PTEN mutation and epidermal growth factor receptor activation regulate vascular endothelial growth factor (VEGF) mRNA expression in human glioblastoma cells by transactivating the proximal VEGF promoter. *Cancer Research* **63**, 236–241 (Jan. 2003).

199. Wen, L. *et al.* Vegf-mediated tight junctions pathological fenestration enhances doxorubicin-loaded glycolipid-like nanoparticles traversing bbb for glioblastoma-targeting therapy. *Drug Delivery* **24**, 1843–1855 (Jan. 2017). doi:10.1080/10717544.2017.1386731.

200. Chen, Z., Ross, J. L. & Hambardzumyan, D. Intravital 2-photon imaging reveals distinct morphology and infiltrative properties of glioblastoma-associated macrophages. *Proceedings of the National Academy of Sciences* **116**, 14254–14259 (July 2019). doi:10.1073/pnas.1902366116.

201. Herting, C. J. *et al.* Tumour-associated macrophage-derived interleukin-1 mediates glioblastoma-associated cerebral oedema. *Brain* **142**, 3834–3851 (Dec. 2019). doi:10.1093/brain/awz331.

202. Lamano, J. B. *et al.* Glioblastoma-Derived IL6 Induces Immunosuppressive Peripheral Myeloid Cell PD-L1 and Promotes Tumor Growth. *Clinical Cancer Research* **25**, 3643–3657 (June 2019). doi:10.1158/1078-0432.CCR-18-2402.

203. Feng, X. *et al.* Loss of CX3CR1 increases accumulation of inflammatory monocytes and promotes gliomagenesis. *Oncotarget* **6**, 15077–15094 (Mar. 2015). doi:10.18632/oncotarget.3730.

204. Chen, Z. *et al.* Monocyte depletion enhances neutrophil influx and proneural to mesenchymal transition in glioblastoma. *Nature Communications* **14**, 1839 (Apr. 2023). doi:10.1038/s41467-023-37361-8.

205. Brat, D. J. *et al.* Pseudopalisades in Glioblastoma Are Hypoxic, Express Extracellular Matrix Proteases, and Are Formed by an Actively Migrating Cell Population. *Cancer Research* **64**, 920–927 (Feb. 2004). doi:10.1158/0008-5472.CAN-03-2073.

206. Markwell, S. M., Ross, J. L., Olson, C. L. & Brat, D. J. Necrotic reshaping of the glioma microenvironment drives disease progression. *Acta Neuropathologica* **143**, 291–310 (Mar. 2022). doi:10.1007/s00401-021-02401-4.

207. Bar, E. E., Lin, A., Mahairaki, V., Matsui, W. & Eberhart, C. G. Hypoxia Increases the Expression of Stem-Cell Markers and Promotes Clonogenicity in Glioblastoma Neurospheres. *The American Journal of Pathology* **177**, 1491–1502 (Sept. 2010). doi:10.2353/ajpath.2010.091021.

208. Sattiraju, A. *et al.* Hypoxic niches attract and sequester tumor-associated macrophages and cytotoxic T cells and reprogram them for immunosuppression. *Immunity* **56**, 1825–1843.e6 (Aug. 2023). doi:10.1016/j.immuni.2023.06.017.

209. Watkins, S. *et al.* Disruption of astrocyte–vascular coupling and the blood–brain barrier by invading glioma cells. *Nature Communications* **5**, 4196 (June 2014). doi:10.1038/ncomms5196.

210. Barthel, L. *et al.* Glioma: molecular signature and crossroads with tumor microenvironment. *Cancer and Metastasis Reviews* **41**, 53–75 (Mar. 2022). doi:10.1007/s10555-021-09997-9.

211. Al-Dalahmah, O. *et al.* Re-convolving the compositional landscape of primary and recurrent glioblastoma reveals prognostic and targetable tissue states. *Nature Communications* **14**, 2586 (May 2023). doi:10.1038/s41467-023-38186-1.

212. Markovic, D. S. *et al.* Gliomas induce and exploit microglial MT1-MMP expression for tumor expansion. *Proceedings of the National Academy of Sciences* **106**, 12530–12535 (July 2009). doi:10.1073/pnas.0804273106.

213. Greenwald, A. C. *et al.* Integrative spatial analysis reveals a multi-layered organization of glioblastoma. *Cell* **187**, 2485–2501.e26 (May 2024). doi:10.1016/j.cell.2024.03.029.

214. Tzec-Interián, J. A., González-Padilla, D. & Góngora-Castillo, E. B. Bioinformatics perspectives on transcriptomics: A comprehensive review of bulk and single-cell RNA sequencing analyses. *Quantitative Biology* **13**, e78 (2025). doi:10.1002/qub2.78.

215. Stark, R., Grzelak, M. & Hadfield, J. RNA sequencing: the teenage years. *Nature Reviews Genetics* **20**, 631–656 (Nov. 2019). doi:10.1038/s41576-019-0150-2.

216. Hu, B., Sajid, M., Lv, R., Liu, L. & Sun, C. A review of spatial profiling technologies for characterizing the tumor microenvironment in immuno-oncology. *Frontiers in Immunology* **13**, (Oct. 2022). doi:10.3389/fimmu.2022.996721.

217. Haas, B. J., Chin, M., Nusbaum, C., Birren, B. W. & Livny, J. How deep is deep enough for RNA-Seq profiling of bacterial transcriptomes?. *BMC Genomics* **13**, 734 (Dec. 2012). doi:10.1186/1471-2164-13-734.

218. Li, X. & Wang, C.-Y. From bulk, single-cell to spatial RNA sequencing. *International Journal of Oral Science* **13**, 1–6 (Nov. 2021). doi:10.1038/s41368-021-00146-0.

219. Lähnemann, D. *et al.* Eleven grand challenges in single-cell data science. *Genome Biology* **21**, 31 (Feb. 2020). doi:10.1186/s13059-020-1926-6.

220. Kim, N., Kang, H., Jo, A., Yoo, S.-A. & Lee, H.-O. Perspectives on single-nucleus RNA sequencing in different cell types and tissues. *Journal of Pathology and Translational Medicine* **57**, 52–59 (Jan. 2023). doi:10.4132/jptm.2022.12.19.

221. Lake, B. B. *et al.* Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* **352**, 1586–1590 (June 2016). doi:10.1126/science.aaf1204.

222. Denisenko, E. *et al.* Systematic assessment of tissue dissociation and storage biases in single-cell and single-nucleus RNA-seq workflows. *Genome Biology* **21**, 130 (June 2020). doi:10.1186/s13059-020-02048-6.

223. Bakken, T. E. *et al.* Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. *PLOS ONE* **13**, e0209648 (Dec. 2018). doi:10.1371/journal.pone.0209648.

224. Im, Y. & Kim, Y. A Comprehensive Overview of RNA Deconvolution Methods and Their Application. *Molecules and Cells* **46**, 99–105 (Feb. 2023). doi:10.14348/molcells.2023.2178.

225. Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E. & Storey, J. D. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* **28**, 882–883 (Mar. 2012). doi:10.1093/bioinformatics/bts034.

226. Nguyen, H., Nguyen, H., Tran, D., Draghici, S. & Nguyen, T. Fourteen years of cellular deconvolution: methodology, applications, technical evaluation and outstanding challenges. *Nucleic Acids Research* **52**, 4761–4783 (May 2024). doi:10.1093/nar/gkae267.

227. Avila Cobos, F., Alquicira-Hernandez, J., Powell, J. E., Mestdagh, P. & De Preter, K. Benchmarking of cell type deconvolution pipelines for transcriptomics data. *Nature Communications* **11**, 5650 (Nov. 2020). doi:10.1038/s41467-020-19015-1.

228. Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods* **12**, 453–457 (May 2015). doi:10.1038/nmeth.3337.

229. Newman, A. M. *et al.* Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nature Biotechnology* **37**, 773–782 (July 2019). doi:10.1038/s41587-019-0114-2.

230. Wang, X., Park, J., Susztak, K., Zhang, N. R. & Li, M. Bulk tissue cell type deconvolution with multi-subject single-cell expression reference. *Nature Communications* **10**, 380 (Jan. 2019). doi:10.1038/s41467-018-08023-x.

231. Momeni, K., Ghorbian, S., Ahmadpour, E. & Sharifi, R. Unraveling the complexity: understanding the deconvolutions of RNA-seq data. *Translational Medicine Communications* **8**, 21 (Sept. 2023). doi:10.1186/s41231-023-00154-8.

232. Peng, X. L., Moffitt, R. A., Torphy, R. J., Volmar, K. E. & Yeh, J. J. De novo compartment deconvolution and weight estimation of tumor samples using DECODER. *Nature Communications* **10**, 4729 (Oct. 2019). doi:10.1038/s41467-019-12517-7.

233. Sturm, G. *et al.* Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. *Bioinformatics* **35**, i436–i445 (July 2019). doi:10.1093/bioinformatics/btz363.

234. Becht, E. *et al.* Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biology* **17**, 218 (Oct. 2016). doi:10.1186/s13059-016-1070-5.

235. Method of the Year 2020: spatially resolved transcriptomics. *Nature Methods* **18**, 1–1 (Jan. 2021). doi:10.1038/s41592-020-01042-x.

236. Ståhl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78–82 (July 2016). doi:10.1126/science.aaf2403.

237. Vickovic, S. *et al.* High-definition spatial transcriptomics for in situ tissue profiling. *Nature Methods* **16**, 987–990 (Oct. 2019). doi:10.1038/s41592-019-0548-y.

238. Rodriques, S. G. *et al.* Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363**, 1463–1467 (Mar. 2019). doi:10.1126/science.aaw1219.

239. Merritt, C. R. *et al.* Multiplex digital spatial profiling of proteins and RNA in fixed tissue. *Nature Biotechnology* **38**, 586–599 (May 2020). doi:10.1038/s41587-020-0472-9.

240. Rudkin, G. T. & Stollar, B. D. High resolution detection of DNA–RNA hybrids in situ by indirect immunofluorescence. *Nature* **265**, 472–473 (Feb. 1977). doi:10.1038/265472a0.

241. Duraiyan, J., Govindarajan, R., Kaliyappan, K. & Palanisamy, M. Applications of immunohistochemistry. *Journal of Pharmacy & Bioallied Sciences* **4**, S307–S309 (Aug. 2012). doi:10.4103/0975-7406.100281.

242. Hu, B., Zhu, J. & Zhao, F. The evolving landscape of spatial proteomics technologies in the AI age. *Fundamental Research,* (Dec. 2024). doi:10.1016/j.fmre.2024.11.023.

243. Method of the Year 2024: spatial proteomics. *Nature Methods* **21**, 2195–2196 (Dec. 2024). doi:10.1038/s41592-024-02565-3.

244. Bodenmiller, B. Highly multiplexed imaging in the omics era: understanding tissue structures in health and disease. *Nature Methods* **21**, 2209–2211 (Dec. 2024). doi:10.1038/s41592-024-02538-6.

245. Chang, Q. *et al.* Imaging Mass Cytometry. *Cytometry Part A* **91**, 160–169 (2017). doi:10.1002/cyto.a.23053.

246. Schürch, C. M. *et al.* Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front. *Cell* **182**, 1341–1359.e19 (Sept. 2020). doi:10.1016/j.cell.2020.07.005.

Chapter 2

GBMdeconvoluteR accurately infers proportions of neoplastic & immune cell populations from bulk glioblastoma transcriptomics data

Shoaib Ajaib[1], Disha Lodha[1,2], Steven Pollock[1], Gemma Hemmings[1], Martina A. Finetti[1], Arief Gusnanto[3], Aruna Chakrabarty[4], Azzam Ismail[4], Erica Wilson[1], Frederick S Varn[5], Bethany Hunter[6], Andrew Filby[6], Asa A. Brockman[7], David McDonald[6], Roel GW Verhaak[5], Rebecca A. Ihrie[7], Lucy F. Stead[1]

[1]Leeds Institute of Medical Research, University of Leeds
[2]EMBL's European Bioinformatics Institute (EMBL-EBI)
[3]School of Mathematics, University of Leeds
[4]Department of Neuropathology, Leeds Teaching Hospitals NHS Trust
[5]The Jackson Laboratory for Genomic Medicine, Farmington, CT, US
[6]Flow Cytometry Core Facility, Newcastle University
[7]Department of Cell & Developmental Biology, Vanderbilt University School of Medicine; Vanderbilt Brain Institute, Vanderbilt-Ingram Cancer Center, Department of Neurological Surgery, Vanderbilt University Medical Center

Abstract

Background: Characterising and quantifying cell types within glioblastoma (GBM) tumours at scale will facilitate a better understanding of the association between the cellular landscape and tumour phenotypes or clinical correlates. We aimed to develop a tool that deconvolutes immune and neoplastic cells within the GBM tumour microenvironment (TME) from bulk RNA sequencing (RNA-seq) data.

Methods: We developed an isocitrate dehydrogenase wild-type (IDHwt) GBM-specific single immune cell reference consisting of B cells, T cells, natural killer cell (NK cell), microglia, tumour-associated macrophages (TAMs), monocytes, mast and dendritic cells (DCs). We used this alongside an existing neoplastic single cell-type reference for astrocyte-like (AC-like), oligodendrocyte progenitor-like (OPC-like), neural progenitor-like (NPC-like) and mesenchymal-like (MES-like) GBM cancer cells to create both marker and gene signature matrix-based deconvolution tools. We applied single-cell resolution imaging mass cytometry (IMC) to ten IDHwt GBM samples, five paired primary and recurrent tumours, to determine which deconvolution approach performed best.

Results: Marker based deconvolution using GBM tissue specific markers was most accurate for both immune cells and cancer cells, so we packaged this approach as GBMdeconvoluteR. We applied GBMdeconvoluteR to bulk GBM RNA-seq data from the cancer genome atlas (TCGA) and recapitulated recent findings from multi-omics single cell studies with regards to associations between MES-like GBM cancer cells and both lymphoid and myeloid lineage immune cells. Furthermore, we expanded upon this to show that these associations are stronger in patients with worse prognosis.

Conclusions: GBMdeconvoluteR accurately quantifies immune and neoplastic cell proportions in IDHwt GBM bulk RNA-seq data and is accessible here: https://gbmdeconvoluter.leeds.ac.uk

2.1 Introduction

GBM brain tumours consist of a multitude of different neoplastic and non-neoplastic cell types[1]. The specific cancer cell subtypes within a GBM are directly influenced by the cellular composition of the microenvironment, which also has a role in shaping the progression of the tumour and its adaption to stressors including treatment[2–4]. It is of paramount importance to accurately characterise the cellular make-up of GBM tumours. This will enable us to understand the phenotypes associated with changing cell landscapes within individual tumours, and to assess correlation between specific cell populations and the efficacy of new treatments, particularly immunotherapies. Whilst single cell and spatial- profiling approaches currently offer the highest resolution of cellular deconvolution, they are technically challenging, and prohibitively costly for larger sample numbers.

Instead, approaches that propose to quantify cell types from bulk tissue RNA-seq data have become increasingly popular[5–9]. These can be split into two main types: those that employ a full cell-type gene expression signature matrix; and those based on marker genes for specific cell types. A widely-adopted implementation of the former approach is CIBERSORTx[9], which was recently used to delineate pan-glioma cell types[3]. However, key studies have shown that the accuracy of any gene expression-based computational deconvolution tool is mostly derived from the signature matrix, or marker genes, underpinning it, which must be derived from the tissue of interest[5,10,11]. We, thus, decided to create a tool that can specifically quantify cancer cell types, as delineated by Neftel et al[2], and immune cell types from bulk IDHwt GBM tumour sequencing data. We developed this tool by amalgamating four independent single-cell GBM datasets to derive signature matrices for use with CIBERSORTx and marker genes for use with MCPcounter. The latter was chosen as it has been benchmarked as one of the most accurate marker gene-based tools available, giving consistently high correlation with ground truths across cell types[12]. We then compared results from these GBM-specific programmes to those from orthogonal cell quantification, using single cell-resolution IMC, on the same IDHwt GBM samples. We included both primary and recurrent GBM samples in our tool development and validation, to enable separate quantification of accuracy in longitudinal samples. We found that the MCPcounter based tool performed best at delineating both immune and neoplastic cancer cell populations and have made this publicly available as an easily accessible, online tool: GBMdeconvoluteR.

2.2 Materials & methods

All statistical analyses were carried out using the R statistical software package version 4.2.0. The name of each test used, and level of significance achieved, is included within the results where the finding from each hypothesis test is confirmed. Plotting was done using ggplot2 (version 3.3.6).

2.2.1 Dataset selection

Four single cell datasets were identified from literature searches (Table 2.1)[13–16]. The inclusion criteria were single-cell RNA sequencing (scRNA-seq) or single-nucleus RNA sequencing (snRNA-seq) expression data from human IDHwt GBM samples. Data had to be available as raw counts.

2.2.2 scRNA-seq data preprocessing

The Seurat R package (version 4.1.1) was used for all pre-processing, integration, clustering, and annotation tasks[17]. Whilst GSE163120 has a single accession code, it contains data from both primary and recurrent sample cells that were sequenced on different platforms so these were processed separately.

2.2.3 Copy-number variant analysis

Single cell datasets were amalgamated. Neoplastic cells were filtered, as has been done previously, by inferring and removing those with large-scale copy number variations (CNVs) such as chromosome 7 gain and chromosome 10 loss (Ch +7/-10) using the inferCNV R package (version 1.3.3)[18,19]. The inferCNV object was created using *CreateInfercnvObject()* taking the raw counts (stored in the *RNA* assay of the Seurat object) for each dataset. Annotations were not provided, instead each dataset was grouped according to sample (i.e. patient). The gene ordering file used was derived using the annotations from Ensembl Genes 91 for Human build 38 (GRCh38), taking the gene name, chromosome, and gene span. The *ref_group_names* argument was set to NULL, to average signal across all cells to define the baseline. The *run()* function was then used to perform InferCNV operations to reveal the CNV signal. A cut-off value of 1 was used for all the datasets apart from GSE163120, where a value of 0.1 was used as suggested by the documentation for InferCNV.

Table 2.1 | Single-cell IDH wildtype GBM datasets used as a reference set for this paper.

Accession	Samples	Platform
GSE141383[13]	Single cell RNAseq of ∼18k cells from 5 primary IDHwt GBM.	Automated microwell array capture and full length RNAseq.
GSE163120[14]	Single cell RNAseq of ∼21k cells from primary and ∼43k cells from recurrent IDHwt GBMs.	10X Genomics GemCode capture and 3' or 5' RNAseq.
GSE135437[15]	Single cell RNAseq of 769 cells from 4 IDHwt GBMs.	Single cell sorting and 3' RNAseq.
GSE138794[16]	Single-cell/nuclei RNA-sequencing of ∼11k single cells from 4 IDHwt primary GBMs.	10X Genomics Chromium capture and 3' RNAseq.

2.2.4 Quality control filtering

Each dataset underwent individual quality control (QC) in which metrics were used to filter out poor quality cells according to dataset-determined thresholds (Table S2.1): the number of reads, or unique molecular identifiers (nUMI_min); the number of non-zero count genes (nGene); the percentage of mitochondrial genes (mitochondial_ratio_min); the percentage of ribosomal genes; and the cell complexity (gene_complexity_min), which is a composite measure derived as:

$$\frac{\log_{10}(\text{nGene})}{\log_{10}(\text{nUMI\_min})}$$

2.2.5 Dataset normalization

Post-filtering, each dataset was normalised individually using *SCTransform*, whilst regressing out dataset-specific confounding sources of variation such as ribosomal/mitochondrial ratio using the *vars.to.regress* function argument. Moreover, due to the disparity in the total number of cells in each dataset, a different number of variable features were passed to the *variable.features.n* function argument. The specific normalisation criteria for each dataset are in Table S2.2.

2.2.6 Dataset integration

The *FindIntegrationAnchors* function was applied to the list of *SCTransform* normalised datasets to identify cross-dataset pairs of cells that were in a matched biological state. These "anchors" were then used with *IntegrateData* to merge all the datasets together[17]. The *normalization.method* argument was set as *SCT* for both *FindIntegrationAnchors* and *IntegrateData*.

2.2.7 Clustering & cell type assignment

Dimensionally reduction was performed on the integrated datasets using principal component analysis (PCA) using *RunPCA* with default settings. This was followed by uniform manifold approximation and projection (UMAP) which was implemented using *RunUMAP* with custom parameters a=0.6 and b=0.75. shared nearest-neighbour (SNN) graphs were constructed based on Euclidean distance using *FindNeighbours*; taking the default k (k=20), the first 30 principal components and using the *rann* method for finding nearest neighbours. Clusters were identified using *FindClusters*, with the smart local moving (SLM) algorithm used for cluster optimization[20].

2.2.8 Cell type annotation

Cell counts per cluster, for each clustering resolution parameter (0.1 – 0.8 in 0.1 increments) were cross tabulated with immune cell type labels transferred from dataset GSE163120. The 0.7 resolution cross-tabulation (Table S2.3) was used, based on cluster robustness and stability[21], to assign cell-type annotation labels to clusters where the majority of cells had labels for either one distinct cell type or and/or where the cells were labelled were unknown. The T cell, NK cell and TAM labelled clusters could not be assigned and were sub-clustered to further resolve them. This constituting isolation of these cells and repeat of the above methodology, from the point of having normalised data, to separate cell types.

2.2.9 Deriving GBM immune & neoplastic cell profiles

Immune cell marker genes were identified from the integrated, clustered and annotated data using the scran R package (version 1.2.2)[22]. The *findMarkers* function was used to identify candidate marker genes by testing for those that were differen-

tially expressed (DE) between pairs of clusters using both *t-test* and *Wilcoxon* rank sum tests. Both "all" and "any" *pval.type* arguments were used to identify genes which were DE between any two clusters and highly ranked/significantly upregulated genes for a given cluster ("all") or significantly upregulated compared with all other clusters ("any"). The *multiMarkerStats* function was then used to combine multiple sets of marker statistics. Neoplastic GBM cell marker genes were taken directly from Neftel *et al.*[2] but were filtered to remove non GBM tumour-intrinsic (TI) genes, to negate the noise that would result from expression of these in the TME[23]. Marker genes for a variety of GBM neoplastic and non-neoplastic cell types have recently been made available as a resource entitled GBMap. We downloaded these directly from the supplementary data of the accompanying preprint for testing within MCPcounter (denoted MCPcounter$_{GBMap}$)[24]. The neoplastic cell markers from GBMap were also filtered to only include GBM TI genes.

2.2.10 CIBERSORTx reference expression profile

The single cell data used to derive the neoplastic expression profiles used with CIBERSORTx was obtained from the Gene Expression Omnibus (GSE131928). These data comprised ∼23,000 cells which were filtered to include only adult GBM samples. Each cell came with a score corresponding to 6 neoplastic cell states: these were converted to four states and then each cell was assigned to a neoplastic cell state or as a hybrid as described in Neftel et al[2]. The neoplastic single cell data was combined with the labelled immune single cells and then randomly downsampled such that the total number of cells in the resulting reference matrix was 5075 and of roughly equal class type (Table 2.2)[25].

2.2.11 Validation samples

Ten human GBM samples were used for validation via bulk RNA-seq and IMC. These were *de novo primary* IDHwt GBM that had been stored in formalin-fixed paraffin-embedded (FFPE) blocks, and the matched locally *recurrent* samples following initial de-bulking surgery and treatment with radiation and temozolomide (TMZ) chemotherapy.

Table 2.2 | Cell types and numbers used to derive the CIBERSORTx signature matrix.

Cell type	No. of cells
AC	458
B cells	458
DC	458
Mast cells	88
MES	458
Microglia	458
Monocytes	458
NK cells	458
NPC	458
OPC	407
T cells	458
TAM	458
Total	**5075**

2.2.12 Ethics statement

Samples were from patients at the Walton Centre, UK, that provided informed consent in writing for the use of their tissue in research. The inclusion of these samples in this project was following approval by the UK National Health Service's Research Ethics Service Committee South Central - Oxford A (Research Ethics Code: 13/SC/0509).

2.2.13 Bulk RNA-seq

Ribonucleic acid (RNA) was extracted from neuropathologist annotated regions containing >60% cancer cells using Qiagen kits (Qiagen, Sussex, UK). Paired-end, 100-base pairs (bp) strand-specific whole transcriptome libraries were prepared using the Illumina NEBNext Ultra Directional RNA Library Prep Kit (New England BioLabs, Herfordshire, UK). Following ribosomal ribonucleic acid (rRNA) depletion with the NEBNext rRNA kit or Ribo-Zero Gold. Libraries were sequenced on an Illumina NextSeq2000. RNA-seq data was processed as previously described[26].

2.2.14 Imaging mass cytometry

Antibody selection

A panel of 33 antibodies targeting neoplastic and immune cell subtypes in GBM was assembled based on literature and manufacturer sources (see Table 2.3 and Table S2.4). Neoplastic markers were selected from overlapping cancer cell signatures reported in three independent single-cell studies, including Neftel *et al.*[2,16,27]. Antibodies were prioritised as follows: pre-conjugated and previously used in IMC of GBM or brain; previously used via bespoke conjugation; carrier-free and validated for immunohistochemistry (IHC) or immunocytochemistry (ICC) in brain or GBM; or simply available in carrier-free format.

A set of panel-wide control tissues was determined: spleen, brain, tonsil, prostate, bone marrow, skin and uterus. Control tissue samples from at least two individuals were amalgamated into a multi-tissue FFPE block. Multi-tissue block sections were used in IHC validation and testing of three antibody concentrations at, above and below those recommended by the manufacturer. Chosen antibody concentrations and control tissue(s) relevant to each antibody are in Table S2.4. Antibody conjugation and staining and IMC took place at the Flow Cytometry Core Facility at Newcastle University. Conjugation was performed using MaxPar metal labelling kits using X8 polymer according to standard manufacturers protocols (with the exception of Gd157 which was obtained by Trace Sciences International and was diluted to 0.1M prior to use with MaxPar reagents). Conjugations were validated by capture on Thermo AbC beads prior to acquisition on a helios mass cytometer.

Sample preparation & mass cytometry

5μm sections, taken consecutively from the same blocks that underwent bulk RNA-seq (see above), were stained with a cocktail of all 33 conjugated antibodies after dewaxing (Xylene) and HIER antigen retrieval in Tris-EDTA (pH9) with 0.5% Tween 20. Sections were incubated for 30 minutes in 0.3μM iridium to counterstain the nuclei prior to air drying. A minimum of three 2mm^2 regions of interest (ROI) were annotated per sample within the area corresponding to that from which RNA was extracted from the adjacent sections. Images were generated on the Hyperion Tissue Imaging cytometer by ablation of the ROI at a 200Hz frequency with a 1μm diameter laser. Raw MCD files were created and exported as ome-tiff from MCD Viewer software (Fluidigm).

Table 2.3 | Antibodies used in IMC.

Marker	Cell category	Type/State	Antibody clone(s)
ANXA1	GBM	Hypoxia MES-like	EPR19342/abcam
ANXA2	GBM	Hypoxia MES-like	MAB3928/RnD
BCAN	GBM	NPC-like	S294A-6/Thermo
CD3	Immune	T cell	Fluidigm/3170019D
CD31	Normal	Vasculature	Fluidigm/EPR3094
CD45	Immune	Pan-immune	Fluidigm/3152016D
CD8	Immune	T cell	SK1/Biolegend
CHI3L1	GBM	MES-like	EPR19078-157/abcam
DNA	All	-	Fluidigm
DLL3	GBM	NPC-like	EPR22592-18/abcam
EZH2	All	Chromatin remodeller	EPR9307(2)/abcam
GFAP	Normal	Astrocyte	ab218309 /abcam
HIF1A	All	Hypoxia	16H4L13/Thermo
HOPX	GBM	AC-like	ab230544
IBA1	Immune	Pan-macrophage	EPR16588 /abcam
JARID2-C	All	Chromatin remodeller	Developed in house
JARID2-N	All	Chromatin remodeller	EPR6357/abcam
Ki67	All	Proliferation	B56/Fluidigm
MOG	Normal	Oligodendrocyte	MA5-24644/Thermo
CD56	Normal	Immature Neuron	HCD56/Biolegend
NeuN	Normal	Mature Neuron	1B7/Biolegend
NKp46	Immune	NK cell	MAB1850/RnD systems
OLIG1	GBM	OPC-like	MAB2417/R&D
P2Y12R	Immune	Microglia	EPR23511-72/abcam
SCD5	GBM	OPC-like	PA5-59963/Thermo
SLC1A3	GBM	AC-like	EPR12686/abcam
SMA	Normal	Vasculature	1A4/R&D
SNAI1	GBM	EMT	AF3639/R&D
SOD2	GBM	MES	EPR2560Y/abcam
SOX2	GBM	GSC	O30-678/Fluidigm
TGFbeta	GBM	GSC	TW4-6H10/Fluidigm
TMEM119	Immune	Microglia	HPA051870/sigma
TNC	GBM	GSC	MAB2138/R&D

Image pre-processing

Following export, the raw data were converted from to ome-tiff format and segmented into single cells using the steinbock pipeline comprised of the following steps[28]. Pixel classification was done using Ilastik (version 1.3.3): Tiff stacks were generated for each of the proteins in the panel and pixels classified into two channels as either nuclear, or background. These were used to train a random forest

classifier, which returned probability masks for each image. The generated probability maps were processed to create single-cell masks using the image analysis software CellProfiler (version 4.1.3). First, probabilities were histogram-equalized (256 bins and kernel size of 17), and then a Gaussian filter was applied to enhance contrast and smooth the probabilities. Subsequently, an Otsu two-class thresholding approach was used to segment nuclear masks. Cell masks were derived from an expansion of nuclear masks using a maximum expansion of 3 pixels. The CellProfiler single cell masks were ultimately overlaid onto the single-cell segmentation masks and single-channel tiff images of all measured channels to extract single-cell marker expression means. The single-cell data was read into R using *read_steinbock* from the imcRtools R package (version 1.2.3) and the expression counts were transformed using an inverse hyperbolic sine function (arcsinh) with *cofactor = 5*. The expression counts were corrected for channel spillover using a non-negative least squares method as previously described[29]. Briefly, each metal-conjugated antibody was spotted on an agarose-coated slide, and this was ablated to generate a background signal which could be used for compensation using the R Bioconductor package CATALYST (version 1.20.1).

Image analysis

All downstream data visualisations, including image and cell segmentation QC were completed using the cytomapper (version 1.8.0) and dittoseq (version 1.8.1) R packages[30]. Batch effect correction of segmented cells was completed using harmony (version 0.1.0)[31]. Cells were clustered based on their similarity in marker expression using the PhenoGraph clustering algorithm (k =45) implemented in Rphenograph (version 0.99.1)[32]. Cluster IDs were mapped on top of UMAP embeddings (*n_neighbors* = 40) derived using the uwot R package (version 0.1.11). Cell type classification was completed using marker enrichment modelling, implemented in the MEM R packages (version 2.0.0), selecting for markers with enrichment scores equal to or greater than 3 (*display.thresh* = 3)[33] for the first clustering, which defined immune cells. Further sub-clustering was required to annotate neoplastic cells with *display.thresh* relaxed to 2 (Table S2.5).

Creating and comparing deconvolution approaches

MCPcounter was run via the R Package (version 1.2.0) in two modes: default mode($MCP_{default}$) used the universal set of 110 immune cell-type marker genes that come provided as standard, meaning no neoplastic cell populations were in-

cluded; GBM mode (MCP_{GBM}) used the GBM-specific neoplastic and immune cell marker genes derived as outlined above. The *Create Signature Matrix* module of CIBERSORTx was run with default parameters and quantile normalization disabled, to create a signature matrix using the single-cell-derived immune and neoplastic expression profiles detailed above. This signature matrix was then used to infer cell fractions of bulk RNA-seq sample mixtures using the CIBERSORTx high-resolution Docker container:

https://hub.docker.com/r/cibersortx/hires.

For all runs, the bulk RNA-seq dataset was input as the *mixture* file and the respective signature matrix was input as the *sigmatrix* file. For all runs, the Batch correction was done in *S-mode* by setting the *rmbatchSmode* parameter to TRUE and the input signature matrix's respective CIBERSORTx-created *source gene expression profile* was input. Finally, absolute mode was set to FALSE for all runs. Cell population quantities inferred from the GBM sample RNA-seq for all expression-based deconvolution approaches were compared with those from the IMC using the Pearson correlation coefficient.

Application to TCGA data

TCGA data was obtained from the Genomics Data Commons Data Portal:

https://portal.gdc.cancer.gov.

The data were filtered on the *data_category* and *data_type* fields to only include *transcriptome profiling* and *Gene Expression Quantification* data, respectively. Further, only primary, IDHwt GBM cases treated with standard/non-standard TMZ chemoradiation were selected. The expression values for the 93 samples were transcripts per million (TPM) normalised counts that were combined into an expression matrix. This matrix was input to GBMdeconvoluteR, which was run using our GBM specific marker genes. Outputted scores were used in correlation analysis using the *cor()* and *cor.test()* functions from base R stats package. The quartiles of overall survival (OS) were calculated and used to extract patients with a worse (OS less than the lower quartile of 8.55 months) or better (OS greater than the lower quartile of 20.55 months) prognosis. Plots were generated using the ggplot2 R package.

Developing GBMdeconvoluteR

GBMdeconvoluteR was developed as an interactive web application using the Shiny R package (version 1.7.1) and packaged as a portable container image using the rocker/shiny:latest base Docker image. The custom image was stored in the Azure Container Registry and deployed using the Azure App Service. All code can be found at:

https://github.com/GliomaGenomics/GBMDeconvoluteR

2.3 Results

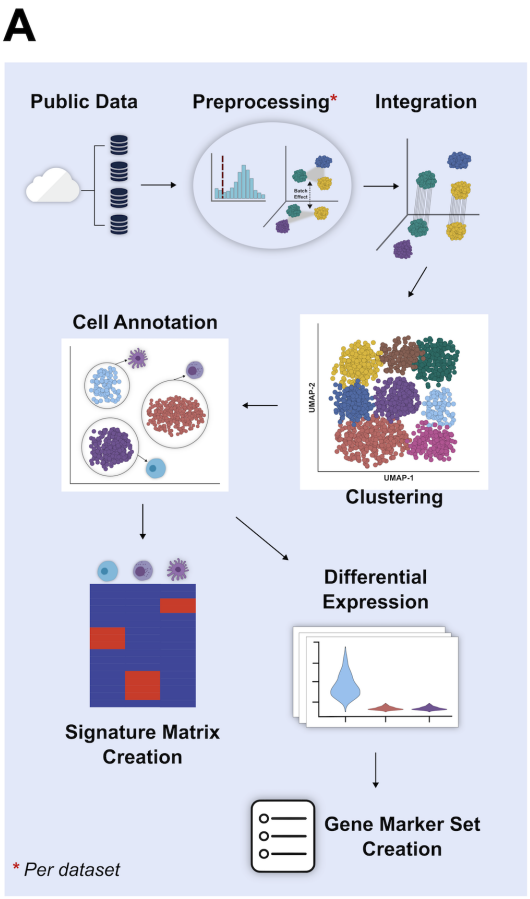
2.3.1 Identifying GBM-specific cell type profiles

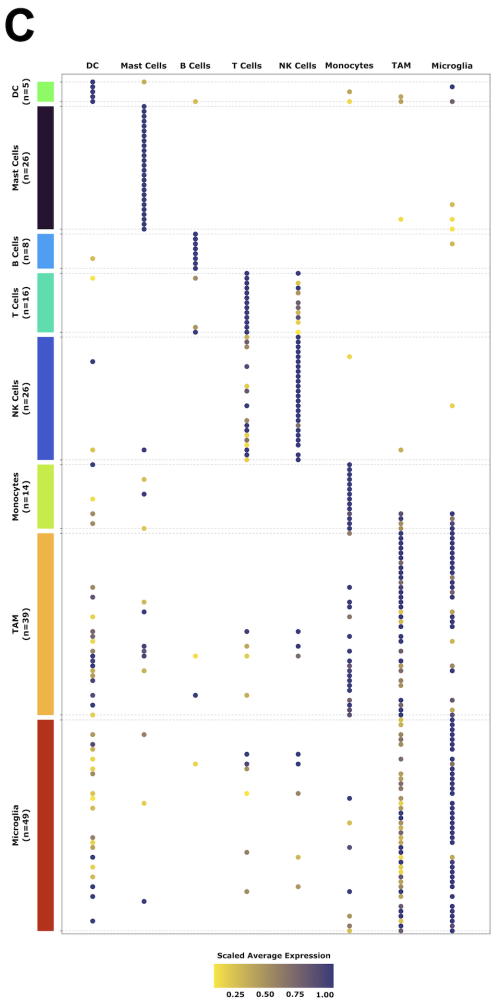
Four independent single cell GBM datasets (Figure 2.1) were used to derive marker genes, or signature gene expression matrices, for GBM tumour-infiltrating immune cells: B cells, T cells, NK cells, microglia, TAMs, monocytes, mast and DCs. Figure 2.1A outlines the process. Datasets underwent pre-processing independently to filter out poor quality cells and copy number analysis to remove neoplastic cells, before being amalgamated. There were significant batch effects owing to different sequencing platforms and originating centres but these were effectively removed using regularized negative binomial regression[34] (Figure 2.1B and Figure S2.1A). One dataset (GSE163120) included the immune cell annotations determined by the original study. This information was used to guide clustering, with optimisation focused first on maximising cluster stability and then on the best separation of pre-annotated cell types[21]. Owing to the difficulty in separating immune types that are known to have similar and overlapping gene expression profiles (namely TAMs and microglia; and NK cell and T cells) cells assigned to any of these groupings were isolated and further sub-clustered, resulting in definitive cluster annotations (Figure 2.1B and Figure S2.1B).

GBM-specific marker genes for each immune cell type were then derived by using differential expression analysis to highlight the top 25 genes, per annotated cluster, that were uniquely or predominantly expressed in that cluster, and visually checking these to identify specific cell type markers corresponding to each immune cell type (Figure 2.1C and Table S2.6). Marker genes for GBM cancer cell subtypes were adopted from Neftel *et al.*[2]. In that study, four neoplastic GBM cell types were delineated from single cell data. We extracted the marker genes that Neftel *et al.* showed to delineate the four subtypes, but then removed those that

are also expressed in the GBM TME, and would therefore confound the results of application to bulk tissue profiles[23] (Table S2.7).

Single cell expression profiles for annotated GBM-associated immune cells, from our combined datasets, or for annotated GBM cancer cell subtypes, from Neftel *et al.*, were amalgamated into a full gene expression matrix. This was then subsampled to produce a total of 5075 single cell gene expression profiles with roughly equal representation of each cell type (Table 2.2).


A
Public Data
Preprocessing*
Integration
Cell Annotation
Clustering
Differential Expression
Signature Matrix Creation
Gene Marker Set Creation
* Per dataset


C
DC
Mast Cells
B Cells
T Cells
NK Cells
Monocytes
TAM
Microglia
Scaled Average Expression
0.25 0.50 0.75 1.00

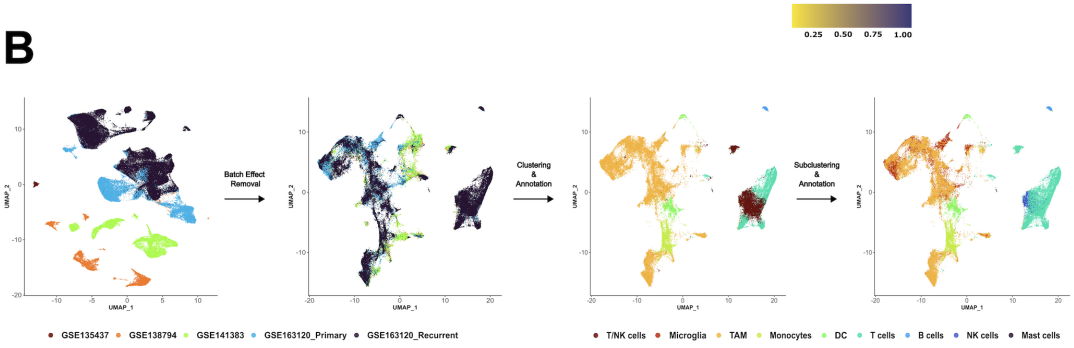
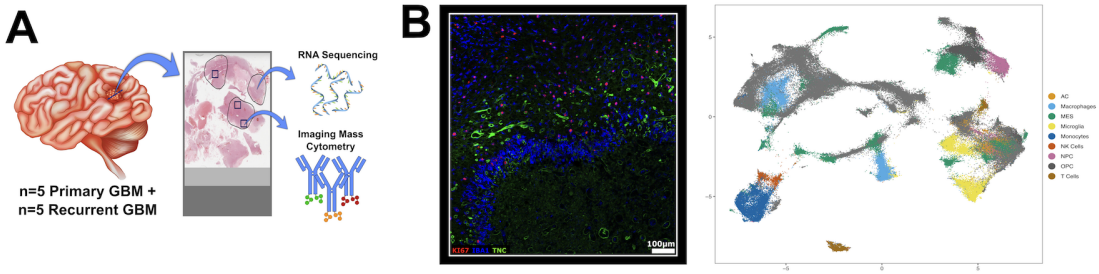

B
Batch Effect Removal
Clustering & Annotation
Subclustering & Annotation
GSE135437
GSE138794
GSE141383
GSE163120_Primary
GSE163120_Recurrent
T/NK cells
Microglia
TAM
Monocytes
DC
T cells
B cells
NK cells
Mast cells

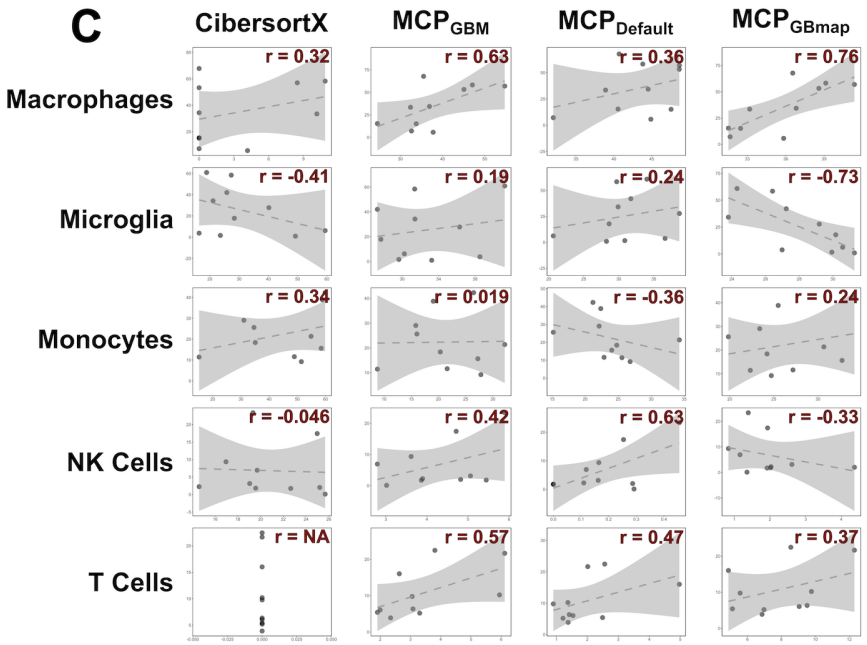
Figure 2.1 | Identifying GBM-specific cell types.

A.) The process adopted to amalgamate several independent single cell GBM datasets and create a GBM-specific immune cell reference signature gene expression matrix (for input to CIBERSORTx) or marker gene set (for input to MCPcounter). **B.)** The inherent batch effects in the amalgamated data are evident in dimensionality reduction plots where clusters initially separated by originating datasets (far left), but were removed by normalisation (middle left and Figure S1A). Initial clustering and cell type assignment of the normalised data was unable to resolve TAM and microglia, and T- and NK cell (middle right) but further sub-clustering enabled these cell types to be further delineated (far right and Figure S1B). **C.)** A dot plot showing the expression of chosen GBM-specific immune cell type markers (y-axis) in each cell type in the amalgamated single cell data (x-axis).

2.3.2 Developing & validating deconvolution approaches

Two gene-expression based computational deconvolution approaches were investigated owing to previous benchmarking studies finding them to be the best full gene expression signature matrix-based approach (CIBERSORTx) and marker gene-based approach (MCPcounter) available[12]. The approaches are distinct and give results with different interpretations. Gene expression signature matrix methods such as CIBERSORTx attempt to quantify cell types in a single sample, enabling comparison of proportions of all cell types within and between samples. Marker gene-based methods like MCPcounter instead score a single cell type for comparison of prevalence across samples; the score from cell type A cannot be compared with cell type B so within-sample comparisons of different cell types is not possible. To ascertain the accuracy of these programmes and determine which performed best, we identified five primary and matched recurrent GBM samples on which to perform both gene expression-based and IMC-based cell type deconvolution (Figure 2.2A and Figure S2.2). The latter is an approach that characterises cells, according to protein expression, at single cell resolution in tissues using up to 40 antibodies (Figure 2.2B). We assembled and validated a panel of antibodies known to distinguish tumour-infiltrating macrophages, microglia, monocytes, NK cell and T cells (Table 2.3 and Table S2.4).


A
RNA Sequencing
Imaging Mass Cytometry
n=5 Primary GBM +
n=5 Recurrent GBM
B
100µm
AC
Macrophages
MES
Microglia
Monocytes
NK Cells
NPC
OPC
T Cells


C
CibersortX
MCP GBM
MCP Default
MCP GBmap
Macrophages
Microglia
Monocytes
NK Cells
T Cells
r = 0.32
r = 0.63
r = 0.36
r = 0.76
r = -0.41
r = 0.19
r = -0.24
r = -0.73
r = 0.34
r = 0.019
r = -0.36
r = 0.24
r = -0.046
r = 0.42
r = 0.63
r = -0.33
r = NA
r = 0.57
r = 0.47
r = 0.37

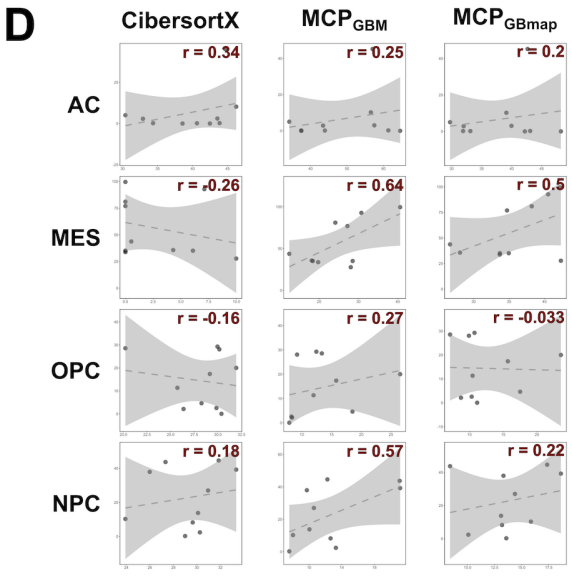

D
CibersortX
MCP GBM
MCP GBmap
AC
MES
OPC
NPC
r = 0.34
r = 0.25
r = 0.2
r = 0.26
r = 0.64
r = 0.5
r = -0.16
r = 0.27
r = -0.033
r = 0.18
r = 0.57
r = 0.22

Figure 2.2 | Validating the deconvolution approach.

A) Schematic showing how patient samples were used for validation. Regions of formalin fixed tissue sections were annotated as high tumour cell content by a neuropathologist (marked in black) and were macro-dissected for RNA-seq. At least three overlapping regions (blue squares) per sample were subjected to IMC on a consecutive section. **B)** Left: A representative image from the IMC for GBM sample 64 with three of the chosen protein markers annotated. Right: The UMAP projection of cell types assigned according to the expression of cell type protein markers quantified by IMC. **C-D)** Scatterplots of gold standard cell proportions quantified by IMC (y-axis) versus those predicted by gene expression based methods (annotated across the top) for immune (C) or neoplastic cancer (D) cell types indicated down the side. The Pearson's correlation coefficient (r) is indicated. The dotted line is the line of best fit and the shaded area denotes the confidence interval. Marker genes for MCPcounter were either default (MCPdefault), GBM-specific according to our research (MCP_{GBM}) or GBM-specific according to GBMap (MCP_{GBMap}) Neoplastic cells: AC-like; OPC-like; NPC-like and MES-like.

2.3.3 Immune cell quantification

MCPcounter can be used in default mode in which in-built canonical immune cells markers are employed. When running the programme in this mode it can only be used for immune cell estimation and we refer to it as $MCP_{default}$. In contrast, the mode using the GBM-tissue specific immune and neoplastic cell markers listed in Table S2.6 and Table S2.7 is denoted MCP_{GBM}. In addition, at the time of preparing this manuscript a larger GBM-specific single cell resource, GBMap, was made available that amalgamated 26 single cell brain and GBM datasets[24]. We, thus, also ran MCPcounter using the GBMap marker genes, denoting this as MCP_{GBMap}. We inspected the concordance between absolute cell proportions predicted by CIBERSORTx, or the relative cell type prevalence scores that resulted from each version of MCPcounter, and the quantification by IMC. We did this for all tumours together (Figure 2.2C and Table S2.8) and for primary and recurrent GBM tumours separately (Figure S2.3A and Table S2.8). Results varied across cell types but MCP_{GBM} performed best overall: it was the only approach to have positive correlations across all cell types (Figure 2.2C and Table S2.8) and had the highest average correlation coefficient (Table S2.8: across all samples, the average Pearson's r was 0.37 between IMC and MCP_{GBM} compared with 0.05 between IMC and CIBERSORTx; 0.27 between IMC and $MCP_{default}$ and 0.06 between IMC and MCP_{GBMap}).

2.3.4 Neoplastic cell quantification

The four GBM cell types described by Neftel *et al.* are delineated by gene expression[2]. Recent studies have shown that such transcriptional cell-type markers often do not translate to protein level markers for use in approaches such as IMC[35,36]. We set out to test this for the GBM neoplastic cell types, specifically. To that end, in our IMC panel we included antibodies against markers of the four neoplastic GBM cell types from Neftel *et al.*, prioritising those that overlapped with markers of GBM cancer cell subsets identified in two independent studies: Wang *et al.*[16] and Couturier *et al.*[27] (Table 2.3 and Table S2.4). These studies also identified GBM cancer cell subsets that were labelled differently but showed good agreement with the Neftel *et al.* study.

Results (Figure 2.2D and Figure S2.3B and Table S2.9) suggest that the protein markers that we selected are capable of delineating neoplastic cell types: performance varied per method and cell type but to the same degree that it did with well-established immune cell protein markers. Again, when judging performance based on correlation with IMC, MCP_{GBM} performed best overall: across all samples, the average Pearson's r was 0.43 between IMC and MCP_{GBM} compared with 0.02 between IMC and CIBERSORTx; and 0.22 between IMC and MCP_{GBMap} (Table S2.9)

2.3.5 Application to TCGA data

Our results show that MCP_{GBM} is able to accurately quantify immune and neoplastic cells in GBM tissue bulk sequencing data. To show how this can be useful in gaining biological and clinical insights from large-scale studies, we applied MCP_{GBM} to bulk RNA-seq data from 93 GBM samples from TCGA. This gave a score per cell type per sample, allowing us to quantify the correlation of cell type prevalence across patients (Figure 2.3A). Recent spatial, multi-omics studies have suggested that different neoplastic GBM cell types associate with, and are programmed by, different environmental niches of GBM tumours[4]. A key finding was that MES-like cancer cells associate with both myeloid and lymphoid compartments, whereas the remaining neoplastic cell types (AC-like, NPC-like and OPC-like cells) are significantly depleted in immune-rich regions. Our results recapitulate these findings: we observed significant, high, positive correlations between MES-like and all immune cells quantified, and significant negative correlations between the remaining neoplastic cell types. This phenomenon was more pronounced for non-MES-like neoplastic cells associated with neuronal de-

velopment (NPC-like and OPC-like cells) than for NPC-like and AC-like cells, also in keeping with the previous findings[4]. Based on the high numbers of samples in TCGA we were able to further separate patients using OS quartiles to extract worse prognosis (OS less than the lower-quartile of 8.55 months) and better prognosis (OS greater than the upper-quartile of 20.55 months) cohorts and compare score distributions (Figure 2.3B) and correlations (Figure 2.3C) in these patient subsets. The prevalence scores of cell types is not significantly different between worse or better prognosis patients (Figure 2.3B). This finding is in agreement with a recent study that defined GBM tumour subtypes based on the TME, but showed no difference in survival between them[37]. However, our results show that the correlations between cell-types are markedly different between better and worse prognosis patients (Figure 2.3C). Patients with worse prognosis have higher and more significant correlations (both negative and positive) between neoplastic and immune cell types. The TME has been shown to shape the neoplastic cell landscape over time in GBM, with more aggressive tumours being linked to greater polarity and classification of neoplastic subtypes[3,4,38]. Our results suggest that, in worse prognosis tumours, neoplastic and immune cells are more tightly associated, potentially through more direct inter-cellular communications, which could be promising therapeutic targets. These preliminary results exemplify how our tool can be used to develop new insights and hypotheses, by being applicable to large scale datasets.

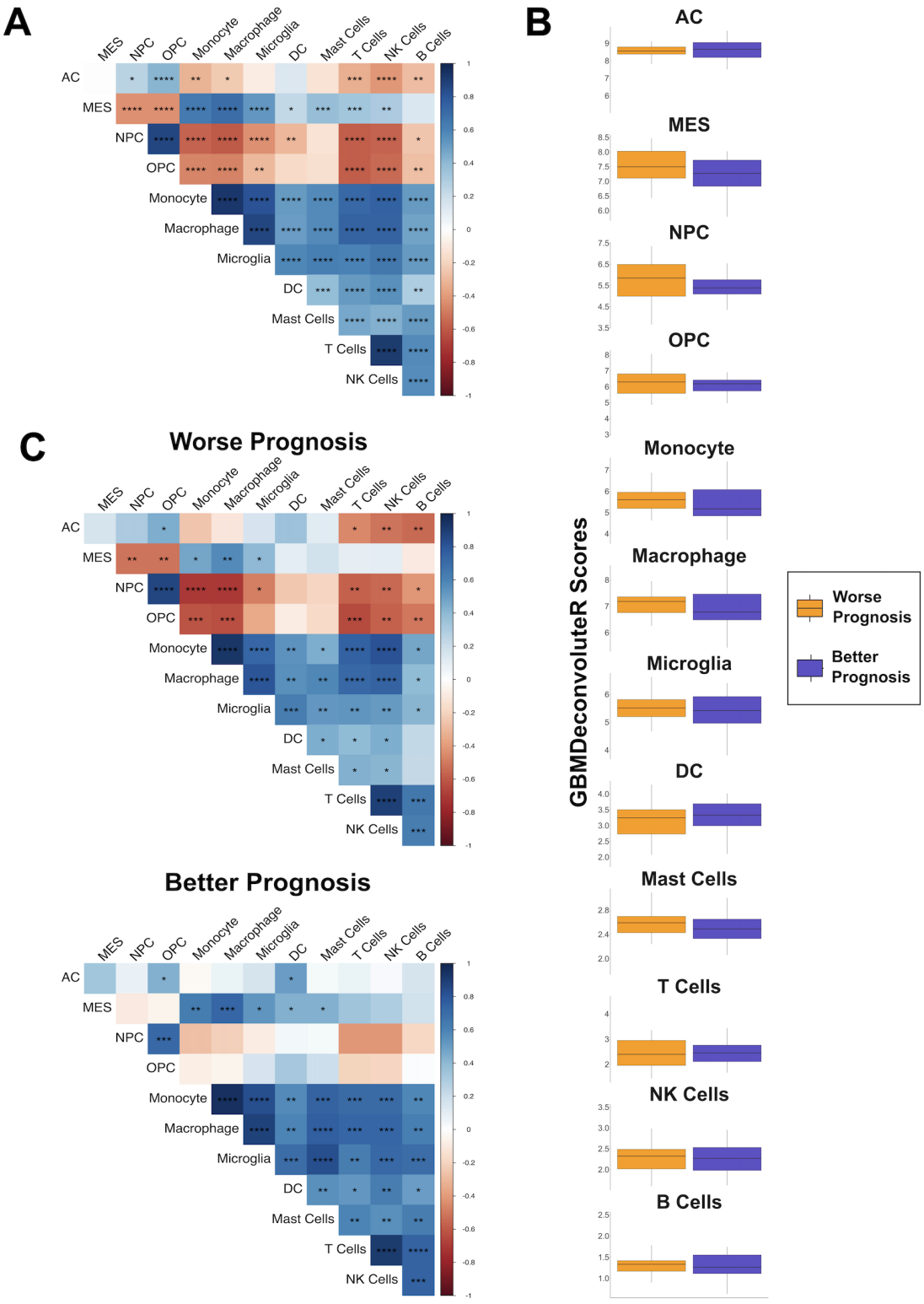

A
MES
NPC
OPC
Monocyte
Macrophage
Microglia
DC
Mast Cells
T Cells
NK Cells
B Cells
AC
C
Worse Prognosis
Better Prognosis
B
GBMDeconvoluteR Scores
Worse Prognosis
Better Prognosis

Figure 2.3 | Application of GBMdeconvoluteR to TCGA data.

MCP_{GBM} was used to score cell types in bulk GBM RNA-seq data from The Cancer Genome Atlas (TCGA). **A)** Heatmap of the correlations between cell type scores across all samples. **B)** Boxplots showing distribution of cell type scores for patients with worse or better prognosis (determined by the lower and upper quartile of overall survival, respectively). **C)** Heatmap of the correlations between cell type scores across samples from patients with worse (left) or better (right) prognosis. Significance is denoted by asterisks: *: $p<0.05$; **: $p<0.01$; ***: $p<0.001$; ****: $p<0.0001$; NS: Not Significant.

2.3.6 Incorporating additional neoplastic GBM cell types and developing GBMdeconvoluteR

To make MCP_{GBM} available to the neuro-oncology community, we have packaged it into an online application called GBMdeconvoluteR. We also give the user the option to use the marker genes from GBMap[24] because, although these did not quantify cell types as accurately as MCP_{GBM}, the GBMap reference set extends the range of GBM non-neoplastic cell types that can be quantified from bulk expression data. GBMdeconvoluteR is, thus, a web-based application that enables users to upload bulk GBM expression profiles and output the relative proportion of immune and neoplastic GBM cells, or using GBMap markers genes as input, to also include normal brain and blood-vessel cells, across multiple samples.

2.4 Discussion

We have developed the first publicly available GBM specific deconvolution tool that can infer both neoplastic and non-neoplastic cell population prevalence from bulk GBM tumour RNA-seq data. This tool was developed by amalgamating four independent, human, single cell sequencing datasets to create tissue specific cell type gene expression reference profiles. The single cell data was from de novo IDHwt GBM either at initial diagnosis (*primary*) or upon *recurrence*. *Recurrent* GBMs have been shown to have altered transcriptional profiles which may impact on the accuracy of the deconvolution results[3], so we included these samples in the tool development and validation. We found that our approach is suitable for deconvoluting recurrent GBM tumours but, in keeping with the aforementioned studies, neoplastic cell deconvolution is not as accurate at the longitudinal time point. Our study confirms, as shown elsewhere, that tissue specific reference

datasets are necessary to achieve maximal accuracy in expression-based deconvolution[5,10,11].

We used IMC to establish a ground truth for cell type characterisation and quantification, which we then compared against the results from the gene-expression-based approaches to evaluate their accuracy and identify the most suitable method. However, it must be noted that the IMC regions, while located within the same tumour, were substantially smaller than the RNA-seq regions (Figure 2.2A and Figure S2.2), and the GBM TME is highly heterogeneous[4]. Moreover, IMC was performed on adjacent (not identical) tissue sections. As such, deviations from perfect correlation reflect not only limitations in deconvolution performance but also genuine biological variation in cell composition. Therefore, while IMC cannot provide an absolute measure of accuracy, it allows for meaningful comparison between deconvolution methods to identify the best-performing approach.

Our study is the first to evaluate whether the marker genes of the four GBM neoplastic cell types, determined by Neftel *et al.* from gene expression data, are preferentially expressed at the protein level. We found a clear association between the protein levels of the selected markers and the gene expression-based quantification.

GBMdeconvoluteR is a publicly available web-application that enables accurate quantification of cell type prevalences in GBM samples using bulk RNA-seq data, using MCPcounter - a marker-gene-based method. The method returns relative, rather than absolute cell type quantification meaning comparisons are possible within cell types across samples, but not within samples across cell types. We applied GBMdeconvoluteR to data from TCGA and were able to confirm recent findings from single cell resolution multi-omics studies, regarding the specific enrichment of MES-like neoplastic cells in immune compartments, and depletion of other GBM cancer cell types. However, because our approach is easily applicable to large-scale sequencing dataset, we could expand upon this further to show that this association is stronger in samples from patients with worst prognosis. This leads to the hypothesis that quantifying immune:neoplastic cell interactions could be prognostic, or that targeting them could be therapeutically beneficial, exemplifying the value in applying GBMdeconvoluteR to gain biological and clinical insights.

In summary, GBMdeconvoluteR can be used to assess associations between cell type quantities and phenotypic, molecular or clinical characteristics with applications for target identification, gaining mechanistic insight or stratifying samples for retrospective therapeutic evaluation or prospective precision medicine approaches.

References

1. Mikkelsen, V. E., Solheim, O., Salvesen, Ø. & Torp, S. H. The histological representativeness of glioblastoma tissue samples. *Acta Neurochirurgica* **163**, 1911–1920 (July 2021). doi:10.1007/s00701-020-04608-y.

2. Neftel, C. *et al.* An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. *Cell* **178**, 835–849.e21 (Aug. 2019). doi:10.1016/j.cell.2019.06.024.

3. Varn, F. S. *et al.* Glioma progression is shaped by genetic evolution and microenvironment interactions. *Cell* **185**, 2184–2199.e16 (June 2022). doi:10.1016/j.cell.2022.04.038.

4. Ravi, V. M. *et al.* Spatially resolved multi-omics deciphers bidirectional tumor-host interdependence in glioblastoma. *Cancer Cell* **40**, 639–655.e13 (June 2022). doi:10.1016/j.ccell.2022.05.009.

5. Racle, J., de Jonge, K., Baumgaertner, P., Speiser, D. E. & Gfeller, D. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. *eLife* **6**, e26476 (Nov. 13, 2017). doi:10.7554/eLife.26476.

6. Aran, D., Hu, Z. & Butte, A. J. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biology* **18**, 220 (Nov. 15, 2017). doi:10.1186/s13059-017-1349-1.

7. Li, B. *et al.* Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. *Genome Biology* **17**, 174 (Aug. 22, 2016). doi:10.1186/s13059-016-1028-7.

8. Becht, E. *et al.* Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biology* **17**, 218 (Oct. 2016). doi:10.1186/s13059-016-1070-5.

9. Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods* **12**, 453–457 (May 2015). doi:10.1038/nmeth.3337.

10. Schelker, M. *et al.* Estimation of immune cell content in tumour tissue using single-cell RNA-seq data. *Nature Communications* **8**, 2032 (Dec. 11, 2017). doi:10.1038/s41467-017-02289-3.

11. Avila Cobos, F., Alquicira-Hernandez, J., Powell, J. E., Mestdagh, P. & De Preter, K. Benchmarking of cell type deconvolution pipelines for transcriptomics data. *Nature Communications* **11**, 5650 (Nov. 2020). doi:10.1038/s41467-020-19015-1.

12. Sturm, G. *et al.* Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. *Bioinformatics* **35**, i436–i445 (July 2019). doi:10.1093/bioinformatics/btz363.

13. Chen, A. X. *et al.* Single-cell characterization of macrophages in glioblastoma reveals MARCO as a mesenchymal pro-tumor marker. *Genome Medicine* **13**, 88 (May 19, 2021). doi:10.1186/s13073-021-00906-x.

14. Pombo Antunes, A. R. *et al.* Single-cell profiling of myeloid cells in glioblastoma across species and disease stage reveals macrophage competition and specialization. *Nature Neuroscience* **24**, 595–610 (Apr. 2021). doi:10.1038/s41593-020-00789-y.

15. Sankowski, R. *et al.* Mapping microglia states in the human brain through the integration of high-dimensional techniques. *Nature Neuroscience* **22**, 2098–2110 (Dec. 2019). doi:10.1038/s41593-019-0532-y.

16. Wang, L. *et al.* The Phenotypes of Proliferating Glioblastoma Cells Reside on a Single Axis of Variation. *Cancer Discovery* **9**, 1708–1719 (Dec. 2019). doi:10.1158/2159-8290.CD-19-0329.

17. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (June 24, 2021). doi:10.1016/j.cell.2021.04.048.

18. Patel, A. P. *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396–1401 (June 2014). doi:10.1126/science.1254257.

19. Tickle, T. I., Georgescu, C., Brown, M. & Haas, B. J. *inferCNV of the Trinity CTAT Project* 2019.

20. Waltman, L. & van Eck, N. J. A smart local moving algorithm for large-scale modularity-based community detection. *The European Physical Journal B* **86**, 471 (Nov. 13, 2013). doi:10.1140/epjb/e2013-40829-0.

21. Patterson-Cross, R. B., Levine, A. J. & Menon, V. Selecting single cell clustering parameter values using subsampling-based robustness metrics. *BMC Bioinformatics* **22**, 39 (Feb. 1, 2021). doi:10.1186/s12859-021-03957-4.

22. Lun, A. T., McCarthy, D. J. & Marioni, J. C. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with bioconductor. *F1000Research* **5**, 2122 (Oct. 31, 2016). doi:10.12688/f1000research.9501.2.

23. Wang, Q. *et al.* Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. *Cancer Cell* **33**, 152 (Jan. 2018). doi:10.1016/j.ccell.2017.12.012.

24. Ruiz-Moreno, C. *et al. Harmonized single-cell landscape, intercellular crosstalk and tumor architecture of glioblastoma* preprint (Cancer Biology, Aug. 2022). doi:10.1101/2022.08.27.505439.

25. Steen, C. B., Liu, C. L., Alizadeh, A. A. & Newman, A. M. Profiling cell type abundance and expression in bulk tissues with CIBERSORTx. *Methods in molecular biology (Clifton, N.J.)* **2117**, 135–157 (2020). doi:10.1007/978-1-0716-0301-7_7.

26. Droop, A. *et al.* How to analyse the spatiotemporal tumour samples needed to investigate cancer evolution: a case study using paired primary and recurrent glioblastoma. *International Journal of Cancer* **142**, 1620–1626 (2018). doi:10.1002/ijc.31184.

27. Couturier, C. P. *et al.* Single-cell RNA-seq reveals that glioblastoma recapitulates a normal neurodevelopmental hierarchy. *Nature Communications* **11**, 3406 (July 2020). doi:10.1038/s41467-020-17186-5.

28. Windhager, J. *et al.* An end-to-end workflow for multiplexed image processing and analysis. *Nature Protocols* **18**, 3565–3613 (Nov. 2023). doi:10.1038/s41596-023-00881-0.

29. Chevrier, S. *et al.* Compensation of Signal Spillover in Suspension and Imaging Mass Cytometry. *Cell Systems* **6**, 612–620.e5 (May 2018). doi:10.1016/j.cels.2018.02.010.

30. Eling, N., Damond, N., Hoch, T. & Bodenmiller, B. Cytomapper: an r/bioconductor package for visualization of highly multiplexed imaging data. *Bioinformatics* **36**, 5706–5708 (Dec. 26, 2020). doi:10.1093/bioinformatics/btaa1061.

31. Korsunsky, I. *et al.* Fast, sensitive, and accurate integration of single cell data with Harmony. *Nature methods* **16**, 1289–1296 (Dec. 2019). doi:10.1038/s41592-019-0619-0.

32. Levine, J. H. *et al.* Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. *Cell* **162**, 184–197 (July 2, 2015). doi:10.1016/j.cell.2015.05.047.

33. Diggins, K. E., Gandelman, J. S., Roe, C. E. & Irish, J. M. Generating quantitative cell identity labels with marker enrichment modeling (MEM). *Current protocols in cytometry* **83**, 10.21.1–10.21.28 (Jan. 18, 2018). doi:10.1002/cpcy.34.

34. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biology* **20**, 296 (Dec. 23, 2019). doi:10.1186/s13059-019-1874-1.

35. Van Deusen, A. L. *et al.* A single-cell mass cytometry-based atlas of the developing mouse brain. *Nature Neuroscience* **28**, 174–188 (Jan. 2025). doi:10.1038/s41593-024-01786-1.

36. Keeler, A. B. *et al.* A developmental atlas of somatosensory diversification and maturation in the dorsal root ganglia by single-cell mass cytometry. *Nature Neuroscience* **25**, 1543–1558 (Nov. 2022). doi:10.1038/s41593-022-01181-8.

37. White, K. *et al.* Identification, validation and biological characterisation of novel glioblastoma tumour microenvironment subtypes: implications for precision immunotherapy. *Annals of Oncology* **34**, 300–314 (Mar. 2023). doi:10.1016/j.annonc.2022.11.008.

38. Wang, L. *et al.* A single-cell atlas of glioblastoma evolution under therapy reveals cell-intrinsic and cell-extrinsic therapeutic targets. *Nature Cancer* **3**, 1534–1552 (Dec. 2022). doi:10.1038/s43018-022-00475-x.

Supplementary material

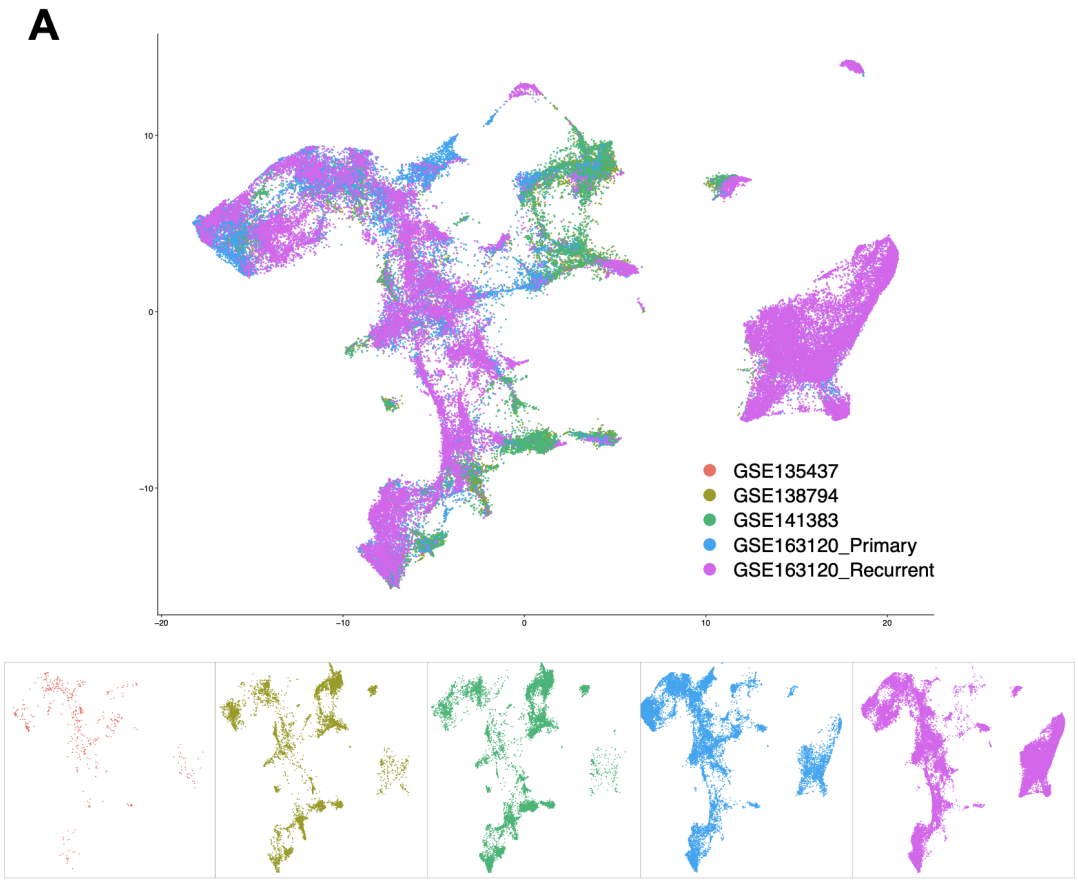

A
GSE135437
GSE138794
GSE141383
GSE163120_Primary
GSE163120_Recurrent
B
T/NK cells
Microglia
unknown
TAM
Monocytes
DC
T cells
B cells
NK cells
Mast cells

Figure S2.1 | Dataset integration and sub-clustering.

A) UMAP projection of single GBM immune cells amalgamated from 4 independent studies, following batch correction. The top repeats the middle left panel of Figure 2.1B but the mappings for each individual dataset are then shown separately below. The data GSE163120 was split in two and processed independently owing to samples being sequenced on different platforms. **B)** The result of isolation and sub-clustering for the T and NK cells (Left) and for the TAMs and microglia (right).

Figure S2.2 | H&E stained GBM tumour sections used in validation.

Haematoxylin and eosin staining (H&E) stained FFPE sections of the five *primary* (P) and matched *recurrent* (R) IDHwt GBM tumours used in the validation of GBMdeconvoluteR. Black circles denote the areas that were macro-dissected from consecutive slides, and from which RNA was extracted for bulk sequencing. Blue squares demarcate the $2mm^2$ regions that underwent characterisation by IMC. Top left to bottom right these samples are: 64P and 64R, 67P and 67R, 71P and 71R, 82P and 82R and 84P and 84R.

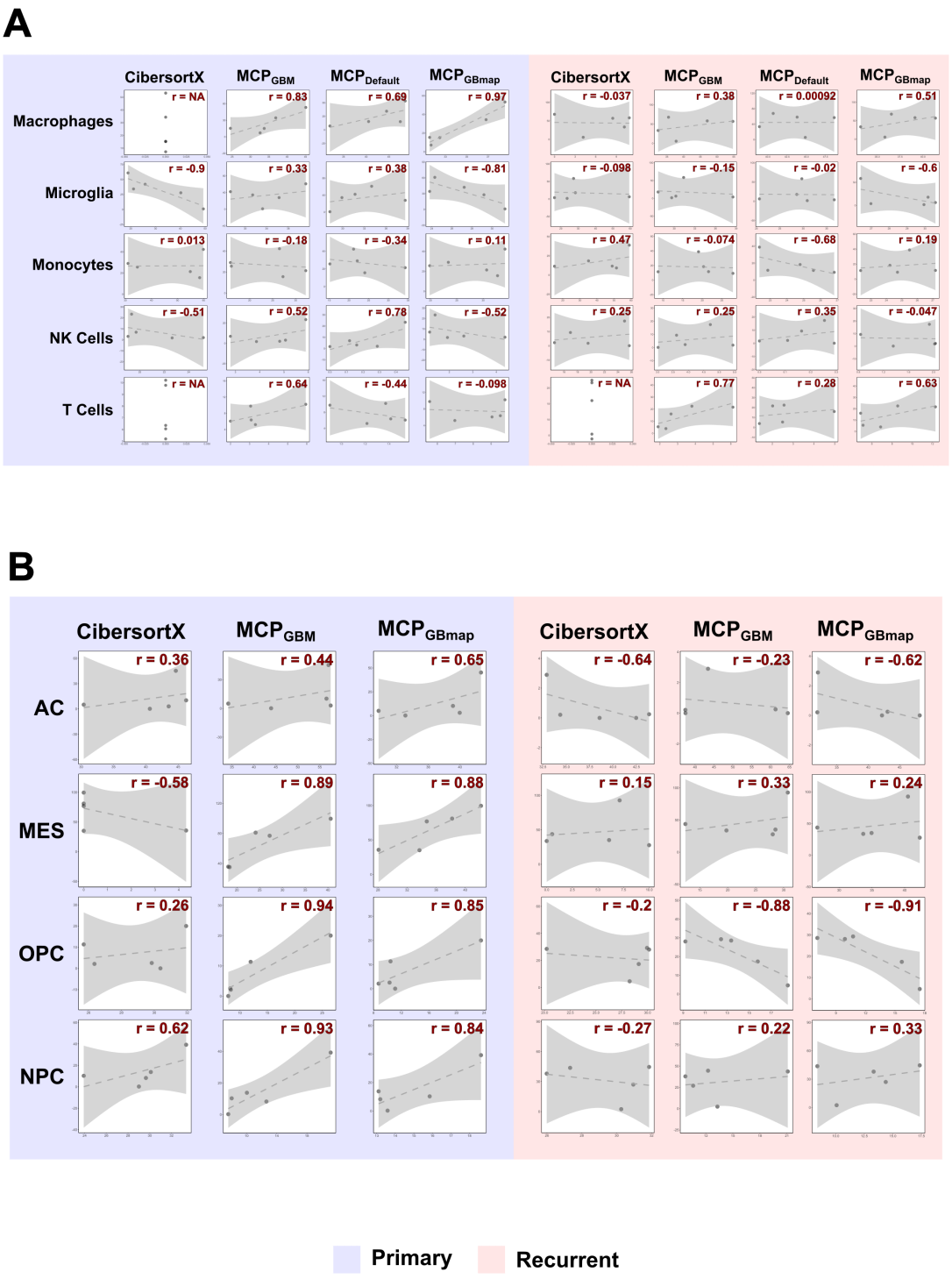

A
CibersortX
MCP GBM
MCP Default
MCP GBmap
Macrophages
Microglia
Monocytes
NK Cells
T Cells
r = NA
r = 0.83
r = 0.69
r = 0.97
r = -0.9
r = 0.33
r = 0.38
r = -0.81
r = 0.013
r = -0.18
r = -0.34
r = 0.11
r = -0.51
r = 0.52
r = 0.78
r = -0.52
r = NA
r = 0.64
r = -0.44
r = -0.098
r = -0.037
r = 0.38
r = 0.00092
r = 0.51
r = -0.098
r = -0.15
r = -0.02
r = -0.6
r = 0.47
r = -0.074
r = -0.68
r = 0.19
r = 0.25
r = 0.25
r = 0.35
r = -0.047
r = NA
r = 0.77
r = 0.28
r = 0.63
B
CibersortX
MCP GBM
MCP GBmap
AC
MES
OPC
NPC
r = 0.36
r = 0.44
r = 0.65
r = -0.58
r = 0.89
r = 0.88
r = 0.26
r = 0.94
r = 0.85
r = 0.62
r = 0.93
r = 0.84
r = -0.64
r = -0.23
r = -0.62
r = 0.15
r = 0.33
r = 0.24
r = -0.2
r = -0.88
r = -0.91
r = -0.27
r = 0.22
r = 0.33
Primary
Recurrent

Figure S2.3 | Concordance between absolute and relative cell type scores.

Scatterplots, in primary or recurrent GBM samples separately, of gold standard cell proportions quantified by IMC (y-axis) versus those predicted by gene expression-based methods (annotated across the top) for immune **(A)** or neoplastic cancer **(B)** cell types indicated down the side. The Pearson's correlation coefficient (r) is indicated. The dotted line is the line of best fit and the shaded area denotes the confidence interval. Marker genes for MCPcounter were either default ($MCP_{Default}$), GBM-specific according to our work (MCP_{GBM}) or GBM-specific according to GBMap (MCP_{GBMap}). Neoplastic cells: AC-like; OPC-like; NPC-like; MES-like.

Table S2.1 | Filter thresholds applied to each single cell dataset used to produce reference profiles for GBM specific cell types.

dataset	nUMI_min	nGene	gene_complexity_min	mitochondial_ratio_min
GSE135437	1500	700	0.8	0.25
GSE138794	2500	1100	0.8	0.25
GSE141383	500	300	0.8	0.25
GSE163120_P	1800	700	0.8	0.25
GSE163120_R	500	300	0.8	0.25

Table S2.2 | Single cell dataset-specific normalisation parameters used prior to data integration.

dataset	variable_features_n	vars_to_regress	return_only_var_genes
GSE135437	3000	mitoRatio, riboRatio	TRUE
GSE138794	3000	-	TRUE
GSE141383	5000	S.Score, G2M.Score	TRUE
GSE163120_P	5000	S.Score, G2M.Score, mitoRatio, riboRatio	TRUE
GSE163120_R	5000	S.Score, G2M.Score, mitoRatio, riboRatio	TRUE

Table S2.3 | Cell type frequencies for the optimal clustering resolution.

Cell type proportions resulting from the clustering of previously annotated (GSE163120) cell types. The optimal clustering resolution of 0.7 was empirically determined based on a sweep of different resolution parameters ranging between 0.1 – 0.8 in increments of 0.1. The N/A field denotes the frequency of cells that were labelled as unknown.

Cluster	B cell	DC	Mast cell	Monocyte	NK cell	T cell	TAM	N/A
0	28	2	1	3	1741	8308	30	187
1	4	10	0	45	2	0	6698	1300
2	1	57	0	245	0	1	5258	1094
3	12	0	3	2	177	6071	17	15
4	3	0	1	30	3	2	4988	1106
5	5	14	1	5	10	21	357	4557
6	1	37	4	2632	4	3	518	1574
7	4	6	5	10	107	660	848	2641
8	11	1	0	80	6	7	2154	1543
9	0	26	1	190	0	1	2810	473
10	1	3	3	585	45	214	317	2240
11	0	21	0	5	0	0	3045	114
12	1	39	1	23	0	5	535	2445
13	0	1	0	4	0	1	2414	139
14	1	5	0	104	0	2	1828	604
15	0	1644	0	176	1	1	496	99
16	3	13	2	24	21	36	919	981
17	4	0	0	0	146	1580	5	16
18	0	0	1	0	155	410	91	949
19	0	0	0	0	9	1279	0	12
20	1	1	0	2	0	0	402	681
21	896	12	0	0	7	4	63	6
22	1	622	1	0	0	0	0	14
23	0	0	0	14	1	1	35	521
24	1	0	0	1	0	0	467	94
25	0	0	0	3	0	0	64	198
26	0	0	80	0	0	1	7	0

Table S2.4 | Details of markers & antibodies used for IMC.

Marker	Category	Type/State	Marker justification (PMIDs)	Manufacturer (antibody clone)	Antibody justification	Concentration (μg/ml)	Metal-isotope	Control tissue
HOPX	cancer	AC-like	31327527; 31554641; 32641768	abcam (ab230544)	-	100	Yb171	Brain, Tonsil
SLC1A3_EAAT1	cancer	AC-like	31327527; 31554641; 32641768	abcam(ab240235)	BSA and azide free	1000	Gd158	Brain
GFAP	normal	astrocyte	25726916	abcam (ab218309)	PMID: 34174183	100	Sm149	Brain
CD56	normal	neuron	28791027	biolegend (318345)	PMID: 28369679	100	Dy162	Brain
IBA1	immune	macrophage	32848611	abcam (ab220815)	PMID: 34174183	200	Eu153	Spleen, Tonsil
NeuN_FOX3	normal	neuron	20452351	biolegend (834502)	PMID: 34174183	400	Sm147	Brain
ANXA_A1	cancer	MES-like	31327527; 31554641; 32641768	abcam (ab222398)	BSA and azide free	-	Yb172	Tonsil
ANXA_A2	cancer	MES-like	31327527; 31554641; 32641768	rndsystems (mab3928)	In PBS with Trehalose	8-25	Er166	Prostate, Tonsil
CHI3L1	cancer	MES-like	31327527; 31554641; 32641768	abcam (ab255864)	BSA and azide free	250	Sm154	Spleen, Brain
SOD2	cancer	MES-like	31327527; 31554641; 32641768	abcam (ab227846)	BSA and azide free, used in IHC	100	Nd146	Prostate
P2Y12R	immune	microglia	32848611	abcam (ab274386)	BSA and azide free	1000	Lu175	Brain
TMEM119	immune	microglia	32848611	sigmaaldrich (HPA051870)	PMID: 31740814	500-1000	Gd155	Brain, Tonsil
NKP46	immune	NK cell	31784984	rndsystems (mab1850)	PMID: 36689332	5-25	Nd144	Spleen
BCAN	cancer	NPC-like	31327527; 31554641; 32641768	thermofisher (MA5-27639)	BSA free, used in ICC	50	Gd160	Brain
DLL3	cancer	NPC-like	31327527; 31554641; 32641768	abcam (ab255694)	BSA and azide free	100	Nd148	Brain
MOG	normal	oligodendrocyte	2649509	rndsystems (mab1850)	-	5-25	Gd157	Brain
OLIG1	cancer	OPC-like	31327527; 31554641; 32641768	rndsystems (mab2417)	carrier free, used in IHC	8-25	Yb174	Skin
SCD5	cancer	OPC-like	31327527; 31554641; 32641768	thermofisher (PA5-59963)	used in IHC	50	Tm169	Brain
CD3	immune	T cell	29768164	fluidigm (3170019D)	PMID: 36689332	75-200	Er170	Spleen, Tonsil
CD8	immune	T cell	29768164	biolegend (344727)	PMID: 28369679	200	Ho165	Spleen, Tonsil
DNA1	DNA intercalator	-	-	fluidigm (201192B)	Preconugated to 191Ir	-	Ir191	-
DNA2	DNA intercalator	-	-	fluidigm (201192B)	Preconugated to 193Ir	-	Ir193	-
CD45	immune	-	12414720	fluidigm (91H029152)	Preconugated to 152Sm	300	Sm152	Spleen, Tonsil
CD31	vasculature	-	27055047	fluidigm (3151025D)	Preconugated to 151Eu		Eu151	Skin, Tonsil, Prostate
SMA	vasculature	-	19929197	rndsystems (mab1420)	used in Cytof	8-25	Dy164	Prostate, Skin, Tonsil
EZH2	-	transcript respressive	23720055	abcam (ab231165)	BSA and azide free	250	Nd145	Tonsil
HIF1A	-	hypoxia	11606368	thermofisher (700505)	PMID: 32868913	400	Dy161	Bone marrow
JARID2_C_Terminus	-	active	30573669	developed in house	-	-	Nd143	Brain
JARID2_N_Terminus	-	repressed	30573669	abcam (ab251123)	BSA free version, validated (by KO)	-	Yb173	Brain
Ki67	-	proliferating	29322240	fluidigm (3168001B)	Preconugated to 168Er	-	Er168	Skin, Tonsil
SNAI1	-	EMT	33806868	rndsystems (af3639)	BSA and azide free, used in IHC	5-15	Tb159	Ubiquitous
SOX2	-	proliferating	30952620	fluidigm (3150019B)	Preconugated to 150Nd	-	Nd150	Brain, Tonsil
TGFBeta	-	quiescent	30952620	fluidigm (3163010B)	Pre-conjugated to 163Dy	-	Dy163	Spleen, Bone marrow, Prostate
TNC	-	quiescent	30952620	rndsystems (mab2138)	used in IF	8-25	Gd156	Uterus

Table S2.5 | MEM labels for the imaging mass cytometry.

Initial Clustering of all cells:		
Assigned Cell Type	*Cluster*	*MEM Labels*
NK Cells	10	NKP46^{+4} CD45^{+4} TGFBeta^{+3}
T Cells	26	CD8^{+5} CD56^{+4} SLC1A3^{+3}
T Cells	2	CD3^{+9} CD45^{+5}
Macrophages	14	IBA1^{+4}
Macrophages	15	IBA1^{+5}
Microglia	6	TMEM119^{+5} IBA1^{+4} CD45^{+3}
Microglia	13	ANXA1^{+5} P2Y12R^{+3}
Monocytes	17	CD31^{+5}
Sub-clustering unassigned cells from above:		
Assigned Cell Type	*Cluster*	*MEM Labels*
Neoplastic sub-clusters	1	SLC1A3^{+5} HOPX^{+5} JARID2_C^{+4} CD3^{+4} SOX2^{+3} BCAN^{+3} ANXA1^{+3} JARID2_N^{+3} OLIG1^{+3} P2Y12R^{+3}
	2	EZH2^{+3} SLC1A3^{+3} JARID2_C^{+2} SOX2^{+2} BCAN^{+2} CD3^{+2} HOPX^{+2} OLIG1^{+2}
	8	ANEXIN_A2^{+2} ANXA1^{+2} OLIG1^{+2}
	10	HIF1A^{+7} JARID2_C^{+4} SCD5^{+3} CD3^{+3} TMEM119^{+2} MOG^{+2} SNAI1^{+2} CD56^{+2} CD8^{+2} HOPX^{+2}
Further clustering cells from neoplastic sub-cluster 1:		
Assigned Cell Type	*Cluster*	*MEM Labels*
AC	9	SLC1A3^{+2}
MES	6	CHI3L1^{+5} TGFBeta^{+2} ANEXIN_A2^{+2}
NPC	11	DLL3^{+4} GFAP^{+3} NKP46^{+2} NeuN^{+2}
OPC	4	SCD5^{+5} IBA1^{+4} SNAI1^{+3} TGFBeta^{+3} MOG^{+2}
Further clustering cells from neoplastic sub-cluster 2:		
Assigned Cell Type	*Cluster*	*MEM Labels*
AC	2	IBA1^{+9} JARID2_N^{+5} CD45^{+3} MOG^{+3} SLC1A3^{+3} EZH2^{+2}
AC	6	P2Y12R^{+4} SLC1A3^{+3} HOPX^{+3} SMA^{+2}
NPC	10	GFAP^{+9} EZH2^{+8} NeuN^{+8} DLL3^{+8} CD45^{+5} NKP46^{+4} CD31^{+4} SNAI1^{+4} JARID2_C^{+3} IBA1^{+3} MOG^{+3} SCD5^{+3}
OPC	4	JARID2_C^{+2} OLIG1^{+2}
Further clustering cells from neoplastic sub-cluster 8:		
Assigned Cell Type	*Cluster*	*MEM Labels*
MES	2	SOX2^{+4} P2Y12R^{+4} ANXA1^{+2}
MES	3	GFAP^{+4} CHI3L1^{+4} SOD2^{+3} DLL3^{+3} ANEXIN_A2^{+3} CD31^{+2} TNC^{+2} ANXA1^{+2}
MES	5	CHI3L1^{+4} SMA^{+2} ANEXIN_A2^{+2}
Further clustering cells from neoplastic sub-cluster 10:		
Assigned Cell Type	*Cluster*	*MEM Labels*
OPC	3	IBA1^{+2} MOG^{+2} OLIG1^{+2}
OPC	2	SCD5^{+2}

Table S2.6 | Final immune markers chosen for MCP_{GBM}.

Marker	Cell type	Also
CD79A	B cells	-
CD79B	B cells	-
CD37	B cells	-
MS4A1	B cells	-
BANK1	B cells	-
BLK	B cells	-
VPREB3	B cells	-
RP5-887A10.1	B cells	-
FCER1A	DC	-
PLD4	DC	-
JAML	DC	-
PKIB	DC	-
CTSH	DC	-
TPSAB1	Mast Cells	-
TPSB2	Mast Cells	-
CPA3	Mast Cells	-
HPGD	Mast Cells	-
HPGDS	Mast Cells	-
HDC	Mast Cells	-
RGS13	Mast Cells	-
MAOB	Mast Cells	-
KRT1	Mast Cells	-
AC004791.2	Mast Cells	-
SLC45A3	Mast Cells	-
SVOPL	Mast Cells	-
CTSG	Mast Cells	-
KIT	Mast Cells	-
VWA5A	Mast Cells	-
ALDH1A1	Mast Cells	-
SDPR	Mast Cells	-
MS4A2	Mast Cells	-
C1orf186	Mast Cells	-
GATA2	Mast Cells	-
LTC4S	Mast Cells	-
RP11-354E11.2	Mast Cells	-
SLC18A2	Mast Cells	-
LAPTM4A	Mast Cells	-
ANXA1	Mast Cells	-
SAMSN1	Mast Cells	-
LYZ	Monocytes	-
VCAN	Monocytes	-
FCN1	Monocytes	-
TIMP1	Monocytes	-
S100A9	Monocytes	-
S100A12	Monocytes	-
S100A6	Monocytes	-
EREG	Monocytes	-
S100A8	Monocytes	-
LINC01272	Monocytes	-
TYROBP	Monocytes	TAMs
CFD	Monocytes	-
SERPINA1	Monocytes	-
NEAT1	Monocytes	-
CD7	NK Cells	-
CCL5	NK Cells	T Cells
KLRB1	NK Cells	-
KLRC1	NK Cells	-
HOPX	NK Cells	-
GZMB	NK Cells	-
APOBEC3G	NK Cells	-
KRT86	NK Cells	-
KLRD1	NK Cells	-
MATK	NK Cells	-
IL2RB	NK Cells	-
HCST	NK Cells	-
XCL2	NK Cells	-
XCL1	NK Cells	-
EVL	NK Cells	T Cells
GNLY	NK Cells	-
TRDC	NK Cells	-
CD247	NK Cells	-
CD2	NK Cells	-
LCK	NK Cells	T Cells
PRF1	NK Cells	-
CST7	NK Cells	-

Marker	Cell type	Also
CTSW	NK Cells	-
STMN1	NK Cells	-
GZMA	NK Cells	T Cells
NKG7	NK Cells	-
IL32	T Cells	NK Cells
SIT1	T Cells	-
CD3D	T Cells	-
GIMAP7	T Cells	NK Cells
SPOCK2	T Cells	-
IL7R	T Cells	-
TRBC1	T Cells	-
TRBC2	T Cells	NK Cells
LCK	T Cells	NK Cells
CCL5	T Cells	NK Cells
CD3G	T Cells	-
CD3E	T Cells	NK Cells
CD2	T Cells	NK Cells
TRAC	T Cells	NK Cells
CD27	T Cells	-
LTB	T Cells	-
FTL	Macrophages	-
C1QC	Macrophages	Microglia
APOE	Macrophages	Microglia
C1QB	Macrophages	Microglia
C1QA	Macrophages	Microglia
APOC1	Macrophages	Microglia
TREM2	Macrophages	-
VSIG4	Macrophages	Microglia
FCGBP	Macrophages	Microglia
PLTP	Macrophages	Microglia
A2M	Macrophages	-
AIF1	Macrophages	-
GPNMB	Macrophages	-
NPC2	Macrophages	-
MT2A	Macrophages	-
MT1E	Macrophages	-
CD9	Macrophages	-
SAT1	Macrophages	-
RP11-552D4.1	Macrophages	-
SPP1	Macrophages	-
ITGA4	Macrophages	-
TGFBI	Macrophages	-
LGALS3BP	Macrophages	-
IFITM2	Macrophages	-
IFITM3	Macrophages	-
TAGLN2	Macrophages	-
FPR3	Macrophages	-
KYNU	Macrophages	-
S100A11	Macrophages	-
S100A6	Macrophages	-
TGM2	Macrophages	-
EMILIN2	Macrophages	-
GDA	Macrophages	-
HP	Macrophages	-
SELL	Macrophages	-
CD163	Macrophages	-
MRC1	Macrophages	-
LYVE1	Macrophages	-
SIGLEC1	Macrophages	-
TMEM119	Microglia	-
TREM2	Microglia	-
CX3CR1	Microglia	-
P2RY12	Microglia	-
ALOX5AP	Microglia	-
ADORA3	Microglia	-
IGSF6	Microglia	-
ACY3	Microglia	-
RASAL3	Microglia	-
TBXAS1	Microglia	-
SASH3	Microglia	-
SUSD3	Microglia	-
P2RY13	Microglia	-
TMIGD3	Microglia	-
GPR34	Microglia	-
OLFML3	Microglia	-
SELPLG	Microglia	-
SPARC	Microglia	-
FOS	Microglia	-

Marker	Cell type	Also
SIGLEC10	Microglia	-
C3	Microglia	-
SLC2A5	Microglia	-
SLC1A3	Microglia	-
PF4	Microglia	-
BIN1	Microglia	-
RNASET2	Microglia	-
RGS10	Microglia	-
SGK1	Microglia	-
F13A1	Microglia	-
IFIT3	Microglia	-
MCM5	Microglia	-
DAB2	Microglia	-
IFNGR1	Microglia	-
APBB1IP	Microglia	-
LILRB4	Microglia	-
CYFIP1	Microglia	-
CXCR2	Microglia	-
SCD2	Microglia	-
PSAT1	Microglia	-
CSF2RA	Microglia	-
CSF1	Microglia	-
CSF1R	Mircoglia	-
CRYBB1	Microglia	-
FCRLS	Microglia	-
MAFB	Microglia	-
PMEPA1	Microglia	-
CD14	Microglia	-
LPL	Microglia	-
CST7	Microglia	-

Table S2.7 | Neoplastic markers (tumour intrinsic) included in MCP_{GBM}.

Marker	Cell type	Tumour intrinsic
CHI3L1	MES	yes
ANXA2	MES	yes
ANXA1	MES	yes
CD44	MES	yes
VIM	MES	yes
MT2A	MES	no
C1S	MES	no
NAMPT	MES	no
EFEMP1	MES	yes
C1R	MES	no
SOD2	MES	yes
IFITM3	MES	no
TIMP1	MES	yes
SPP1	MES	yes
A2M	MES	yes
S100A11	MES	yes
MT1X	MES	yes
S100A10	MES	yes
FN1	MES	yes
LGALS1	MES	yes
S100A16	MES	yes
CLIC1	MES	yes
MGST1	MES	no
RCAN1	MES	yes
TAGLN2	MES	yes
NPC2	MES	yes
SERPING1	MES	no
C8orf4	MES	yes
EMP1	MES	yes
APOE	MES	no
CTSB	MES	yes
C3	MES	no
LGALS3	MES	yes
MT1E	MES	no
EMP3	MES	yes
SERPINA3	MES	no
ACTN1	MES	yes
PRDX6	MES	yes
IGFBP7	MES	no
SERPINE1	MES	yes
PLP2	MES	yes
MGP	MES	yes
CLIC4	MES	yes
GFPT2	MES	yes
GSN	MES	yes
NNMT	MES	yes
TUBA1C	MES	yes
GJA1	MES	yes
TNFRSF1A	MES	yes
WWTR1	MES	yes
HILPDA	MES	no
ADM	MES	yes
DDIT3	MES	yes
NDRG1	MES	yes
HERPUD1	MES	yes
DNAJB9	MES	no
TRIB3	MES	yes
ENO2	MES	yes
AKAP12	MES	yes
SQSTM1	MES	yes
ATF3	MES	yes
NRN1	MES	yes
SLC2A1	MES	yes
BNIP3	MES	no
INSIG2	MES	yes
IGFBP3	MES	yes
PPP1R15A	MES	yes
PLOD2	MES	yes
GBE1	MES	yes
SLC2A3	MES	yes
FTL	MES	yes
WARS	MES	yes
ERO1L	MES	yes

Marker	Cell type	Tumour intrinsic
XPOT	MES	yes
HSPA5	MES	yes
GDF15	MES	yes
EPAS1	MES	yes
LDHA	MES	yes
P4HA1	MES	yes
SERTAD1	MES	yes
PFKP	MES	no
PGK1	MES	yes
EGLN3	MES	no
SLC6A6	MES	yes
CA9	MES	yes
BNIP3L	MES	yes
RPL21	MES	no
TRAM1	MES	yes
UFM1	MES	yes
ASNS	MES	no
GOLT1B	MES	yes
ANGPTL4	MES	yes
SLC39A14	MES	yes
CDKN1A	MES	yes
HSPA9	MES	yes
TNC	MES	yes
ZFP36L1	MES	yes
PDLIM4	MES	yes
WDR13	MES	yes
NPNT	MES	no
HSPB1	MES	yes
SGCE	MES	yes
GABARAPL1	MES	no
PON2	MES	yes
CYR61	MES	no
COL1A2	MES	yes
SAT1	MES	yes
ATP1A2	MES	yes
CHL1	MES	yes
CST3	AC	yes
S100B	AC	yes
SLC1A3	AC	yes
HEPN1	AC	no
HOPX	AC	yes
MT3	AC	no
SPARCL1	AC	no
MLC1	AC	yes
GFAP	AC	no
FABP7	AC	yes
BCAN	AC	yes
PON2	AC	yes
METTL7B	AC	yes
SPARC	AC	yes
GATM	AC	yes
RAMP1	AC	yes
PMP2	AC	yes
AQP4	AC	no
DBI	AC	yes
EDNRB	AC	yes
PTPRZ1	AC	yes
CLU	AC	yes
PMP22	AC	yes
ATP1A2	AC	yes
S100A16	AC	yes
HEY1	AC	yes
PCDHGC3	AC	yes
TTYH1	AC	yes
NDRG2	AC	no
PRCP	AC	yes
ATP1B2	AC	yes
AGT	AC	yes
PLTP	AC	yes
GPM6B	AC	no
F3	AC	yes
RAB31	AC	yes
PPAP2B	AC	yes
ANXA5	AC	yes
TSPAN7	AC	no
CPNE1	AC	yes
CSPG4	AC	yes
PCMTD2	AC	yes

Marker	Cell type	Tumour intrinsic
PDZD2	AC	yes
METL7B	AC	no
EGFR-AS1	AC	no
MT2A	AC	no
SOCS2	AC	yes
NEAT1	AC	no
ATP13A4	AC	yes
MEOX2	AC	yes
GPR37L1	AC	yes
CD82	AC	yes
SEC61G	AC	yes
EGFR	AC	yes
BCAN	OPC	yes
PLP1	OPC	no
GPR17	OPC	yes
FIBIN	OPC	yes
LHFPL3	OPC	yes
OLIG1	OPC	yes
PSAT1	OPC	yes
SCRG1	OPC	yes
OMG	OPC	no
APOD	OPC	no
SIRT2	OPC	yes
TNR	OPC	no
THY1	OPC	yes
PHYHIPL	OPC	yes
SOX2-OT	OPC	no
NKAIN4	OPC	yes
LPPR1	OPC	no
PTPRZ1	OPC	yes
VCAN	OPC	yes
DBI	OPC	yes
PMP2	OPC	yes
CNP	OPC	yes
TNS3	OPC	yes
LIMA1	OPC	yes
CA10	OPC	yes
PCDHGC3	OPC	yes
CNTN1	OPC	yes
SCD5	OPC	yes
P2RX7	OPC	yes
CADM2	OPC	no
TTYH1	OPC	yes
FGF12	OPC	yes
TMEM206	OPC	yes
NEU4	OPC	yes
FXYD6	OPC	yes
RNF13	OPC	yes
RTKN	OPC	yes
GPM6B	OPC	no
LMF1	OPC	yes
ALCAM	OPC	yes
PGRMC1	OPC	yes
HRASLS	OPC	yes
BCAS1	OPC	no
RAB31	OPC	yes
PLLP	OPC	yes
FABP5	OPC	yes
NLGN3	OPC	yes
SERINC5	OPC	yes
EPB41L2	OPC	yes
GPR37L1	OPC	yes
PDGFRA	OPC	yes
RPL32	OPC	yes
LSM3	OPC	yes
CUTA	OPC	yes
UBE2E2	OPC	no
RPL34	OPC	yes
UGDH	OPC	yes
TXN	OPC	yes
RPL45A	OPC	no
MPZL1	OPC	yes
CUEDC2	OPC	yes
PIP4K2A	OPC	no
CTHRC1	OPC	yes
H19	OPC	no
H2AFZ	OPC	yes
PLAT	OPC	yes

Marker	Cell type	Tumour intrinsic
COX7A2	OPC	yes
VDAC2	OPC	yes
SOX10	OPC	yes
MIF	OPC	no
CDR1	OPC	no
STMN2	OPC	no
RP1	OPC	no
RPL35	OPC	yes
RPS24	OPC	yes
COL9A1	OPC	yes
PGAM1	OPC	no
RPL21	OPC	no
SNHG5	OPC	no
BEX4	OPC	no
GAS5	OPC	no
PPA1	OPC	no
GAD1	OPC	no
LINC00643	OPC	no
TCEAL2	OPC	no
NME1	OPC	yes
BEX1	OPC	no
ITM2A	OPC	no
DLL3	NPC	yes
DLL1	NPC	yes
SOX4	NPC	yes
TUBB3	NPC	yes
HES6	NPC	yes
TAGLN3	NPC	no
NEU4	NPC	yes
MARCKSL1	NPC	yes
CD24	NPC	no
STMN1	NPC	yes
TCF12	NPC	yes
BEX1	NPC	no
OLIG1	NPC	yes
MAP2	NPC	yes
FXYD6	NPC	yes
PTPRS	NPC	yes
MLLT11	NPC	no
NPPA	NPC	yes
BCAN	NPC	yes
MEST	NPC	yes
ASCL1	NPC	yes
BTG2	NPC	yes
DCX	NPC	yes
NXPH1	NPC	yes
HN1	NPC	yes
PFN2	NPC	yes
SCG3	NPC	yes
MYT1	NPC	yes
CHD7	NPC	yes
GPR56	NPC	yes
TUBA1A	NPC	yes
PCBP4	NPC	yes
ETV1	NPC	yes
SHD	NPC	yes
TNR	NPC	no
AMOTL2	NPC	yes
DBN1	NPC	yes
HIP1	NPC	yes
ABAT	NPC	yes
ELAVL4	NPC	yes
LMF1	NPC	yes
GRIK2	NPC	yes
SERINC5	NPC	yes
TSPAN13	NPC	yes
ELMO1	NPC	no
GLCCI1	NPC	yes
SEZ6L	NPC	yes
LRRN1	NPC	yes
SEZ6	NPC	yes
SOX11	NPC	yes
STMN2	NPC	no
RND3	NPC	yes
HMP19	NPC	no
MIAT	NPC	no
NSG1	NPC	no
DLX6-AS1	NPC	no

Marker	Cell type	Tumour intrinsic
NREP	NPC	no
FNBP1L	NPC	yes
STMN4	NPC	no
DLX5	NPC	no
MAP1B	NPC	yes
RBFOX2	NPC	yes
IGFBPL1	NPC	no
TMEM161B-AS1	NPC	no
DPYSL3	NPC	yes
Sep-03	NPC	no
PKIA	NPC	yes
ATP1B1	NPC	no
DYNC1I1	NPC	no
CD200	NPC	no
SNAP25	NPC	no
PAK3	NPC	no
NDRG4	NPC	yes
KIF5A	NPC	no
UCHL1	NPC	no
ENO2	NPC	yes
KIF5C	NPC	no
DDAH2	NPC	yes
TUBB2A	NPC	no
LBH	NPC	yes
LOC150568	NPC	no
TCF4	NPC	yes
GNG3	NPC	no
NFIB	NPC	yes
DPYSL5	NPC	yes
CRABP1	NPC	no
NFIX	NPC	yes
CEP170	NPC	yes
BLCAP	NPC	no
VIPR2	NPC	yes
PSPH	NPC	yes
NTRK2	NPC	yes
GSTM1	NPC	yes
CDKN2A	NPC	no
CDK4	NPC	yes
DCTN2	NPC	yes
B4GALNT1	NPC	yes
DDIT3	NPC	yes
ALCAM	NPC	yes
DGKB	NPC	no
LINC00689	NPC	no
PHYHIPL	NPC	yes
RARRES2	NPC	no

Table S2.8 | Concordance between immune cell scores and IMC.

Pearson's correlation coefficients between IMC and the immune cell scores from different cell deconvolution methods used in this study. the best correlated method per sample type is shown highlighted.

Cell Type	Method	Sample Type	Pearson's r	Average r
Macrophages	CIBERSORTx	All	0.3165	
Microglia	CIBERSORTx	All	-0.4051	
Monocytes	CIBERSORTx	All	0.3437	0.052165
NK Cells	CIBERSORTx	All	-0.04644	
T Cells	CIBERSORTx	All	N/A	
Macrophages	$MCP_{Default}$	All	0.3564	
Microglia	$MCP_{Default}$	All	0.2434	
Monocytes	$MCP_{Default}$	All	-0.3598	0.26788
NK Cells	$MCP_{Default}$	All	0.6306	
T Cells	$MCP_{Default}$	All	0.4688	
Macrophages	MCP_{GBM}	All	0.628	
Microglia	MCP_{GBM}	All	0.193	
Monocytes	MCP_{GBM}	All	0.01866	0.365652
NK Cells	MCP_{GBM}	All	0.4209	
T Cells	MCP_{GBM}	All	0.5677	
Macrophages	MCP_{GBMap}	All	0.7614	
Microglia	MCP_{GBMap}	All	-0.7328	
Monocytes	MCP_{GBMap}	All	0.2441	0.06302
NK Cells	MCP_{GBMap}	All	-0.3304	
T Cells	MCP_{GBMap}	All	0.3728	
Macrophages	CIBERSORTx	Primary	N/A	
Microglia	CIBERSORTx	Primary	-0.8962	
Monocytes	CIBERSORTx	Primary	0.01298	-0.46394
NK Cells	CIBERSORTx	Primary	-0.5086	
T Cells	CIBERSORTx	Primary	N/A	
Macrophages	$MCP_{Default}$	Primary	0.6936	
Microglia	$MCP_{Default}$	Primary	0.3849	
Monocytes	$MCP_{Default}$	Primary	-0.337	0.21644
NK Cells	$MCP_{Default}$	Primary	0.7782	
T Cells	$MCP_{Default}$	Primary	-0.4375	
Macrophages	MCP_{GBM}	Primary	0.8253	0.42766
Microglia	MCP_{GBM}	Primary	0.3283	0.42766
Monocytes	MCP_{GBM}	Primary	-0.1758	0.42766
NK Cells	MCP_{GBM}	Primary	0.5174	0.42766
T Cells	MCP_{GBM}	Primary	0.6431	0.42766
Macrophages	MCP_{GBMap}	Primary	0.9697	
Microglia	MCP_{GBMap}	Primary	-0.8095	
Monocytes	MCP_{GBMap}	Primary	0.1084	-0.068824
NK Cells	MCP_{GBMap}	Primary	-0.5152	
T Cells	MCP_{GBMap}	Primary	-0.09752	
Macrophages	CIBERSORTx	Recurrent	-0.03685	
Microglia	CIBERSORTx	Recurrent	-0.09832	
Monocytes	CIBERSORTx	Recurrent	0.4736	0.1467575
NK Cells	CIBERSORTx	Recurrent	0.2486	
T Cells	CIBERSORTx	Recurrent	N/A	
Macrophages	$MCP_{Default}$	Recurrent	0.0009185	
Microglia	$MCP_{Default}$	Recurrent	-0.02028	
Monocytes	$MCP_{Default}$	Recurrent	-0.6759	-0.0142923
NK Cells	$MCP_{Default}$	Recurrent	0.3468	
T Cells	$MCP_{Default}$	Recurrent	0.277	
Macrophages	MCP_{GBM}	Recurrent	0.3758	0.23498
Microglia	MCP_{GBM}	Recurrent	-0.1501	0.23498
Monocytes	MCP_GB	Recurrent	-0.0743	0.23498
NK Cells	MCP_{GBM}	Recurrent	0.2496	0.23498
T Cells	MCP_{GBM}	Recurrent	0.7739	
Macrophages	MCP_{GBMap}	Recurrent	0.5093	
Microglia	MCP_{GBMap}	Recurrent	-0.5961	
Monocytes	MCP_{GBMap}	Recurrent	0.1857	0.135942
NK Cells	MCP_{GBMap}	Recurrent	-0.04689	
T Cells	MCP_{GBMap}	Recurrent	0.6277	

Table S2.9 | Concordance between neoplastic cell scores and IMC.

Pearson's correlation coefficients between IMC and the neoplastic cell scores from different cell deconvolution methods used in this study. the best correlated method per sample type is shown highlighted.

Cell Type	Method	Sample Type	Pearson's r	Average r
AC	CIBERSORTx	All	0.3372	0.024025
MES	CIBERSORTx	All	-0.2594	
NPC	CIBERSORTx	All	-0.1613	
OPC	CIBERSORTx	All	0.1796	
AC	MCP_{GBM}	All	0.2503	0.43265
MES	MCP_{GBM}	All	0.6448	0.43265
NPC	MCP_{GBM}	All	0.2662	0.43265
OPC	MCP_{GBM}	All	0.5693	0.43265
AC	MCP_{GBMap}	All	0.2011	0.2231125
MES	MCP_{GBMap}	All	0.5003	
NPC	MCP_{GBMap}	All	-0.03285	
OPC	MCP_{GBMap}	All	0.2239	
AC	CIBERSORTx	Primary	0.3563	0.164775
MES	CIBERSORTx	Primary	-0.5799	
NPC	CIBERSORTx	Primary	0.263	
OPC	CIBERSORTx	Primary	0.6197	
AC	MCP_{GBM}	Primary	0.4354	0.7992
MES	MCP_{GBM}	Primary	0.8923	
NPC	MCP_{GBM}	Primary	0.9434	
OPC	MCP_{GBM}	Primary	0.9257	
AC	MCP_{GBMap}	Primary	0.6451	0.805275
MES	MCP_{GBMap}	Primary	0.8815	0.805275
NPC	MCP_{GBMap}	Primary	0.8501	0.805275
OPC	MCP_{GBMap}	Primary	0.8444	0.805275
AC	CIBERSORTx	Recurrent	-0.6441	-0.24
MES	CIBERSORTx	Recurrent	0.1488	
NPC	CIBERSORTx	Recurrent	-0.1952	
OPC	CIBERSORTx	Recurrent	-0.2695	
AC	MCP_{GBM}	Recurrent	-0.2283	-0.13695
MES	MCP_{GBM}	Recurrent	0.3344	-0.13695
NPC	MCP_{GBM}	Recurrent	-0.8784	-0.13695
OPC	MCP_{GBM}	Recurrent	0.2245	-0.13695
AC	MCP_{GBMap}	Recurrent	-0.6178	-0.23965
MES	MCP_{GBMap}	Recurrent	0.2415	
NPC	MCP_{GBMap}	Recurrent	-0.9085	
OPC	MCP_{GBMap}	Recurrent	0.3262	

Chapter 3

Deconvoluting cellular landscape changes in IDH-wildtype glioblastoma tumours pre- & post-treatment

Shoaib Ajaib[1], Antonia Barry[1,2], Steven Pollock[1], Sophie Williams[2], Arief Gusnanto[3], Gemma Hemmings[1], Aruna Chakrabarty[4], Azzam Ismail[4], Erica Wilson[1], Lucy F. Stead[1]

[1]Leeds Institute of Medical Research, University of Leeds
[2]School of Medicine and Population Health, University of Sheffield
[3]School of Mathematics, University of Leeds
[4]Department of Neuropathology, Leeds Teaching Hospitals NHS Trust

Abstract

Background: Isocitrate dehydrogenase wild-type (IDHwt) glioblastoma (GBM) is an aggressive, incurable brain malignancy characterised by significant cellular heterogeneity that contributes to treatment resistance. Recently, we proposed a patient classification of GBM tumours into distinct Up and Down responder subtypes based on their transcriptional responses to standard treatment.

Methods: Here we used cell deconvolution to identify cell type proportions using longitudinal, bulk RNA sequencing (RNA-seq) data from 219, paired pre- and post-treatment IDHwt GBM tumour samples.

Results: We find significant reductions in vascular cells and increases in normal neurons, oligodendrocytes, and plasma B cells through treatment. Further spatial proteomics and transcriptomics validated showed that plasma B cells significantly co-localised with oligodendrocytes. Stratifying patients into Up and Down responders we observe subtype-specific cellular dynamics: Up responders show elevated plasma B cells, neurons, and oligodendrocytes post-treatment, indicating a more proneural (PN) phenotype involving oligodendrocyte remyelination. Down responders align with a more mesenchymal (MES) phenotype, marked by increased hypoxia, decreased vasculature that does not associate with any other cell type and also more myeloid lineage cells.

Conclusions: Responder-specific cellular landscape changes through treatment reveal potential mechanisms of GBM treatment resistance and therapeutic targets.

3.1 Introduction

IDHwt GBM is the most common and aggressive type of primary brain malignancy[1]. Despite aggressive standard treatment of surgical resection, radiation and chemotherapy, it remains incurable. This failure arises due to a combination of genetic, epigenetic, developmental and microenvironmental factors that contribute to GBM tumour heterogeneity[1,2].

Building on studies that examined inter-tumour transcriptional heterogeneity within GBM[3,4], single-cell RNA sequencing (scRNA-seq) and single-nucleus RNA sequencing (snRNA-seq) have revealed that GBM cells exist in various transcriptionally defined neoplastic states[2,4–6]. These states mimic distinct neural cell types and are influenced by epigenetic mechanisms[7]. Additionally, these states are associated with properties like proliferation, stemness and neurodevelopmental programs[8]; PN-to-MES axis[9], cellular specialization and metabolism[10]; the tumour microenvironment (TME) and injury responses[11].

Previously, we developed a high-throughput cell deconvolution tool to quantify the cellular composition of IDHwt GBM tumours *in-silico*[12]. More recently, we analysed bulk RNA-seq data from pairs of pre- and post-treatment, IDHwt GBM tumours and identified a subset of genes which were differentially regulated amongst two patients groups[13]. These changes were mainly linked to neoplastic cells and suggested subtype-specific adaptive treatment resistance mechanisms and distinct TME alterations were at play, leading us to define two responder subtypes: Up and Down.

In this study, we applied cell deconvolution to significantly enlarged cohorts of bulk, longitudinally paired IDHwt GBM tumour samples to understand how the cellular landscape changes through treatment. We used cell type markers for immune cells, normal brain cells, including vasculature components, and markers defining neoplastic GBM cells. For the latter we used two schemas as described by (1) Neftel *et al.*[2] and (2) Wang *et al.*[9]. The first study defined 4 GBM cell types: oligodendrocyte progenitor-like (OPC-like), astrocyte-like (AC-like), neural progenitor-like (NPC-like), and mesenchymal-like (MES-like). The MES-like signatures were further divided into hypoxia-independent mesenchymal-like (MES1) and hypoxia-dependent mesenchymal-like (MES2) categories. Similarly, NPC-like signatures were subdivided into oligodendrocyte-lineage associated neural progenitor-like (NPC1) and neuronal-lineage associated neural progenitor-like (NPC2), based on the inclusion of oligodendrocyte progenitor cell (OPC)-related versus neuronal lineage genes, respectively. Conversely, the second study delin-

eated a bidirectional transcriptional axis with cells on a spectrum between PN and MES (herein referred to as Wang *et al.* mesenchymal (MES3))[9]. We also expanded our previous work by stratifying these IDHwt patients into responder subtypes, and further examined the cell type changes specific to each subgroup[13].

Finally, we explored whether information contained within the cell type scores was sufficient to distinguish tumour samples as pre- or post-treatment and identify different responder subtypes. Our ultimate goal was to uncover biologically and clinically meaningful insights that could relate to distinct treatment resistance mechanisms which could be targeted therapeutically.

3.2 Materials & methods

All analyses were performed using the R statistical software package (v4.2.1) and Python (v3.11)[14]. Further details of packages and specific versions are provided in the sections below. The name of each statistical test used, and level of significance achieved, is included within the results where the finding from each hypothesis test is confirmed. All plots were generated using ggplot2 (v3.4.4)[15].

3.2.1 Data availability

Data used within this study, that isn't already publicly available, will be made available at the time of publication. Publicly available data obtained from longitudinal bulk tumour studies is available as described in Table S3.1.

3.2.2 Code Availability

All code used to generate the results detailed in this paper can be found at:

https://github.com/GliomaGenomics/GBM_Deconvolution_Analysis

3.2.3 Ethics statement

Tumour samples used in this study were obtained from patients at the Walton Centre, UK, who provided informed consent in writing for the use of their tissue in research. The inclusion of these samples in this project was following approval

by the UK National Health Services (NHSs) Research Ethics Service Committee South Central - Oxford A (Research Ethics Code: 13/SC/0509).

3.2.4 Sample collection & processing

Collection and processing of longitudinal GBM samples for the Stead cohort were acquired from The Walton Centre, Lancashire Teaching Hospitals, and Leeds Teaching Hospitals NHS Foundation Trusts via the Brain Tumour Northwest Tissue Banks and the Leeds Neuropathology Research Tissue Bank. In addition, tissue samples were obtained from Cambridge University Hospitals NHS Foundation Trust as part of the brain archive information network UK (BRAIN UK)[16]. These samples were processed as previously described[17]. Briefly, formalin-fixed paraffin-embedded (FFPE) blocks were sectioned and the first and last sections were haematoxylin and eosin staining (H&E)-stained and underwent neuropathologist review to identify areas of $> 60\%$ tumour. Regions of overlap were macro-dissected from the intervening sections and ribonucleic acid (RNA) was extracted using All Prep deoxyribonucleic acid (DNA)/RNA FFPE Kit (catalogue #80,234) from Qiagen (UK).

3.2.5 Imaging mass cytometry

The imaging mass cytometry (IMC) antibody selection, image processing and analysis was performed as previously described[12].

3.2.6 RNA-seq data acquisition & processing

All RNA extracted in-house underwent ribosomal ribonucleic acid (rRNA) depletion using the NEBNext rRNA Depletion Kit (Human/Mouse/Rat) and then strand-directional, whole transcriptome library preparation using NEBNext Ultra™ II Directional RNA Library Prep Kit for Illumina©, both from New England Biolabs (UK). Libraries were sequenced on Illumina next-generation sequencers as 100 base pairs (bp) paired-end reads. Raw RNA data was acquired from several published studies following negotiation of data transfer agreements, where necessary[13,18–23]. Discovery, Stead cohort sample FASTQ data were trimmed of low-quality bases, *Phred* threshold = 20, and adapters via Trim Galore v0.4.3, wrapping Cutadapt v1.8.3[24]. Trimmed reads were quality checked using FASTQC[25] and then aligned to the human reference genome GRCh38.13 using STAR v020201 in

two-pass mode with a maximum of 5 multireads[26]. Gene and transcript count and gene expression were quantified via CuffQuant v2.2.1 taking directional specifics of the library as input, using probabilistic weighting of multi-reads and quantifying against the GENCODEv27 human genome annotation with haplotypes and scaffolds included[27,28]. Discovery, glioma longitudinal analysis consortium (GLASS) cohort RNA-seq data was acquired as pre-processed transcripts per million (TPM) counts via the GLASS portal at https://www.synapse.org/glass[18]. The validation, European organisation for research and treatment of cancer (EORTC) cohort RNA-seq data was processed according to the methodology previously described - the raw counts were converted from fragments per kilobase million (FPKM) to TPM counts[19]. All TPM counts were then summed per gene and filtered to remove genes with zero expression and variance across all samples. We also removed non-protein coding genes, to mitigate against the different library preparation methods used by the cohorts.

3.2.7 Batch correction

Stead and GLASS cohort patient samples were obtained for multiple sources and therefore were batch corrected to remove this unwanted source of variation using Combat-seq from the SVA package (v3.35.2)[29]. This was visualised before and after correction using the base R principal component analysis (PCA) function, *prcomp()* and taking the first two principal components (Figure S3.1).

3.2.8 Estimating tumour purity

For each patient sample we estimated the purity (proportion of malignant cells) of the tumour using GBMPurity: a deep learning model specifically designed to estimate the purity of IDHwt *primary* GBM from bulk RNA-seq[30]. We excluded any samples which did not meet the minimum purity threshold determined by the model.

3.2.9 Cell deconvolution

Cell type deconvolution was performed using GBMDeconvoluteR, as previously described[12]. A custom marker list was supplied to the programme, comprising of IDHwt GBM tumour-specific immune cell markers from the Ajaib *et al.* gene marker list, supplemented with immune and brain resident cell types from the

Ruiz-Moreno *et al.* gene marker list[31]. Additional GBM malignant cell, neoplastic cell-state markers were also added, taking the original six markers described by Neftel *et al.*[2] and the two described by Wang *et al.*[9]. All off the neoplastic markers used were originally derived using single cell data and therefore, were refined using a previously defined methodology[2], in order to exclude markers which were not applicable to bulk RNA-seq data. Moreover, all non-tumour-intrinsic (TI) markers, as defined by Wang *et al.* were also filtered out before performing cell deconvolution[4].

3.2.10 Differential cell type scores

All cell type scores were compared using a two-sided, paired, *Wilcoxon signed-rank test* implemented in the base R, *wilcox.test()* function[32]. Comparisons were initially performed using data from all of the samples and also individually for each of the cohorts (Table S3.1): the data were also stratified across up and Down responder types. All comparison p values were adjusted (q-values) using the false discovery rate (FDR) method to correct for multiple testing[33]. Comparisons were visualised using EnhancedVolcano R package (v1.13.2), taking the $\log_2$-fold changes in cell-type scores and the corrected $-\log_{10}$ p values.

3.2.11 Cell type score correlations

Correlation analysis was conducted with the base R *cor()* function, using the Spearman Rank correlation test to account for non-linear associations[34] and the significance of each correlation tested using the *cor.test()* function. Visualisation of the correlations were generated using the ggcorrplot R package (v0.1.4.1). Multiple correlation coefficient thresholds, ranging from 0 to 0.7, were tested to capture diverse levels of correlation strength and significance.

3.2.12 Testing cell-cell interactions

Cell-cell spatial interactions were tested using the *testInteractions* from imcRtools with the previously described method[35]. Briefly, co-detection by indexing (CODEX) spatial interaction graph data (from Greenwald *et al.*[36]) edges were summed and aggregated across individual patients and then divided by the number of cells of type A that had at least one neighbour of type B. These observed interactions were compared against a derived null distribution, describing the in-

teractions formed under spatial randomness. Significance was determined using two one-tailed *permutation* tests:

$$P_{AB} = \begin{cases} 1, & C_{\text{obs}} = 0; \\ \frac{\sum(C_{\text{perm}\geq(\leq)}C_{\text{obs}})+1}{N_{\text{perm}}+1}, & \text{otherwise}, \end{cases}$$

where C_{perm} is the number of cell pairs (A, B) in each permutation, C_{obs} is the actual number of cell pairs (A, B) given a defined distance, and N_{perm} is the number of permutations.

3.2.13 Survival analysis

All survival analyses were conducted using the base R survival package (v3.5.7). Kaplan-Meier survival curves were generated using the *survfit()* function, tested for significance using the *log-rank* test[37] and visualised using ggplot2 the survminer R package (v0.4.9). *Cox* regression models were generated using the *coxph()* function, assessing the relation between progression-free survival (PFS) and overall survival (OS) in months. The non-linearity of each exploratory variable was examined using Martingale residuals[38], and the proportional hazards (PH) assumptions for each covariate were assessed using the base R *cox.zph()* function, correlating the corresponding set of scaled Schoenfeld residuals with time[39]. Multivariate *Cox* models were initially constructed incorporating each individual cell type score predictor variable and two known confounder variables (age and O^6-methylguanine DNA methyltransferase (MGMT) expression) and then subsequently, using all the cell type score predictors. Visualisation of the *Cox* regression results was done using the forestploter R package (v1.1.1).

3.2.14 Dimensionality reduction

PCA was performed in 10 dimensions taking a vector of all cell type scores using the *prcomp()* R function and visualised using the factoextra R package (v1.0.7). Further dimensionality reduction was performed in two dimensions for clustering and visualisation purposes using the pairwise controlled manifold approximation (PaCMAP) algorithm from the pacmap python package (v0.7.0) using the default parameters[40].

3.2.15 Unsupervised clustering

K-means clustering was performed using the base R *kmeans()* function with the numbers of clusters (k) set between 2-20. The optimal number of k clusters was determined using the total within-cluster sum of squares (WCSS) method for each dataset and then visualised using the PaCMAP reduced dimensions. Unsupervised cluster survival curves were compared for PFS and OS as detailed in the survival analysis subsection 3.2.13.

3.2.16 Supervised classification

The MLR3 R package (v0.17.2) was used to perform supervised classification[41]. Briefly, multiple binary classification algorithms were benchmarked, predicting the probabilities of outcomes: surgery (primary or recurrent) and responder type (up or down). Each benchmarked model was then evaluated using the classification accuracy and classification error metrics and also compared against a featureless classifier. The initial benchmarking was done using all the data and then following this, the Stead and GLASS cohorts were re-trained using the top performing model from the benchmarking process and validated using the EORTC cohort. A *10-fold* cross-validation resampling strategy was employed for all model training and benchmarking steps. The best performing binary classifier was evaluated using receiver operator curve-area under the curve (ROC-AUC) analysis and visualised using the *autoplot()* function. The most important features of the best performing model were evaluated using the permutation feature importance (PFI) method[42], implemented in the iml R package (v0.11.1)[43].

Table 3.1 | Summary of clinical metadata.

Summary of the clinical metadata for the paired GBM tumour samples used in the study. Data are shown for the full cohort ("All samples") as well as for a filtered subset of samples with tumour purity $\geq 40\%$. For each group, metadata are provided across three contributing datasets: Stead, GLASS and EORTC. PFS: progression-free survival; OS: overall survival; IQR: interquartile range; GLASS: glioma longitudinal analysis consortium; EORTC: European organisation for research and treatment of cancer.

All samples				
	All	*Stead*	*GLASS*	*EORTC*
tumour pairs (up/down responders)	219 (137/82)	98 (60/38)	23 (14/9)	98 (63/35)
sex (male/female/unknown)	75/37/107	58/31/9	17/ 6/0	0/0/98
median age in years (IQR)	57 (49 - 63)	58 (48 - 63)	60 (50 - 66)	55 (50 - 62)
median PFS in months (IQR)	13 (8 - 21)	13 (9 - 22)	14 (8 - 24)	12 (8 - 19)
median OS in months (IQR)	22 (17 - 33)	21 (17 - 33)	23 (18 - 33)	22 (17 - 33)
Tumour purity ($\geq$ 40%) filtered samples				
	All	*Stead*	*GLASS*	*EORTC*
tumour pairs (up/down responders)	124 (79/45)	66 (40/26)	9 (6/3)	49 (33/16)
sex (male/female/unknown)	47/23/ 54	41/20/5	6/ 3/0	0/0/ 49
median age in years (IQR)	56 (50 - 63)	57 (49 - 63)	52 (49 - 63)	56 (51 - 62)
median PFS in months (IQR)	15 (10 - 24)	14 (10 - 23)	16 (16 - 32)	15 (9 - 24)
median OS in months (IQR)	23 (17 - 34)	19 (17 - 31)	34 (22- 44)	24 (17 - 34)

3.3 Results

3.3.1 Overview of cohorts

We restricted our study to *de novo*, IDHwt GBM tumours, where all patients had received standard treatment comprising of radiation and temozolomide (TMZ). All patient samples were longitudinally paired (primary and recurrent) and only included the first recurrence. In total, we analysed 219 GBM tumour pairs across three independent datasets, detailed in Table S3.1 and also summarised in Table 3.1: Stead[13]; GLASS[18]; and EORTC[19]. The RNA-seq data from the EORTC and Stead cohorts were processed locally, while the GLASS data were processed separately 14. The data were sub-divided into discovery (Stead and GLASS) and validation EORTC cohorts to evaluate consistent longitudinal changes: we applied batch correction to the discovery cohort to account for multi-centre origins of these data (Figure S3.1). Further, we excluded tumour sample pairs which had less than 40% tumour purity[30], retaining 124 GBM tumour pairs (Table S3.1 and Figure 3.1A) for analysis to ensure accurate responder subtype assignment based on cancer cell content[13].

3.3.2 Longitudinal cell type score changes through treatment

Differential cell-type changes

We first investigated how the prevalence of each cell type changed through treatment, in a pair-wise manner across both discovery and validation cohorts (Figure 3.1B and Table S3.2). We found cells making up the vasculature component were decreased: endothelial cells ($q = 4.46 \times 10^{-7}$) and mural cells ($q = 7.87 \times 10^{-4}$). Decreased vasculature, specifically endothelial cells, following treatment in GBM has previously been reported[19] and is thought to be a result of the recurrent tumour not having had enough time to re-vascularise following surgical resection.

We also found normal brain cells: neurons ($q = 7.76 \times 10^{-4}$) and oligodendrocytes ($q = 2.32 \times 10^{-5}$), significantly increased following treatment. Hoogstrate *et al.* previously analysed transcriptional changes in paired IDHwt GBM tumours and found that deceasing tumour purity was accompanied by increased expression of neurons and oligodendrocytes[19]. They and others have also reported increases in tumour-associated macrophages (TAMs) and T cells[19,44]. Plasma B cells were

the only immune cell-type we found that was significantly increased across both cohorts ($q = 2.43 \times 10^{-5}$). Notably, whilst plasma B cells are increasing, B cell populations on the whole are found to be decreasing through treatment across both cohorts, albeit not significantly.

To validate the transcriptional cell type changes, at the protein level, we analysed IMC using 5 patient samples from our discovery cohort (Table S3.1 and Figure 3.1C). In agreement with our deconvolution results, we found that endothelial cell proportions were decreasing ($p < 0.01$) and normal brain cell populations were all increasing: neurons ($p < 0.01$); oligodendrocytes ($p < 0.05$) and astrocytes ($p < 0.01$). This was also apparent when visualising the respective cell-type protein marker abundances (Figure 3.1C). Astrocytes are the most abundant cell type found in brain microenvironment and become activated in response to brain injury, in order to form a protective barrier that limits the extent of tissue damage within the delicate brain tissue and also promotes axonal regrowth[45]. This may explain the astrocyte increase observed in IMC data post-treatment, though the effect, whilst still present, is less pronounced at the transcript level across a larger cohort, where astrocytes already constitute a major cell population. Conversely, the immune cell increases observed are likely a consequence of wound healing response to treatment, predominantly surgical resection[46].

Cell type correlations

The GBM TME is shaped by both the properties of individual cell types and their interactions, which together drive treatment resistance through enhanced proliferation, invasion, and migration[47]. To study this interplay, we analysed correlations in cell type score changes (recurrent – primary) across discovery and validation cohorts (Figure 3.1D).

We observed strong positive correlations amongst all immune cells except for plasma B cells, which correlated more strongly with normal brain cells: oligodendrocytes ($r = 0.91$), astrocytes ($r = 0.69$), and neurons ($r = 0.56$).

Plasma B cells are terminally differentiated B cells that secrete Immunoglobulin M (IgM) early in the immune response and Immunoglobulin G (IgG) later following memory B cell differentiation[48]. Notably, oligodendrocyte progenitor cells express high-affinity receptors for the fragment crystallizable region s (Fc regions) of both antibody types, promoting their maturation[49]. Disrupting IgM signalling during neonatal stages has been shown to impair axonal myelination in later development[48,49].

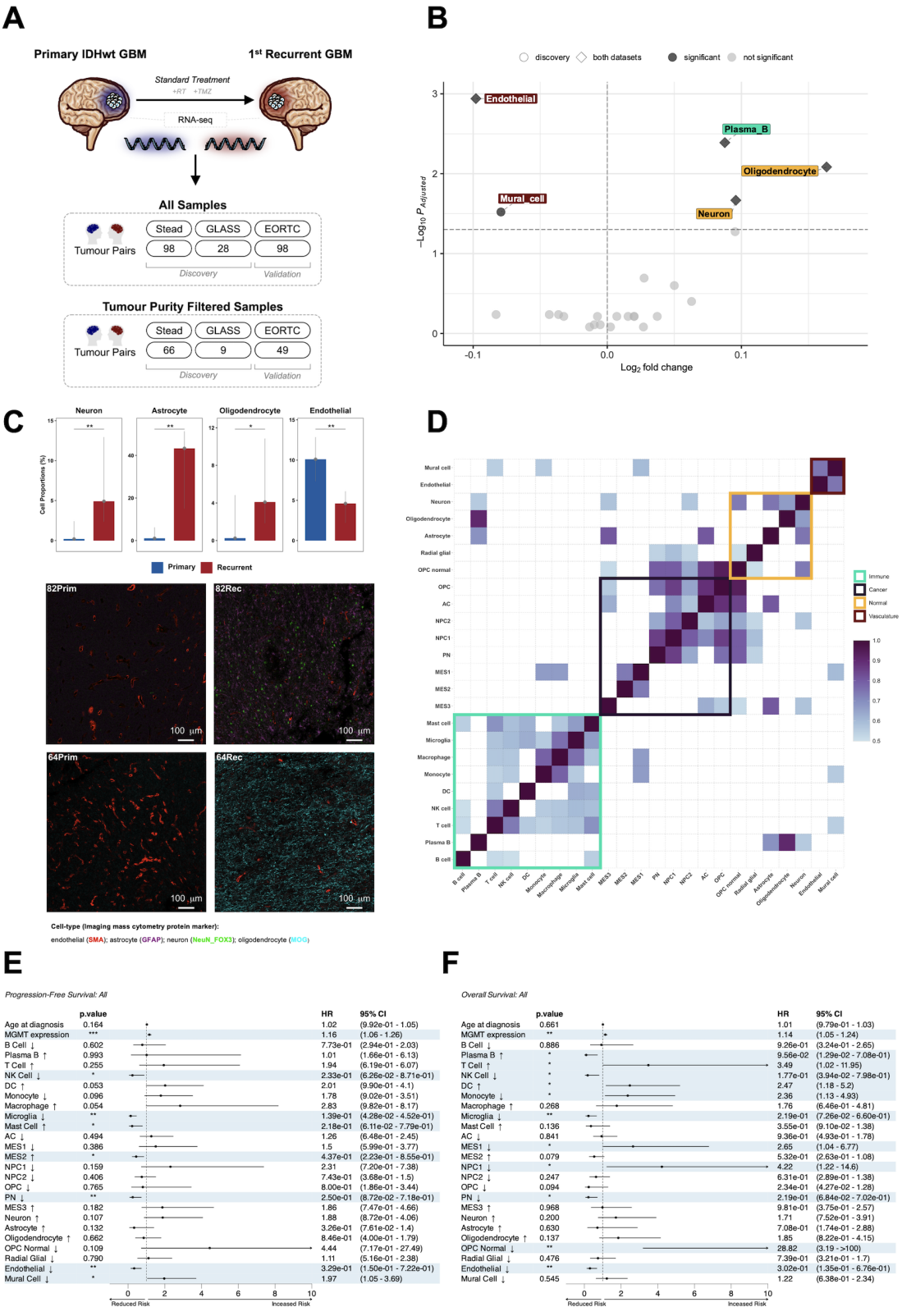

A
Primary IDHwt GBM
1st Recurrent GBM
Standard Treatment
+RT +TMZ
RNA-seq
All Samples
Stead GLASS EORTC
Tumour Pairs 98 28 98
Discovery Validation
Tumour Purity Filtered Samples
Stead GLASS EORTC
Tumour Pairs 66 9 49
Discovery Validation
B
discovery both datasets significant not significant
Endothelial
Plasma_B
Oligodendrocyte
Mural_cell
Neuron
−Log10 P Adjusted
Log2 fold change
C
Neuron Astrocyte Oligodendrocyte Endothelial
Cell Proportions (%)
Primary Recurrent
82Prim
82Rec
64Prim
64Rec
100 µm
Cell-type (Imaging mass cytometry protein marker):
endothelial (SMA); astrocyte (GFAP); neuron (NeuN_FOX3); oligodendrocyte (MOG)
D
Immune
Cancer
Normal
Vasculature
E
Progression-Free Survival: All
p.value HR 95% CI
Age at diagnosis 0.164 1.02 (9.92e-01 - 1.05)
MGMT expression *** 1.16 (1.08 - 1.26)
B Cell ↓ 0.602 7.73e-01 (2.94e-01 - 2.03)
Plasma B ↑ 0.993 1.01 (1.66e-01 - 6.13)
T Cell ↑ 0.255 1.94 (6.19e-01 - 6.07)
NK Cell ↓ * 2.33e-01 (6.26e-02 - 8.71e-01)
DC ↑ 0.053 2.01 (9.90e-01 - 4.1)
Monocyte ↓ 0.096 1.78 (9.02e-01 - 3.51)
Macrophage ↑ 0.054 2.83 (9.82e-01 - 8.17)
Microglia ↓ ** 1.39e-01 (4.28e-02 - 4.52e-01)
Mast Cell ↑ * 2.18e-01 (6.11e-02 - 7.79e-01)
AC ↓ 0.494 1.26 (6.48e-01 - 2.45)
MES1 ↓ 0.386 1.5 (5.99e-01 - 3.77)
MES2 ↑ * 4.37e-01 (2.23e-01 - 8.55e-01)
NPC1 ↓ 0.159 2.31 (7.20e-01 - 7.38)
NPC2 ↓ 0.406 7.43e-01 (3.68e-01 - 1.5)
OPC ↓ 0.765 8.00e-01 (1.86e-01 - 3.44)
PN ↓ ** 2.50e-01 (8.72e-02 - 7.18e-01)
MES3 ↑ 0.182 1.86 (7.47e-01 - 4.66)
Neuron ↑ 0.107 1.88 (8.72e-01 - 4.06)
Astrocyte ↑ 0.132 3.26e-01 (7.61e-02 - 1.4)
Oligodendrocyte ↑ 0.662 8.46e-01 (4.00e-01 - 1.79)
OPC Normal ↓ 0.109 4.44 (7.17e-01 - 27.49)
Radial Glial ↓ 0.790 1.11 (5.16e-01 - 2.38)
Endothelial ↓ ** 3.29e-01 (1.50e-01 - 7.22e-01)
Mural Cell ↓ * 1.97 (1.05 - 3.69)
Reduced Risk
Increased Risk
F
Overall Survival: All
p.value HR 95% CI
Age at diagnosis 0.661 1.01 (9.79e-01 - 1.03)
MGMT expression ** 1.14 (1.05 - 1.24)
B Cell ↓ 0.886 9.26e-01 (3.24e-01 - 2.65)
Plasma B ↑ * 9.56e-02 (1.29e-02 - 7.08e-01)
T Cell ↑ * 3.49 (1.02 - 11.95)
NK Cell ↓ * 1.77e-01 (3.94e-02 - 7.98e-01)
DC ↑ * 2.47 (1.18 - 5.2)
Monocyte ↓ * 2.36 (1.13 - 4.93)
Macrophage ↑ 0.268 1.76 (6.46e-01 - 4.81)
Microglia ↓ ** 2.19e-01 (7.26e-02 - 6.60e-01)
Mast Cell ↑ 0.136 3.55e-01 (9.10e-02 - 1.38)
AC ↓ 0.841 9.36e-01 (4.93e-01 - 1.78)
MES1 ↓ * 2.65 (1.04 - 6.77)
MES2 ↑ 0.079 5.32e-01 (2.63e-01 - 1.08)
NPC1 ↓ * 4.22 (1.22 - 14.6)
NPC2 ↓ 0.247 6.31e-01 (2.89e-01 - 1.38)
OPC ↓ 0.094 2.34e-01 (4.27e-02 - 1.28)
PN ↓ * 2.19e-01 (6.84e-02 - 7.02e-01)
MES3 ↑ 0.968 9.81e-01 (3.75e-01 - 2.57)
Neuron ↑ 0.200 1.71 (7.52e-01 - 3.91)
Astrocyte ↑ 0.630 7.08e-01 (1.74e-01 - 2.88)
Oligodendrocyte ↑ 0.137 1.85 (8.22e-01 - 4.15)
OPC Normal ↓ ** 28.82 (3.19 - >100)
Radial Glial ↓ 0.476 7.39e-01 (3.21e-01 - 1.7)
Endothelial ↓ ** 3.02e-01 (1.35e-01 - 6.76e-01)
Mural Cell ↓ 0.545 1.22 (6.38e-01 - 2.34)
Reduced Risk
Increased Risk

Figure 3.1 | Cell type score dynamics and their associations with survival across discovery and validation cohorts.

A) Schematic representation illustrating the paired patient study design and the number of patient samples used across each of the discovery (Stead and GLASS) and validation (EORTC) cohorts. The number of paired samples remaining after filtering for low tumour purity is also shown. **B)** Volcano plots showing the changes in GBMdeconvoluteR cell type scores across paired, primary and recurrent samples in the discovery cohort. The y-axis represents the $-\log_{10}$ FDR-corrected p-values and the x-axis shows $\log_2$-fold changes. Significant cell type changes ($p < 0.05$) are highlighted (by cell category) and labelled. The point shapes indicate whether changes are also significant in the validation cohort. **C)** Top: boxplots showing the prevalence of cell type score changes between primary and recurrent surgery, based on IMC data. Bottom: representative IMC images showing the protein marker abundance used to identify normal brain cells including vasculature components. **D)** Heatmap of spearman rank correlation coefficients between changes in cell-type scores (recurrent – primary) in the discovery cohort. The tile fill colour reflects the direction and strength of the correlation coefficients and only strong correlation coefficients ($\geq$(-)0.5) that are significant ($p < 0.05$) across both discovery and validation cohorts are shown. **E-F)** Forest plots showing the association between change (recurrent – primary) in cell-type score through treatment; age at diagnosis; and MGMT expression in relation to PFS **(E)** and OS **(F)** across all patient samples. Arrows after each cell-type denote the median change for that cell type across comparison samples: up arrow - increased through treatment; down arrow – decreased through treatment. Hazard ratios (HRs) are shown as points, with 95% confidence intervals represented by horizontal bars. Predictors significantly associated with survival outcomes ($p < 0.05$) are shaded blue. Significance thresholds: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; $^{****}p < 0.0001$.

The vasculature cell types (endothelial and mural cells) are strongly correlated ($r = 0.75$) with each other. Endothelial cells do not correlate with any other cell types, though mural cells correlate with T cells ($r = 0.55$), monocytes ($r = 0.56$) and the MES1 ($r = 0.60$). The immunosuppressive role of TAMs and their role in inducing a MES-like state in GBM are also well documented[50,51], as are the influences of increased monocytic lineage cells in impacting on OS in tumours that have a high amount of tertiary lymphoid structures (TLSs) present[52].

Amongst normal brain cells we find two patterns of association, one between normal glial cells and neurons: astrocytes ($r = 0.75$), oligodendrocytes ($r = 0.67$) and those between non-malignant neoplastic cell types, radial glial cells and oligodendrocyte progenitor cells ($r = 0.51$) which also strongly correlate with the malignant neoplastic cell types.

The malignant neoplastic cell types defined by Neftel *et al.* also separate into two groups along a PN to MES axis where MES-like neoplastic cells (MES1 and

MES2) and neural/glial progenitor-like cells (NPC1/NPC2, OPC-like and AC-like) are correlated separately. Neftel *et al.* reported NPC2 associated more with neuronal lineage genes and also we find this to be the case with NPC2 correlation with normal neurons (r = 0.57) across both cohorts. Interestingly, whilst we see a similar separation between MES3 and PN neoplastic cell types defined by Wang *et al.*, the MES3 is more correlated with normal glial cells such as astrocytes (r = 0.81), neurons (r = 0.58) and the AC-like neoplastic cell type (r = 0.63), rather than the two Neftel mesenchymal neoplastic cell states.

Taken together these results suggest that, whilst there are some consistent changes in individual cell types through treatment, coordinated changes in groups of cells between primary and recurrent GBM are more prevalent. Furthermore, this coordination occurs within cell subsets (immune, neoplastic, etc.) but also implicates associations between cells of different groupings that may infer functional relationships. An example of this being the coordinated upregulation of oligodendrocytes and plasma B cells.

Association of cell type changes with survival outcomes

In addition to identifying cell changes and correlations through treatment we also wanted to assess the impact of such changes on patient survival. Therefore, we modelled OS and PFS risk associated with the change (recurrent – primary scores) in each cell type, alongside two confounding predictors known to affect GBM survival[53]: age at diagnosis and the expression of MGMT in the primary tumours.

For this analysis we combined the discovery and validation data to increase statistical power and identify the most robust cell type change predictors. We find that higher MGMT expression is associated with both an increased risk of PFS (HR = 1.16, 95% confidence interval (CI): 1.06 - 1.26) and OS (HR = 1.14, 95% CI: 1.05 - 1.24) resulting in a worse patient outcome, consistent with its use as a prognostic marker in GBM[54].

Aberrant vasculature is a hallmark of GBM progression where it facilitates hypoxic niches and creates hostile immune environments[55]. We find that decreases in endothelial cells through treatment are associated with better PFS (HR = 0.33, 95% CI: 0.15 - 0.72) and OS (HR = 0.30, 95% CI: 0.14 - 0.68) outcomes, whilst mural cell decreases through treatment are associated with a worse PFS outcome (HR = 1.97, 95% CI: 1.05 - 3.69). Although, both vascular cell components are significantly decreasing through treatment (Figure 3.1B), we note that mural cells are significantly correlated (Figure 3.1D) with T cells, monocytes and MES1. Mu-

ral cells collectively refer to smooth muscle cells and pericytes, the latter of which facilitate GBM cell migration to infiltrating tumour edges[56]. These pericytes express mesenchymal stem cell (MSC) molecular markers and when activated under inflammatory conditions are involved in both angiogenesis and immune-response mediation[57]. Additionally, aside from supporting vasculature, pericytes are integral to neural proliferation by maintaining of the perivascular region (PVR) which comprises neural progenitors, astrocytes, endothelial cells, pericytes, and extracellular matrix components[57].

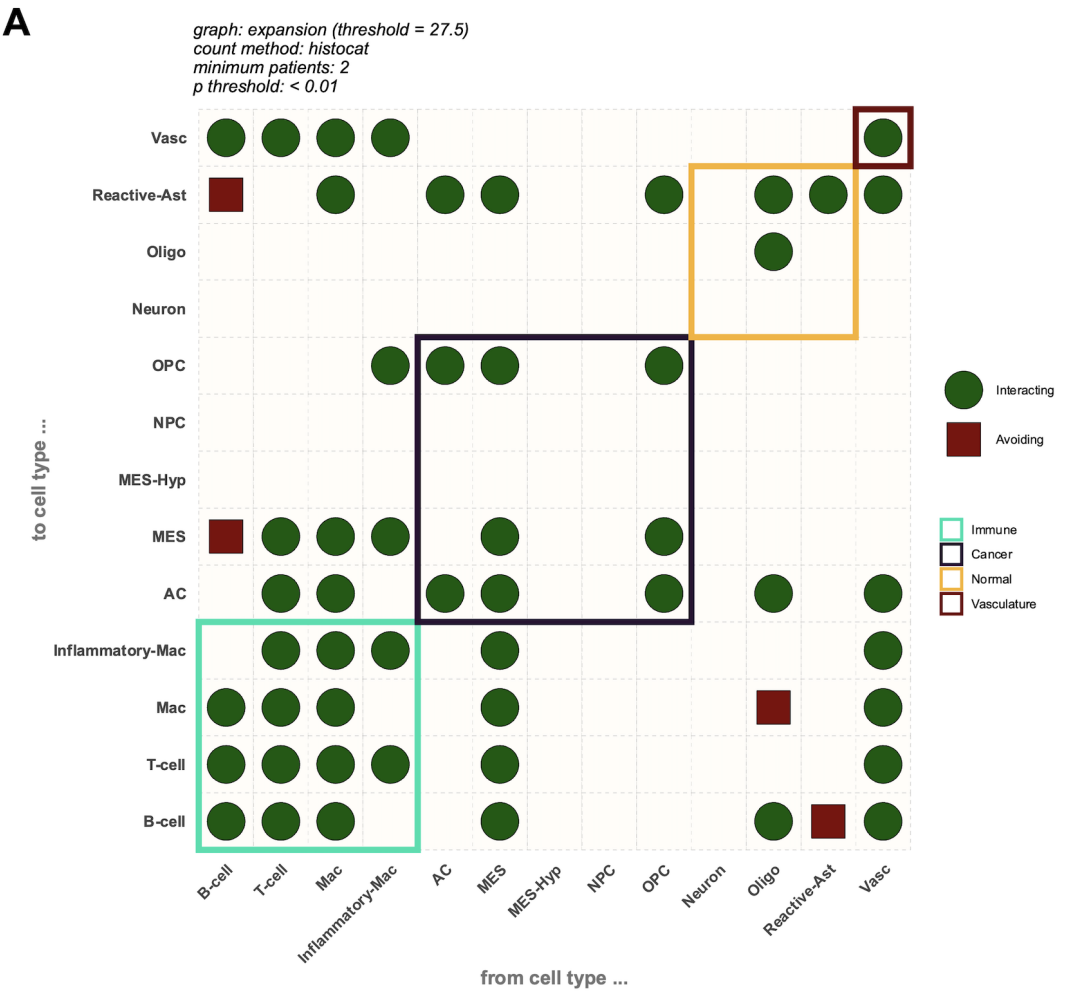
Decreases in specific immune cell populations within the GBM TME through treatment are associated with improved patient outcomes. Notably, overall reductions in microglia correlate with better PFS (HR = 0.14, 95% CI: 4.28×10^{-2} - 0.45) and OS (HR = 0.22, 95% CI: 7.26×10^{-2} - 0.67). Similarly, declining natural killer cell (NK cell) populations, are associated with better PFS (HR = 0.23, 95% CI: 6.26×10^{-2} - 0.87) and OS (HR = 0.18, 95% CI: 3.94×10^{-2} - 0.80) outcomes.

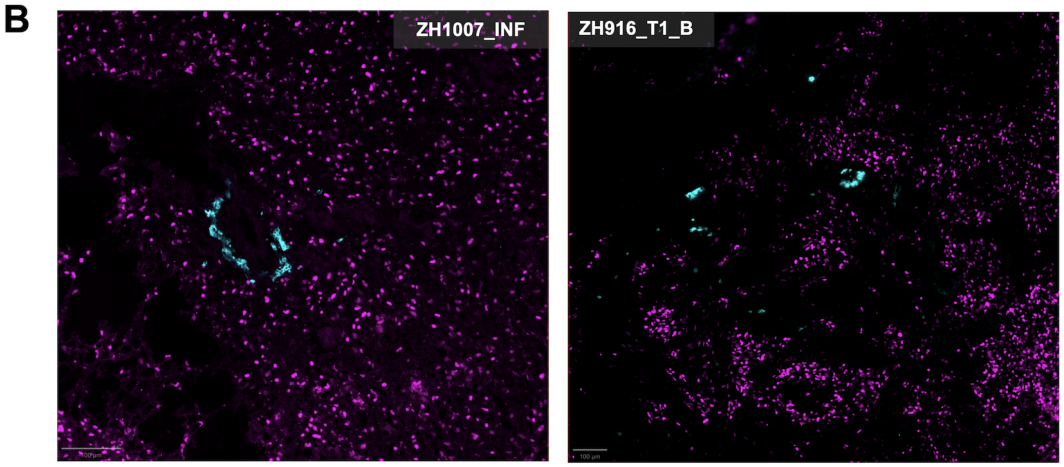
Other immune cells also show treatment-associated changes in prevalence that correlates with OS, but not consistently with PFS. Importantly, these associations with OS depend on the direction of the cell type changes through treatment (Figure 3.1F). This variability reflects the heterogeneity of immune cell populations in the GBM TME, where they can be protective when activated, but also tumour-promoting when co-opted or suppressed, through interactions with neoplastic GBM cells and other components of the TME.

Validating plasma B cell co-localisation with oligodendrocytes

Our data indicate potential functional interplay between plasma B cells and oligodendrocytes in GBM tissue, which appears to be increasingly important through treatment. To see if we could validate the co-localisation of these cell types in GBM tissue, we analysed the cell interactions in two publicly available IDHwt patient samples generated via spatial proteomics (CODEX)[36]. This also provides a further layer of validation of cell type identification using gene versus protein expression. B cells were identified using the pan-B cell marker CD19, which is also expressed on activated plasma B cells. We found that oligodendrocytes significantly interact with B cells (*permutation test*, $p = 9.99 \times 10^{-4}$) in samples where B cells were present (Table S3.3 and Figure 3.2A). These interactions were also confirmed by visual inspection of the protein marker expression in both cell types (Figure 3.2B).

We also utilised spot-based spatial transcriptomics to examine the spatial distribution of cells from the tumour core and the tumour margin. We found that plasma B cells co-localise in the same spatial region as oligodendrocytes in both regions (Figure S3.2 and Figure S3.3). Notably, these niches also comprised very low expression of all other immune cell types, with the exception of macrophages which also co-localised with plasma B cells in the tumour core (Figure S3.2.





B cell (CD19); oligodendrocyte (OLIG2 & APOD)

Figure 3.2 | Spatial proteomic (CODEX) cell-cell interactions.

A).Dotplot showing the cell-cell interactions based on spatial proteomic analysis (CODEX) of two patients. The interactions were determined for each cell type and compared against a null distribution which defined cell positions under spatial randomness. Interactions were only counted in instances where cells of type A that had at least one neighbour of type B: the plot shows the most significant ($p < 0.01$) interactions that were present across both patient samples. Shape and point colours both denote the type of significant cell-cell interaction, and the tile highlights denote the cell category each cell type belongs to. **B)** Representative CODEX images across two patient samples showing the protein marker abundance used to identify oligodendrocytes and B cells with the GBM TME.

3.3.3 Longitudinal cell type changes across responder types

Differential cell type changes

We recently described how IDHwt GBM tumours can be stratified into two responder types based on transcriptional changes through treatment: Up and Down responders[13]. Further, we showed that the biology underpinning each responder type implicated distinct treatment resistance mechanisms. To expand on this work, we investigated cellular landscape changes from primary to recurrence in Up (Figure 3.3A and Table S3.4) and Down (Figure 3.3B and Table S3.5) responder types separately, using our significantly expanded cohorts. Again, we focus here on findings from the discovery cohort that were independently confirmed in the validation cohort.

Post-treatment changes in cell type compositions reveal distinct patterns between Up and Down responders (Figure 3.3). In Up responders, we find that the increases in plasma B cells ($q = 9.97 \times 10^{-6}$), neurons ($q = 2.81 \times 10^{-9}$), and oligodendrocytes ($q = 4.22 \times 10^{-7}$), which we previously observed when analysing all samples together (Figure 3.1B), are maintained. In contrast, these changes are absent in Down responders, which only show significant decreases in endothelial cells ($q = 5.64 \times 10^{-3}$) across both cohorts, following treatment.

Strikingly, neurons show opposing effects, being significantly increased over time in the Up responders (Figure 3.3A), and decreased in the Down responders (Figure 3.3B). The increasing prevalence of neurons in the Up responders occurs alongside validated increases in oligodendrocytes and plasma B cells, which are cells required for myelination necessary to ensheathe newly formed axons. Remyeli-

nation has been extensively studied in multiple sclerosis (a demyelinating disease of the central nervous system) where astrocytes and microglia also facilitate remyelination by interacting with OPCs[58]. We find similar patterns with divergent changes in microglia, astrocytes, and oligodendrocyte populations (including both normal and neoplastic progenitors) depending on responder classification. Taken together, this further supports the idea that Up responders uniquely display the previously identified 'neuronal signalling' phenotype that is enriched at recurrence in GBM in a subset of tumours[13,18].

Cell type correlations

We also observe distinct correlation patterns between Up and Down responders, which are consistent across both the discovery and validation cohorts (Figure 3.3C and Figure 3.3D). These indicate subtype-specific shifts in cellular landscape through treatment.

In Up responders plasma B cells show a strong correlation with oligodendrocytes ($r = 0.86$) and mast cells ($r = 0.59$). In contrast plasma B cells are only associated with astrocytes ($r = 0.7$) in Down responders. There are also marked differences between B cell populations between responder types: Up responders are strongly associated with multiple immune cell types, whereas in Down responders plasma B cells only correlate with mast cells ($r = 0.5$).

Mast cells are myeloid lineage-derived cells which possess high affinity receptors for Immunoglobulin E (IgE) and become sensitised upon binding, leading to degranulation and secretion of cytokines. They play a critical role in linking the innate and adaptive immune responses by mediating the recruitment and activation of macrophages, triggering inflammation, vascular dilation and coordinating further adaptive immune responses. In Up responders mast cells are associated with both lymphoid and myeloid immune cells, suggestive of immune infiltration and the involvement of adaptative immunity. Conversely, in Down responders, mast cells are only significantly associated with NK cells ($r = 0.7$), suggesting a more immunologically restricted environment in which more innate immune mechanisms dominate, possibly in conjunction with the suppression of adaptative immunity.

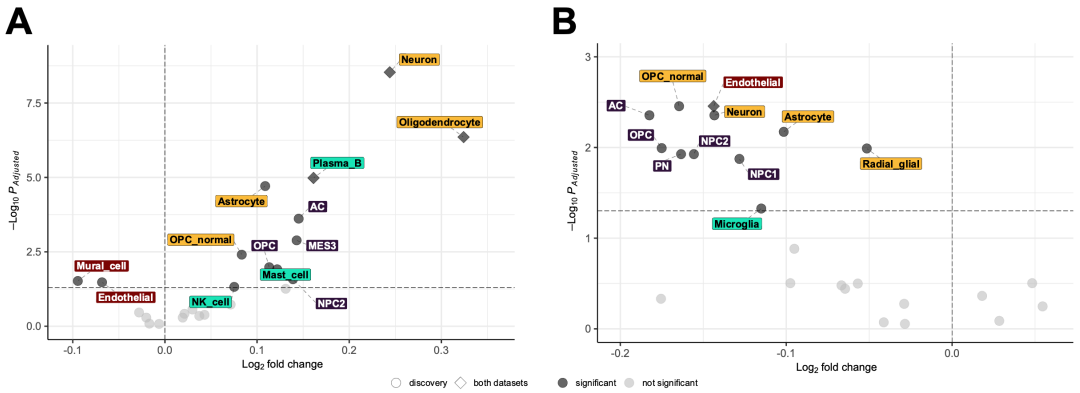
Although vasculature components decrease across both responder types (Figure 3.3A-B), we still find distinct cell type associations present across both cohorts. In Up responders mural cells and endothelial cells are strongly correlated whereas in Down responders, no significant association is observed. We previously reported that Down responders were associated with a more aggressive MES phe-

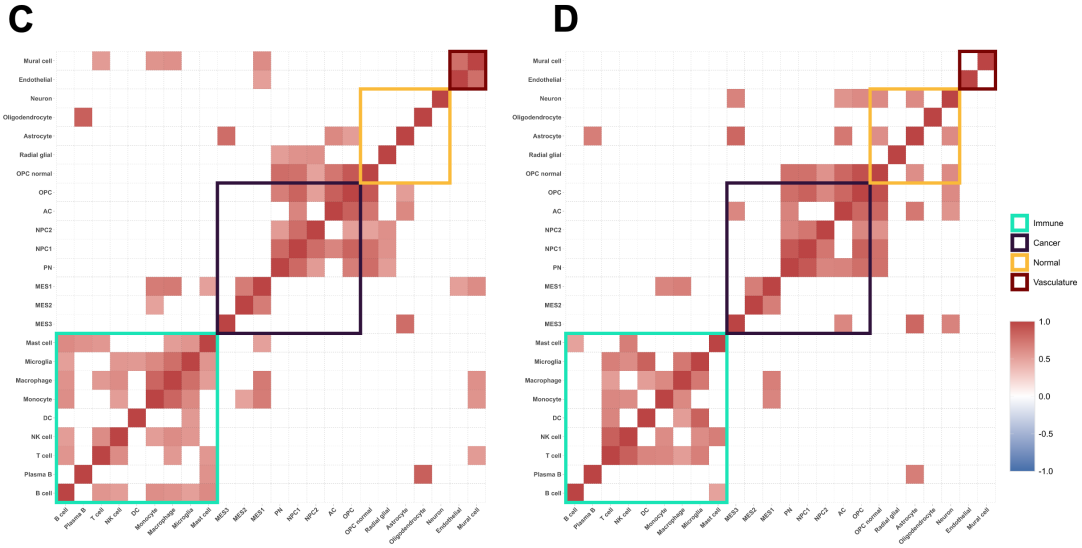
notype with hypoxia as a distinguishing feature[13]. Whilst it is expected that surgical intervention removes significant amounts of vasculature, the specific loss of the mural-endothelial cell association is seen when there is a significant breakdown of the blood-brain barrier (BBB). These data suggest that such disruption of the vasculature is much more prominent in Down responders compared with Up responders. Additionally, in Up responders we find mural cells strongly associate with immune cells: T cells (r = 0.54), macrophages (r = 0.60), monocytes (r = 0.58), but these associations are not significant in Down responders.

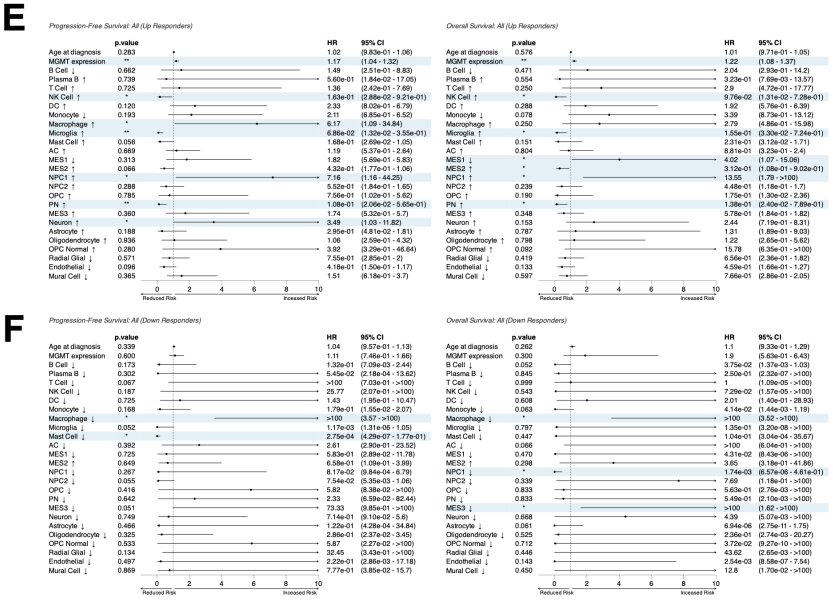
Normal neurons are significantly associated with normal OPCs (r = 0.63) and neurodevelopmental malignant cells including, AC-like (r = 0.58) and OPC-like (r = 0.63) in Down responders. These associations align with recent finding from Greenwald *et al.* who reported a five-layer model of spatial organisation in structured regions within IDHwt GBM tumours, driven by hypoxia[36]. Taken together they suggest that Down responders may exhibit a more aggressive phenotype, with greater infiltration into healthy brain parenchyma. Supporting this, we find that the MES3 malignant cell type that is associated with hypoxia and mesenchymal transition is also significantly correlated with normal neurons (r = 0.67) and AC-like cells (r = 0.65) in Down responders across both cohorts. In contrast, in Up responders, MES3 is associated with astrocytes (r = 0.79), pointing to potential differences in spatial organisation and cell interactions between responder types.

In Up responders the PN cell state is associated with all the key neurodevelopmental malignant cell types and non-malignant cell types including OPC-like (r = 0.79) and radial glial (r = 0.53), with the exception of the OPC-like neoplastic cell state. By contrast, in Down responders the PN state shows significant associations with OPC-like (r = 0.67) and loses all the significant radial glial cell associations that were present in Up responders: PN (r = 0.53), NPC1 (r = 0.59) and NPC1 (r = 0.6).

We previously reported that Up responders may be able to resist treatment by converting to more quiescent, proneural phenotype and/or through differentiation into normal neuronal circuits[13]. These data suggest radial glial cells may play a key role in facilitating such a process. Radial glial cells are the first neural stem cells to appear during embryonic development and have been reported to be present in adult GBM. They have a capacity for self-renewal, existing between quiescent and cycling states and have been shown to be able to exit dormancy through cytokine-mediated activation[59].


A
B
Neuron
Oligodendrocyte
Plasma_B
Astrocyte
AC
OPC_normal
OPC
MES3
Mural_cell
Mast_cell
Endothelial
NK_cell
NPC2
OPC_normal
Endothelial
AC
Neuron
Astrocyte
OPC
NPC2
PN
NPC1
Radial_glial
Microglia
−Log10 P Adjusted
Log2 fold change
discovery
both datasets
significant
not significant


C
D
Immune
Cancer
Normal
Vasculature


E
Progression-Free Survival: All (Up Responders)
Overall Survival: All (Up Responders)
Reduced Risk
Increased Risk

F
Progression-Free Survival: All (Down Responders)
Overall Survival: All (Down Responders)
Reduced Risk
Increased Risk

Figure 3.3 | Cell type changes in Up and Down responders across discovery and validation cohorts.

A-B) Volcano plots showing the changes in GBMdeconvoluteR cell type scores across paired patient samples for Up responder types **(A)** and Down responder types **(B)** in the discovery cohort. The y-axis represents the the $-\log_{10}$ FDR-corrected p-values and the x-axis shows $\log_2$-fold changes. Significant cell type changes ($p < 0.05$) are highlighted (by cell category) and labelled. The point shapes indicate whether changes are also significant in the validation cohort. **C-D)** Heatmap of Spearman rank correlations between change (recurrent – primary) in cell-type scores across Up **(C)** and Down **(D)** responders in the discovery cohort. The tile fill colours represent the direction and strength of each correlation coefficient. The tile fill colour reflects the direction and strength of the correlation coefficients and only strong correlation coefficients $\geq$(-)0.5) that are significant ($p < 0.05$) across both discovery and validation cohorts are shown. **E-F)** Forest plots showing predictors: change (recurrent – primary) in cell-type score through treatment; age at diagnosis; and MGMT protein expression and their associations with PFS (left) and OS (right), across Up responders **(E)** Down responders **(F)**. Arrows after each cell-type denote the median change for that cell type across all comparison samples: up arrow - increased through treatment; down arrow – decreased through treatment. HRs are denoted as points, with horizontal bars representing the 95% confidence intervals. The shaded rows indicate predictors which have a significant ($p < 0.05$) association with survival. Significance thresholds: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; $^{****}p < 0.0001$.

Association between cell types & survival outcomes

Our recent publication reported that the responder type patient stratification demarcated differences in treatment resistance mechanisms between patients along a PN to MES axis[13]. Here we find (Figure 3.1E-F) that an overall decrease in PN cells through treatment is associated with a better PFS (HR = 0.25, 95% CI: 8.72×10^{-2} - 0.72) and OS (HR = 0.22, 95% CI: 6.84×10^{-2} - 0.70). While not statistically significant, the MES3 cell type, which represents the mesenchymal end of the PN–MES axis used in our stratification[9,13], shows an opposite trend suggesting poorer outcomes that is more pronounced for PFS.

To investigate how responder type-specific, cell type changes associated with survival outcomes we combined the discovery and validation cohorts to improve statistical power (Figure 3.3E and Figure 3.3F).

In Up responder, increasing NK cells were correlated with reduced risk in both PFS (HR = 0.16, 95% CI: 2.88×10^{-2} - 0.92) and OS (HR = 9.76×10^{-2}, 95% CI: 1.31×10^{-2} - 0.73). Similarly increasing microglia were associated with better

PFS (HR = 6.86×10^{-2}, 95% CI: 1.32×10^{-2} - 0.36) and OS (HR = 0.16, 95% CI: 3.30×10^{-2} - 0.72), suggesting they are in a pro-inflammatory phenotype.

In contrast, increased macrophages in Up responders are associated with a significantly higher PFS (HR = 6.17, 95% CI: 1.09 - 34.84), pointing toward a more immunosuppressive phenotype. Conversely, in Down responders decreasing macrophages are significantly associated with an increased PFS (HR = >100, 95% CI: 3.57 - >100) and OS (HR = >100, 95% CI: 3.52 - >100), suggestive of a more pro-inflammatory phenotype.

We also find that, in Up responders, increases in PN cells through treatment are associated with better PFS (HR = 0.10, 95% CI: 2.06×10^{-2} - 0.57) and OS (HR = 0.14, 95% CI: 2.40×10^{-2} - 0.79). The opposite effect in MES3 across Up responders is seen where they are associated with a worse PFS, though this isn't significant and also not observed for OS.

3.3.4 Supervised classification using cell type scores accurately predicts surgery type

To further evaluate the importance and significance of cell type changes through treatment, we next investigated whether patterns within the scores themselves could predict patient surgery type.

We benchmarked twelve different supervised binary classification algorithms, combining all data together to increase statistical power and reduce the effects of sample size on predictive performance (Figure 3.4A). Several models were able to achieve ∼70% accuracy, with a corresponding low classification error via *10-fold* cross-validation (Figure 3.4B), indicating that the cell type scores contained meaningful discriminatory information for distinguishing between surgery types. Among the top-performing models, the linear discriminant analysis (LDA) model, achieved a training accuracy of 69.28% (95% CI: ±17.94%) and a test accuracy of 70.27%, with an area under the ROC-AUC of 0.77 (Figure 3.4C). When stratifying patients by responder type, the classification performance improved. For Up responders accuracy was 70.83%, with an ROC-AUC of 0.81 (Figure 3.4D). In Down responders the model achieved an 85.71% accuracy with a ROC-AUC of 0.92, suggesting that the cell type score changes patterns in these patients is strongly associated with surgical outcome (85.71%, ROC-AUC 0.92; Figure 3.4E). We also attempted to classify the responder types (Up and Down)using the cell types scores, but this approach did not achieve a better classification accuracy compared with the featureless model(Figure S3.6).

Having achieved a good classification accuracy for surgery type, we then assessed which cell types were driving this performance by examining the model feature (cell type) importance (Figure 3.4C). Both plasma B cells and normal B cells emerge as important cell types distinguishing between surgery types using all data, suggesting this signal may have a robust biological underpinning. Additionally, myeloid lineage cells comprising macrophages, microglia, and monocytes were also highly important for classifying whether a GBM tumour was sampled pre- or post-treatment. Proneural-associated cell type such as PN and normal neurons also contributed substantially to predictive performance. Collectively, these top features correspond with cell types linked to either proneural or mesenchymal phenotypes, pointing to an underlying biological axis of variation influencing surgical outcome. To explore this further, we looked at the feature importance within the responder type-stratified models.

In Up responders, we see both normal and malignant neurodevelopmental cell types such as astrocytes, neurons, OPC-like and NPC2 are the most influential features (Figure 3.4D). Interestingly, the previously noted prominence of B and plasma B cells from the model using the combined patient data (Figure 3.4C) does not hold. However, radial glial cells, which were less important, become highly important features in Up responders, further supporting their role in the proneural-associated treatment resistance.

The Down responder (Figure 3.4E) stratified model shows important contribution from myeloid lineage cells, previously shown to align with more mesenchymal phenotype, namely macrophage, monocytes, and microglia[2,9,13]. Intriguingly, we find that whilst the Wang *et al.* MES3 is highly important for surgery prediction in Down responders, the Neftel *et al.* MES1 and MES2 cell states did not appear as strong predictors.

Endothelial cells were important for prediction in both responder types but had a greater impact on the predictive performance in Down responders, where the classification error loss was nearly double that seen in Up responders ($\sim$3 vs $\sim$1.7).

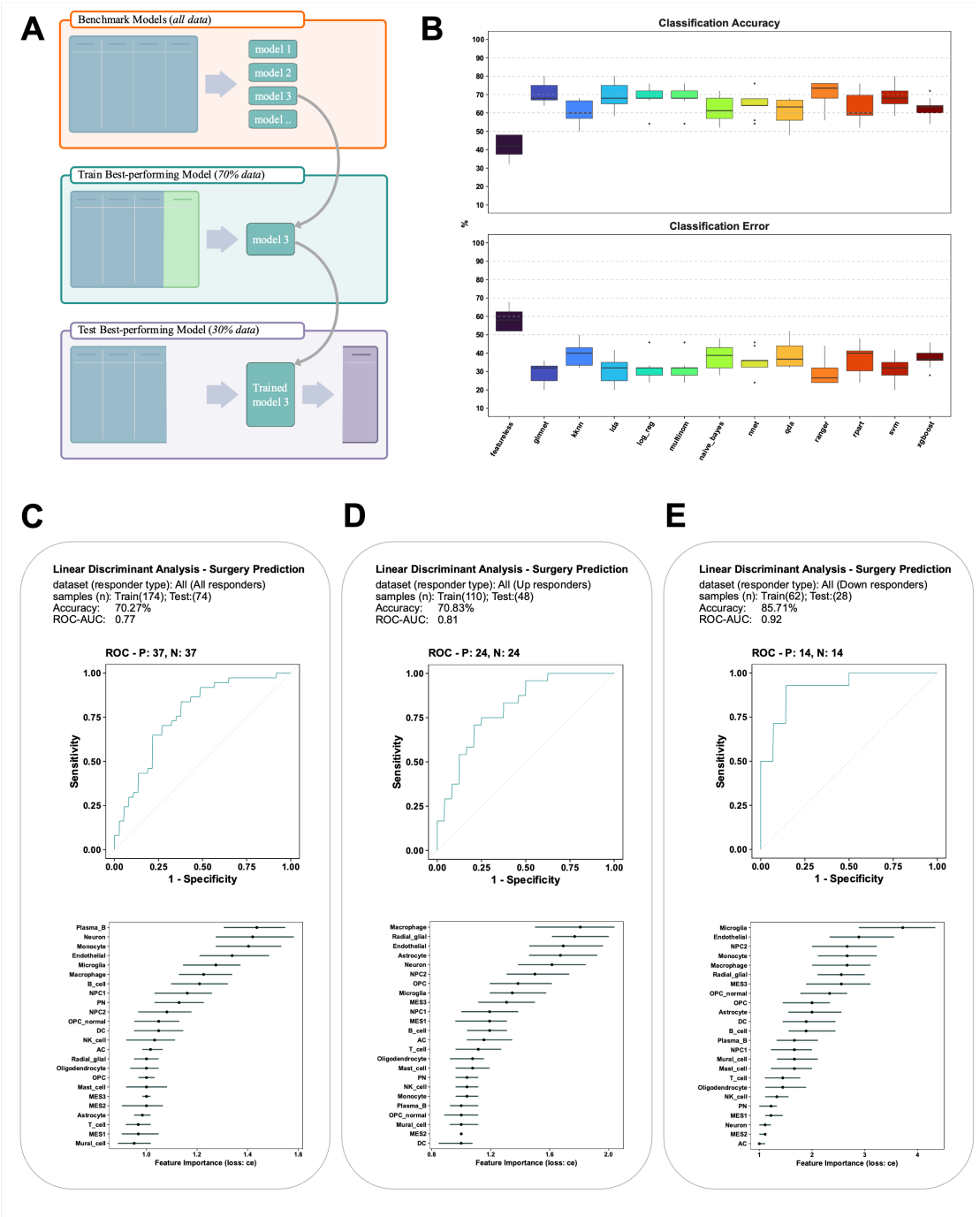

A
Benchmark Models (all data)
model 1
model 2
model 3
model ..
Train Best-performing Model (70% data)
model 3
Test Best-performing Model (30% data)
Trained model 3
B
Classification Accuracy
Classification Error
%
featureless
gbmnet
ksvm
lda
log_reg
multinom
naive_bayes
nnet
qda
ranger
rpart
svm
xgbtree
C
Linear Discriminant Analysis - Surgery Prediction
dataset (responder type): All (All responders)
samples (n): Train(174); Test:(74)
Accuracy: 70.27%
ROC-AUC: 0.77
ROC - P: 37, N: 37
Sensitivity
1 - Specificity
Feature Importance (loss: ce)
D
Linear Discriminant Analysis - Surgery Prediction
dataset (responder type): All (Up responders)
samples (n): Train(110); Test:(48)
Accuracy: 70.83%
ROC-AUC: 0.81
ROC - P: 24, N: 24
Sensitivity
1 - Specificity
Feature Importance (loss: ce)
E
Linear Discriminant Analysis - Surgery Prediction
dataset (responder type): All (Down responders)
samples (n): Train(62); Test:(28)
Accuracy: 85.71%
ROC-AUC: 0.92
ROC - P: 14, N: 14
Sensitivity
1 - Specificity
Feature Importance (loss: ce)

Figure 3.4 | Supervised classification of surgery type using cell type scores.

A) Schematic detailing the process for selecting an optimal machine learning model to predict surgery type. The best performing model was trained on all patient samples and on Up/Down responder patient samples, using 70/30 (*train/test*) split. **B)** Model benchmarking results: boxplots showing the classification accuracy (top) and classification error (bottom) across each machine learning classifier. The benchmarking was conducted using a *10-fold* cross-validation resampling strategy. A featureless model, returning the mean label frequencies was used as a baseline. **C-E)** Model performance (top) and feature importance (bottom) of the best-performing model (LDA) applied to all patient samples **(C)**; Up responder samples **(D)**; and down responder patient samples **(E)**. The model performance is shown by the ROC-AUC which indicates the model's ability to predict surgery types. The model feature importance is determined using the classification error (CE) loss of independently shuffling the values of each feature. Higher increases in CE for a given feature are proportional to its relative importance for the classifier. The median CE loss over all the shuffling repetitions is shown as a point and the confidence intervals are denoted by the length of the horizontal lines. All features are shown ranked in order of highest (top) importance. Models: generalised linear model with elastic net (glmnet); k-nearest neighbour (KNN); LDA; logistic regression (Log-reg); multinomial loglinear model (multinorm); neural network (nnet); quadratic discriminant analysis (QDA); random forest (ranger); classification tree (rpart); support vector machine (SVM); extreme gradient boosting (xgboost).

3.3.5 Unsupervised clustering reveals distinct clusters associated with worse survival outcomes

Building on insights from the supervised classification, we also applied unsupervised machine learning techniques and clustered the cell type scores across surgery using the combined data, to uncover latent cell type patterns.

When examining resulting clusters stratified by primary and recurrent surgeries (Figure 3.5A and Figure 3.5B), we identified one primary sample cluster (cluster 3) that was comprised entirely of Up responders (Figure 3.5A). This cluster is characterised by high levels of cancer cells (Figure 3.5C) and is particularly enriched in AC-like, OPC-like, NPC1/NPC2, and PN. It also features low expression of macrophages, microglia, monocytes and MES1/MES2. All other clusters featured a mixed responder-type distribution, though the overall trend reflected the previously reported distribution, with approximately two-thirds of samples being Up responders and one third Down responders[13].

Although the total number of clusters differed between surgery we found two comparable immune-rich clusters: primary clusters 5 (Figure 3.5C), and recurrent clusters 8 (Figure 3.5D). Both clusters are characterised by high immune and vascular expression, along with medium cancer and normal brain expression. The primary cluster is distinguished by lower expression of astrocytes, OPC-like and MES-like, whereas the recurrent cluster has lower PN and NPC1 expression. When comparing the survival outcomes of all unsupervised clusters we found no significant global difference in either OS or PFS (Figure 3.5E). However, pairwise comparisons revealed a significant OS difference between primary clusters 4 and 5 ($q = 0.01$, Figure 3.5F). The cluster associated with a worse prognosis (cluster 4) is distinguished by high normal brain expression, particularly astrocytes, and reduced lymphoid immune cell infiltration.

Since the Up and Down responder subtype remains stable through treatment, we also performed unsupervised clustering of the primary surgery samples within each responder (Figure 3.6). We identified two comparable clusters: Up cluster 5 (Figure 3.6A) and Down cluster 2 (Figure 3.6B). These clusters were marked by similar aggregated expression profiles (Figure 3.6C-D): medium immune/cancer cells and high vasculature/normal brain cells. Despite their overall similarities, the clusters differed in their immune composition. Down cluster 2 was primarily driven by elevated macrophage expression, whereas Up responder cluster 5 comprised high expression of B cells, plasma B cells, microglia, and mast cells.

We found no significant global differences in OS or PFS when stratifying primary samples by responder type (Figure 3.6E). Though, pairwise comparisons revealed that Up responder cluster 7 had a significantly better PFS outcome ($q = 4.8 \times 10^{-2}$) compared with Up cluster 8. Up cluster 7 is enriched for both cancer and normal brain cell types, with the cancer population dominated by OPC-like, NPC1/NPC1, and PN cells, and the normal brain population by oligodendrocytes and radial glial cells.

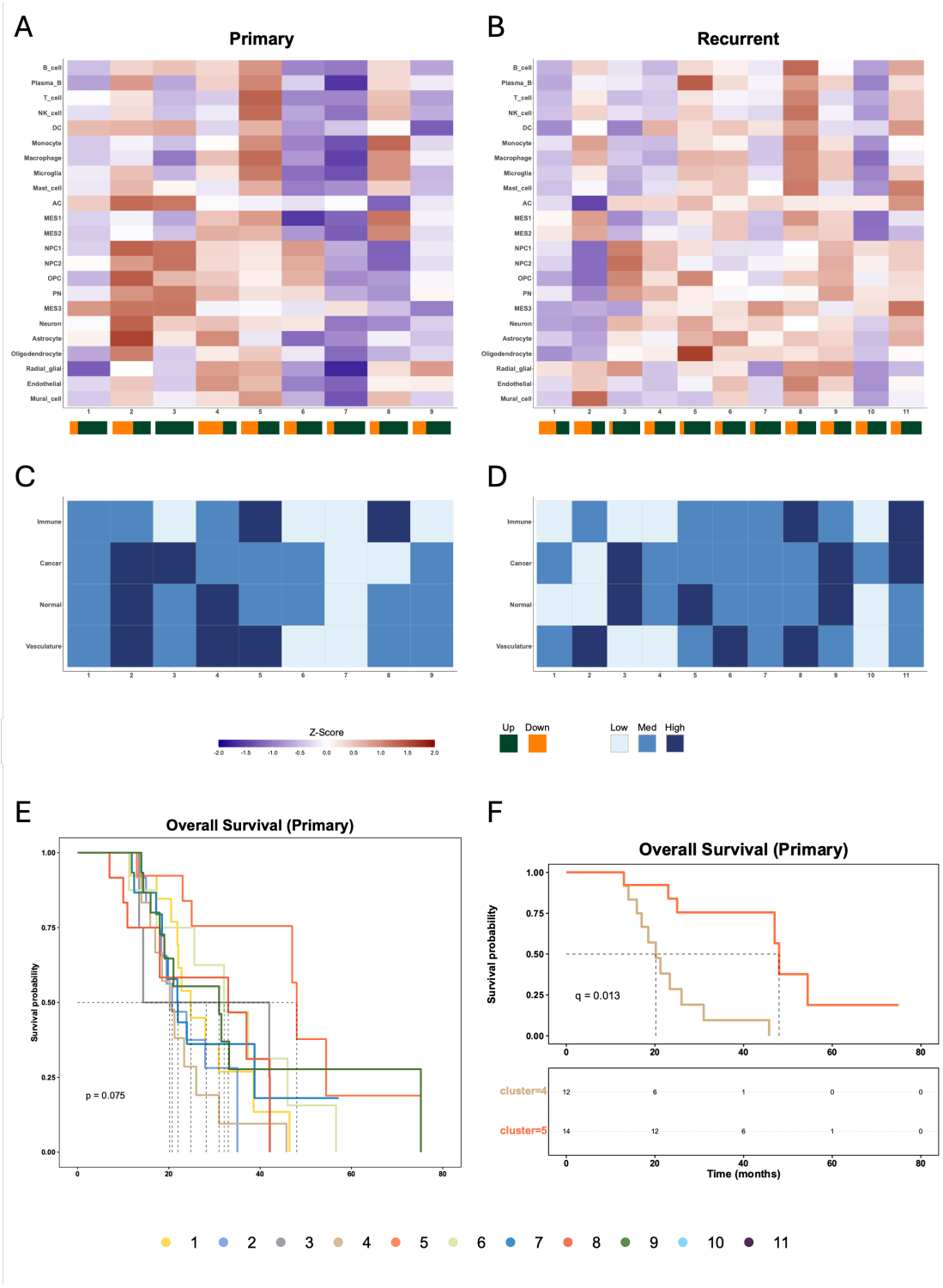

A
Primary
B
Recurrent
B_cell
Plasma_B
T_cell
NK_cell
DC
Monocyte
Macrophage
Microglia
Mast_cell
AC
MES1
MES2
NPC1
NPC2
OPC
PN
MES3
Neuron
Astrocyte
Oligodendrocyte
Radial_glial
Endothelial
Mural_cell
C
D
Immune
Cancer
Normal
Vasculature
Z-Score
Up
Down
Low
Med
High
E
Overall Survival (Primary)
Survival probability
p = 0.075
F
Overall Survival (Primary)
q = 0.013
cluster=4
cluster=5
Time (months)
1
2
3
4
5
6
7
8
9
10
11

Figure 3.5 | Unsupervised clustering of cell type scores across all patients and by surgery type.

A-B) Heatmap of cluster median expression: *z-score* scaled expression of cell type scores across each of the unsupervised *k-means* clusters in primary samples **(A)** and recurrent samples **(B)**. The bar plots below the x-axis denote the percentage of up and Down responder present within each cluster. **C-D)** Heatmaps showing the relative enrichment of each cell type class across each of the clusters in primary samples **(C)** and recurrent samples **(D)**. **E)** Kaplan-Meier survival curve showing the overall survival associated with each of the primary sample *k-means* clusters. The dashed lines denote the median survival associated with each cluster and the p-value represents *Peto's log-rank test* significance indicating if there are difference across all groups simultaneously. **F)** *post-hoc* pairwise comparison of two clusters and their associated overall survival in primary patient samples. The q value shown is *Peto's log-rank test* significance, corrected for multiple testing (*Bonferroni* correction). The table shown below the survival plot details the number of patients surviving at each given time point.

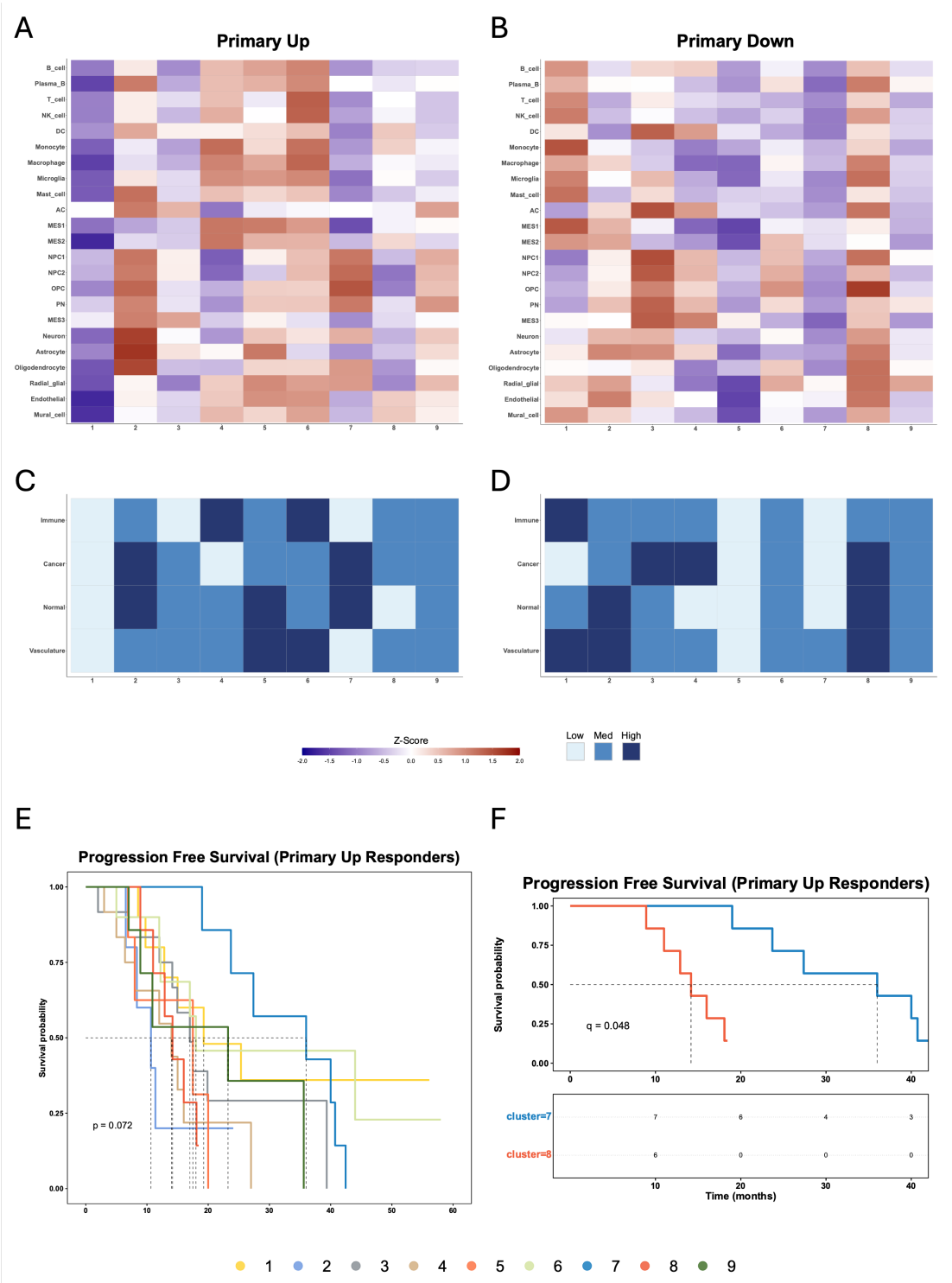

A
Primary Up
B
Primary Down
B_cell
Plasma_B
T_cell
NK_cell
DC
Monocyte
Macrophage
Microglia
Mast_cell
AC
MES1
MES2
NPC1
NPC2
OPC
PN
MES3
Neuron
Astrocyte
Oligodendrocyte
Radial_glial
Endothelial
Mural_cell
C
D
Immune
Cancer
Normal
Vasculature
Z-Score
Low
Med
High
E
Progression Free Survival (Primary Up Responders)
Survival probability
p = 0.072
F
Progression Free Survival (Primary Up Responders)
Survival probability
q = 0.048
cluster=7
cluster=8
Time (months)
1
2
3
4
5
6
7
8
9

Figure 3.6 | Unsupervised clustering of cell type scores in primary patient samples, split across responder-types.

A-B) Heatmap of cluster median expression: *z-score* scaled expression of cell type scores across each of the unsupervised *k-means* clusters in primary, Up responder samples **(A)** and primary, down samples **(B)**. **C-D)** Heatmaps showing the relative enrichment of each cell type class across each of the clusters in primary, Up responders **(C)** and primary, Down responders **(D)**. **E)** Kaplan-Meier survival curve showing the progression-free survival associated with each of the primary, Up responder sample *k-means* clusters. The dashed lines denote the median survival associated with each cluster and the p-value represents *Peto's log-rank* significance indicating if there are difference across all groups simultaneously. **E)** *post-hoc* pairwise comparison of two clusters and their associated overall survival in primary, Up responder patient samples. The q value shown is *Peto's log-rank* significance, corrected for multiple testing (*Bonferroni* correction). The table shown below the survival plot details the number of patients surviving at each given time point.

3.4 Discussion

In this study we used bulk RNA-seq in conjunction with cell type deconvolution to profile the cellular composition of a large, paired cohort of IDHwt GBM patient samples, to better understand how the cell populations change through standard therapy. We note, our data is skewed towards younger adult patients (<60 years of age), which may explain why age at diagnosis does not appear significant when we consider survival associations. Older patients are likely under-represented as they are less likely to undergo de-bulking surgery, and as a result, paired tumour samples are less frequently available for this group. This under-representation is also reflected in the incomplete clinical metadata, particularly in relation to overall survival, which limits the statistical power to assess survival associations within finer stratifications.

Tumour purity is a critical factor when interpreting bulk RNA-seq of GBM tumours, given the heterogeneity in non-malignant cells components across samples. Such heterogeneity can significantly affect the classification of intrinsic tumour features such as the responder types we have previously reported, and therefore it should be carefully considered when drawing conclusions about deconvoluted data from GBM.

Overall, we find that there is a distinct pattern in which increases in oligodendrocyte through treatment are associated with parallel increases in plasma B cells. Moreover, these associations are supported not only by transcriptional correla-

tions, but also at the protein and spatial-level, where we find they occur more frequently than is expected by chance. A potential mechanistic link between these interactions lies in the role of oligodendrocytes in supporting normal neuronal function through homeostasis of axon myelination. While this relationship is not widely reported in GBM, it is well established in the context of multiple sclerosis (MS), which is a demyelinating disease of the central nervous system (CNS)[60–62]. In MS, B cells comprise a relatively minor cell population but contribute to disease progression by producing autoreactive antibodies against myelin upon being abnormally activated by myelin-reactive T cells[61,62]. These mechanisms also involve the activation of brain-resident astrocytes and microglia which facilitate oligodendrocyte damage resulting in the patterns of de-myelination seen in MS[61]. In contrast to their pathological role of de-myelination in MS, other studies have shown that immunoglobulins produced by differentiated B cells can positively influence OPC cell differentiation and promote remyelination[49]. Our data supports the latter in GBM as we find that increases in plasma B cells through treatment are associated with a reduction in OS risk, suggesting that plasma B cell mediated OPC differentiation and re-myelination. We find there is a divergence in B cell dynamics through treatment where increasing plasma B cells are accompanied by decreases in non-activated B cells. This may reflect a maturation process in which B cells transition into more activated plasma B cells and reduce their need to interact with other immune cells. We do also observe a lack of broader immune cell associations between plasma B cells at the protein and spatial levels, indicating that plasma B cells in GBM may acting independently within the GBM TME.

We also find consistent correlations between mast cells and both the lymphoid and myeloid lineage immune cells. The role of mast cells in cancer and specifically neuro-inflammation is well-documented[63]. Mast cells function throughout the early stages of the innate immune response, initiating and amplifying the initial response. Due to this role bridging innate and adaptive immunity, they can be both anti- and pro-tumorigenic, depending on the release of specific cytokines and mediators[64]. Additionally, mast cells also directly influence neurogenesis, neurodegeneration and the permeability of the BBB[63,64]. In our data, we do not find any strong evidence of direct interactions between masts cells and glial, neural, or vascular cells. However, we do find that increasing mast cells through treatment are associated with a better PFS, suggesting in the context of GBM they are playing a more pro-inflammatory role. Given their pivotal role in coordinating the immune response, modulating mast cell activity may be considered a novel therapeutic strategy to enhance the efficacy of current immunotherapies in GBM.

Our findings support the view that PN to MES axis represents the primary source

of cell state variation between malignant GBM cells, consistent with previous studies[13]. However, we note key differences exist between the mesenchymal-like cell states described by Neftel *et al.* and Wang *et al.*[2,9]. The MES3 cell state, defined by Wang *et al.*, shows strong correlations with astrocytes, AC-like malignant cells, normal OPCs and OPC-like cells. This likely reflects the bi-directional model proposed by Wang *et al.*, where AC-like cells are considered intermediate or transitional stages, rather than a discrete subtypes. Further, supporting this, recent refinements of the Neftel MES-like states identified an astrocyte-related variant (MES-Ast), which is specifically enriched with more T cells abundance in addition to the hypoxia and myeloid lineage profiles previously reported[65]. This highlights the continuous and context-dependent nature of malignant GBM cell states. It also reinforces that case for more granular delineation of malignant GBM cell phenotypes, as by only considering the most variable markers present along a bi-directional axis, we may be obscuring meaningful biological differences.

We previously reported a responder type stratification and here in this study, we further added to this model by looking in detail at the cell type changes associated with each of the responder type through treatment[13]. We found that NK cell infiltration through treatment was significantly increased in Up responders, here we add to this finding by reporting that this increase is associated with a reduced PFS and OS when looking at this larger cohort.

Although vasculature decreases overall through treatment, we find that the reduction is associated with different cell types and its composition between responder types. In our initial paper stratifying patients, we found that there was a greater reduction in mural cells in Down responder compared with Up responder. Here we find that mural cells decreases are associated with worse PFS when looking at all patients, but also that these cells are not correlated with and other cell types including endothelial cells in Down responders.

Radial glial cells emerge as a potential mechanism by which Up responders adopt a more quiescent state, supporting treatment resistance through neural stem-like plasticity.

The plasma B cells associations with oligodendrocytes detailed above are only found in Up responders; in Down responders we find that they are associated more with astrocytes, suggesting differing immune-glial interactions that need to be investigated further. Interestingly, reactive astrocytes were found to be actively avoiding B cells when we look at significant cell-cell interactions at the protein-level across two patient samples (Figure S3.2A).

Using supervised classification of surgery type, based on deconvolved cell-type scores, we accurately predicted patient surgery status. Further, we were also able to identify important cell types aligning with the PN to MES axis, when further stratifying by responder types. These findings reinforce the biological importance and utility of our responder type stratifications as a model for understanding treatment resistance in GBM.

In this study we have revealed some prominent changes in the cellular landscape of GBM, through treatment and aligned these to our previously proposed responder-type stratification, uncovering key cell type associations that may underpin treatment resistance. Notably, we highlight the potential role of radial glial cells in facilitating the proneural phenotype observed in Up responders, and the distinct involvement of antibody-producing plasma B cells, which show responder-specific associations. Further experimental validation to characterise these interactions and underlying pathways may uncover new therapeutic targets.

References

1. Louis, D. N. *et al.* The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathologica* **131**, 803–820 (June 2016). doi:10.1007/s00401-016-1545-1.

2. Neftel, C. *et al.* An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. *Cell* **178**, 835–849.e21 (Aug. 2019). doi:10.1016/j.cell.2019.06.024.

3. Verhaak, R. G. W. *et al.* Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in *PDGFRA*, *IDH1*, *EGFR*, and *NF1*. *Cancer Cell* **17**, 98–110 (Jan. 2010). doi:10.1016/j.ccr.2009.12.020.

4. Wang, Q. *et al.* Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. *Cancer Cell* **32**, 42–56.e6 (July 2017). doi:10.1016/j.ccell.2017.06.003.

5. Patel, A. P. *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396–1401 (June 2014). doi:10.1126/science.1254257.

6. Sottoriva, A. *et al.* Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proceedings of the National Academy of Sciences* **110**, 4009–4014 (Mar. 2013). doi:10.1073/pnas.1219747110.

7. Johnson, K. C. *et al.* Single-cell multimodal glioma analyses identify epigenetic regulators of cellular plasticity and environmental stress response. *Nature genetics* **53**, 1456–1468 (Oct. 2021). doi:10.1038/s41588-021-00926-8.

8. Castellan, M. *et al.* Single-cell analyses reveal YAP/TAZ as regulators of stemness and cell plasticity in Glioblastoma. *Nature cancer* **2**, 174–188 (Feb. 2021). doi:10.1038/s43018-020-00150-z.

9. Wang, L. *et al.* The Phenotypes of Proliferating Glioblastoma Cells Reside on a Single Axis of Variation. *Cancer Discovery* **9**, 1708–1719 (Dec. 2019). doi:10.1158/2159-8290.CD-19-0329.

10. Garofano, L. *et al.* Pathway-based classification of glioblastoma uncovers a mitochondrial subtype with therapeutic vulnerabilities. *Nature cancer* **2**, 141–156 (Feb. 2021). doi:10.1038/s43018-020-00159-4.

11. Richards, L. M. *et al.* Gradient of Developmental and Injury Response transcriptional states defines functional vulnerabilities underpinning glioblastoma heterogeneity. *Nature Cancer* **2**, 157–173 (Feb. 2021). doi:10.1038/s43018-020-00154-9.

12. Ajaib, S. *et al.* GBMdeconvoluteR accurately infers proportions of neoplastic and immune cell populations from bulk glioblastoma transcriptomics data. *Neuro-Oncology* **25**, 1236–1248 (July 2023). doi:10.1093/neuonc/noad021.

13. Tanner, G. *et al.* IDHwt glioblastomas can be stratified by their transcriptional response to standard treatment, with implications for targeted therapy. *Genome Biology* **25**, 45 (Feb. 2024). doi:10.1186/s13059-024-03172-3.

14. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2022).

15. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag New York, 2016).

16. Nicoll, J. A. R., Bloom, T., Clarke, A., Boche, D. & Hilton, D. BRAIN UK: Accessing NHS tissue archives for neuroscience research. *Neuropathology and Applied Neurobiology* **48**, e12766 (2022). doi:10.1111/nan.12766.

17. Conway, C. *et al.* Elucidating drivers of oral epithelial dysplasia formation and malignant transformation to cancer using RNAseq. *Oncotarget* **6**, 40186–40201 (Oct. 2015). doi:10.18632/oncotarget.5529.

18. Varn, F. S. *et al.* Glioma progression is shaped by genetic evolution and microenvironment interactions. *Cell* **185**, 2184–2199.e16 (June 2022). doi:10.1016/j.cell.2022.04.038.

19. Hoogstrate, Y. *et al.* Transcriptome analysis reveals tumor microenvironment changes in glioblastoma. *Cancer Cell* **41**, 678–692.e7 (Apr. 2023). doi:10.1016/j.ccell.2023.02.019.

20. Körber, V. *et al.* Evolutionary Trajectories of IDHWT Glioblastomas Reveal a Common Path of Early Tumorigenesis Instigated Years ahead of Initial Diagnosis. *Cancer Cell* **35**, 692–704.e12 (Apr. 2019). doi:10.1016/j.ccell.2019.02.007.

21. Kim, J. *et al.* Spatiotemporal Evolution of the Primary Glioblastoma Genome. *Cancer Cell* **28**, 318–328 (Sept. 2015). doi:10.1016/j.ccell.2015.07.013.

22. Wang, J. *et al.* Clonal Evolution of Glioblastoma under Therapy. *Nature genetics* **48**, 768–776 (July 2016). doi:10.1038/ng.3590.

23. Kim, E. L. *et al.* Intratumoral Heterogeneity and Longitudinal Changes in Gene Expression Predict Differential Drug Sensitivity in Newly Diagnosed and Recurrent Glioblastoma. *Cancers* **12**, 520 (Feb. 2020). doi:10.3390/cancers12020520.

24. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (May 2011). doi:10.14806/ej.17.1.200.

25. Andrews, S. *et al. FastQC* Jan. 2012.

26. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (Jan. 2013). doi:10.1093/bioinformatics/bts635.

27. Trapnell, C. *et al.* Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nature biotechnology* **31**, 10.1038/nbt.2450 (Jan. 2013). doi:10.1038/nbt.2450.

28. Frankish, A. *et al.* GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Research* **47**, D766–D773 (Jan. 2019). doi:10.1093/nar/gky955.

29. Zhang, Y., Parmigiani, G. & Johnson, W. E. ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genomics and Bioinformatics* **2**, lqaa078 (Sept. 2020). doi:10.1093/nargab/lqaa078.

30. Thomas, M. P. H., Ajaib, S., Tanner, G., Bulpitt, A. J. & Stead, L. F. GBMPurity: A Machine Learning Tool for Estimating Glioblastoma Tumour Purity from Bulk RNA-seq Data. *Neuro-Oncology,* noaf026 (Feb. 2025). doi:10.1093/neuonc/noaf026.

31. Ruiz-Moreno, C. *et al. Harmonized single-cell landscape, intercellular crosstalk and tumor architecture of glioblastoma* preprint (Cancer Biology, Aug. 2022). doi:10.1101/2022.08.27.505439.

32. Wilcoxon, F. Individual Comparisons by Ranking Methods. *Biometrics Bulletin* **1**, 80–83 (1945). doi:10.2307/3001968.

33. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* **57**, 289–300 (Jan. 1995). doi:10.1111/j.2517-6161.1995.tb02031.x.

34. Spearman, C. The Proof and Measurement of Association between Two Things. *The American Journal of Psychology* **15**, 72–101 (1904). doi:10.2307/1412159.

35. Schapiro, D. *et al.* histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data. *Nature Methods* **14**, 873–876 (Sept. 2017). doi:10.1038/nmeth.4391.

36. Greenwald, A. C. *et al.* Integrative spatial analysis reveals a multi-layered organization of glioblastoma. *Cell* **187**, 2485–2501.e26 (May 2024). doi:10.1016/j.cell.2024.03.029.

37. Peto, R. & Peto, J. Asymptotically Efficient Rank Invariant Test Procedures. *Journal of the Royal Statistical Society. Series A (General)* **135**, 185–207 (1972). doi:10.2307/2344317.

38. Therneau, T. M., Grambsch, P. M. & Fleming, T. R. Martingale-Based Residuals for Survival Models. *Biometrika* **77**, 147–160 (1990). doi:10.2307/2336057.

39. SCHOENFELD, D. Partial residuals for the proportional hazards regression model. *Biometrika* **69**, 239–241 (Apr. 1982). doi:10.1093/biomet/69.1.239.

40. Wang, Y., Huang, H., Rudin, C. & Shaposhnik, Y. Understanding How Dimension Reduction Tools Work: An Empirical Approach to Deciphering t-SNE, UMAP, TriMap, and PaCMAP for Data Visualization. *Journal of Machine Learning Research* **22**, 1–73 (2021).

41. Lang, M. *et al.* Mlr3: A modern object-oriented machine learning framework in R. *Journal of Open Source Software* **4**, 1903 (Dec. 2019). doi:10.21105/joss.01903.

42. Breiman, L. Random Forests. *Machine Learning* **45**, 5–32 (Oct. 2001). doi:10.1023/A:1010933404324.

43. Molnar, C., Casalicchio, G. & Bischl, B. Iml: An R package for Interpretable Machine Learning. *Journal of Open Source Software* **3**, 786 (June 2018). doi:10.21105/joss.00786.

44. Loussouarn, D. *et al.* Spatial Distribution of Immune Cells in Primary and Recurrent Glioblastoma: A Small Case Study. *Cancers* **15**, 3256 (June 2023). doi:10.3390/cancers15123256.

45. Becerra-Calixto, A. & Cardona-Gómez, G. P. The Role of Astrocytes in Neuroprotection after Brain Stroke: Potential in Cell Therapy. *Frontiers in Molecular Neuroscience* **10**, 88 (Apr. 2017). doi:10.3389/fnmol.2017.00088.

46. Giulian, D., Chen, J., Ingeman, J., George, J. & Noponen, M. The role of mononuclear phagocytes in wound healing after traumatic injury to adult mammalian brain. *The Journal of Neuroscience* **9**, 4416–4429 (Dec. 1989). doi:10.1523/JNEUROSCI.09-12-04416.1989.

47. Schiffer, D., Annovazzi, L., Casalone, C., Corona, C. & Mellai, M. Glioblastoma: Microenvironment and Niche Concept. *Cancers* **11**, 5 (Dec. 2018). doi:10.3390/cancers11010005.

48. Nakahara, J., Seiwa, C., Shibuya, A., Aiso, S. & Asou, H. Expression of Fc receptor for immunoglobulin M in oligodendrocytes and myelin of mouse central nervous system. *Neuroscience Letters* **337**, 73–76 (Feb. 2003). doi:10.1016/s0304-3940(02)01312-5.

49. Li, Y. *et al.* Immunoglobulin directly enhances differentiation of oligodendrocyte-precursor cells and remyelination. *Scientific Reports* **13**, 9394 (June 2023). doi:10.1038/s41598-023-36532-3.

50. Sautès-Fridman, C., Petitprez, F., Calderaro, J. & Fridman, W. H. Tertiary lymphoid structures in the era of cancer immunotherapy. *Nature Reviews Cancer* **19**, 307–325 (June 2019). doi:10.1038/s41568-019-0144-6.

51. Hara, T. *et al.* Interactions between cancer cells and immune cells drive transitions to mesenchymal-like states in glioblastoma. *Cancer Cell* **39**, 779–792.e11 (June 2021). doi:10.1016/j.ccell.2021.05.002.

52. White, K. *et al.* Identification, validation and biological characterisation of novel glioblastoma tumour microenvironment subtypes: implications for precision immunotherapy. *Annals of Oncology* **34**, 300–314 (Mar. 2023). doi:10.1016/j.annonc.2022.11.008.

53. Delgado-López, P. D. & Corrales-García, E. M. Survival in glioblastoma: a review on the impact of treatment modalities. *Clinical and Translational Oncology* **18**, 1062–1071 (Nov. 2016). doi:10.1007/s12094-016-1497-x.

54. He, Z.-C. *et al.* Lower MGMT expression predicts better prognosis in proneural-like glioblastoma. *International Journal of Clinical and Experimental Medicine* **8**, 20287–20294 (Nov. 2015).

55. Zhang, D. *et al.* PHGDH-mediated endothelial metabolism drives glioblastoma resistance to chimeric antigen receptor T cell immunotherapy. *Cell Metabolism* **35**, 517–534.e8 (Mar. 2023). doi:10.1016/j.cmet.2023.01.010.

56. Caspani, E. M., Crossley, P. H., Redondo-Garcia, C. & Martinez, S. Glioblastoma: A Pathogenic Crosstalk between Tumor Cells and Pericytes. *PLOS ONE* **9**, e101402 (July 2014). doi:10.1371/journal.pone.0101402.

57. Pombero, A., Garcia-Lopez, R. & Martinez, S. Brain mesenchymal stem cells: physiology and pathological implications. *Development, Growth & Differentiation* **58**, 469–480 (2016). doi:10.1111/dgd.12296.

58. Enrich-Bengoa, J., Manich, G., Dégano, I. R. & Perálvarez-Marín, A. Deciphering the Genetic Crosstalk between Microglia and Oligodendrocyte Precursor Cells during Demyelination and Remyelination Using Transcriptomic Data. *International Journal of Molecular Sciences* **23**, 14868 (Jan. 2022). doi:10.3390/ijms232314868.

59. Wang, R. *et al.* Adult Human Glioblastomas Harbor Radial Glia-like Cells. *Stem Cell Reports* **14**, 338–350 (Feb. 2020). doi:10.1016/j.stemcr.2020.01.007.

60. Torres Iglesias, G. *et al.* Dual role of peripheral B cells in multiple sclerosis: emerging remote players in demyelination and novel diagnostic biomarkers. *Frontiers in Immunology* **14**, (Aug. 2023). doi:10.3389/fimmu.2023.1224217.

61. Comi, G. *et al.* Role of B Cells in Multiple Sclerosis and Related Disorders. *Annals of Neurology* **89**, 13–23 (2021). doi:10.1002/ana.25927.

62. Friese, M. A., Schattling, B. & Fugger, L. Mechanisms of neurodegeneration and axonal dysfunction in multiple sclerosis. *Nature Reviews Neurology* **10**, 225–238 (Apr. 2014). doi:10.1038/nrneurol.2014.37.

63. Hendriksen, E., van Bergeijk, D., Oosting, R. S. & Redegeld, F. A. Mast cells in neuroinflammation and brain disorders. *Neuroscience & Biobehavioral Reviews* **79**, 119–133 (Aug. 2017). doi:10.1016/j.neubiorev.2017.05.001.

64. Wang, X. *et al.* Mast cells: a double-edged sword in inflammation and fibrosis. *Frontiers in Cell and Developmental Biology* **12**, (Sept. 2024). doi:10.3389/fcell.2024.1466491.

65. Chanoch-Myers, R., Wider, A., Suva, M. L. & Tirosh, I. Elucidating the diversity of malignant mesenchymal states in glioblastoma by integrative analysis. *Genome Medicine* **14**, 106 (Sept. 2022). doi:10.1186/s13073-022-01109-8.

Supplementary material

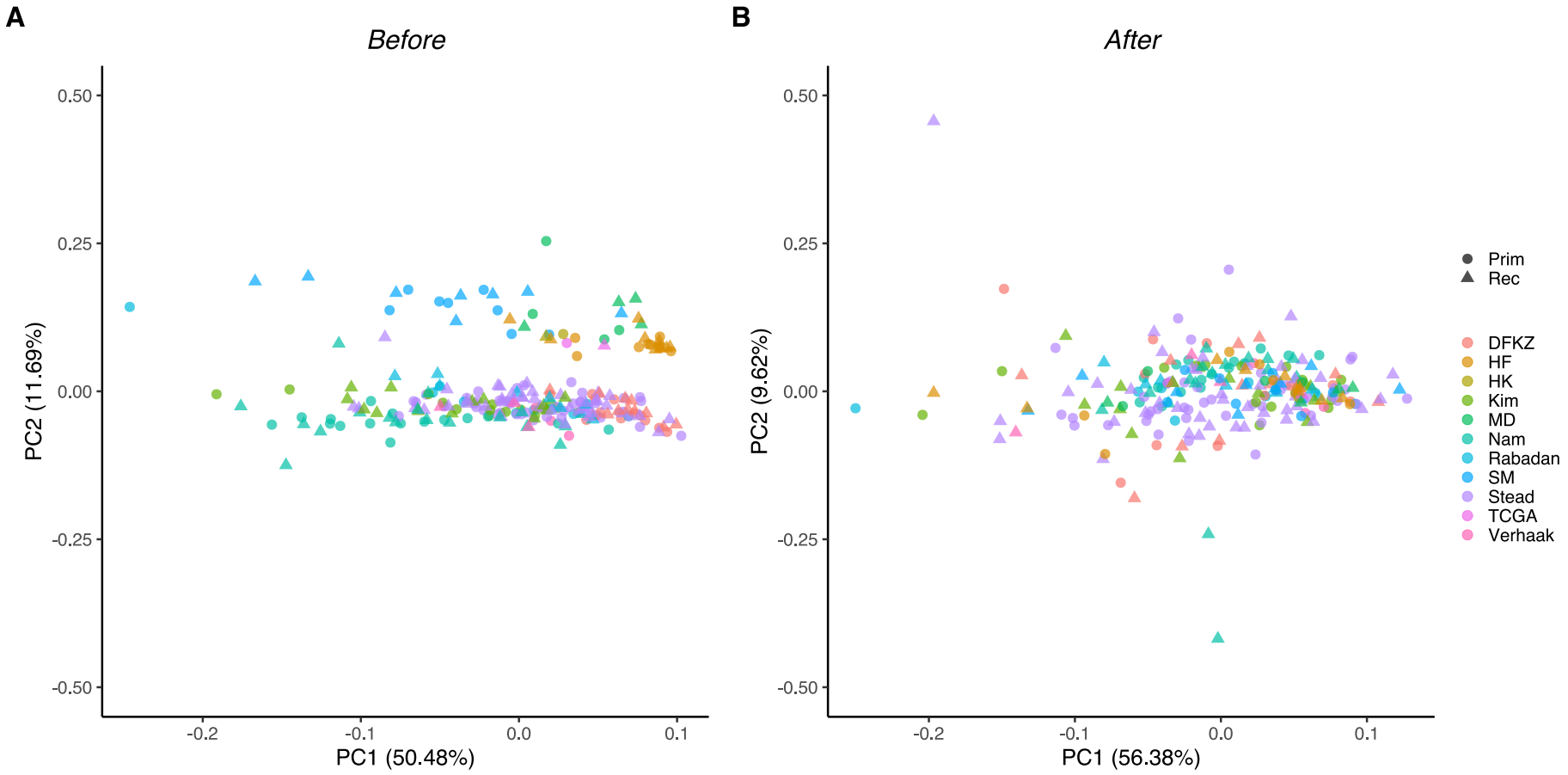

A
Before
B
After
PC2 (11.69%)
PC1 (50.48%)
PC2 (9.62%)
PC1 (56.38%)
0.50
0.25
0.00
-0.25
-0.50
-0.2
-0.1
0.0
0.1
Prim
Rec
DFKZ
HF
HK
Kim
MD
Nam
Rabadan
SM
Stead
TCGA
Verhaak

Figure S3.1 | PCA illustrating batch effects.

Each point represents a sample, coloured by the centre of origin and the shape of the denotes the sample surgery type with primary tumours (Prim) depicted as circles and recurrent tumours (Rec) as triangles. Before batch correction **(A)**, samples cluster based on their centre of origin, indicating significant batch effects. After correction using the Combat-Seq R package **(B)**, these effects are mitigated. PC1 and PC2 percentages indicate the proportion of variance explained by each respective principal component.

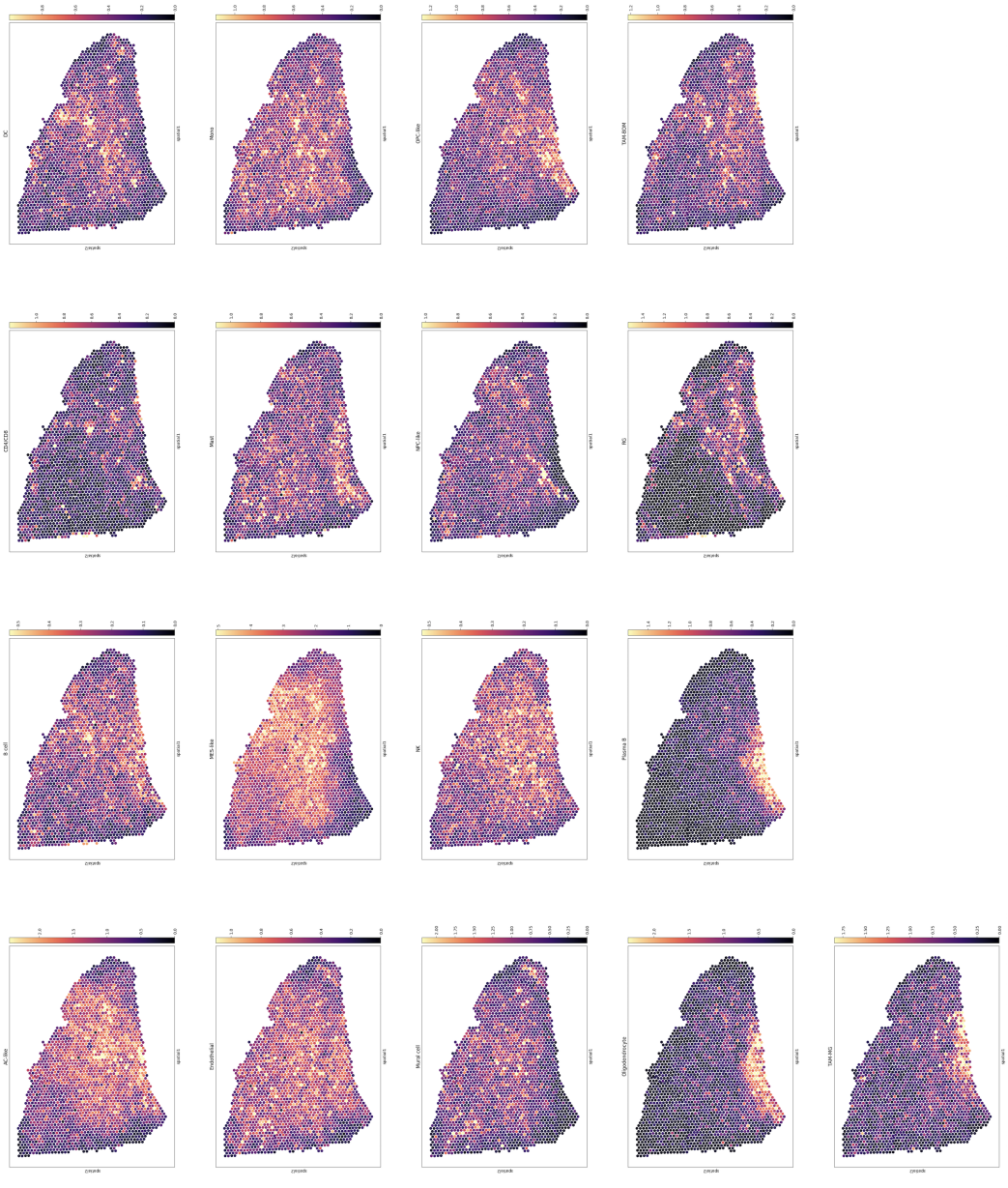

DC
Mono
cDC like
TAM-BDM
CD4/CD8
Mast
aDC like
M2
B cell
MES like
NK
Plasma B
AC like
Endothelial
Mural cell
Oligodendrocyte
TAM-MG

Figure S3.2 | 10X Visium spatial transcriptomics of a GBM tumour core sample.

Cell type abundance estimates at each spatial spot location using *Cell2location* from one GBM patient samples taken from the tumour core region.

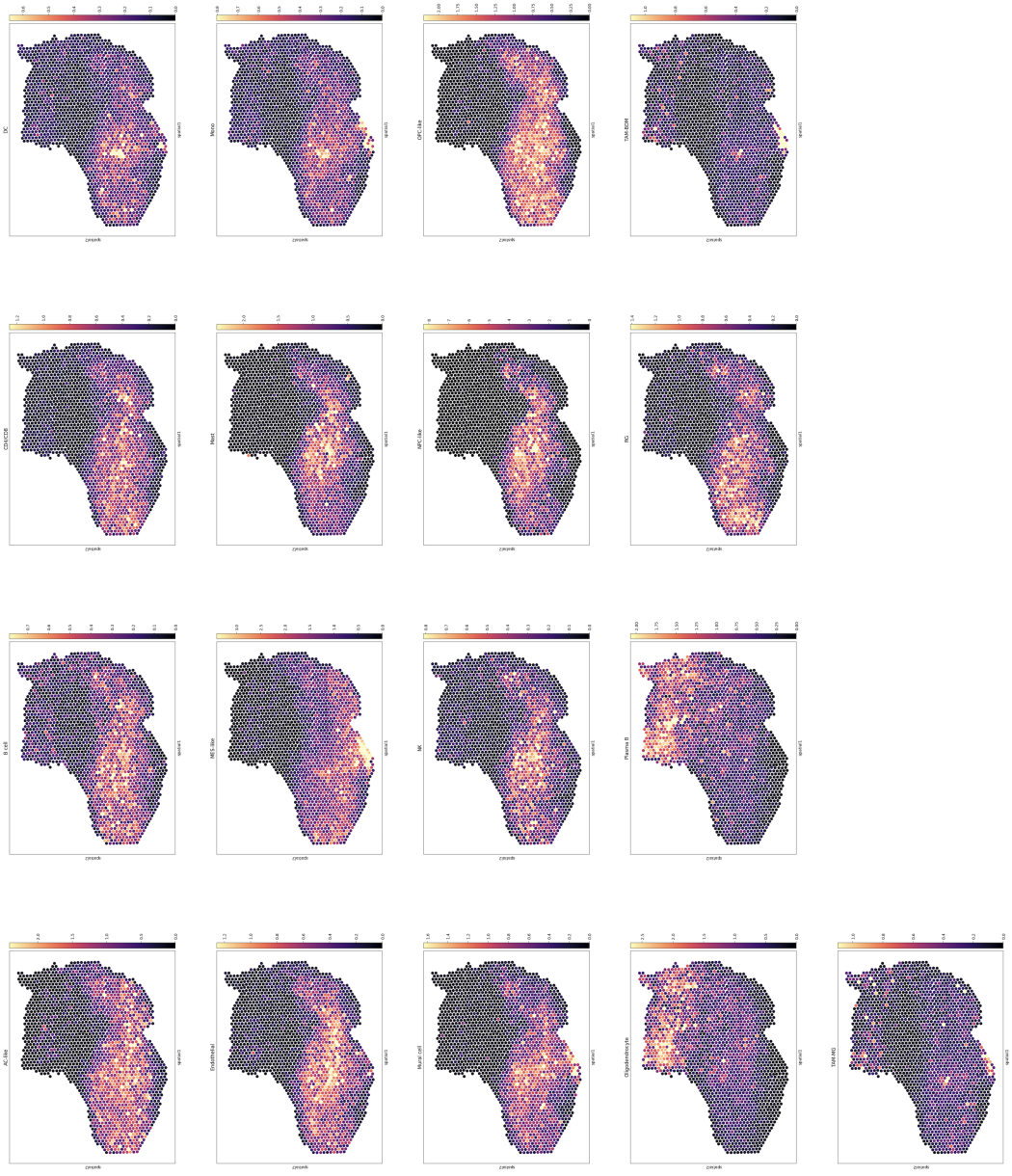


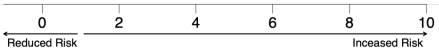
Figure S3.3 | 10X Visium spatial transcriptomics of a GBM tumour edge sample.

Cell type abundance estimates at each spatial spot location using *Cell2location* from one GBM patient samples taken from the tumour edge region.

A

all_up: OS

	p.value	HR	95% CI
diagnosis_age	0.576	1.01	(9.71e−01 – 1.05)
MGMT_ExpP	**	1.22	(1.08 – 1.37)
delta_B_cell	0.471	2.04	(2.93e−01 – 14.2)
delta_Plasma_B	0.554	3.23e−01	(7.69e−03 – 13.57)
delta_T_cell	0.250	2.9	(4.72e−01 – 17.77)
delta_NK_cell	*	9.76e−02	(1.31e−02 – 7.28e−01)
delta_DC	0.288	1.92	(5.76e−01 – 6.39)
delta_Monocyte	0.078	3.39	(8.73e−01 – 13.12)
delta_Macrophage	0.250	2.79	(4.86e−01 – 15.98)
delta_Microglia	*	1.55e−01	(3.30e−02 – 7.24e−01)
delta_Mast_cell	0.151	2.31e−01	(3.12e−02 – 1.71)
delta_AC	0.804	8.81e−01	(3.23e−01 – 2.4)
delta_MES1	*	4.02	(1.07 – 15.06)
delta_MES2	*	3.12e−01	(1.08e−01 – 9.02e−01)
delta_NPC1	*	13.55	(1.79 – >100)
delta_NPC2	0.239	4.48e−01	(1.18e−01 – 1.7)
delta_OPC	0.190	1.75e−01	(1.30e−02 – 2.36)
delta_PN	*	1.38e−01	(2.40e−02 – 7.89e−01)
delta_MES3	0.348	5.78e−01	(1.84e−01 – 1.82)
delta_Neuron	0.153	2.44	(7.19e−01 – 8.31)
delta_Astrocyte	0.787	1.31	(1.89e−01 – 9.03)
delta_Oligodendrocyte	0.798	1.22	(2.65e−01 – 5.62)
delta_OPC_normal	0.092	15.78	(6.35e−01 – >100)
delta_Radial_glial	0.419	6.56e−01	(2.36e−01 – 1.82)
delta_Endothelial	0.133	4.59e−01	(1.66e−01 – 1.27)
delta_Mural_cell	0.597	7.66e−01	(2.86e−01 – 2.05)



B

all_down: OS

	p.value	HR	95% CI
diagnosis_age	0.262	1.1	(9.33e−01 – 1.29)
MGMT_ExpP	0.300	1.9	(5.63e−01 – 6.43)
delta_B_cell	0.052	3.75e−02	(1.37e−03 – 1.03)
delta_Plasma_B	0.845	2.50e−01	(2.32e−07 – >100)
delta_T_cell	0.999	1	(1.09e−05 – >100)
delta_NK_cell	0.543	7.29e−02	(1.57e−05 – >100)
delta_DC	0.608	2.01	(1.40e−01 – 28.93)
delta_Monocyte	0.063	4.14e−02	(1.44e−03 – 1.19)
delta_Macrophage	*	>100	(3.52 – >100)
delta_Microglia	0.797	1.35e−01	(3.20e−08 – >100)
delta_Mast_cell	0.447	1.04e−01	(3.04e−04 – 35.67)
delta_AC	0.066	>100	(6.04e−01 – >100)
delta_MES1	0.470	4.31e−02	(8.43e−06 – >100)
delta_MES2	0.298	3.65	(3.18e−01 – 41.86)
delta_NPC1	*	1.74e−03	(6.57e−06 – 4.61e−01)
delta_NPC2	0.339	7.69	(1.18e−01 – >100)
delta_OPC	0.833	5.63e−01	(2.76e−03 – >100)
delta_PN	0.833	5.49e−01	(2.10e−03 – >100)
delta_MES3	*	>100	(1.62 – >100)
delta_Neuron	0.668	4.39	(5.07e−03 – >100)
delta_Astrocyte	0.061	6.94e−06	(2.75e−11 – 1.75)
delta_Oligodendrocyte	0.525	2.36e−01	(2.74e−03 – 20.27)
delta_OPC_normal	0.712	3.72e−02	(9.27e−10 – >100)
delta_Radial_glial	0.446	43.62	(2.65e−03 – >100)
delta_Endothelial	0.143	2.54e−03	(8.58e−07 – 7.54)
delta_Mural_cell	0.450	12.8	(1.70e−02 – >100)

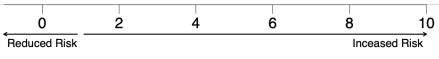
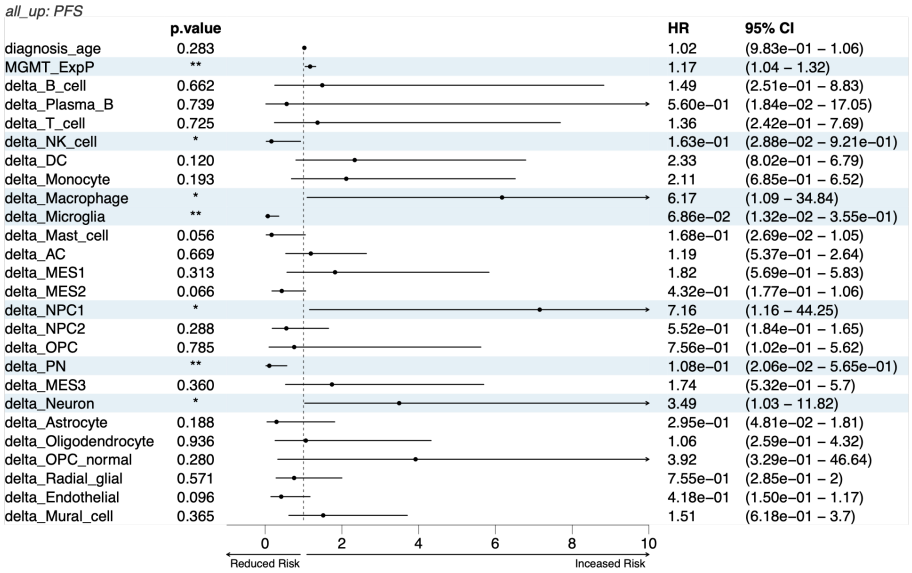


Figure S3.4 | Cell type score associated OS across all responder.

A) Forest plots showing the association between OS and the change (Δ score: recurrent – primary) in cell type scores, age, and MGMT expression for Up responders **(A)** and Down responders **(B)**. HRs are denoted as points, with horizontal bars representing the 95% confidence intervals. The shaded rows indicate predictors which have a significant ($p < 0.05$) association with survival. Significance thresholds: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$.

A


all_up: PFS
p.value
HR
95% CI
diagnosis_age 0.283 1.02 (9.83e−01 – 1.06)
MGMT_ExpP ** 1.17 (1.04 – 1.32)
delta_B_cell 0.662 1.49 (2.51e−01 – 8.83)
delta_Plasma_B 0.739 5.60e−01 (1.84e−02 – 17.05)
delta_T_cell 0.725 1.36 (2.42e−01 – 7.69)
delta_NK_cell * 1.63e−01 (2.88e−02 – 9.21e−01)
delta_DC 0.120 2.33 (8.02e−01 – 6.79)
delta_Monocyte 0.193 2.11 (6.85e−01 – 6.52)
delta_Macrophage * 6.17 (1.09 – 34.84)
delta_Microglia ** 6.86e−02 (1.32e−02 – 3.55e−01)
delta_Mast_cell 0.056 1.68e−01 (2.69e−02 – 1.05)
delta_AC 0.669 1.19 (5.37e−01 – 2.64)
delta_MES1 0.313 1.82 (5.69e−01 – 5.83)
delta_MES2 0.066 4.32e−01 (1.77e−01 – 1.06)
delta_NPC1 * 7.16 (1.16 – 44.25)
delta_NPC2 0.288 5.52e−01 (1.84e−01 – 1.65)
delta_OPC 0.785 7.56e−01 (1.02e−01 – 5.62)
delta_PN ** 1.08e−01 (2.06e−02 – 5.65e−01)
delta_MES3 0.360 1.74 (5.32e−01 – 5.7)
delta_Neuron * 3.49 (1.03 – 11.82)
delta_Astrocyte 0.188 2.95e−01 (4.81e−02 – 1.81)
delta_Oligodendrocyte 0.936 1.06 (2.59e−01 – 4.32)
delta_OPC_normal 0.280 3.92 (3.29e−01 – 46.64)
delta_Radial_glial 0.571 7.55e−01 (2.85e−01 – 2)
delta_Endothelial 0.096 4.18e−01 (1.50e−01 – 1.17)
delta_Mural_cell 0.365 1.51 (6.18e−01 – 3.7)
0
2
4
6
8
10
Reduced Risk
Inceased Risk

B

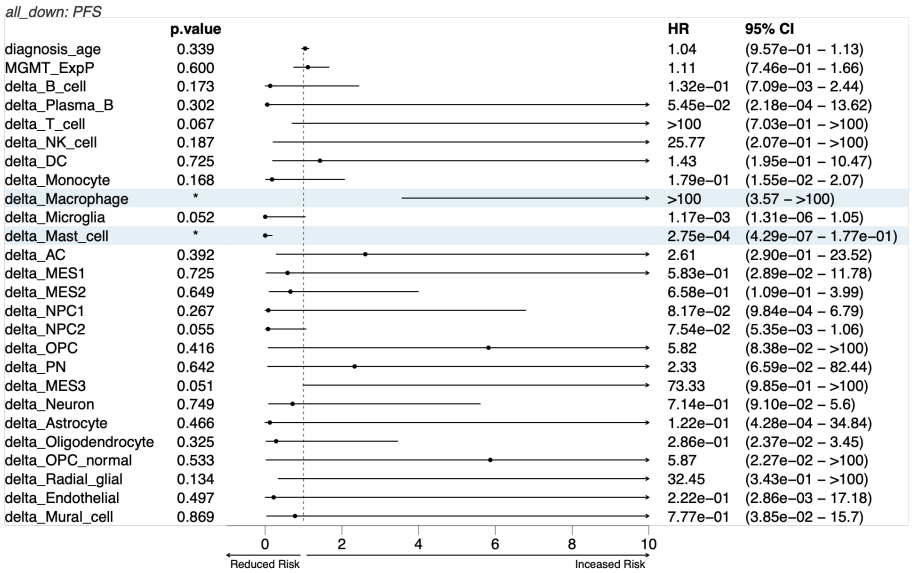

all_down: PFS
p.value
HR
95% CI
diagnosis_age 0.339 1.04 (9.57e−01 – 1.13)
MGMT_ExpP 0.600 1.11 (7.46e−01 – 1.66)
delta_B_cell 0.173 1.32e−01 (7.09e−03 – 2.44)
delta_Plasma_B 0.302 5.45e−02 (2.18e−04 – 13.62)
delta_T_cell 0.067 >100 (7.03e−01 – >100)
delta_NK_cell 0.187 25.77 (2.07e−01 – >100)
delta_DC 0.725 1.43 (1.95e−01 – 10.47)
delta_Monocyte 0.168 1.79e−01 (1.55e−02 – 2.07)
delta_Macrophage * >100 (3.57 – >100)
delta_Microglia 0.052 1.17e−03 (1.31e−06 – 1.05)
delta_Mast_cell * 2.75e−04 (4.29e−07 – 1.77e−01)
delta_AC 0.392 2.61 (2.90e−01 – 23.52)
delta_MES1 0.725 5.83e−01 (2.89e−02 – 11.78)
delta_MES2 0.649 6.58e−01 (1.09e−01 – 3.99)
delta_NPC1 0.267 8.17e−02 (9.84e−04 – 6.79)
delta_NPC2 0.055 7.54e−02 (5.35e−03 – 1.06)
delta_OPC 0.416 5.82 (8.38e−02 – >100)
delta_PN 0.642 2.33 (6.59e−02 – 82.44)
delta_MES3 0.051 73.33 (9.85e−01 – >100)
delta_Neuron 0.749 7.14e−01 (9.10e−02 – 5.6)
delta_Astrocyte 0.466 1.22e−01 (4.28e−04 – 34.84)
delta_Oligodendrocyte 0.325 2.86e−01 (2.37e−02 – 3.45)
delta_OPC_normal 0.533 5.87 (2.27e−02 – >100)
delta_Radial_glial 0.134 32.45 (3.43e−01 – >100)
delta_Endothelial 0.497 2.22e−01 (2.86e−03 – 17.18)
delta_Mural_cell 0.869 7.77e−01 (3.85e−02 – 15.7)
0
2
4
6
8
10
Reduced Risk
Inceased Risk

Figure S3.5 | Cell type score associated PFS across all responder.

A) Forest plots showing the association between PFS and the change (Δ score: recurrent – primary) in cell type scores, age, and MGMT expression for Up responders **(A)** and Down responders **(B)**. HRs are denoted as points, with horizontal bars representing the 95% confidence intervals. The shaded rows indicate predictors which have a significant ($p < 0.05$) association with survival. Significance thresholds: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$.

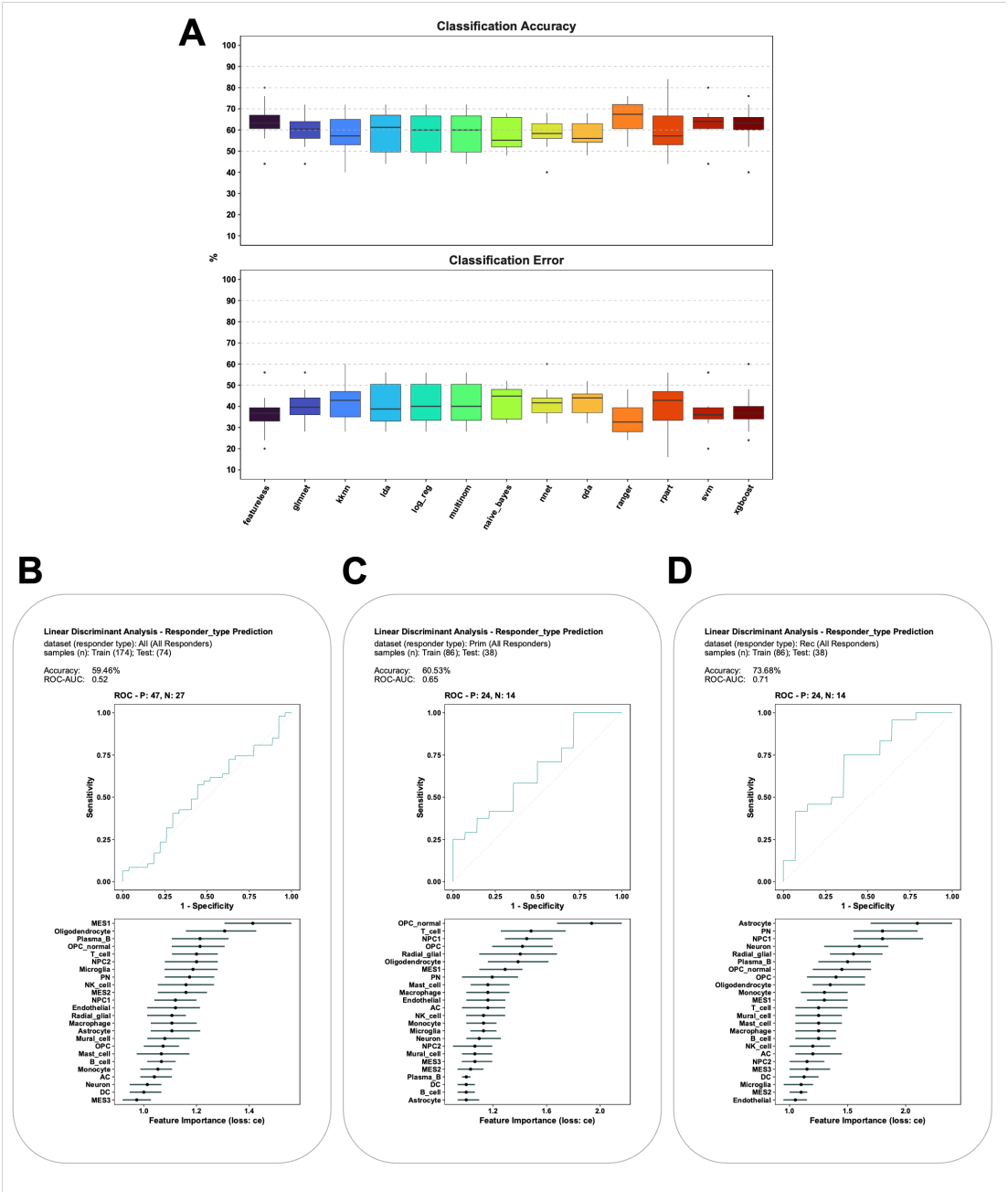

A
Classification Accuracy
Classification Error
%
featureless
glmnet
kknn
lda
log_reg
multinom
naive_bayes
nnet
qda
ranger
rpart
svm
xgboost
B
C
D
Linear Discriminant Analysis - Responder_type Prediction
dataset (responder type): All (All Responders)
samples (n): Train (174); Test: (74)
Accuracy: 59.46%
ROC-AUC: 0.52
ROC - P: 47, N: 27
Sensitivity
1 - Specificity
Feature Importance (loss: ce)
Linear Discriminant Analysis - Responder_type Prediction
dataset (responder type): Prim (All Responders)
samples (n): Train (86); Test: (38)
Accuracy: 60.53%
ROC-AUC: 0.65
ROC - P: 24, N: 14
Sensitivity
1 - Specificity
Feature Importance (loss: ce)
Linear Discriminant Analysis - Responder_type Prediction
dataset (responder type): Rec (All Responders)
samples (n): Train (86); Test: (38)
Accuracy: 73.68%
ROC-AUC: 0.71
ROC - P: 24, N: 14
Sensitivity
1 - Specificity
Feature Importance (loss: ce)

Figure S3.6 | Supervised classification of responder types using cell type scores.

Model benchmarking results: boxplots showing the classification accuracy (top) and classification error (bottom) across each machine learning classifier. The benchmarking was conducted using a *10-fold* cross-validation resampling strategy. A featureless model, returning the mean label frequencies was used as a baseline. **B-D)** Model performance (top) and feature importance (bottom) of the best-performing model (LDA) applied to all patient samples **(B)**; primary surgery samples **(C)**; and recurrent surgery samples **(D)**. The model performance is shown by the ROC-AUC which indicates the model's ability to predict surgery types. The model feature importance is determined using the CE loss of independently shuffling the values of each feature. Higher increases in CE for a given feature are proportional to its relative importance for the classifier. The median CE loss over all the shuffling repetitions is shown as a point and the confidence intervals are denoted by the length of the horizontal lines. All features are shown ranked in order of highest (top) importance. Models: glmnet; KNN; LDA; Log-reg; multinorm; nnet; QDA; ranger; rpart; SVM; xgboost.

Table S3.1 | Clinical and molecular data of patients included in the study.

The publication id corresponds to the PubMed identifier. The age field contains the age of the patient at diagnosis, and the status indicates the survival outcome: alive (0.00) or deceased (1.00). The blue shaded rows denote samples which were experimentally validated using spatial proteomics (IMC).

Patient	Cohort	Dataset	Publication id	Responder	Sex	Age	PFS	OS	Status	MGMT exp	Purity (pair)
Camb10	Discovery	Stead	38326875	up	Male	63	12.00	-	1.00	5.59	62%
Camb12	Discovery	Stead	38326875	up	Male	24	24.00	-	1.00	5.81	24%
Camb13	Discovery	Stead	38326875	down	Male	55	24.00	-	1.00	2.58	76%
Camb2	Discovery	Stead	This Publication	up	Male	33	36.00	-	1.00	10.58	-
Camb4	Discovery	Stead	38326875	up	Male	52	8.00	-	1.00	6.54	69%
Camb7	Discovery	Stead	38326875	up	Male	49	18.00	-	1.00	6.87	54%
Camb8	Discovery	Stead	38326875	down	Male	62	24.00	-	1.00	4.55	77%
Camb9	Discovery	Stead	38326875	down	Male	40	24.00	-	1.00	3.76	81%
H043_4PGF	Discovery	Stead	30905762	down	Male	42	4.11	35.97	1.00	3.93	31%
H043_63R6	Discovery	Stead	30905762	down	Male	47	10.22	17.06	0.00	10.22	45%
H043_BU96	Discovery	Stead	30905762	up	Male	69	1.51	16.70	0.00	8.55	49%
H043_D9MRCY	Discovery	Stead	30905762	down	Male	50	9.63	13.94	1.00	2.42	79%
H043_DSX2	Discovery	Stead	30905762	up	Male	60	10.92	19.50	0.00	7.19	74%
H043_GESMJV	Discovery	Stead	30905762	up	Male	50	6.67	19.10	0.00	9.27	50%
H043_GKS176	Discovery	Stead	30905762	down	Male	70	6.84	26.89	0.00	4.92	43%
H043_LNWEGT	Discovery	Stead	30905762	down	Male	57	12.62	26.04	0.00	10.29	52%
H043_N7LCPV	Discovery	Stead	30905762	up	Male	60	9.90	16.67	0.00	7.70	72%
H043_PWC258	Discovery	Stead	30905762	up	Female	48	21.93	38.40	0.00	1.52	73%
H043_XACH	Discovery	Stead	30905762	down	Female	77	32.68	58.62	0.00	1.82	84%
H043_ZMHY	Discovery	Stead	30905762	up	Female	60	10.65	18.31	1.00	6.47	74%
HF2919	Discovery	Stead	26373279	up	Male	64	11.00	18.00	1.00	1.82	45%
HF2998	Discovery	Stead	26373279	up	Female	67	1.00	5.00	1.00	13.38	23%
HF3162	Discovery	Stead	26373279	up	Male	52	7.00	9.00	0.00	9.95	61%
Imperial14	Discovery	Stead	This Publication	up	Female	70	29.16	-	0.00	3.61	44%
Imperial3	Discovery	Stead	This Publication	down	Female	59	22.44	-	0.00	14.01	19%
K06	Discovery	Stead	26373279	down	Female	37	19.00	47.00	1.00	8.16	51%
K12	Discovery	Stead	26373279	up	Male	50	16.00	33.00	1.00	1.10	43%
K14	Discovery	Stead	26373279	up	Male	60	27.00	33.00	1.00	4.29	36%
K15	Discovery	Stead	26373279	down	Female	34	29.00	32.00	1.00	3.74	25%
K16	Discovery	Stead	26373279	up	Male	51	13.00	32.00	1.00	7.31	38%
K17	Discovery	Stead	This Publication	up	Male	29	20.00	31.00	1.00	7.65	-
K20	Discovery	Stead	26373279	up	Female	38	13.00	25.00	1.00	3.22	16%
K21	Discovery	Stead	26373279	up	Female	32	12.00	25.00	1.00	5.82	41%
K22	Discovery	Stead	26373279	up	Male	51	14.00	23.00	1.00	3.86	28%
K24	Discovery	Stead	26373279	up	Male	74	-	19.00	1.00	0.73	48%
K26	Discovery	Stead	26373279	down	Female	58	12.00	17.00	1.00	5.76	81%
K27	Discovery	Stead	26373279	up	Female	72	7.00	16.00	1.00	7.27	59%

Patient	Cohort	Dataset	Publication id	Responder	Sex	Age	PFS	OS	Status	MGMT exp	Purity (pair)
K30	Discovery	Stead	26373279	down	Male	74	6.00	14.00	1.00	6.66	67%
K32	Discovery	Stead	26373279	up	Male	44	5.00	13.00	1.00	10.24	47%
K34	Discovery	Stead	26373279	up	Male	45	7.00	12.00	1.00	8.03	28%
K36	Discovery	Stead	26373279	up	Male	49	3.00	10.00	1.00	10.47	40%
K37	Discovery	Stead	26373279	up	Male	73	7.00	8.00	1.00	3.74	27%
Leeds101	Discovery	Stead	38326875	up	Male	59	27.00	37.00	1.00	3.48	81%
Leeds105	Discovery	Stead	38326875	down	Male	61	22.00	-	1.00	6.85	70%
Leeds106	Discovery	Stead	38326875	up	Male	62	19.00	21.00	1.00	6.76	65%
Leeds107	Discovery	Stead	This Publication	down	Male	40	9.23	-	0.00	0.00	10%
Leeds28	Discovery	Stead	38326875	up	Female	33	18.00	30.00	0.00	4.75	13%
Preston22	Discovery	Stead	38326875	up	Male	66	8.00	14.00	1.00	3.32	72%
Preston27	Discovery	Stead	38326875	up	Male	67	17.00	28.00	1.00	0.00	60%
Preston34	Discovery	Stead	38326875	up	Male	47	15.00	18.00	1.00	16.16	42%
Preston36	Discovery	Stead	38326875	up	Female	45	14.00	18.00	1.00	4.18	54%
R001	Discovery	Stead	27270107	up	-	70	5.00	7.00	1.00	1.41	40%
R002	Discovery	Stead	27270107	up	-	42	10.00	18.00	1.00	7.65	36%
R003	Discovery	Stead	27270107	up	-	67	9.00	17.00	1.00	6.29	0%
R005	Discovery	Stead	27270107	down	-	58	5.00	13.00	1.00	9.28	67%
R006	Discovery	Stead	27270107	up	-	59	11.00	18.00	1.00	15.19	4%
R007	Discovery	Stead	27270107	down	-	69	12.00	21.00	1.00	6.10	40%
R008	Discovery	Stead	27270107	down	-	46	10.00	15.00	1.00	5.67	79%
R009	Discovery	Stead	This Publication	up	-	67	40.00	46.00	1.00	2.79	-
R010	Discovery	Stead	27270107	up	-	50	17.00	23.00	1.00	10.97	43%
s123	Discovery	Stead	This Publication	up	Male	28	32.00	-	0.00	3.75	-
s139	Discovery	Stead	32102350	down	Male	59	1.00	-	0.00	3.09	67%
s151	Discovery	Stead	32102350	up	Male	68	23.00	-	0.00	0.79	29%
s181	Discovery	Stead	32102350	down	Male	57	4.00	-	0.00	9.73	35%
s245	Discovery	Stead	32102350	up	Male	62	4.00	-	0.00	3.73	12%
s259	Discovery	Stead	32102350	down	Male	56	26.00	-	0.00	2.08	46%
s279	Discovery	Stead	32102350	down	Male	45	9.00	-	0.00	5.44	41%
s295	Discovery	Stead	32102350	down	Male	75	6.00	-	0.00	14.12	26%
s309	Discovery	Stead	32102350	down	Female	74	14.00	-	0.00	7.79	49%
s326	Discovery	Stead	32102350	down	Female	63	8.00	-	0.00	2.25	48%
s341	Discovery	Stead	32102350	down	Female	56	31.00	-	0.00	0.56	69%
s357	Discovery	Stead	32102350	up	Female	72	5.00	-	0.00	9.99	23%
s383	Discovery	Stead	32102350	down	Female	75	9.00	-	0.00	14.22	38%
s402	Discovery	Stead	32102350	down	Female	49	3.00	-	0.00	10.51	1%
Walton10	Discovery	Stead	38326875	up	Male	55	17.00	23.00	1.00	9.29	35%
Walton19	Discovery	Stead	38326875	down	Female	55	11.00	26.00	1.00	9.01	82%
Walton2	Discovery	Stead	38326875	up	Female	36	23.00	42.00	1.00	14.19	31%
Walton43	Discovery	Stead	38326875	up	Male	45	12.20	20.19	1.00	9.55	59%
Walton46	Discovery	Stead	38326875	up	Male	43	6.90	21.24	1.00	8.32	50%
Walton5	Discovery	Stead	38326875	down	Male	53	22.00	31.00	1.00	13.57	51%
Walton50	Discovery	Stead	38326875	up	Male	62	9.21	-	0.00	10.66	13%
Walton55	Discovery	Stead	38326875	down	Male	63	14.53	24.62	1.00	9.68	37%

Patient	Cohort	Dataset	Publication id	Responder	Sex	Age	PFS	OS	Status	MGMT exp	Purity (pair)
Walton59	Discovery	Stead	38326875	down	Male	64	26.47	-	0.00	3.71	71%
Walton62	Discovery	Stead	38326875	up	Female	57	13.35	-	0.00	2.35	71%
Walton63	Discovery	Stead	38326875	up	Male	49	12.89	19.07	1.00	2.36	66%
Walton64	Discovery	Stead	38326875	up	Male	57	22.55	48.76	1.00	6.29	1%
Walton67	Discovery	Stead	38326875	up	Male	51	40.73	75.19	1.00	4.26	62%
Walton71	Discovery	Stead	38326875	down	Female	60	17.10	39.12	1.00	6.36	39%
Walton8	Discovery	Stead	38326875	down	Female	60	26.00	41.00	0.00	18.41	17%
Walton82	Discovery	Stead	38326875	up	Female	59	23.24	33.21	1.00	7.78	66%
Walton84	Discovery	Stead	38326875	up	Female	72	17.49	45.76	1.00	2.53	75%
Walton101	Discovery	Stead	This Publication	down	Female	52	14.60	18.50	1.00	1.59	72%
Walton3	Discovery	Stead	This Publication	up	Female	60	44.00	48.00	1.00	0.22	-
Walton33	Discovery	Stead	This Publication	up	Female	58	16.00	31.50	1.00	0.23	60%
Walton35	Discovery	Stead	This Publication	up	Male	61	30.50	-	0.00	0.00	64%
Walton40	Discovery	Stead	This Publication	up	Male	59	13.30	-	0.00	1.53	39%
Walton48	Discovery	Stead	This Publication	down	Female	66	10.00	15.92	1.00	4.71	54%
Walton91	Discovery	Stead	This Publication	down	Male	40	30.57	35.00	1.00	13.01	61%
GLSS-HF-2548	Discovery	GLASS	35649412	up	Male	52	32.00	40.00	0.00	1.55	46%
GLSS-HF-3050	Discovery	GLASS	35649412	down	Male	63	16.00	22.00	0.00	13.95	43%
GLSS-HF-57AE	Discovery	GLASS	35649412	down	Female	59	4.00	11.00	0.00	11.82	28%
GLSS-HF-6504	Discovery	GLASS	35649412	up	Male	64	8.00	26.00	0.00	8.53	29%
GLSS-HF-6658	Discovery	GLASS	35649412	up	Male	60	5.00	17.00	0.00	11.25	40%
GLSS-HF-9A7A	Discovery	GLASS	35649412	up	Male	49	58.00	75.00	0.00	1.78	53%
GLSS-HF-B972	Discovery	GLASS	35649412	up	Male	49	7.00	21.00	0.00	12.57	58%
GLSS-HF-DF35	Discovery	GLASS	35649412	up	Male	64	10.00	13.00	0.00	9.44	38%
GLSS-HF-EE74	Discovery	GLASS	35649412	up	Male	69	13.00	21.00	0.00	18.65	31%
GLSS-HK-0003	Discovery	GLASS	35649412	up	Female	67	13.00	25.00	0.00	22.34	0%
GLSS-MD-0022	Discovery	GLASS	35649412	down	Male	68	4.00	10.00	0.00	9.16	19%
GLSS-MD-0023	Discovery	GLASS	35649412	up	Male	65	8.00	12.00	0.00	13.20	6%
GLSS-MD-0026	Discovery	GLASS	35649412	down	Male	77	34.00	44.00	0.00	2.95	41%
GLSS-MD-0035	Discovery	GLASS	35649412	down	Male	44	20.00	22.00	0.00	2.74	20%
GLSS-SM-R063	Discovery	GLASS	35649412	down	Female	34	30.00	33.00	0.00	4.70	26%
GLSS-SM-R065	Discovery	GLASS	35649412	up	Female	38	14.00	26.00	0.00	4.36	16%
GLSS-SM-R066	Discovery	GLASS	35649412	down	Male	51	14.00	23.00	0.00	5.20	22%
GLSS-SM-R068	Discovery	GLASS	35649412	up	Male	74	16.00	19.00	0.00	0.96	47%
GLSS-SM-R080	Discovery	GLASS	35649412	down	Male	60	27.00	33.00	0.00	5.72	38%
GLSS-SM-R083	Discovery	GLASS	35649412	down	Male	50	16.00	34.00	0.00	1.80	44%
GLSS-SM-R088	Discovery	GLASS	35649412	up	Female	30	12.00	25.00	0.00	8.21	43%
GLSS-SM-R093	Discovery	GLASS	35649412	up	Male	73	7.00	8.00	0.00	5.36	27%
TCGA-06-0125	Discovery	GLASS	35649412	up	Female	63	27.00	48.00	0.00	1.54	62%
AAG	Validation	EORTC	36898379	down	-	44	3.55	12.46	1.00	7.60	32%
AAM	Validation	EORTC	36898379	up	-	63	8.91	13.48	1.00	1.21	43%
AAN	Validation	EORTC	36898379	up	-	50	14.96	17.33	1.00	1.76	63%
AAP	Validation	EORTC	36898379	up	-	48	17.56	59.24	1.00	0.84	21%
AAS	Validation	EORTC	36898379	down	-	37	8.68	16.70	1.00	7.81	21%
AAT	Validation	EORTC	36898379	up	-	63	27.39	41.98	1.00	0.78	66%

Patient	Cohort	Dataset	Publication id	Responder	Sex	Age	PFS	OS	Status	MGMT exp	Purity (pair)
AAU	Validation	EORTC	36898379	up	-	53	21.27	42.05	1.00	8.36	37%
AAV	Validation	EORTC	36898379	up	-	61	17.49	46.42	1.00	2.55	67%
AAW	Validation	EORTC	36898379	up	-	41	3.42	14.73	1.00	5.18	20%
AAX	Validation	EORTC	36898379	up	-	43	2.01	24.79	1.00	0.58	66%
ACA	Validation	EORTC	36898379	up	-	65	4.83	8.12	1.00	0.63	35%
ADA	Validation	EORTC	36898379	down	-	56	11.93	19.79	1.00	4.50	52%
AFA	Validation	EORTC	36898379	down	-	54	16.31	27.95	1.00	0.66	59%
AHA	Validation	EORTC	36898379	up	-	51	6.41	10.92	1.00	3.87	54%
AIA	Validation	EORTC	36898379	down	-	65	9.21	14.37	1.00	4.58	24%
AMA	Validation	EORTC	36898379	up	-	62	56.12	57.21	0.00	1.16	51%
AOA	Validation	EORTC	36898379	down	-	65	14.27	19.96	0.00	3.86	38%
AZA	Validation	EORTC	36898379	up	-	64	17.62	32.88	1.00	1.15	20%
AZB	Validation	EORTC	36898379	down	-	54	17.23	20.52	1.00	5.66	65%
AZC	Validation	EORTC	36898379	down	-	60	16.70	23.34	1.00	6.48	57%
AZE	Validation	EORTC	36898379	down	-	56	53.26	74.76	0.00	0.52	39%
AZF	Validation	EORTC	36898379	up	-	46	29.85	41.00	1.00	0.52	30%
BAA	Validation	EORTC	36898379	down	-	75	7.69	18.54	1.00	0.69	0%
BAB	Validation	EORTC	36898379	up	-	52	23.70	32.09	1.00	1.01	61%
BAC	Validation	EORTC	36898379	up	-	58	10.98	26.66	1.00	5.31	39%
BAD	Validation	EORTC	36898379	up	-	62	12.82	21.96	1.00	0.31	62%
BAE	Validation	EORTC	36898379	down	-	77	8.52	14.40	1.00	2.38	53%
BAH	Validation	EORTC	36898379	up	-	45	12.92	46.19	1.00	4.71	39%
BAI	Validation	EORTC	36898379	up	-	46	28.90	43.04	1.00	3.61	28%
BAK	Validation	EORTC	36898379	up	-	65	17.39	21.73	1.00	0.66	24%
BAL	Validation	EORTC	36898379	down	-	59	16.80	38.30	1.00	1.71	0%
BAM	Validation	EORTC	36898379	down	-	51	6.71	17.33	1.00	2.48	56%
BAR	Validation	EORTC	36898379	down	-	54	7.50	18.38	1.00	1.67	7%
BAT	Validation	EORTC	36898379	up	-	52	10.88	14.37	1.00	4.16	77%
BAW	Validation	EORTC	36898379	down	-	64	28.37	37.38	1.00	1.00	69%
CAC	Validation	EORTC	36898379	up	-	45	8.88	18.94	1.00	1.96	30%
CAF	Validation	EORTC	36898379	down	-	28	9.07	22.82	1.00	6.46	32%
CAO	Validation	EORTC	36898379	up	-	50	19.89	30.90	1.00	2.15	68%
CBA	Validation	EORTC	36898379	up	-	57	3.58	6.25	1.00	4.03	3%
CBE	Validation	EORTC	36898379	up	-	54	9.76	17.19	1.00	1.19	63%
CBM	Validation	EORTC	36898379	up	-	55	8.91	11.61	1.00	6.30	47%
CBP	Validation	EORTC	36898379	up	-	61	5.95	17.95	1.00	8.85	31%
CBR	Validation	EORTC	36898379	up	-	60	42.44	11.28	1.00	5.87	66%
CBV	Validation	EORTC	36898379	up	-	50	10.13	45.47	1.00	5.48	8%
CDA	Validation	EORTC	36898379	up	-	66	20.65	29.26	1.00	1.91	24%
CDD	Validation	EORTC	36898379	up	-	56	3.25	20.32	0.00	2.51	32%
DAC	Validation	EORTC	36898379	up	-	64	-	19.69	1.00	2.97	37%
EAC	Validation	EORTC	36898379	down	-	53	31.20	54.44	1.00	1.57	63%
EAD	Validation	EORTC	36898379	up	-	64	32.75	35.97	0.00	0.78	81%
EAE	Validation	EORTC	36898379	up	-	48	24.10	33.53	0.00	4.36	58%
EAI	Validation	EORTC	36898379	up	-	49	6.90	12.39	1.00	4.10	52%

Patient	Cohort	Dataset	Publication id	Responder	Sex	Age	PFS	OS	Status	MGMT exp	Purity (pair)
EAJ	Validation	EORTC	36898379	down	-	64	6.38	22.78	1.00	6.44	63%
EAO	Validation	EORTC	36898379	up	-	61	12.03	22.06	1.00	2.20	51%
EAT	Validation	EORTC	36898379	down	-	49	11.51	13.45	1.00	8.49	27%
EAU	Validation	EORTC	36898379	up	-	64	14.17	21.86	1.00	4.00	50%
EAW	Validation	EORTC	36898379	up	-	61	10.59	25.02	1.00	6.54	19%
EAY	Validation	EORTC	36898379	up	-	68	5.88	17.00	1.00	7.43	28%
EAZ	Validation	EORTC	36898379	down	-	57	33.04	38.56	1.00	1.35	74%
EBB	Validation	EORTC	36898379	up	-	76	25.35	38.76	1.00	2.39	42%
EBC	Validation	EORTC	36898379	down	-	58	12.20	25.78	0.00	1.43	58%
EBF	Validation	EORTC	36898379	up	-	49	27.78	38.37	0.00	2.75	50%
EBG	Validation	EORTC	36898379	up	-	56	16.87	28.73	0.00	0.96	56%
EBH	Validation	EORTC	36898379	down	-	57	8.81	16.57	1.00	9.44	17%
EBR	Validation	EORTC	36898379	down	-	50	7.66	9.96	1.00	7.38	35%
EBV	Validation	EORTC	36898379	up	-	52	19.73	28.54	1.00	3.59	14%
ECA	Validation	EORTC	36898379	up	-	66	8.68	21.60	1.00	2.54	29%
ECD	Validation	EORTC	36898379	down	-	54	16.47	25.68	1.00	3.30	29%
ECE	Validation	EORTC	36898379	up	-	64	7.50	14.07	1.00	4.01	33%
ECI	Validation	EORTC	36898379	up	-	70	8.55	11.90	1.00	7.86	71%
ECK	Validation	EORTC	36898379	down	-	63	11.87	18.51	1.00	4.54	49%
ECN	Validation	EORTC	36898379	up	-	62	5.82	10.52	0.00	4.24	32%
FAB	Validation	EORTC	36898379	up	-	44	18.44	24.69	0.00	2.83	80%
FAF	Validation	EORTC	36898379	down	-	24	12.92	28.87	1.00	0.57	40%
FAG	Validation	EORTC	36898379	down	-	51	15.02	29.62	0.00	1.92	66%
FAI	Validation	EORTC	36898379	down	-	55	10.95	18.67	1.00	4.40	36%
FAJ	Validation	EORTC	36898379	up	-	54	39.35	56.61	1.00	0.51	48%
FAK	Validation	EORTC	36898379	up	-	56	8.35	16.44	0.00	6.90	47%
FAM	Validation	EORTC	36898379	up	-	62	22.09	46.45	1.00	1.87	0%
FAN	Validation	EORTC	36898379	down	-	50	23.08	42.08	1.00	0.94	66%
FAP	Validation	EORTC	36898379	down	-	62	7.07	12.16	1.00	6.56	35%
GAG	Validation	EORTC	36898379	down	-	55	16.83	50.83	1.00	3.52	31%
GAH	Validation	EORTC	36898379	up	-	65	19.27	23.97	1.00	2.92	45%
GAI	Validation	EORTC	36898379	up	-	60	6.54	19.50	1.00	5.10	49%
GAK	Validation	EORTC	36898379	up	-	62	11.38	20.71	1.00	1.71	56%
GAL	Validation	EORTC	36898379	up	-	59	20.88	27.72	0.00	1.94	39%
GAM	Validation	EORTC	36898379	down	-	66	18.28	20.91	1.00	7.50	37%
GAP	Validation	EORTC	36898379	up	-	41	8.35	15.91	1.00	5.25	42%
GAR	Validation	EORTC	36898379	up	-	55	18.08	25.58	1.00	1.47	46%
HAB	Validation	EORTC	36898379	up	-	52	7.76	37.25	1.00	1.13	15%
HAD	Validation	EORTC	36898379	up	-	36	9.07	24.39	0.00	11.84	26%
HAE	Validation	EORTC	36898379	up	-	49	6.12	16.01	1.00	4.28	32%
HAF	Validation	EORTC	36898379	up	-	46	7.76	11.24	1.00	6.26	32%
JAE	Validation	EORTC	36898379	down	-	46	6.38	29.42	0.00	1.54	72%
JAH	Validation	EORTC	36898379	up	-	43	35.61	42.15	1.00	2.11	59%
JAK	Validation	EORTC	36898379	up	-	45	14.14	21.86	1.00	3.97	59%
JAL	Validation	EORTC	36898379	down	-	52	6.18	14.66	1.00	7.73	28%

Patient	Cohort	Dataset	Publication id	Responder	Sex	Age	PFS	OS	Status	MGMT exp	Purity (pair)
JAN	Validation	EORTC	36898379	down	-	61	-	23.87	1.00	1.95	70%
KAD	Validation	EORTC	36898379	up	-	51	14.17	29.95	1.00	2.27	20%

Table S3.2 | Cell type score changes across all patient samples.

Comparisons were performed using all data and separately for the discovery and validation cohorts. Statistical significance was assessed using the paired, *Wilcoxon test* and p values shown were adjusted using the FDR method. The change denotes the direction of the fold-change for each cell type score through treatment from primary to recurrence. The rows are ordered by $\log_2$-fold change and shaded rows denote cell types which were statistically significant ($p < 0.05$) across one or more cohorts.

	Discovery			**Validation**			**All**		
	p value	**$\log_2$FC**	**Change**	**p value**	**$\log_2$FC**	**Change**	**p value**	**$\log_2$FC**	**Change**
Oligodendrocyte	7.90×10^{-3}	0.16	up	2.52×10^{-3}	0.21	up	2.32×10^{-5}	0.18	up
Plasma B	3.91×10^{-3}	0.09	up	1.72×10^{-2}	0.11	up	2.43×10^{-5}	0.10	up
Neuron	2.06×10^{-2}	0.10	up	3.84×10^{-2}	0.08	up	7.76×10^{-4}	0.09	up
Mast Cell	5.10×10^{-2}	0.10	up	7.50×10^{-1}	-0.02	down	2.57×10^{-1}	0.05	up
MES3	2.41×10^{-1}	0.05	up	9.74×10^{-1}	-0.02	down	3.82×10^{-1}	0.02	up
DC	5.86×10^{-1}	0.04	up	6.81×10^{-1}	0.00	down	7.50×10^{-1}	0.02	up
T Cell	3.81×10^{-1}	0.06	up	4.71×10^{-1}	-0.07	down	7.74×10^{-1}	0.02	up
Astrocyte	1.94×10^{-1}	0.03	up	9.92×10^{-1}	-0.01	down	2.57×10^{-1}	0.01	up
MES2	5.86×10^{-1}	0.02	up	9.92×10^{-1}	-0.01	down	7.02×10^{-1}	0.01	up
Macrophage	5.86×10^{-1}	0.01	up	9.91×10^{-1}	0.00	down	7.02×10^{-1}	0.00	up
NPC2	8.29×10^{-1}	0.03	up	7.64×10^{-1}	-0.04	down	9.32×10^{-1}	0.00	down
OPC	8.29×10^{-1}	0.00	up	9.91×10^{-1}	-0.02	down	9.27×10^{-1}	-0.01	down
OPC Normal	8.29×10^{-1}	-0.01	down	9.66×10^{-1}	-0.02	down	7.71×10^{-1}	-0.02	down
MES1	7.45×10^{-1}	-0.01	down	4.37×10^{-1}	-0.03	down	4.83×10^{-1}	-0.02	down
AC	5.86×10^{-1}	0.02	up	1.12×10^{-1}	-0.08	down	7.02×10^{-1}	-0.02	down
NPC1	7.45×10^{-1}	-0.01	down	6.41×10^{-1}	-0.05	down	5.86×10^{-1}	-0.02	down
Radial Glial	5.86×10^{-1}	-0.01	down	1.91×10^{-1}	-0.06	down	2.47×10^{-1}	-0.02	down
NK Cell	5.86×10^{-1}	0.02	up	8.47×10^{-2}	-0.12	down	6.61×10^{-1}	-0.03	down
Microglia	5.86×10^{-1}	-0.03	down	3.80×10^{-1}	-0.04	down	2.68×10^{-1}	-0.04	down
Monocyte	5.56×10^{-1}	-0.04	down	4.37×10^{-1}	-0.04	down	2.57×10^{-1}	-0.04	down
PN	5.56×10^{-1}	-0.04	down	4.37×10^{-1}	-0.07	down	2.57×10^{-1}	-0.06	down
Mural Cell	2.89×10^{-2}	-0.08	down	4.88×10^{-2}	-0.08	down	7.87×10^{-4}	-0.08	down
B Cell	5.56×10^{-1}	-0.08	down	2.32×10^{-1}	-0.13	down	2.00×10^{-1}	-0.10	down
Endothelial	1.10×10^{-3}	-0.10	down	1.57×10^{-3}	-0.13	down	4.46×10^{-7}	-0.11	down

Table S3.3 | Patient-specific cell-cell interactions from spatial proteomics (CODEX) data.

The results are separated by patient and show the cell-cell interactions compared to a null model of spatial randomness. Statistical significance and direction is determined using a permutation test, with p-values indicating interactions more or less likely than random: 1 (significant positive interactions); -1 (significant avoidance interactions); 0 (neutral and/or non- statistically significant interactions).

Patient	From cell	To cell	Observed count	Permutations (>)	Permutations (<)	p value	Significance/Direction
ZH1007	AC	AC	2.39	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	AC	B-cell	0.00	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH1007	AC	Inflammatory-Mac	1.25	8.89×10^{-1}	1.16×10^{-1}	1.16×10^{-1}	0
ZH1007	AC	MES	2.20	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	AC	MES-Hyp	-	-	-	-	-
ZH1007	AC	Mac	1.79	2.00×10^{-3}	9.99×10^{-1}	2.00×10^{-3}	1
ZH1007	AC	NPC	-	-	-	-	-
ZH1007	AC	Neuron	-	-	-	-	-
ZH1007	AC	OPC	2.43	3.00×10^{-3}	9.98×10^{-1}	3.00×10^{-3}	1
ZH1007	AC	Oligo	1.29	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH1007	AC	Reactive-Ast	1.42	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	AC	T-cell	1.05	5.88×10^{-1}	4.22×10^{-1}	4.22×10^{-1}	0
ZH1007	AC	Vasc	3.01	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	B-cell	AC	0.00	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH1007	B-cell	B-cell	5.53	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	B-cell	Inflammatory-Mac	0.00	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH1007	B-cell	MES	1.00	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH1007	B-cell	MES-Hyp	-	-	-	-	-
ZH1007	B-cell	Mac	4.45	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	B-cell	NPC	-	-	-	-	-
ZH1007	B-cell	Neuron	-	-	-	-	-
ZH1007	B-cell	OPC	2.67	1.06×10^{-1}	9.04×10^{-1}	1.06×10^{-1}	0
ZH1007	B-cell	Oligo	1.00	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH1007	B-cell	Reactive-Ast	0.00	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH1007	B-cell	T-cell	4.39	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	B-cell	Vasc	3.63	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Inflammatory-Mac	AC	1.15	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH1007	Inflammatory-Mac	B-cell	0.00	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH1007	Inflammatory-Mac	Inflammatory-Mac	4.82	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Inflammatory-Mac	MES	2.45	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Inflammatory-Mac	MES-Hyp	-	-	-	-	-
ZH1007	Inflammatory-Mac	Mac	1.74	1.70×10^{-2}	9.84×10^{-1}	1.70×10^{-2}	0
ZH1007	Inflammatory-Mac	NPC	-	-	-	-	-
ZH1007	Inflammatory-Mac	Neuron	-	-	-	-	-
ZH1007	Inflammatory-Mac	OPC	2.89	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Inflammatory-Mac	Oligo	2.00	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Inflammatory-Mac	Reactive-Ast	1.16	2.97×10^{-1}	7.05×10^{-1}	2.97×10^{-1}	0

Patient	From cell	To cell	Observed count	Permutations (>)	Permutations (<)	p value	Significance/Direction
ZH1007	Inflammatory-Mac	T-cell	1.16	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Inflammatory-Mac	Vasc	3.69	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	MES	AC	1.59	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	MES	B-cell	3.00	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	MES	Inflammatory-Mac	2.87	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	MES	MES	4.20	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	MES	MES-Hyp	-	-	-	-	-
ZH1007	MES	Mac	2.19	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	MES	NPC	-	-	-	-	-
ZH1007	MES	Neuron	-	-	-	-	-
ZH1007	MES	OPC	2.63	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	MES	Oligo	1.31	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH1007	MES	Reactive-Ast	1.22	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	MES	T-cell	1.23	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	MES	Vasc	2.96	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	MES-Hyp	AC	-	-	-	-	-
ZH1007	MES-Hyp	B-cell	-	-	-	-	-
ZH1007	MES-Hyp	Inflammatory-Mac	-	-	-	-	-
ZH1007	MES-Hyp	MES	-	-	-	-	-
ZH1007	MES-Hyp	MES-Hyp	-	-	-	-	-
ZH1007	MES-Hyp	Mac	-	-	-	-	-
ZH1007	MES-Hyp	NPC	-	-	-	-	-
ZH1007	MES-Hyp	Neuron	-	-	-	-	-
ZH1007	MES-Hyp	OPC	-	-	-	-	-
ZH1007	MES-Hyp	Oligo	-	-	-	-	-
ZH1007	MES-Hyp	Reactive-Ast	-	-	-	-	-
ZH1007	MES-Hyp	T-cell	-	-	-	-	-
ZH1007	MES-Hyp	Vasc	-	-	-	-	-
ZH1007	Mac	AC	1.63	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Mac	B-cell	5.11	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Mac	Inflammatory-Mac	2.07	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Mac	MES	2.98	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Mac	MES-Hyp	-	-	-	-	-
ZH1007	Mac	Mac	2.27	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Mac	NPC	-	-	-	-	-
ZH1007	Mac	Neuron	-	-	-	-	-
ZH1007	Mac	OPC	2.12	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH1007	Mac	Oligo	1.47	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH1007	Mac	Reactive-Ast	1.28	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Mac	T-cell	1.54	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Mac	Vasc	3.93	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	NPC	AC	-	-	-	-	-
ZH1007	NPC	B-cell	-	-	-	-	-
ZH1007	NPC	Inflammatory-Mac	-	-	-	-	-
ZH1007	NPC	MES	-	-	-	-	-

Patient	From cell	To cell	Observed count	Permutations (>)	Permutations (<)	p value	Significance/Direction
ZH1007	NPC	MES-Hyp	-	-	-	-	-
ZH1007	NPC	Mac	-	-	-	-	-
ZH1007	NPC	NPC	-	-	-	-	-
ZH1007	NPC	Neuron	-	-	-	-	-
ZH1007	NPC	OPC	-	-	-	-	-
ZH1007	NPC	Oligo	-	-	-	-	-
ZH1007	NPC	Reactive-Ast	-	-	-	-	-
ZH1007	NPC	T-cell	-	-	-	-	-
ZH1007	NPC	Vasc	-	-	-	-	-
ZH1007	Neuron	AC	-	-	-	-	-
ZH1007	Neuron	B-cell	-	-	-	-	-
ZH1007	Neuron	Inflammatory-Mac	-	-	-	-	-
ZH1007	Neuron	MES	-	-	-	-	-
ZH1007	Neuron	MES-Hyp	-	-	-	-	-
ZH1007	Neuron	Mac	-	-	-	-	-
ZH1007	Neuron	NPC	-	-	-	-	-
ZH1007	Neuron	Neuron	-	-	-	-	-
ZH1007	Neuron	OPC	-	-	-	-	-
ZH1007	Neuron	Oligo	-	-	-	-	-
ZH1007	Neuron	Reactive-Ast	-	-	-	-	-
ZH1007	Neuron	T-cell	-	-	-	-	-
ZH1007	Neuron	Vasc	-	-	-	-	-
ZH1007	OPC	AC	1.85	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	OPC	B-cell	1.14	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	OPC	Inflammatory-Mac	2.80	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	OPC	MES	2.47	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	OPC	MES-Hyp	-	-	-	-	-
ZH1007	OPC	Mac	1.64	9.36×10^{-1}	6.49×10^{-2}	6.49×10^{-2}	0
ZH1007	OPC	NPC	-	-	-	-	-
ZH1007	OPC	Neuron	-	-	-	-	-
ZH1007	OPC	OPC	3.04	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	OPC	Oligo	1.49	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH1007	OPC	Reactive-Ast	1.25	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	OPC	T-cell	1.28	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	OPC	Vasc	2.81	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Oligo	AC	1.66	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Oligo	B-cell	1.50	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Oligo	Inflammatory-Mac	1.39	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Oligo	MES	1.97	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Oligo	MES-Hyp	-	-	-	-	-
ZH1007	Oligo	Mac	1.55	9.99×10^{-4}	1.00	9.99×10^{-4}	-1
ZH1007	Oligo	NPC	-	-	-	-	-
ZH1007	Oligo	Neuron	-	-	-	-	-
ZH1007	Oligo	OPC	2.03	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH1007	Oligo	Oligo	2.23	9.99×10^{-4}	1.00	9.99×10^{-4}	1

Patient	From cell	To cell	Observed count	Permutations (>)	Permutations (<)	p value	Significance/Direction
ZH1007	Oligo	Reactive-Ast	1.25	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Oligo	T-cell	1.14	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Oligo	Vasc	2.90	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Reactive-Ast	AC	1.72	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Reactive-Ast	B-cell	0.00	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH1007	Reactive-Ast	Inflammatory-Mac	1.20	9.90×10^{-1}	1.10×10^{-2}	1.10×10^{-2}	0
ZH1007	Reactive-Ast	MES	2.21	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Reactive-Ast	MES-Hyp	-	-	-	-	-
ZH1007	Reactive-Ast	Mac	1.64	7.26×10^{-1}	2.75×10^{-1}	2.75×10^{-1}	0
ZH1007	Reactive-Ast	NPC	-	-	-	-	-
ZH1007	Reactive-Ast	Neuron	-	-	-	-	-
ZH1007	Reactive-Ast	OPC	2.05	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH1007	Reactive-Ast	Oligo	1.58	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH1007	Reactive-Ast	Reactive-Ast	1.37	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Reactive-Ast	T-cell	1.29	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Reactive-Ast	Vasc	2.81	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	T-cell	AC	1.57	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	T-cell	B-cell	4.25	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	T-cell	Inflammatory-Mac	2.35	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	T-cell	MES	2.68	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	T-cell	MES-Hyp	-	-	-	-	-
ZH1007	T-cell	Mac	2.27	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	T-cell	NPC	-	-	-	-	-
ZH1007	T-cell	Neuron	-	-	-	-	-
ZH1007	T-cell	OPC	2.04	9.97×10^{-1}	4.00×10^{-3}	4.00×10^{-3}	-1
ZH1007	T-cell	Oligo	1.48	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH1007	T-cell	Reactive-Ast	1.10	8.00×10^{-1}	2.13×10^{-1}	2.13×10^{-1}	0
ZH1007	T-cell	T-cell	3.16	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	T-cell	Vasc	3.75	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Vasc	AC	1.35	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Vasc	B-cell	2.95	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Vasc	Inflammatory-Mac	1.67	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Vasc	MES	1.99	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Vasc	MES-Hyp	-	-	-	-	-
ZH1007	Vasc	Mac	1.80	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Vasc	NPC	-	-	-	-	-
ZH1007	Vasc	Neuron	-	-	-	-	-
ZH1007	Vasc	OPC	1.64	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH1007	Vasc	Oligo	1.34	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH1007	Vasc	Reactive-Ast	1.21	2.00×10^{-3}	9.99×10^{-1}	2.00×10^{-3}	1
ZH1007	Vasc	T-cell	1.39	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH1007	Vasc	Vasc	9.23	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	AC	AC	7.33	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	AC	B-cell	1.17	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	AC	Inflammatory-Mac	1.53	1.00	9.99×10^{-4}	9.99×10^{-4}	-1

Patient	From cell	To cell	Observed count	Permutations (>)	Permutations (<)	p value	Significance/Direction
ZH916	AC	MES	2.56	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	AC	MES-Hyp	2.61	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	AC	Mac	3.15	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	AC	NPC	1.70	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	AC	Neuron	1.41	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	AC	OPC	3.61	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	AC	Oligo	1.09	3.00×10^{-3}	9.99×10^{-1}	3.00×10^{-3}	1
ZH916	AC	Reactive-Ast	1.54	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	AC	T-cell	1.31	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	AC	Vasc	2.22	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	B-cell	AC	2.63	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	B-cell	B-cell	4.31	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	B-cell	Inflammatory-Mac	2.35	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	B-cell	MES	2.39	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	B-cell	MES-Hyp	1.00	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	B-cell	Mac	4.98	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	B-cell	NPC	0.00	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	B-cell	Neuron	1.00	1.00	9.01×10^{-1}	9.01×10^{-1}	0
ZH916	B-cell	OPC	1.00	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	B-cell	Oligo	1.00	9.52×10^{-1}	9.89×10^{-1}	9.52×10^{-1}	0
ZH916	B-cell	Reactive-Ast	0.00	1.00	4.00×10^{-3}	4.00×10^{-3}	-1
ZH916	B-cell	T-cell	5.28	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	B-cell	Vasc	5.20	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Inflammatory-Mac	AC	1.81	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Inflammatory-Mac	B-cell	1.53	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Inflammatory-Mac	Inflammatory-Mac	3.95	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Inflammatory-Mac	MES	3.75	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Inflammatory-Mac	MES-Hyp	3.33	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Inflammatory-Mac	Mac	3.50	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Inflammatory-Mac	NPC	2.06	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Inflammatory-Mac	Neuron	1.08	2.00×10^{-3}	9.99×10^{-1}	2.00×10^{-3}	1
ZH916	Inflammatory-Mac	OPC	2.75	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Inflammatory-Mac	Oligo	1.00	1.00	8.74×10^{-1}	8.74×10^{-1}	0
ZH916	Inflammatory-Mac	Reactive-Ast	1.63	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Inflammatory-Mac	T-cell	2.48	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Inflammatory-Mac	Vasc	5.22	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	MES	AC	2.77	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	MES	B-cell	1.38	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	MES	Inflammatory-Mac	1.98	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	MES	MES	6.99	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	MES	MES-Hyp	2.52	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	MES	Mac	3.82	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	MES	NPC	2.03	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	MES	Neuron	1.25	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	MES	OPC	3.01	9.99×10^{-4}	1.00	9.99×10^{-4}	1

Patient	From cell	To cell	Observed count	Permutations (>)	Permutations (<)	p value	Significance/Direction
ZH916	MES	Oligo	1.04	4.00×10^{-3}	9.97×10^{-1}	4.00×10^{-3}	1
ZH916	MES	Reactive-Ast	1.33	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	MES	T-cell	1.43	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	MES	Vasc	2.18	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	MES-Hyp	AC	1.62	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	MES-Hyp	B-cell	1.00	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	MES-Hyp	Inflammatory-Mac	2.06	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	MES-Hyp	MES	3.03	9.99×10^{-4}	1.00	9.99×10^{-4}	-1
ZH916	MES-Hyp	MES-Hyp	6.39	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	MES-Hyp	Mac	2.05	9.99×10^{-4}	1.00	9.99×10^{-4}	-1
ZH916	MES-Hyp	NPC	1.41	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	MES-Hyp	Neuron	1.23	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	MES-Hyp	OPC	2.50	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	MES-Hyp	Oligo	1.00	1.00	8.59×10^{-1}	8.59×10^{-1}	0
ZH916	MES-Hyp	Reactive-Ast	1.88	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	MES-Hyp	T-cell	1.25	6.89×10^{-1}	3.12×10^{-1}	3.12×10^{-1}	0
ZH916	MES-Hyp	Vasc	1.93	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	Mac	AC	2.86	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Mac	B-cell	1.95	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Mac	Inflammatory-Mac	2.25	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Mac	MES	4.15	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Mac	MES-Hyp	2.56	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Mac	Mac	6.57	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Mac	NPC	2.02	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Mac	Neuron	1.48	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Mac	OPC	2.88	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Mac	Oligo	1.02	8.29×10^{-2}	9.19×10^{-1}	8.29×10^{-2}	0
ZH916	Mac	Reactive-Ast	1.48	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Mac	T-cell	2.22	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Mac	Vasc	3.64	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	NPC	AC	1.60	8.69×10^{-2}	9.15×10^{-1}	8.69×10^{-2}	0
ZH916	NPC	B-cell	0.00	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	NPC	Inflammatory-Mac	1.70	4.60×10^{-1}	5.41×10^{-1}	4.60×10^{-1}	0
ZH916	NPC	MES	2.96	9.99×10^{-4}	1.00	9.99×10^{-4}	-1
ZH916	NPC	MES-Hyp	1.61	9.99×10^{-1}	2.00×10^{-3}	2.00×10^{-3}	-1
ZH916	NPC	Mac	2.91	9.99×10^{-4}	1.00	9.99×10^{-4}	-1
ZH916	NPC	NPC	3.72	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	NPC	Neuron	1.25	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	NPC	OPC	3.83	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	NPC	Oligo	1.00	9.92×10^{-1}	9.80×10^{-1}	9.80×10^{-1}	0
ZH916	NPC	Reactive-Ast	1.13	1.20×10^{-2}	9.93×10^{-1}	1.20×10^{-2}	0
ZH916	NPC	T-cell	1.04	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	NPC	Vasc	2.10	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	Neuron	AC	1.71	3.90×10^{-2}	9.62×10^{-1}	3.90×10^{-2}	0
ZH916	Neuron	B-cell	2.00	9.99×10^{-4}	1.00	9.99×10^{-4}	1

Patient	From cell	To cell	Observed count	Permutations (>)	Permutations (<)	p value	Significance/Direction
ZH916	Neuron	Inflammatory-Mac	1.81	1.53×10^{-1}	8.48×10^{-1}	1.53×10^{-1}	0
ZH916	Neuron	MES	2.71	9.99×10^{-4}	1.00	9.99×10^{-4}	-1
ZH916	Neuron	MES-Hyp	2.18	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Neuron	Mac	3.13	8.69×10^{-1}	1.32×10^{-1}	1.32×10^{-1}	0
ZH916	Neuron	NPC	1.25	1.30×10^{-2}	9.95×10^{-1}	1.30×10^{-2}	0
ZH916	Neuron	Neuron	2.39	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Neuron	OPC	1.90	4.70×10^{-2}	9.54×10^{-1}	4.70×10^{-2}	0
ZH916	Neuron	Oligo	1.00	6.47×10^{-1}	9.96×10^{-1}	6.47×10^{-1}	0
ZH916	Neuron	Reactive-Ast	2.00	4.00×10^{-3}	1.00	4.00×10^{-3}	1
ZH916	Neuron	T-cell	1.22	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Neuron	Vasc	1.50	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	OPC	AC	3.86	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	OPC	B-cell	1.00	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	OPC	Inflammatory-Mac	1.72	7.19×10^{-2}	9.29×10^{-1}	7.19×10^{-2}	0
ZH916	OPC	MES	3.57	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	OPC	MES-Hyp	3.39	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	OPC	Mac	3.03	9.99×10^{-4}	1.00	9.99×10^{-4}	-1
ZH916	OPC	NPC	2.69	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	OPC	Neuron	1.47	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	OPC	OPC	6.06	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	OPC	Oligo	1.07	3.00×10^{-3}	9.98×10^{-1}	3.00×10^{-3}	1
ZH916	OPC	Reactive-Ast	1.66	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	OPC	T-cell	1.15	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	OPC	Vasc	1.95	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	Oligo	AC	3.20	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Oligo	B-cell	2.00	8.99×10^{-3}	9.99×10^{-1}	8.99×10^{-3}	1
ZH916	Oligo	Inflammatory-Mac	1.14	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	Oligo	MES	3.77	1.70×10^{-2}	9.84×10^{-1}	1.70×10^{-2}	0
ZH916	Oligo	MES-Hyp	8.00	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Oligo	Mac	2.50	9.99×10^{-4}	1.00	9.99×10^{-4}	-1
ZH916	Oligo	NPC	1.00	9.92×10^{-1}	7.36×10^{-1}	7.36×10^{-1}	0
ZH916	Oligo	Neuron	1.11	1.30×10^{-2}	9.88×10^{-1}	1.30×10^{-2}	0
ZH916	Oligo	OPC	3.44	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Oligo	Oligo	1.43	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Oligo	Reactive-Ast	1.50	3.00×10^{-3}	9.99×10^{-1}	3.00×10^{-3}	1
ZH916	Oligo	T-cell	1.43	9.69×10^{-2}	9.13×10^{-1}	9.69×10^{-2}	0
ZH916	Oligo	Vasc	1.60	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	Reactive-Ast	AC	1.00	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	Reactive-Ast	B-cell	0.00	1.00	4.00×10^{-3}	4.00×10^{-3}	-1
ZH916	Reactive-Ast	Inflammatory-Mac	1.68	5.70×10^{-1}	4.32×10^{-1}	4.32×10^{-1}	0
ZH916	Reactive-Ast	MES	2.00	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	Reactive-Ast	MES-Hyp	1.88	7.69×10^{-2}	9.24×10^{-1}	7.69×10^{-2}	0
ZH916	Reactive-Ast	Mac	2.33	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	Reactive-Ast	NPC	1.29	1.30×10^{-2}	9.88×10^{-1}	1.30×10^{-2}	0
ZH916	Reactive-Ast	Neuron	1.40	1.40×10^{-2}	9.87×10^{-1}	1.40×10^{-2}	0

Patient	From cell	To cell	Observed count	Permutations (>)	Permutations (<)	p value	Significance/Direction
ZH916	Reactive-Ast	OPC	3.01	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Reactive-Ast	Oligo	1.00	5.89×10^{-1}	9.97×10^{-1}	5.89×10^{-1}	0
ZH916	Reactive-Ast	Reactive-Ast	2.27	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Reactive-Ast	T-cell	1.31	2.42×10^{-1}	7.66×10^{-1}	2.42×10^{-1}	0
ZH916	Reactive-Ast	Vasc	1.40	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	T-cell	AC	2.02	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	T-cell	B-cell	2.34	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	T-cell	Inflammatory-Mac	2.65	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	T-cell	MES	3.53	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	T-cell	MES-Hyp	3.22	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	T-cell	Mac	4.59	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	T-cell	NPC	1.63	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	T-cell	Neuron	1.22	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	T-cell	OPC	2.63	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	T-cell	Oligo	1.00	1.00	9.46×10^{-1}	9.46×10^{-1}	0
ZH916	T-cell	Reactive-Ast	2.83	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	T-cell	T-cell	5.07	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	T-cell	Vasc	6.12	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Vasc	AC	4.38	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Vasc	B-cell	1.71	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Vasc	Inflammatory-Mac	2.18	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Vasc	MES	2.35	9.99×10^{-4}	1.00	9.99×10^{-4}	-1
ZH916	Vasc	MES-Hyp	1.30	1.00	9.99×10^{-4}	9.99×10^{-4}	-1
ZH916	Vasc	Mac	3.95	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Vasc	NPC	1.48	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Vasc	Neuron	1.23	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Vasc	OPC	1.98	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Vasc	Oligo	1.14	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Vasc	Reactive-Ast	1.75	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Vasc	T-cell	2.54	9.99×10^{-4}	1.00	9.99×10^{-4}	1
ZH916	Vasc	Vasc	1.15×10^{1}	9.99×10^{-4}	1.00	9.99×10^{-4}	1

Table S3.4 | Cell type score changes across Up responder patients.

Comparisons were performed using all Up responder patient samples and also separately for Up responder in each the discovery and validation cohorts. Statistical significance was assessed using the paired, *Wilcoxon test* and p values shown were adjusted using the FDR method. The change filed denotes the main direction of the fold-change for each cell type score through treatment from primary to recurrence. The rows are ordered by $\log_2$-FC and shaded rows denote cell types which were statistically significant ($p < 0.05$) across one or more cohorts.

	Discovery			**Validation**			**All**		
	p value	**$\log_2$FC**	**Change**	**p value**	**$\log_2$FC**	**Change**	**p value**	**$\log_2$FC**	**Change**
Oligodendrocyte	4.22×10^{-7}	0.32	up	5.05×10^{-5}	0.32	up	1.14×10^{-9}	0.32	up
Neuron	2.81×10^{-9}	0.24	up	5.05×10^{-5}	0.17	up	1.13×10^{-10}	0.21	up
Plasma B	9.97×10^{-6}	0.16	up	7.82×10^{-5}	0.20	up	1.86×10^{-9}	0.18	up
T Cell	5.70×10^{-2}	0.13	up	8.52×10^{-1}	0.08	up	8.85×10^{-2}	0.11	up
Mast Cell	1.32×10^{-2}	0.12	up	8.52×10^{-1}	0.07	up	2.32×10^{-2}	0.10	up
MES3	1.25×10^{-3}	0.14	up	4.07×10^{-1}	0.04	up	1.03×10^{-3}	0.09	up
NPC2	2.81×10^{-2}	0.14	up	8.52×10^{-1}	0.01	up	4.60×10^{-2}	0.09	up
Astrocyte	1.87×10^{-5}	0.11	up	2.12×10^{-1}	0.03	up	4.31×10^{-6}	0.08	up
OPC	1.14×10^{-2}	0.11	up	8.52×10^{-1}	0.02	up	2.40×10^{-2}	0.08	up
DC	4.19×10^{-1}	0.04	up	8.52×10^{-1}	0.09	up	4.18×10^{-1}	0.06	up
OPC Normal	3.94×10^{-3}	0.08	up	8.49×10^{-1}	0.02	up	9.27×10^{-3}	0.06	up
AC	2.33×10^{-4}	0.15	up	2.06×10^{-1}	-0.06	down	6.75×10^{-2}	0.06	up
Macrophage	2.81×10^{-1}	0.03	up	3.45×10^{-1}	0.07	up	8.85×10^{-2}	0.04	up
NPC1	1.92×10^{-1}	0.07	up	8.52×10^{-1}	-0.03	down	4.57×10^{-1}	0.03	up
NK Cell	4.95×10^{-2}	0.08	up	8.52×10^{-1}	-0.06	down	3.65×10^{-1}	0.03	up
Microglia	5.16×10^{-1}	0.02	up	8.52×10^{-1}	0.03	up	4.57×10^{-1}	0.03	up
PN	4.55×10^{-1}	0.04	up	8.52×10^{-1}	-0.03	down	7.38×10^{-1}	0.01	up
MES2	8.37×10^{-1}	-0.01	down	8.52×10^{-1}	0.02	up	9.44×10^{-1}	0.00	up
Monocyte	8.21×10^{-1}	-0.02	down	8.52×10^{-1}	0.02	up	9.44×10^{-1}	0.00	down
Radial Glial	3.88×10^{-1}	0.02	up	1.86×10^{-1}	-0.07	down	5.93×10^{-1}	-0.01	down
B Cell	5.16×10^{-1}	-0.02	down	9.30×10^{-1}	0.00	up	5.93×10^{-1}	-0.01	down
MES1	3.52×10^{-1}	-0.03	down	8.56×10^{-1}	-0.01	down	4.18×10^{-1}	-0.02	down
Endothelial	3.49×10^{-2}	-0.07	down	8.28×10^{-2}	-0.10	down	2.53×10^{-3}	-0.08	down
Mural Cell	3.17×10^{-2}	-0.10	down	1.86×10^{-1}	-0.08	down	5.75×10^{-3}	-0.09	down

Table S3.5 | Cell type score changes across Down responder patients.

Comparisons were performed using all Down responder patient samples and also separately for Down responder in each the discovery and validation cohorts. Statistical significance was assessed using the paired, *Wilcoxon test* and p values shown were adjusted using the FDR method. The change filed denotes the main direction of the fold-change for each cell type score through treatment from primary to recurrence. The rows are ordered by $\log_2$-FC and shaded rows denote cell types which were statistically significant ($p < 0.05$) across one or more cohorts.

	Discovery			Validation			All		
	p value	**$\log_2$FC**	**Change**	**p value**	**$\log_2$FC**	**Change**	**p value**	**$\log_2$FC**	**Change**
MES2	3.24×10^{-1}	0.05	up	4.42×10^{-1}	-0.05	down	4.61×10^{-1}	0.02	up
MES1	4.40×10^{-1}	0.02	up	1.92×10^{-1}	-0.08	down	7.04×10^{-1}	-0.01	down
Mast Cell	5.69×10^{-1}	0.06	up	1.79×10^{-1}	-0.19	down	7.02×10^{-1}	-0.03	down
Plasma B	5.35×10^{-1}	-0.03	down	2.61×10^{-1}	-0.07	down	1.92×10^{-1}	-0.04	down
Radial Glial	1.15×10^{-2}	-0.05	down	5.88×10^{-1}	-0.03	down	4.46×10^{-2}	-0.05	down
DC	8.20×10^{-1}	0.03	up	1.79×10^{-1}	-0.19	down	6.06×10^{-1}	-0.05	down
Macrophage	8.82×10^{-1}	-0.03	down	6.21×10^{-2}	-0.15	down	1.19×10^{-1}	-0.06	down
Mural Cell	3.25×10^{-1}	-0.06	down	1.79×10^{-1}	-0.10	down	5.57×10^{-2}	-0.07	down
Oligodendrocyte	3.24×10^{-1}	-0.10	down	6.69×10^{-1}	-0.03	down	2.00×10^{-1}	-0.08	down
Monocyte	3.39×10^{-1}	-0.07	down	8.21×10^{-2}	-0.16	down	2.78×10^{-2}	-0.10	down
Astrocyte	8.05×10^{-3}	-0.10	down	1.42×10^{-1}	-0.11	down	1.35×10^{-3}	-0.11	down
MES3	1.37×10^{-1}	-0.10	down	1.79×10^{-1}	-0.14	down	2.01×10^{-2}	-0.11	down
NPC1	1.42×10^{-2}	-0.13	down	4.36×10^{-1}	-0.09	down	2.06×10^{-2}	-0.11	down
NK Cell	3.68×10^{-1}	-0.07	down	5.28×10^{-2}	-0.25	down	2.78×10^{-2}	-0.12	down
Neuron	5.64×10^{-3}	-0.14	down	1.86×10^{-1}	-0.10	down	1.35×10^{-3}	-0.13	down
Microglia	4.97×10^{-2}	-0.12	down	4.81×10^{-2}	-0.20	down	1.35×10^{-3}	-0.14	down
OPC Normal	3.50×10^{-3}	-0.17	down	2.24×10^{-1}	-0.11	down	1.13×10^{-3}	-0.14	down
T Cell	8.49×10^{-1}	-0.04	down	5.28×10^{-2}	-0.39	down	1.10×10^{-1}	-0.15	down
OPC	1.15×10^{-2}	-0.18	down	4.13×10^{-1}	-0.10	down	1.04×10^{-2}	-0.15	down
NPC2	1.28×10^{-2}	-0.16	down	2.29×10^{-1}	-0.15	down	5.37×10^{-3}	-0.15	down
AC	5.64×10^{-3}	-0.18	down	4.21×10^{-1}	-0.12	down	1.61×10^{-3}	-0.16	down
Endothelial	5.64×10^{-3}	-0.14	down	1.75×10^{-2}	-0.20	down	3.22×10^{-5}	-0.16	down
PN	1.28×10^{-2}	-0.16	down	2.29×10^{-1}	-0.16	down	8.64×10^{-3}	-0.16	down
B Cell	4.71×10^{-1}	-0.18	down	6.03×10^{-2}	-0.38	down	4.80×10^{-2}	-0.24	down

Chapter 4

Spatial profiling of longitudinal glioblastoma reveals consistent changes in cellular architecture, post-treatment

Shoaib Ajaib[1], Steven Pollock[1], Gemma Hemmings[1], Arief Gusnanto[2], Aruna Chakrabarty[3], Azzam Ismail[3], Erica Wilson[1], Bethany Hunter[4], Andrew Filby[4], David McDonald[4], Asa A. Brockman[5], Rebecca A. Ihrie[6], Lucy F. Stead[1]

[1]Leeds Institute of Medical Research, University of Leeds
[2]School of Mathematics, University of Leeds
[3]Department of Neuropathology, Leeds Teaching Hospitals NHS Trust
[4]Flow Cytometry Core Facility, Newcastle University, Newcastle upon Tyne, UK
[5]Department of Cell & Developmental Biology, Vanderbilt University School of Medicine; Vanderbilt Brain Institute, Vanderbilt-Ingram Cancer Center, Department of Neurological Surgery, Vanderbilt University Medical Center, Nashville, Tennessee, United States
[6]Paediatrics – Section of Child Neurology, University of Colorado Anschutz Medical Campus

Abstract

Background: glioblastoma (GBM), the most aggressive adult brain cancer, comprises a complex tumour microenvironment (TME) with diverse cellular interactions that drive progression and pathobiology. The aim of this study was to understand how these spatial patterns and interactions evolve with treatment.

Methods: To explore these relationships, we employed imaging mass cytometry (IMC) to measure the expression of 34 protein markers, enabling the identification of GBM-specific cell types and their interactions at single-cell protein level in paired pre- and post-treatment GBM samples from five patients.

Results: We find a significant post-treatment increase in normal brain cells alongside a reduction in vascular cells. Moreover, despite minimal overall change in cellular diversity, interactions among astrocytes, oligodendrocytes, and vascular cells increase post-treatment, suggesting reorganisation of the TME. The GBM TME cells form spatially organized layers driven by hypoxia pre-treatment, but this influence diminishes post-treatment, giving way to less organised layers with organisation driven by reactive astrocytes and lymphocytes.

Conclusions: These findings provide insight into treatment-induced shifts in GBMs cellular landscape, highlighting aspects of the evolving TME that appear to facilitate recurrence and are, therefore, potential therapeutic targets.

4.1 Introduction

Isocitrate dehydrogenase wild-type (IDHwt) GBM is the most common and aggressive form of adult diffuse glioma, with a median survival of ∼15 months[1]. Standard treatment consists of surgical resection followed by radiation and chemotherapy with temozolomide (TMZ)[2]. However, tumour recurrence is inevitable due to: a) the infiltrative nature of primary GBM, which precludes complete surgical removal; and b) significant intra- and inter-tumour heterogeneity, which enables residual cells to resist chemoradiation and continue proliferating[3,4]. Characterising how unresected GBM cells respond to treatment can highlight potential mechanisms of treatment resistance that could be additionally targeted with combined therapies.

It is known that IDHwt GBM cells exhibit plasticity across four neoplastic cell states along a proneural (PN) to mesenchymal (MES) axis[5–7]: neural progenitor-like (NPC-like), oligodendrocyte progenitor-like (OPC-like), astrocyte-like (AC-like), and mesenchymal-like (MES-like). However, these neoplastic cells do not function in isolation. In their updated hallmarks of cancer, Hanahan and Weinberg remarked that any understanding of tumours "must encompass the contributions of the (TME)"[8]. In GBM the TME comprises a diverse array of tumour cells and also complex network of immune cells, stromal cells, and vascular elements, that play a critical role in GBM progression and treatment resistance, acting as a dynamic ecosystem that influences tumour behaviour and therapeutic response[9].

To truly understand GBM tumour response to treatment, therefore, requires characterisation at single-cell level in ways that incorporate information about interactions with the TME. This is now possible through the use of spatial molecular profiling technologies[10]. Such approaches have recently been applied to GBM tumours, revealing niches containing specific neoplastic cells and distinct immune-associated programs[11–13]. These niches have also been shown to organize into structured layers, beyond what is visible via conventional microscopy and histopathology, and are associated with cellular states such as hypoxia[11].

These findings describe consistent organizational patterns across GBM tumours suggesting that neoplastic phenotypes are driven by environmental interactions. However, one crucial aspect remains unexplored: how the spatial patterns and interactions within the TME are impacted by treatment to enable some neoplastic cells to survive. To begin to address this, we analysed multiplex IMC data[14], from five paired pre- and post-treatment IDHwt GBM patient samples, focusing on protein-level changes that reveal alterations in cellular prevalence and states.

4.2 Materials & methods

4.2.1 Data availability

The data used in this study (including the raw high-dimensional TIFF images, spillover correction files; cell-object segmentation masks; patient and sample metadata; phenotype labelled single-cell data and the cell-object spatial information) are deposited and available online at Zenodo:

https://doi.org/10.5281/zenodo.14679442

4.2.2 Code availability

The analysis code that was used to process the IMC data and produce the results in this paper can be found on GitHub:

https://github.com/GliomaGenomics/GBM_IMC_Analysis

4.2.3 Ethics statement

Tumour samples used in this study were obtained from patients at the Walton Centre, UK, that provided informed consent in writing for the use of their tissue in research. The inclusion of these samples in this project was following approval by the UK National Health Service (NHS) Research Ethics Service Committee South Central - Oxford A (Research Ethics Code: 13/SC/0509).

4.2.4 Imaging mass cytometry analysis

Data acquisition and initial processing of samples is as previously detailed[15]. Briefly, for each of the five patients with paired tumour samples from initial surgery and first recurrence, we selected three spatially spatially distinct 1mm^2 regions of interests (ROIs) per time-point, resulting in a total of 30 ROIs (Table S4.1, Table S4.2 and Figure S4.1). An antibody panel of 34 proteins was designed and control tissues were used to validate each antibody (Table S4.3). Following this 5μm tissue sections were stained with a cocktail of all conjugated antibodies and each ROI was ablated at a 200Hz frequency with a 1μm diameter laser. The raw MCD files were exported as exported in OME-TIFF format. All IMC image

processing and downstream analysis steps were performed using the R statistical software package ($\geq$ version 4.3.0) and Python (version 3.11.3). Cell-cell interactions, neighbourhoods and contexts were generated using functions from imcRtools (v1.10.0). All plots were generated using ggplot2 (version 3.5.1).

4.2.5 Cell segmentation

Steinbock[16] (v0.13.5) converted raw MCD files into multi-channel TIFF images (n=30), which were processed with tifffile (v2023.4.12). Nuclear (Ir191, Ir193) and cytoplasmic (Sm149, Eu153, Dy164, Yb171) channels were combined into single RGB images using pandas (v2.0.3) and numpy (v1.24.0), followed by random cropping into 100μm^2 sections. Cellpose (v2.0) was used for segmentation as previously described[17]. Random RGB crops were segmented with the pre-trained cytoplasm model (*cyto2*), iteratively re-annotated, and refined by updating the model and adjusting the mean pixel diameter. The final model was applied to full 1000μm^2 ROIs using the *cellpose_train* command with default settings. Signal intensities were extracted using the *measure intensities* function in Steinbock: pixels were aggregated by computing the mean intensity across channels. Additional spatial features were calculated using the *measure regionprops* function (Figure S4.2).

4.2.6 Single-cell & image processing

The quantified single cell data and images were processed as previously detailed[15]. Briefly, segmented cell expression counts were transformed using an inverse hyperbolic sine function (arcsinh) with *cofactor* = 5. Following this, single-cells were integrated to correct for unwanted sources of variation present across each patient using harmony (Figure S4.3 and Figure S4.4).

4.2.7 Cell phenotyping

Cells with high expression (>90th percentile) across multiple markers (>50% of marker) were excluded from phenotyping. Expression counts were *z-score* normalised, was ranked from low (1) to high (20) based on expression levels. Labels were assigned using a logical gating approach: labels were only assigned if the marker expression rank exceeded a user-defined threshold (Table S4.4). Where phenotyping based on marker rank alone was insufficient, marker-specific thresh-

olds were applied, and intersecting populations were combined to generate final cell type labels. Cells were classified as having low (< -1.2) or high (> 1.2) hypoxia (HIF1A$^+$) and epithelial-to-mesenchymal transition (EMT) (SNAI1$^+$) based on their *z-score* normalised expression counts.

4.2.8 Measurement of intra-patient heterogeneity

Shannon entropy (H) was used to measure the compositional diversity of cell types across ROIs. To account for differences in cell frequencies across ROIs, 1,000 cells were randomly sub-sampled per group (i), in each of ten iterations. In each round, Shannon entropy was calculated using cell type frequencies (P_c) as:

$$H_i = -\sum_c P_c \log_2(P_c)$$

The *Wilcoxon* rank-sum test was used to compare Shannon entropy (H) between primary and recurrent surgeries, both across all data and within each patient (n=5).

4.2.9 Spatial interaction graphs

Spatial interaction graphs were generated for each ROI (n=30) using the *buildSpatialGraph* function from imcRtools with the Delaunay triangulation method. To reduce spurious connections, edges were pruned by setting the *max_dist* argument to 50μm.

4.2.10 Testing cell-cell interactions

Cell-cell spatial interactions were tested using the *testInteractions* from imcRtools with the previously described *histoCAT* method[18]. Briefly, co-detection by indexing (CODEX) spatial interaction graph data edges were summed and aggregated across individual patient/surgery ROIs and then divided by the number of cells of type A that had at least one neighbour of type B. These observed interactions were compared against a derived null distribution, describing the interactions formed under spatial randomness. Significance was determined using two one-tailed *permutation* tests:

$$P_{AB} = \begin{cases} 1, & C_{\text{obs}} = 0; \\ \frac{\sum(C_{\text{perm}\geq(\leq)}C_{\text{obs}})+1}{N_{\text{perm}}+1}, & \text{otherwise}, \end{cases}$$

where C_{perm} is the number of cell pairs (A, B) in each permutation, C_{obs} is the actual number of cell pairs (A, B) given a defined distance, and N_{perm} is the number of permutations.

4.2.11 Cell neighbourhoods & spatial contexts

Cellular neighbourhoods (CNs) and Spatial contexts (SCs) were defined using a previously established method[19]. Briefly, for each cell in each ROI (n=30), the proportion of neighbouring cell types was aggregated using the *aggregateNeighbors* function from imcRtools. These proportions were then clustered using k-means with k=12, determined through a parameter sweep and visual inspection.

CNs were further aggregated and SCs identified using the *detectSpatialContext* function from imcRtools with default parameters. The *filterSpatialContext* function from imcRtools was used to further refine SCs, retaining the most dominant: 1.) SCs present in >3 patients; and 2.) SCs comprising >5% of total cells within each surgery type.

4.3 Results

4.3.1 Identifying & labelling cell types in GBM

To assess the spatial evolution of GBM tumours through treatment, we collected tumour samples from five patients who had undergone surgical resections of both primary and recurrent IDHwt GBM. Each primary tumour developed *de novo*, and all patients received radiation, chemotherapy with TMZ and had a local recurrence. For patient information see Table S4.1. Three spatially distinct $1mm^2$ (ROIs) were selected for each tumour sample, following immunohistochemical staining for key markers of proliferation (Ki67), hypoxia (HIF1A) and immune cells (CD45), to capture intra-tumour heterogeneity and avoid the bias of examining only a single small region (Figure 4.1A and Table S4.2).

We designed a panel of 34 protein markers to identify GBM-specific cell types (neoplastic, immune, and normal brain cells) along with markers of cell states such as proliferation and hypoxia (see Table S4.2). Using a deep learning-based image segmentation approach, we assigned cell type labels to each segmented object and also subsequently grouped cells into four categories: immune, cancer, normal brain, and vasculature (Figure 4.1B). Approximately 107,000 cells were labelled across all samples (Figure 4.1C-D) after applying batch effect correction to account for variability between individual patients and to ensure that expression profiles were comparable (Figure S4.1 and Figure S4.2).

A comparison of cell categories across each ROI (Figure 4.1E) showed surprisingly consistent within-sample distributions, confirming that there is intra-tumour TME heterogeneity but that this is not as significant as inter-tumour TME heterogeneity. ANOVA analysis (see Table S4.5) confirmed that the effect of patient & surgery was significant for all cell types ($p < 0.001$), indicating considerable inter-tumour heterogeneity. In contrast, intra-tumour heterogeneity, represented by differences across ROIs, was not significant for any cell category suggesting that intra-tumour TME variability is less pronounced compared to inter-tumour heterogeneity. Therefore, we combined the three ROIs per sample, prior to subsequent downstream analyses, to increase the number of cells per sample whilst minimizing sampling bias from specific regions.

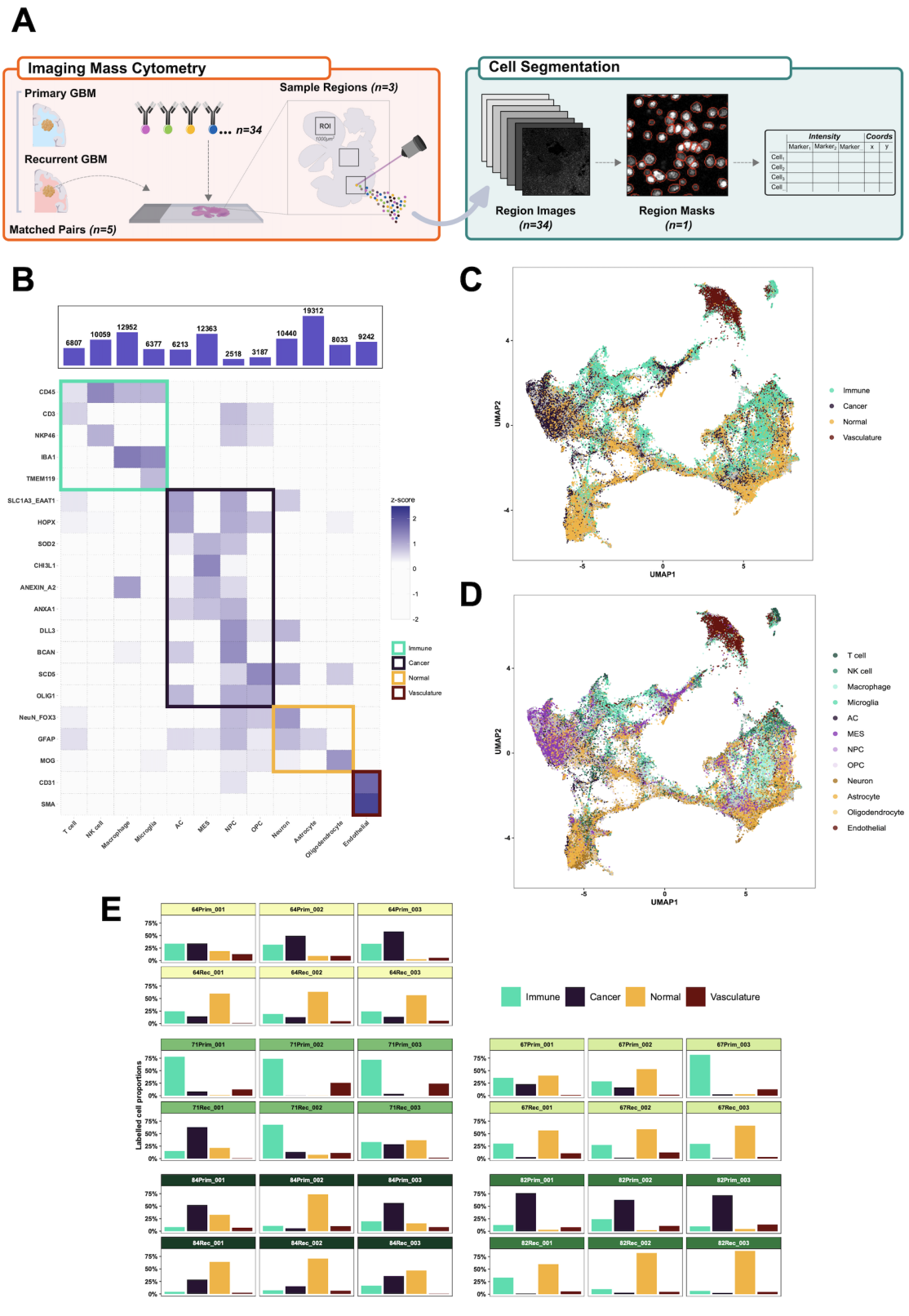

A
Imaging Mass Cytometry
Primary GBM
Recurrent GBM
Matched Pairs (n=5)
... n=34
Sample Regions (n=3)
ROI
Cell Segmentation
Region Images (n=34)
Region Masks (n=1)
Intensity
Coords
Marker1
Marker2
Marker
x
y
B
6807
10059
12952
6377
6213
12363
2518
3187
10440
19312
8033
9242
CD45
CD3
NKP46
IBA1
TMEM119
SLC1A3_EAAT1
HOPX
SOD2
CHI3L1
ANEXIN_A2
ANXA1
DLL3
BCAN
SCD5
OLIG1
NeuN_FOX3
GFAP
MOG
CD31
SMA
z-score
Immune
Cancer
Normal
Vasculature
T cell
NK cell
Macrophage
Microglia
AC
MES
NPC
OPC
Neuron
Astrocyte
Oligodendrocyte
Endothelial
C
UMAP1
UMAP2
Immune
Cancer
Normal
Vasculature
D
UMAP1
UMAP2
T cell
NK cell
Macrophage
Microglia
AC
MES
NPC
OPC
Neuron
Astrocyte
Oligodendrocyte
Endothelial
E
Labelled cell proportions
Immune
Cancer
Normal
Vasculature
64Prim_001
64Prim_002
64Prim_003
64Rec_001
64Rec_002
64Rec_003
71Prim_001
71Prim_002
71Prim_003
71Rec_001
71Rec_002
71Rec_003
67Prim_001
67Prim_002
67Prim_003
67Rec_001
67Rec_002
67Rec_003
84Prim_001
84Prim_002
84Prim_003
84Rec_001
84Rec_002
84Rec_003
82Prim_001
82Prim_002
82Prim_003
82Rec_001
82Rec_002
82Rec_003

Figure 4.1 | Cell segmentation and phenotyping overview.

A) Schematic detailing the IMC process for one patient sample, including the downstream analysis steps comprising of object segmentation and marker abundance quantification. **B)** Heatmap of protein marker abundances (rows) for each of the labelled cell types (columns). The tile colours denote the scaled (*z-score*) marker intensities, and the tile highlight colours represent the four main cell categories. **C-D)** Uniform manifold approximation and projection (UMAP) of all (patients and surgeries) cell objects identified following segmentation, batch correction and phenotyping: cells are coloured by cell category **(C)** and cell type **(D)**. **E)** Proportion of labelled cell categories (columns) across each ROI. The facets are grouped by patient/surgery and each of the facet header colours denote an individual patient.

4.3.2 Alterations in cellular prevalence through treatment in GBM

We first assessed how the prevalence of each cell category changed through treatment, between primary and recurrent samples (Figure 4.2A and Table S4.6). Whilst a reduction in the percentage of both immune and neoplastic cells was observed, from primary to recurrence, only the decrease in vasculature cells was significant (*Wilcoxon* p = 0.09, 0.06 and 9.88×10^{-3} respectively). The only significant increase was in the proportion of normal brain cells (*Wilcoxon* p = 4.52×10^{-4}).

Drilling down into how specific cell types change through treatment showed that no individual immune cell type exhibited significant changes (Figure 4.2B and Table S4.7). Similarly, no individual cancer cell types altered in a consistent direction, although the AC-like cancer cell type showed the largest and most consistent decrease (*Wilcoxon* p = 0.065). All normal brain cell types showed significant increases during treatment, with astrocytes exhibiting a particularly notable increase from primary to recurrence (*Wilcoxon* p = 3.02×10^{-3}), that was consistent across each patient. Of note, astrocytes appear to be the most prevalent normal brain cell type overall, consistent with reports of their high prevalence in both normal brain and the GBM TME[20,21]. These changes agree with those from our larger cohort studies where we performed deconvolution from bulk RNA sequencing (RNA-seq), validating our approach[22,23].

The limited number of consistent, significant changes in the prevalence of cell category or type over time highlights the variability in immune and neoplastic cell categories, post-treatment, across patients (Figure 4.2A-B). We, therefore decided

to systematically evaluate how cell diversity changes through treatment, both overall and at an individual patient level, to determine if any consistent patterns emerge.

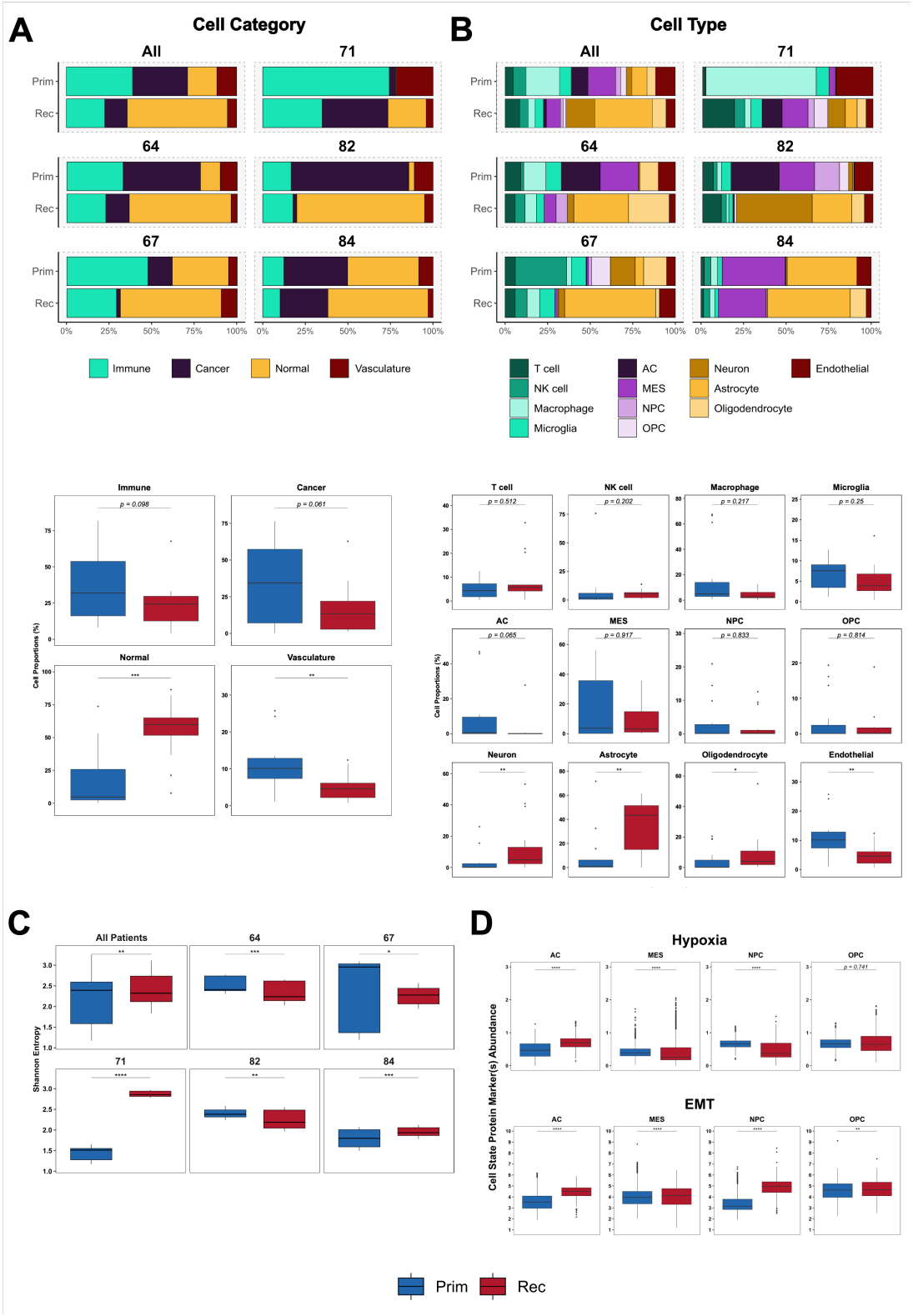


Figure 4.2 | Changes in GBM cell categories and types through treatment.

A) Top: stacked bar charts showing the labelled cell category prevalences across all patients (top left) and also separately for each individual patient. Bottom: boxplots showing the distribution of each cell category proportion across all patients grouped by primary and recurrent surgeries. **B)** Top: stacked bar charts showing the labelled cell type prevalences across all patients (top left) and also separately for each individual patient. Bottom: boxplots showing the distribution of each cell type proportion across all patients grouped by primary and recurrent surgeries. **C)** Boxplots showing the distribution of *Shannon's entropy* values grouped by surgery and split across all patients (top left) and also for each individual patient. **D)** Boxplots, grouped by surgery showing the distribution of protein marker abundance for markers which define hypoxia (top) and the EMT (bottom).
The black horizontal boxplot lines represent the median and the upper and lower box bounds denote the 25^{th} and 75^{th} quantiles, respectively.AC-like; MES-like; NPC-like; OPC-like; EMT. Significant p values are represented as: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; $^{****}p < 0.0001$.

4.3.3 Alterations in cellular diversity through treatment in GBM

To inspect cellular diversity in our samples, we quantified the Shannon's entropy (H) for each one (Figure 4.2C and Table S4.8). A high Shannon's entropy value indicates a tumour with many different cell types of similar frequency, whereas low entropy suggests that the tumour is dominated by few(er) cell types. This metric thus serves as a good proxy for assessing intra-tumour cellular heterogeneity for each sample, for example pre- and post-treatment.

We found that, overall, Shannon's entropy significantly decreased from primary to recurrence (*Wilcoxon* $q = 3.93 \times 10^{-3}$, Figure 4.2C), suggesting that cell distributions become less diverse, likely owing to certain cell types becoming more dominant within the distribution at recurrence. Linking this back to the results in Figure 4.2A and Figure 4.2B, this appears to be driven by the greater abundance of normal brain cells, and especially astrocytes, in the recurrent tumours. However, analysis of individual patients revealed variability in how cellular heterogeneity changed over time. Two patients (71 and 84) had significantly increased diversity through treatment (*Wilcoxon* $q = 8.45 \times 10^{-17}$ and 7.10×10^{-4}, respectively, Figure 4.2C). In patient 84, this increase was primarily driven by the appearance of oligodendrocytes at recurrence, which weren't present in the primary tumour (Figure 4.2B). Conversely, for patient 71, the increase in entropy was associated

with a reduction of dominating macrophages in the primary and presence of a larger neoplastic and normal brain cell fraction at recurrence (Figure 4.2B).

Given a lack of consistent trends in how treatment affects cell type prevalence or dominance, we proceeded to investigate whether changes in cell state could indicate how treatment shapes cancer cell phenotypes.

4.3.4 Alterations in neoplastic cellular states through treatment in GBM

The MES phenotype in GBM cancer cells is characterised by high proliferative and metastatic potential, often leading to a poorer prognosis compared to PN subtypes[24–27]. Moreover, elevated hypoxia and the expression of EMT genes, typically involved in neural tube formation or wound healing, have been shown to be closely linked to the MES cell state[28].

In our IMC panel we included antibodies against proteins indicating hypoxia (HIF1A) and EMT (SNAI1 & TGFBeta) to assess the proportion of each of the four identified neoplastic cancer cell types that are in these cellular states, and how they changed through treatment. We found that significantly more AC-like cancer cells expressed hypoxia markers post-treatment (*Wilcoxon* $p = 4.98 \times 10^{-115}$), whilst significantly fewer MES-like and NPC-like cells did (*Wilcoxon* $p = 9.62 \times 10^{-125}$ and $p = 5.49 \times 10^{-70}$, respectively) (Figure 4.2D and Table S4.9).

All four neoplastic cell types had a significantly higher proportion of cells expressing markers of EMT post-treatment, with the largest effect sizes observed in AC-like and NPC-like cells (*Wilcoxon* $p = 5.29 \times 10^{-161}$ and $p = 2.32 \times 10^{-183}$, respectively).

The power of our approach is not just in inspecting paired longitudinal GBM samples at single cell resolution, but also describing how treatment alters the cellular landscape in terms of spatial context. Hence, we moved on to looking at in situ cellular interactions, i.e. cells directly adjacent to one another.

4.3.5 Alterations in cellular interactions through treatment in GBM

We evaluated pairwise cell-cell adjacency, serving as an indicator of cell interaction partners, to assess whether any were significantly more likely (cells are "interact-

ing") or less likely (cells are "avoiding") compared to the null hypothesis of spatial randomness (Figure 4.3A and Table S4.10).

We performed this analysis on the primary and recurrent samples separately to see which significant findings were time-point dependent. Many cell types predominantly interacted with themselves in the primary tumours (Figure 4.3B). This is in keeping with previous spatial analysis of GBM that used a "spot-based" technology, that is not resolved at the single-cell level but rather aggregated over a small defined area (spot), which found that signal from the majority of spots seemed to emanate from a single cell type[11]. Our expansion to recurrent samples shows that these "self" interactions remained consistent through treatment (Figure 4.3B). Two additional, clear observations from our results are that there are no cells significantly avoiding one another, and there are many more recurrence-specific, significant cell-cell interactions than primary-specific ones (27 versus 10). Hence, despite finding an overall reduction in cell diversity at recurrence (Figure 4.2C), there are more interactions between differing cell types, suggesting that these are non-random and, thus, phenotypically important.

Neoplastic cells

Amongst the GBM cancer cell types, MES-like cells formed the highest number of significant interactions with other cell types. MES-like interactions with immune cells remained consistent between paired samples but interactions with normal brain cells were increased at recurrence.

Vasculature

Despite decreasing through treatment (Figure 4.2A), endothelial cells still formed significant interactions at both time points (Figure 4.3C). Unique to the primary tumours were significant interactions from the endothelial cells to the microglia (*permutation test*, $p = 9.99 \times 10^{-4}$) and MES-like cancer cells (*permutation test*, $p = 9.99 \times 10^{-4}$). MES-like cells interacting with myeloid lineage cells (e.g., macrophages and microglia) have been shown to lead to a highly proliferative state, increasing angiogenesis and contributing to a more invasive phenotype, which may explain these findings in the primary tumour[29]. Moreover, these interactions have also been shown to induce chemo-resistance in GBM, which has the potential to be addressed therapeutically[30]. The interactions from the endothelial cells to the macrophages were particular to recurrent tumours (*permutation test*, $p = 9.99 \times 10^{-4}$). These findings could be visualised in the IMC data (Figure 4.3C)

where a clear reduction in endothelial cells over time coincided with changes in the cells interacting with the remaining vasculature. It has previously been shown that bone derived macrophages populate a GBM tumour post-treatment, via the vascular system, which may explain this result and further indicate that therapies which hijack this infiltration could be effective for preventing or prolonging GBM recurrence[31]. Interactions from all normal brain cells to endothelial cells were also specific to the recurrent tumour. The post-treatment increase in normal brain cell abundance within the resected tissue (Figure 4.2A) may reflect the brain's wound healing response, with neuronal and glial cells re-populating the void left by surgery[32].

Normal brain cells

Significant interactions from and to oligodendrocytes almost universally occurred in the recurrent tumours, barring those from oligodendrocytes to MES-like cancer cells, which were primary-specific (Figure 4.3B). This could be observed in the IMC visualisations (Figure 4.3D). The prevalence of oligodendrocytes increases from primary to recurrent (Figure 4.2B) suggesting that this population did not simply expand *in situ* but rather infiltrated the recurrent TME. In GBM, oligodendrocyte lineage cells have commonly been reported to reside at tumour border niches including the invasion front/resection border where they co-localise with macrophages/microglia[33]. Moreover, oligodendrocytes have been shown to support GBM tumourigenicity and migration by promoting angiogenesis in GBM[34,35]. We also found evidence supporting the model of interactions, as microglia and endothelial cells were significantly interacting with oligodendrocytes at recurrence (Figure 4.3D).

Of all the normal brain cells, astrocytes were found to significantly interact most frequently and significantly with the cancer cells, though this is mostly specifically at recurrence. In fact, aside from "self" interactions which were consistent through treatment, normal astrocytes only formed significant interactions during recurrence.

Crosstalk between microglia and macrophages is known to induce reactive astrocyte phenotypes, which are crucial for the brain's wound healing process - a key aspect in GBM[36,37]. Moreover, the MES phenotypes, as described by Wang *et al.* and Neftel *et al.*, have been shown to overlap significantly with the presence of reactive astrocytes, indicating that these cells may migrate to injury sites after resection as part of the healing process[32]. In our samples we found significant interactions between normal neurons and astrocytes, suggesting the activation of

cellular programmes that could restore normal tissue function (Figure 4.3E and Figure 4.3F).

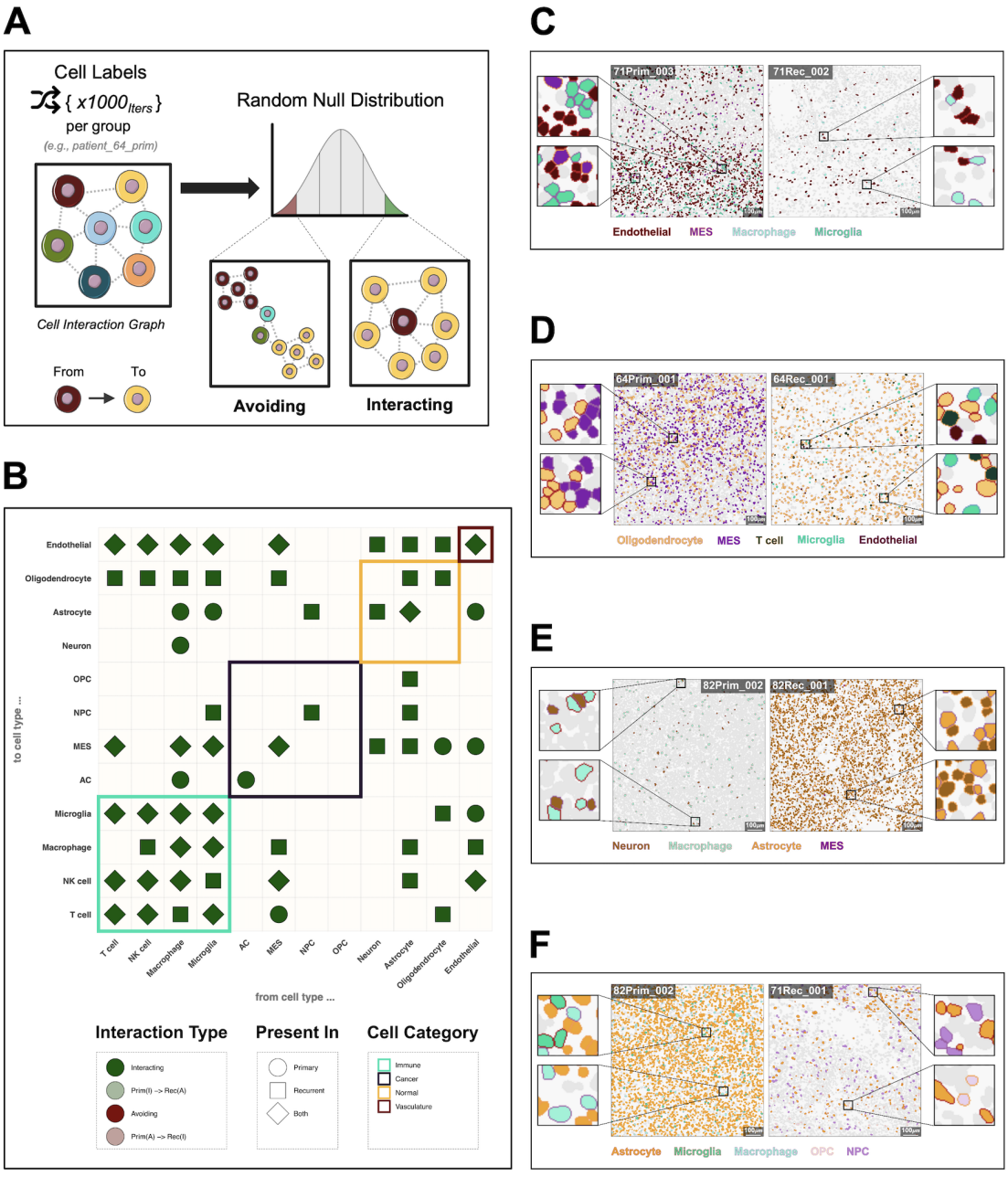


Figure 4.3 | Identifying significant GBM TME cell-cell interactions through treatment.

A) Schematic showing the process for testing if cell types interact more or less frequently than random. Each cell-cell interaction (edge) from an interaction graph is counted and averaged across a defined group (e.g., patient, surgery etc.). These average cell-cell interaction are then divided by the number of cells of type A that have at least one neighbour of type B. Finally, each observed cell-cell interaction count is compared against a null distribution that is generated by shuffling the cell-type labels 1000 times (1000 iterations) and counting the interactions between two specific cell types, giving the interaction counts under spatial randomness. Two cell types are *"avoiding"* when there are the significantly fewer interactions compared to random expectation for a given p values threshold. Conversely, when there are significantly more interactions between the two cell types they are *"interacting"*. **B)** Dotplot showing the significant ($p < 0.01$) cell-cell interactions which are present across a minimum of three patients. Shape denotes whether a specific cell-cell interaction is significant across either primary, recurrent or both surgeries. Point colours denote the type of significant interaction, i.e., interacting/avoiding and also cases where the interaction type changes through surgery. The tile highlights denote the cell category of each cell type. **C-F)** Representative IMC images showing single-cell segmentation masks coloured by the corresponding cell type labels for primary (left) and recurrent (right) surgery regions of interest.

4.3.6 Alterations in cellular neighbourhoods through treatment in GBM

Cell interactions within the GBM TME are heavily influenced by the spatial context, as GBM tumours consist of distinct anatomical regions[38]. To generalise groups of interacting cell types we defined CNs using a *nearest neighbours* approach (Figure 4.4A). This method defined 12 distinct CNs that provided a different level of structure from that observed based just on individual cells (as exemplified in Figure 4.4B). As expected, owing to the fact that each cell significantly associates with itself in both the primary and recurrent tumours (Figure 4.3B) we found that most CNs are dominated by a specific type (Figure 4.4C). CNs capture multiple cells in close proximity (Figure 4.4A) so are akin to the information captured by spot-based spatial technologies such as the 10X Visium platform. Our finding of dominance of a given cell type in each defined CN agrees with Greenwald *et al.*'s recently published results from application of the Visium platform to primary GBM samples[11]. Extending these results using IMC, which provides single cell resolution, we can further see that this dominance rarely equates to more than 50% of the cell types in a given CN, meaning there is clear ad-mixture and heterogeneity

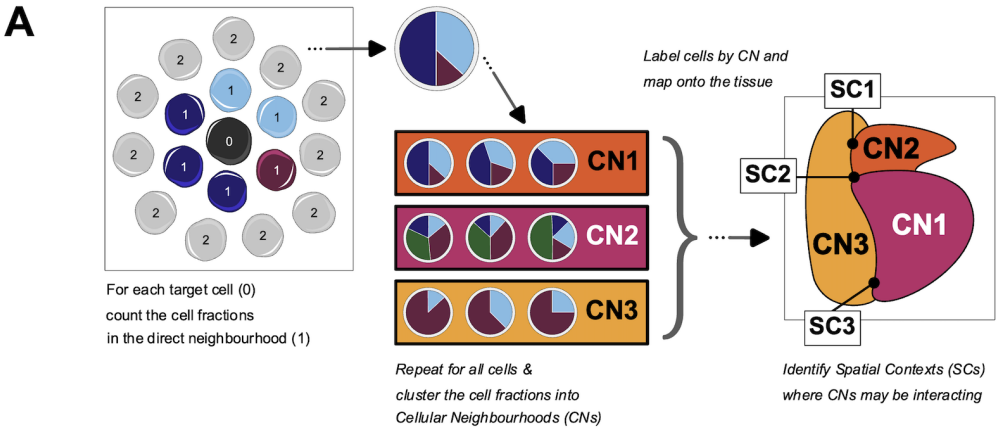
in interacting cells even when signal from one type predominates (Figure 4.4C).

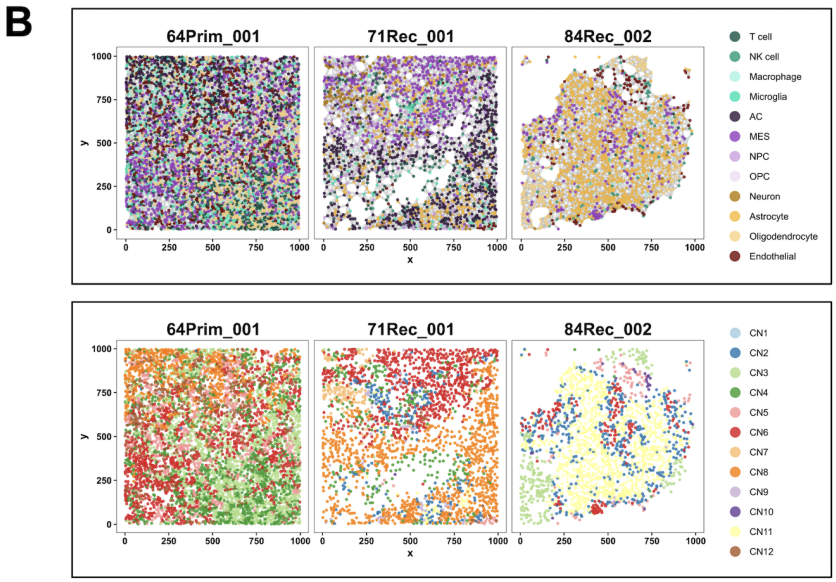
Greenwald *et al.* proceeded to cluster their spot-based gene expression profiles into 16 "metaprograms" (MPs). Our CNs map to these MPs (Table 4.1 and Table S4.12), though with some differences due to the level of cellular resolution and the differences in dimensionality and modality between the two studies. Specifically, Visium spots capture signals from 1-35 cells, so some MPs result from more than just nearest neighbours; and MPs are derived from gene expression (typically 7000 parameters our CNs derive from protein expression (34 parameters). It is worth noting that we aligned both CN4 (T-cell dominated) and CN10 (natural killer cell (NK cell) dominated) with the T-cell MP, owing to the functional similarities between T- and NK-cells[39].

Having aligned with previous findings from GBM tumours at a single time-point, we wished to see how the prevalence of CNs change over time. We see that certain CNs increased in abundance from primary to recurrent tumours, and others decrease (Figure 4.4D). Primary samples were enriched in neighbourhoods that included immune cells, particularly macrophages (CN1 and CN12) and lymphocytes (CN10), vasculature (CN5), hypoxic MES-like (CN6) and AC-like cancer cells (CN8). In contrast, the recurrent surgery samples were enriched in neighbourhoods dominated by normal brain cells: astrocytes (CN9 and CN11); neurons (CN7); and oligodendrocytes (CN3). Interestingly, we found that whilst hypoxic mesenchymal-driven CN6 decreased, astrocytic like mesenchymal-driven CN2 was increased from primary to recurrence.

Ultimately these results reconfirm what was seen when looking at cell type or category prevalences in isolation (Figure 4.2) i.e. that immune cell-driven (CN1, CN10 and CN12), vascular-cell driven (CN5) and cancer cell-driven (CN6 and CN8) neighbourhoods decreased from primary to recurrence, whereas normal brain cell driven (CN2, CN3, CN7, CN9 and CN11) neighbourhoods increased. CN4, which was dominated by T-cells, changed least in prevalence over time.

Greenwald *et al.*'s seminal finding was that, in some primary GBMs, MPs form organised layers that result in a global tumour architecture, which is seemingly driven by the presence of hypoxic niches. We, therefore, proceeded to investigate whether this organisation was evident in our primary samples and whether it was maintained post-treatment[11].


A
For each target cell (0)
count the cell fractions
in the direct neighbourhood (1)
CN1
CN2
CN3
Repeat for all cells &
cluster the cell fractions into
Cellular Neighbourhoods (CNs)
Label cells by CN and
map onto the tissue
SC1
SC2
SC3
CN1
CN2
CN3
Identify Spatial Contexts (SCs)
where CNs may be interacting


B
64Prim_001
71Rec_001
84Rec_002
T cell
NK cell
Macrophage
Microglia
AC
MES
NPC
OPC
Neuron
Astrocyte
Oligodendrocyte
Endothelial
CN1
CN2
CN3
CN4
CN5
CN6
CN7
CN8
CN9
CN10
CN11
CN12

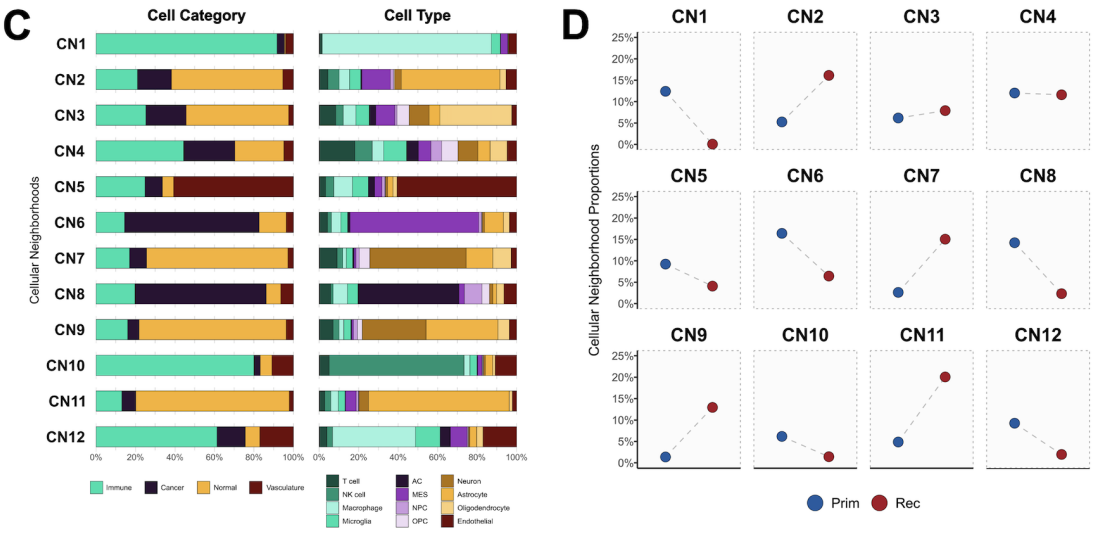

C
Cell Category
Cell Type
Cellular Neighborhoods
Immune
Cancer
Normal
Vasculature
D
Cellular Neighborhood Proportions
Prim
Rec

Figure 4.4 | Identifying distinct GBM TME CNs present across samples through treatment.

A) Schematic showing the process of identifying CNs: each cells direct neighbourhood (as defined by an interaction graph) cell fraction is aggregated and clustered across each patient/surgery. The resulting CN cluster labels are then mapped to each cell object. The SCs are then defined as locations where the most dominant CNs interactions are also interacting. **B)** Plots showing the single-cell spatial locations of three representative patient/surgery sample regions of interest visualised as nodes on a 2-dimensional plane, with cell-cell interactions shown in the form of undirected edges between nodes (top). The nodes are coloured according to the cell type label (top) and also by the cellular neighbourhoods they belong to (bottom). **C)** Stacked bar charts showing the proportion of each cell category (left) and cell type (right) that is present across each CN (rows). **D)** Dot plots showing the relative proportion of cells in each of the CNs (facets) across each surgery type.

Table 4.1 | Mapping of previously defined spatial GBM meta-programs to the CNs defined in this study.

Greenwald et al. Meta-program (MP)	Meta-program Description	CN (Figure 4.4C)
MES-Hyp	Hypoxic mesenchymal cancer cells	CN6
MES-Ast	Astrocytic-like mesenchymal cancer cells	CN2
MES	Mesenchymal (other) cancer cells	
OPC	Oligodendrocyte progenitor cell-like cancer cells	CN3
AC	Astrocytic like cancer cells	CN8
NPC	Neural progenitor cell-like cancer cells	
Oligo	Oligodendrocytes	CN3
Neuron	Neurons	CN7 & CN9
Reactive Ast	Reactive astrocytes	CN9 & CN11
Inflammatory Mac	Inflammatory macrophages	CN12
Mac	Macrophage and microglia	CN1
T-cell	T-cells	CN4 & CN10
B-cell	B-cells	
Vasc	Vasculature	CN5
Chromatin-Reg	Chromatin regulation	
Prolif-Metab	Proliferation and metabolism	

4.3.7 Alterations in spatial organisation through treatment in GBM

To better understand higher order structuring of our CNs, we classified SCs; locations where distinct CNs were found to consistently interact (Figure 4.4A). When considering the most dominant CN interactions present across primary and recurrent surgeries our results reproduce similar ordered layers to those reported in Greenwald *et al.*[11]. However, the prevalence and importance of states which make up the layers differs greatly through treatment, as revealed by the structure and parameters of the calculated CN interaction networks (Figure 4.5).

In the primary samples the most influential and prevalent cellular neighbourhoods were those characterised by layers 1 and layers 3, which denote the hypoxic/necrotic core niche and the angiogenesis-immune hub, respectively (Figure 4.5A). These layers were comprised of CNs with high network centrality scores across all three measures (degree, closeness and betweenness), indicating their importance for communication between other neighbourhoods and layers. This concurs with previous findings suggesting that hypoxia potentially drives the presence of the organised layers owing to phenotypic consequences of reduced oxygen, especially at the tumour core[11].

Conversely, in recurrent samples there were many more significant interactions between CNs in different layers (Figure 4.5B) in agreement with our findings from pairwise cellular interaction analysis (Figure 4.3B). Additionally, the most influential and prevalent CNs in recurrent samples were mostly in layers 2 and 5, which represented the hypoxia-adjacent and normal infiltrative brain layers (Figure 4.5B). This suggests a reduced global structure with less well organised layers, potentially owing to a reduction in the presence of hypoxic niches in recurrent versus primary GBM.

Worth noting is that T-cell dominated CN4, which remained the most stably prevalent between primary and recurrent samples (Figure 4.4D), in combination with CN10 (together these CNs align to the previously denoted T-cell MP: Table 4.1) maintain high network parameters in both primary and recurrent GBMs (Figure 4.5), implying they are important in driving spatial contexts both pre- and post-treatment.

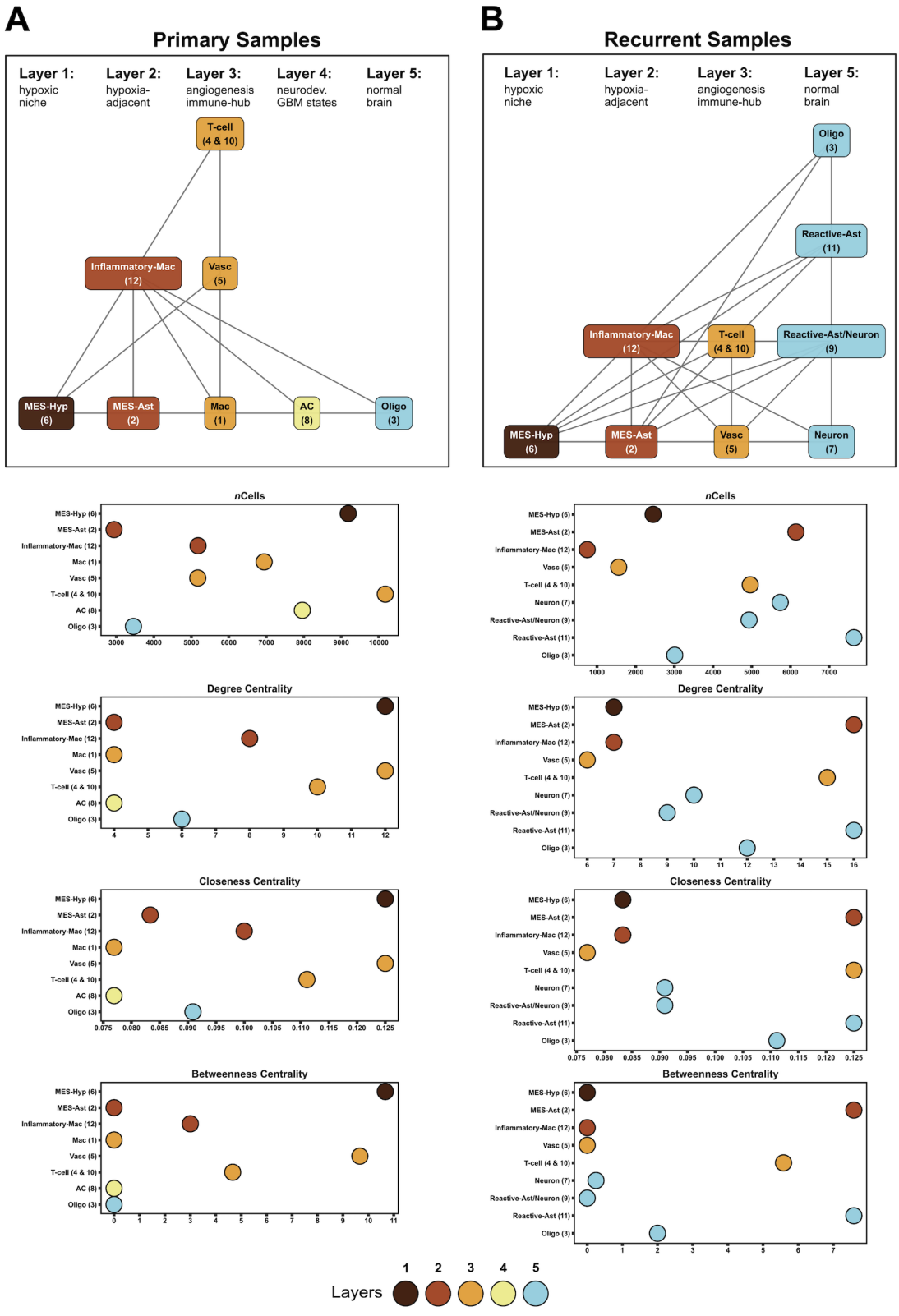

A
Primary Samples
Layer 1: hypoxic niche
Layer 2: hypoxia-adjacent
Layer 3: angiogenesis immune-hub
Layer 4: neurodev. GBM states
Layer 5: normal brain
T-cell (4 & 10)
Inflammatory-Mac (12)
Vasc (5)
MES-Hyp (6)
MES-Ast (2)
Mac (1)
AC (8)
Oligo (3)
B
Recurrent Samples
Layer 1: hypoxic niche
Layer 2: hypoxia-adjacent
Layer 3: angiogenesis immune-hub
Layer 5: normal brain
Oligo (3)
Reactive-Ast (11)
Inflammatory-Mac (12)
T-cell (4 & 10)
Reactive-Ast/Neuron (9)
MES-Hyp (6)
MES-Ast (2)
Vasc (5)
Neuron (7)
nCells
Degree Centrality
Closeness Centrality
Betweenness Centrality
Layers
1 2 3 4 5

Figure 4.5 | Spatial organisation of CNs across surgeries.

A-B) Top: network graphs with nodes labelled according to the cell meta-programs identified in Greenwald *et al.*, 2024 along with their corresponding CNs (shown in the brackets below). The edges represent the most dominant interactions present across primary surgery **(A)** and recurrent surgery **(B)** ROIs, respectively. Bottom: dot plots showing the number of cells present across each surgery specific CN and also three network-specific centrality measures: degree; closeness and betweenness. Both network graphs and their corresponding metrics are coloured and ordered according to the structured GBM spatial layers described in Greenwald *et al.*, 2024.

4.4 Discussion

This was a small study (n=5), and technical limitations of protein-based expression profiling confined the number of markers that could be used to assign cell types (n=34). This meant that some cells (e.g. B cells) were not included, thereby reducing our ability to conclusively discern between neoplastic and normal brain cell types, which they closely related[6]. Notwithstanding, protein markers exhibit less stochastic expression and lower signal dropout (false negatives) compared to single-cell sequencing approaches.

Our results reveal an influx of normal brain cells into the GBM microenvironment post-treatment, alongside a reduction in vascular cells (Figure 4.2A). The latter is expected, as surgery aims to de-bulk the highly vascularised tumour core[40,41]. However, the reduction in endothelial cells in recurrent GBM suggests a reduced functional reliance on vasculature, which may explain the failure of angiogenic therapies like bevacizumab (Avastin) in clinical trials[42].

Several large cohort studies which deconvoluted cellular signals from paired GBM using bulk RNA-seq and single-cell RNA sequencing (scRNA-seq) have reported an increased presence of oligodendrocytes at recurrence[22,43,44]. Herein, we confirm this finding (Figure 4.2B) and further show that oligodendrocytes integrate into the GBM TME, as their interactions with other cells significantly increase at recurrence (Figure 4.3B). Oligodendrocytes are essential for cerebral homeostasis and regulate neuronal activity via axon myelination[33]. We characterised oligodendrocytes using the myelin oligodendrocyte glycoprotein (MOG) marker (Figure 4.1B), suggesting that the increases we observe relate to myelinating oligodendrocytes integrating into the tissue and implying a potential functional role for myelination within the GBM TME post-treatment. Interestingly, the increased in-

teractions and integration of oligodendrocytes we observed in recurrent tumours, primarily involve non-neoplastic cells. Among cancer cells, recurrence-specific interactions involving oligodendrocytes are restricted to MES-like neoplastic cells (Figure 4.3B). Oligodendrocytes have been shown to upregulate the invasive capacity of GBM cancer cells via Angiopoietin-2 signalling, and MES-like are the most invasive neoplastic subtype[34,45]. We found that all neoplastic GBM cells showed increased EMT at recurrence (Figure 4.2D). Moreover, the CN dominated by oligodendrocytes (CN3) had higher closeness and degree centrality at recurrence (Figure 4.5), indicating greater connectivity and interaction with other CNs. We propose that the role of oligodendrocytes in driving post-treatment recovery of GBM is worthy of further exploration.

Astrocytes exhibit the largest increase of normal brain cells (Figure 4.2B) within the recurrent GBM TME, and also the highest number of recurrence-specific interactions, particularly with neoplastic cells (Figure 4.3B). Within the healthy brain parenchyma, astrocytes are crucial for neuronal cell homeostasis and also help drive the brain's injury response by acquiring a reactive phenotype. Consistent with this role, CNs that map to the previously defined reactive astrocytic metaprogramme (CN9 and CN11) and the astrocytic mesenchymal metaprogramme (CN2), are increased at recurrence (Figure 4.4D and Table 4.1). Astrocytes also exhibit resistance to apoptosis triggered by death receptors during inflammation, such as apoptosis antigen 1 and TNF-related apoptosis-inducing ligands (FAS, TRAIL), indicating their resilience under inflammatory conditions. Together, this suggests a phenotypic response within the (infiltrating) astrocytic population of the TME that could serve to protect neoplastic cells.

Previous research has indicated that there is a shift towards a more mesenchymal state in bulk tumours at recurrence[22]. Single cell analyses have further refined this understanding showing that, whilst some GBMs show increase in MES-like cancer cells post-treatment, others show increases in more proneural (OPC-like and NPC-like) cells at recurrence[46,47]. We also find that that there is no significant, consistent change in neoplastic cell types at recurrence, but instead, a universal increase in EMT markers across all the neoplastic cell (Figure 4.2D). This potentially explains the shift to mesenchymal expression signatures observed from bulk tumour profiling[22].

In keeping with our previous findings, AC-like cancer cells reduce most consistently at recurrence[23]. However, the remaining AC-like cells had elevated levels of hypoxia (Figure 4.2B and Figure 4.2D), while these decreased within MES-like and NPC-like cell populations. Hypoxia can induce a reactive astrocyte pheno-

type within the TME, which may extend to AC-like cancer cells, potentially even promoting plastic conversion to this neoplastic subtype[48,49].

Overall, we find no consistent changes in cellular diversity between primary and recurrent GBM (Figure 4.2C), suggesting that while cellular heterogeneity is maintained, post-treatment GBM tumours have greater interactions between differing cell types (greater admixture, Figure 4.3 and Figure 4.4). A recent spatial profiling study of primary GBM tumours by Greenwald *et al.* concluded that hypoxia drives organisation of a GBM architecture, composed of layers[11]. Our findings concur with theirs for primary GBM but expand further, revealing that this layering is less structured post-treatment (Figure 4.5). The decrease in CN6, which maps to their hypoxic MES-like cancer cell metaprogramme (Table 4.1), but increase in CN2, which maps to their astrocytic MES cancer cell metaprogramme (Table 4.1), at recurrence suggests that an overall reduction in hypoxia post-treatment, could drive this increased disorder. This influx and integration of normal brain cells in the GBM TME at recurrence corresponds with these cells becoming much more influential in terms of the interaction between cellular layers, particularly CN11 which map to the reactive astrocyte metaprogramme of Greenwald *et al.* (Figure 4.5B).

Whilst lymphocyte abundance remains unchanged between primary and recurrent GBM (Figure 2B), neighbourhoods (CN4 and CN10) mapping to the T-cell metaprogramme (Table 4.1) become much more influential in the recurrent GBM (Figure 4.5B). T-cells and tertiary lymphoid structures (TLSs) (regions enriched in lymphocytes, resembling CN4 and CN10) have been shown to increase in subsets of paired primary and recurrent GBM[46,50],which has renewed interest in understanding their potential role in immunotherapy. In support of this, activated T-cells have been shown to associate specifically with astrocytic MES which is the subtype we also find increased recurrent tumours[24].

This study highlights prominent post-treatment changes in the GBM cellular landscape and offers novel insight into the importance of specific interactions between GBM cancer cells and the TME during tumour survival and regrowth. Further characterisation of these interactions could identify therapeutic targets to overcome treatment resistance and better treat recurrent tumours.

References

1. Delgado-López, P. D. & Corrales-García, E. M. Survival in glioblastoma: a review on the impact of treatment modalities. *Clinical and Translational Oncology* **18**, 1062–1071 (Nov. 2016). doi:10.1007/s12094-016-1497-x.

2. Stupp, R. *et al.* Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma. *New England Journal of Medicine* **352**, 987–996 (Mar. 2005). doi:10.1056/NEJMoa043330.

3. Barthel, F. P. *et al.* Longitudinal molecular trajectories of diffuse glioma in adults. *Nature* **576**, 112–120 (Dec. 2019). doi:10.1038/s41586-019-1775-1.

4. Körber, V. *et al.* Evolutionary Trajectories of IDHWT Glioblastomas Reveal a Common Path of Early Tumorigenesis Instigated Years ahead of Initial Diagnosis. *Cancer Cell* **35**, 692–704.e12 (Apr. 2019). doi:10.1016/j.ccell.2019.02.007.

5. Wang, L. *et al.* The Phenotypes of Proliferating Glioblastoma Cells Reside on a Single Axis of Variation. *Cancer Discovery* **9**, 1708–1719 (Dec. 2019). doi:10.1158/2159-8290.CD-19-0329.

6. Couturier, C. P. *et al.* Single-cell RNA-seq reveals that glioblastoma recapitulates a normal neurodevelopmental hierarchy. *Nature Communications* **11**, 3406 (July 2020). doi:10.1038/s41467-020-17186-5.

7. Neftel, C. *et al.* An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. *Cell* **178**, 835–849.e21 (Aug. 2019). doi:10.1016/j.cell.2019.06.024.

8. Hanahan, D. & Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **144**, 646–674 (Mar. 2011). doi:10.1016/j.cell.2011.02.013.

9. Quail, D. F. & Joyce, J. A. The Microenvironmental Landscape of Brain Tumors. *Cancer Cell* **31**, 326–341 (Mar. 2017). doi:10.1016/j.ccell.2017.02.009.

10. Bressan, D., Battistoni, G. & Hannon, G. J. The dawn of spatial omics. *Science (New York, N.Y.)* **381**, eabq4964 (Aug. 2023). doi:10.1126/science.abq4964.

11. Greenwald, A. C. *et al.* Integrative spatial analysis reveals a multi-layered organization of glioblastoma. *Cell* **187**, 2485–2501.e26 (May 2024). doi:10.1016/j.cell.2024.03.029.

12. Manoharan, V. T. *et al.* Spatiotemporal modeling reveals high-resolution invasion states in glioblastoma. *Genome Biology* **25**, 264 (Oct. 2024). doi:10.1186/s13059-024-03407-3.

13. Ravi, V. M. *et al.* Spatially resolved multi-omics deciphers bidirectional tumor-host interdependence in glioblastoma. *Cancer Cell* **40**, 639–655.e13 (June 2022). doi:10.1016/j.ccell.2022.05.009.

14. Giesen, C. *et al.* Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nature Methods* **11**, 417–422 (Apr. 2014). doi:10.1038/nmeth.2869.

15. Ajaib, S. *et al.* GBMdeconvoluteR accurately infers proportions of neoplastic and immune cell populations from bulk glioblastoma transcriptomics data. *Neuro-Oncology* **25**, 1236–1248 (July 2023). doi:10.1093/neuonc/noad021.

16. Windhager, J. *et al.* An end-to-end workflow for multiplexed image processing and analysis. *Nature Protocols* **18**, 3565–3613 (Nov. 2023). doi:10.1038/s41596-023-00881-0.

17. Pachitariu, M. & Stringer, C. Cellpose 2.0: how to train your own model. *Nature Methods* **19**, 1634–1641 (Dec. 2022). doi:10.1038/s41592-022-01663-4.

18. Schapiro, D. *et al.* histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data. *Nature Methods* **14**, 873–876 (Sept. 2017). doi:10.1038/nmeth.4391.

19. Bhate, S. S., Barlow, G. L., Schürch, C. M. & Nolan, G. P. Tissue schematics map the specialization of immune tissue motifs and their appropriation by tumors. *Cell Systems* **13**, 109–130.e6 (Feb. 2022). doi:10.1016/j.cels.2021.09.012.

20. Zhang, Y., Parmigiani, G. & Johnson, W. E. ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genomics and Bioinformatics* **2**, lqaa078 (Sept. 2020). doi:10.1093/nargab/lqaa078.

21. Gieryng, A., Pszczolkowska, D., Walentynowicz, K. A., Rajan, W. D. & Kaminska, B. Immune microenvironment of gliomas. *Laboratory Investigation* **97**, 498–518 (May 2017). doi:10.1038/labinvest.2017.19.

22. Varn, F. S. *et al.* Glioma progression is shaped by genetic evolution and microenvironment interactions. *Cell* **185**, 2184–2199.e16 (June 2022). doi:10.1016/j.cell.2022.04.038.

23. Tanner, G. *et al.* IDHwt glioblastomas can be stratified by their transcriptional response to standard treatment, with implications for targeted therapy. *Genome Biology* **25**, 45 (Feb. 2024). doi:10.1186/s13059-024-03172-3.

24. Chanoch-Myers, R., Wider, A., Suva, M. L. & Tirosh, I. Elucidating the diversity of malignant mesenchymal states in glioblastoma by integrative analysis. *Genome Medicine* **14**, 106 (Sept. 2022). doi:10.1186/s13073-022-01109-8.

25. Gao, Z. *et al.* PDIA3P1 promotes Temozolomide resistance in glioblastoma by inhibiting C/ebpbeta degradation to facilitate proneural-to-mesenchymal transition. *Journal of Experimental & Clinical Cancer Research : CR* **41**, 223 (July 2022). doi:10.1186/s13046-022-02431-0.

26. Yoon, S.-J. *et al.* Tumor Mesenchymal Stem-Like Cell as a Prognostic Marker in Primary Glioblastoma. *Stem Cells International* **2016**, 6756983 (2016). doi:10.1155/2016/6756983.

27. Carro, M. S. *et al.* The transcriptional network for mesenchymal transformation of brain tumours. *Nature* **463**, 318–325 (Jan. 2010). doi:10.1038/nature08712.

28. Barriere, G., Fici, P., Gallerani, G., Fabbri, F. & Rigaud, M. Epithelial Mesenchymal Transition: a double-edged sword. *Clinical and Translational Medicine* **4**, e14 (2015). doi:10.1186/s40169-015-0055-4.

29. Spinelli, C. *et al.* Mesenchymal glioma stem cells trigger vasectasia—distinct neovascularization process stimulated by extracellular vesicles carrying EGFR. *Nature Communications* **15**, 2865 (Apr. 2024). doi:10.1038/s41467-024-46597-x.

30. Cheng, J. *et al.* Myeloid cells coordinately induce glioma cell-intrinsic and cell-extrinsic pathways for chemoresistance via GP130 signaling. *Cell Reports Medicine* **5**, 101658 (Aug. 2024). doi:10.1016/j.xcrm.2024.101658.

31. Andreou, T. *et al.* Hematopoietic stem cell gene therapy targeting tgf-beta enhances the efficacy of irradiation therapy in a preclinical glioblastoma model. *Journal for Immunotherapy of Cancer* **9**, e001143 (Mar. 2021). doi:10.1136/jitc-2020-001143.

32. Brooks, L. J., Ragdale, H. S., Hill, C. S., Clements, M. & Parrinello, S. Injury programs shape glioblastoma. *Trends in Neurosciences* **45**, 865–876 (Nov. 2022). doi:10.1016/j.tins.2022.08.006.

33. Hide, T. *et al.* Oligodendrocyte Progenitor Cells and Macrophages/Microglia Produce Glioma Stem Cell Niches at the Tumor Border. *eBioMedicine* **30**, 94–104 (Apr. 2018). doi:10.1016/j.ebiom.2018.02.024.

34. Kawashima, T. *et al.* Oligodendrocytes Up-regulate the Invasive Activity of Glioblastoma Cells via the Angiopoietin-2 Signaling Pathway. *Anticancer Research* **39**, 577–584 (Feb. 2019). doi:10.21873/anticanres.13150.

35. Huang, Y. *et al.* Oligodendrocyte Progenitor Cells Promote Neovascularization in Glioma by Disrupting the Blood–Brain Barrier. *Cancer Research* **74**, 1011–1021 (Feb. 2014). doi:10.1158/0008-5472.CAN-13-1072.

36. Hortega, P. d. R. & Penfield, W. *Cerebral Cicatrix: The Reaction of Neuroglia and Microglia to Brain Wounds* (Johns Hopkins Hospital, 1927).

37. Escartin, C. *et al.* Reactive astrocyte nomenclature, definitions, and future directions. *Nature neuroscience* **24**, 312 (Feb. 2021). doi:10.1038/s41593-020-00783-4.

38. Puchalski, R. B. *et al.* An anatomic transcriptional atlas of human glioblastoma. *Science* **360**, 660–663 (May 2018). doi:10.1126/science.aaf2666.

39. Narni-Mancinelli, E., Vivier, E. & Kerdiles, Y. M. The 'T-cell-ness' of NK cells: unexpected similarities between NK cells and T cells. *International Immunology* **23**, 427–431 (July 2011). doi:10.1093/intimm/dxr035.

40. Maddison, K. *et al.* Vasculogenic Mimicry Occurs at Low Levels in Primary and Recurrent Glioblastoma. *Cancers* **15**, 3922 (Aug. 2023). doi:10.3390/cancers15153922.

41. Ahir, B. K., Engelhard, H. H. & Lakka, S. S. Tumor Development and Angiogenesis in Adult Brain Tumor: Glioblastoma. *Molecular Neurobiology* **57**, 2461–2478 (2020). doi:10.1007/s12035-020-01892-8.

42. Wick, W. *et al.* LB-05phase III TRIAL EXPLORING THE COMBINATION OF BEVACIZUMAB AND LOMUSTINE IN PATIENTS WITH FIRST RECURRENCE OF a GLIOBLASTOMA: THE EORTC 26101 TRIAL. *Neuro-Oncology* **17**, v1–v1 (Nov. 1, 2015). doi:10.1093/neuonc/nov306.

43. Hoogstrate, Y. *et al.* Transcriptome analysis reveals tumor microenvironment changes in glioblastoma. *Cancer Cell* **41**, 678–692.e7 (Apr. 2023). doi:10.1016/j.ccell.2023.02.019.

44. Wang, L. *et al.* A single-cell atlas of glioblastoma evolution under therapy reveals cell-intrinsic and cell-extrinsic therapeutic targets. *Nature Cancer* **3**, 1534–1552 (Dec. 2022). doi:10.1038/s43018-022-00475-x.

45. Zhong, J., Paul, A., Kellie, S. J. & O'Neill, G. M. Mesenchymal Migration as a Therapeutic Target in Glioblastoma. *Journal of Oncology* **2010**, 430142 (2010). doi:10.1155/2010/430142.

46. Van de Walle, T. *et al.* Tertiary Lymphoid Structures in the Central Nervous System: Implications for Glioblastoma. *Frontiers in Immunology* **12**, 724739 (Sept. 2021). doi:10.3389/fimmu.2021.724739.

47. Stead, L. F. Treating glioblastoma often makes a MES. *Nature Cancer* **3**, 1446–1448 (Dec. 2022). doi:10.1038/s43018-022-00471-1.

48. Pantazopoulou, V., Jeannot, P., Rosberg, R., Berg, T. J. & Pietras, A. Hypoxia-Induced Reactivity of Tumor-Associated Astrocytes Affects Glioma Cell Properties. *Cells* **10**, 613 (Mar. 2021). doi:10.3390/cells10030613.

49. Rosberg, R. *et al.* Hypoxia-induced complement component 3 promotes aggressive tumor growth in the glioblastoma microenvironment. *JCI Insight* **9**, (Oct. 2024). doi:10.1172/jci.insight.179854.

50. White, K. *et al.* Identification, validation and biological characterisation of novel glioblastoma tumour microenvironment subtypes: implications for precision immunotherapy. *Annals of Oncology* **34**, 300–314 (Mar. 2023). doi:10.1016/j.annonc.2022.11.008.

Supplementary material

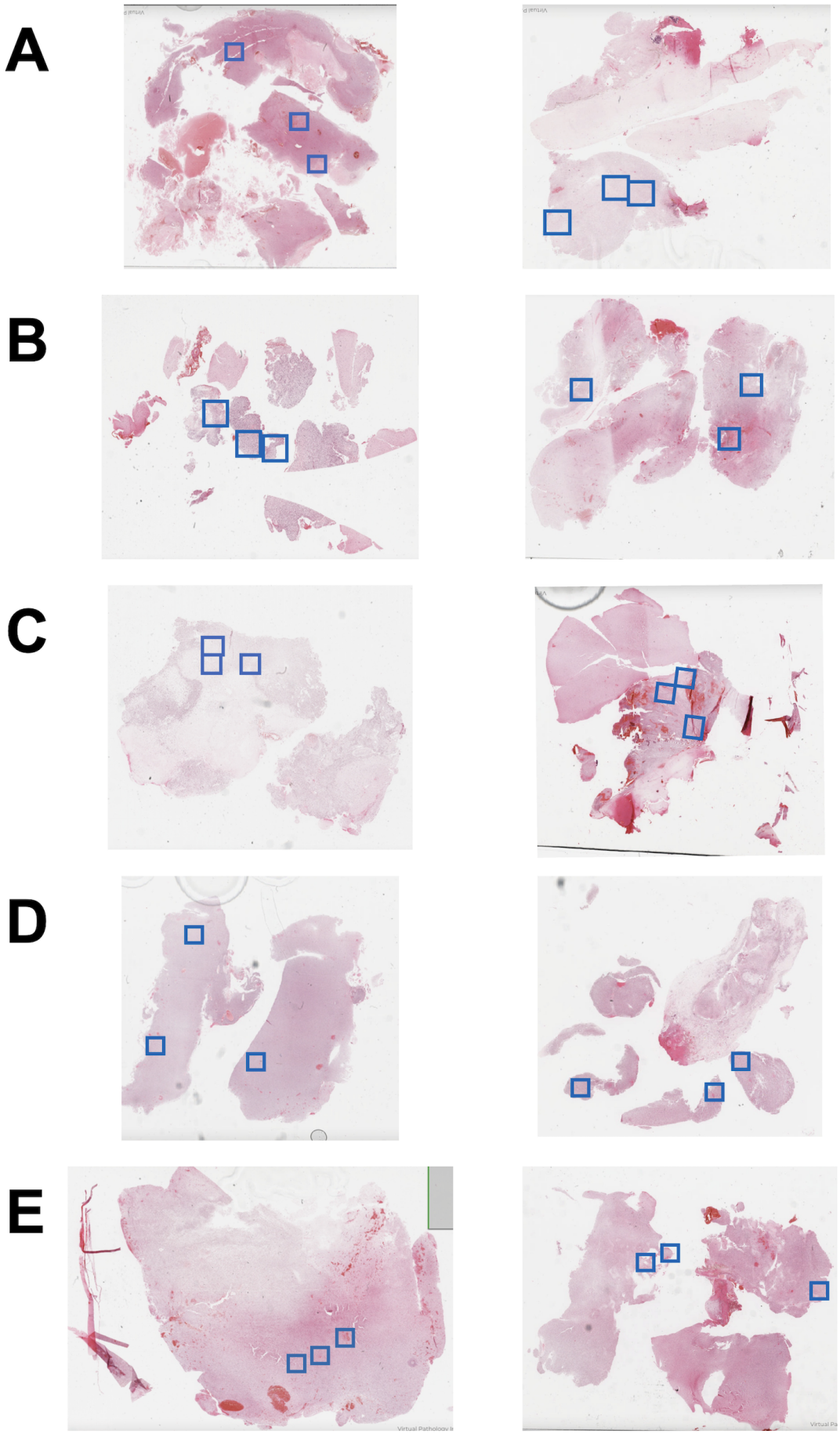

A
B
C
D
E

Figure S4.1 | Haematoxylin and eosin staining (H&E) stained formalin-fixed paraffin-embedded (FFPE) sections of the five matched, primary (P) and recurrent (R) IDHwt GBMtumour samples used in the study.

The blue outlines demarcate the 1mm^2 ROIs that underwent IMC. The FFPE sections correspond (from left to right) to the following patient/surgeries: **A)** 64P and 64R; **B)** 67P and 67R; **C)** 71P and 71R; **D)** 82P and 82R; **E)** 84P and 84R.

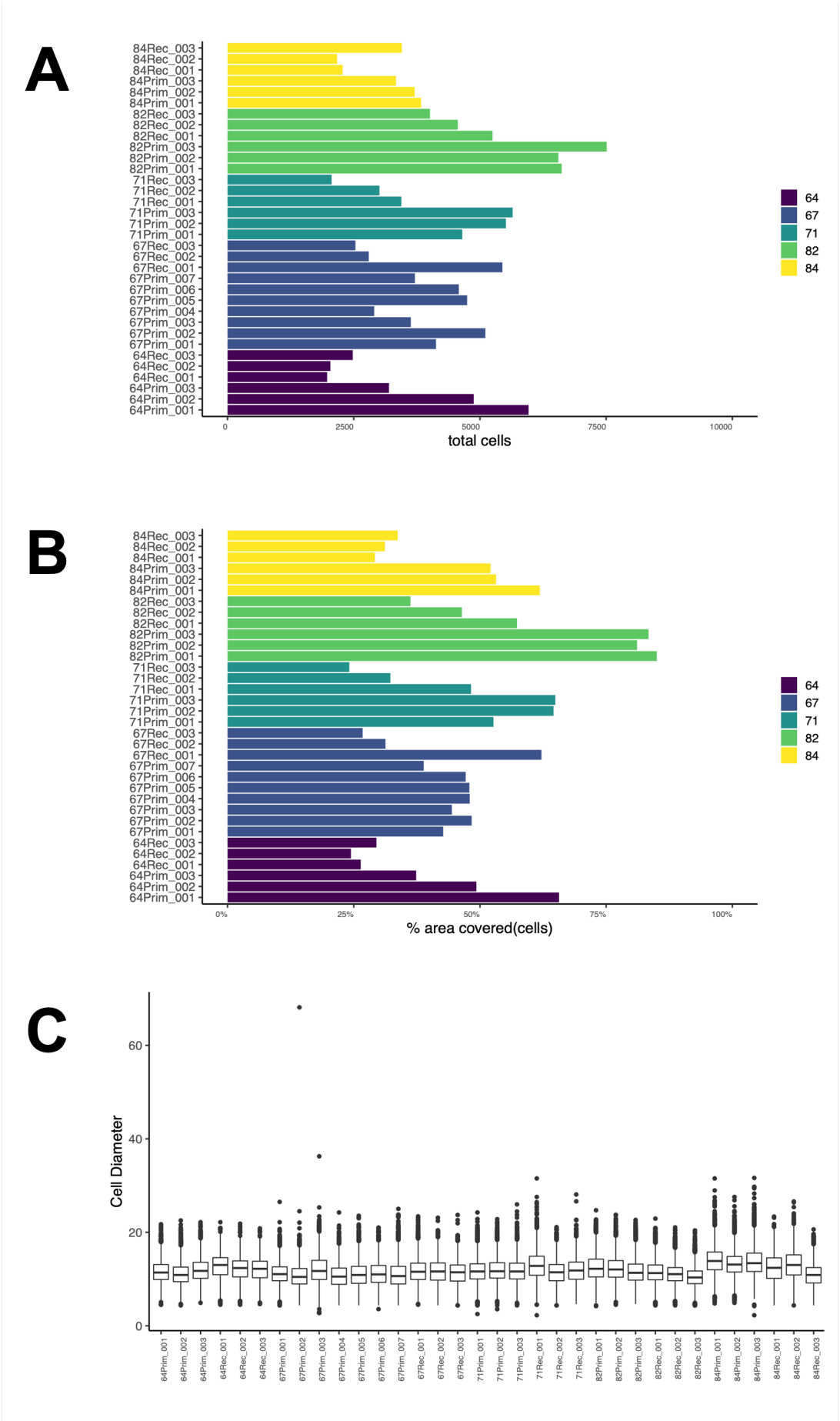

A
84Rec_003
84Rec_002
84Rec_001
84Prim_003
84Prim_002
84Prim_001
82Rec_003
82Rec_002
82Rec_001
82Prim_003
82Prim_002
82Prim_001
71Rec_003
71Rec_002
71Rec_001
71Prim_003
71Prim_002
71Prim_001
67Rec_003
67Rec_002
67Rec_001
67Prim_007
67Prim_006
67Prim_005
67Prim_004
67Prim_003
67Prim_002
67Prim_001
64Rec_003
64Rec_002
64Rec_001
64Prim_003
64Prim_002
64Prim_001
0
2500
5000
7500
10000
total cells
64
67
71
82
84
B
0%
25%
50%
75%
100%
% area covered(cells)
C
Cell Diameter
0
20
40
60

Figure S4.2 | Segmented single-cell object metrics.

A) The total number of cells present in each ROI coloured by patient. **B)** The percentage area covered by the segmented cell objects present across each ROI coloured by patient. **C)** The distribution of cell diameters corresponding to the segmented cell objects present in each patient/surgery ROIs.

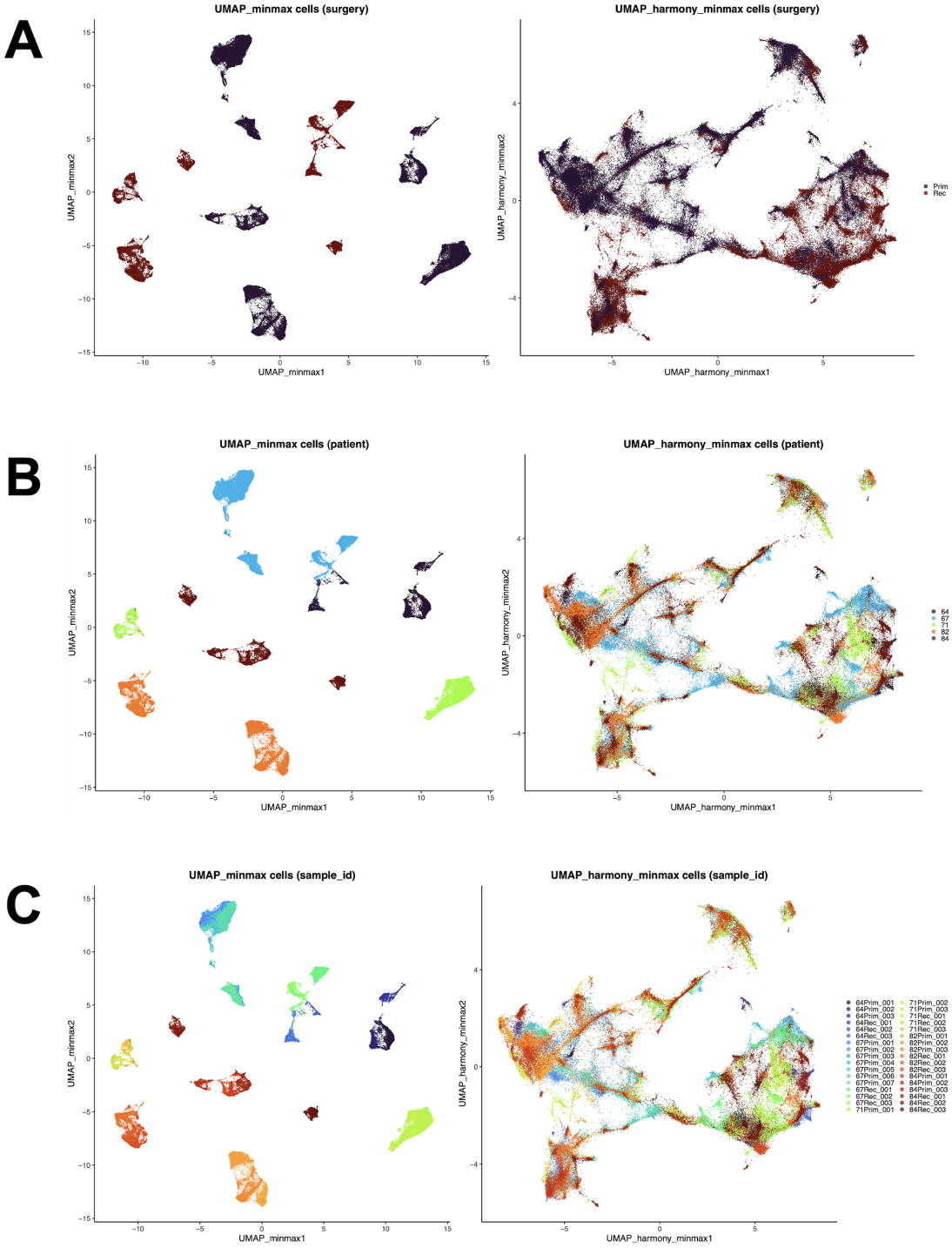

A
UMAP_minmax cells (surgery)
UMAP_harmony_minmax cells (surgery)
UMAP_minmax1
UMAP_minmax2
UMAP_harmony_minmax1
UMAP_harmony_minmax2
Prim
Rec
B
UMAP_minmax cells (patient)
UMAP_harmony_minmax cells (patient)
64
67
71
82
84
C
UMAP_minmax cells (sample_id)
UMAP_harmony_minmax cells (sample_id)
64Prim_001
64Prim_002
64Prim_003
64Rec_001
64Rec_002
64Rec_003
67Prim_001
67Prim_002
67Prim_003
67Prim_004
67Prim_005
67Prim_006
67Prim_007
67Rec_001
67Rec_002
67Rec_003
71Prim_001
71Prim_002
71Prim_003
71Rec_001
71Rec_002
71Rec_003
82Prim_001
82Prim_002
82Prim_003
82Rec_001
82Rec_002
82Rec_003
84Prim_001
84Prim_002
84Prim_003
84Rec_001
84Rec_002
84Rec_003

Figure S4.3 | UMAP of all segmented single-cell objects before and after batch correction of *min-max* scaled protein marker abundances.

The UMAP projections are coloured according to known sources of sample variation: **A)** surgery type; **B)** patient from which samples were obtained; **C)** patient and surgery-specific ROI.

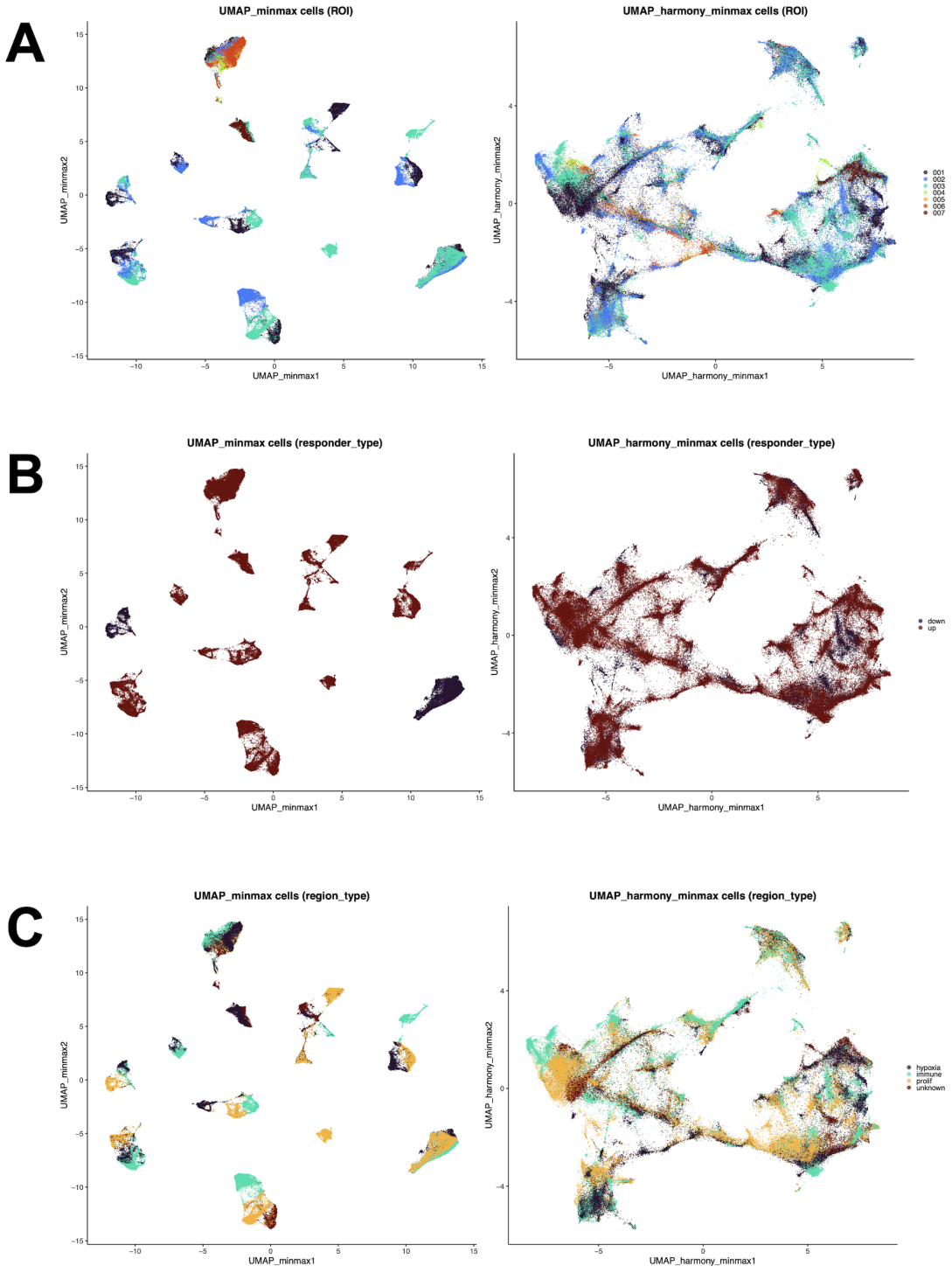

A
UMAP_minmax cells (ROI)
UMAP_harmony_minmax cells (ROI)
UMAP_minmax2
UMAP_minmax1
UMAP_harmony_minmax2
UMAP_harmony_minmax1
001
002
003
004
005
006
007
B
UMAP_minmax cells (responder_type)
UMAP_harmony_minmax cells (responder_type)
down
up
C
UMAP_minmax cells (region_type)
UMAP_harmony_minmax cells (region_type)
hypoxia
immune
prolif
unknown

Figure S4.4 | UMAP of all segmented single-cell objects before and after batch correction of *min-max* scaled protein marker abundances.

The UMAP projections are coloured according to known sources of sample variation: **A)** ROI; **B)** responder types as defined in Tanner *et al.*[23]; **C)** immunohistochemistry (IHC) annotation of each ROI based on areas of high hypoxia, proliferation, and immune cell infiltration.

Table S4.1 | Clinical and molecular data of patients included in the study.

The *Tumour Loc* field denotes the primary tumour brain region. The Age is that at initial diagnosis. The progression-free survival (PFS) and the overall survival (OS) are given in months and the Status field denotes: 1=deceased;0=alive. The responder field corresponds to the responder-type stratification.

Patient	Sex	Tumour Loc	Age	PFS	OS	Status	Responder
64	Male	Frontal	57	22.55	48.76	1.00	up
67	Male	Temporal	51	40.73	75.19	1.00	up
71	Female	Temporal	60	17.10	39.12	1.00	down
82	Female	Parietal	59	23.24	33.21	1.00	up
84	Female	Parietal	72	17.49	45.76	1.00	up

Table S4.2 | IMC ROIs analysed in this study.

Image id	Patient	Surgery	ROI	IHC annotation
64Prim_001	64	primary	001	prolif
64Prim_002	64	primary	002	hypoxia
64Prim_003	64	primary	003	immune
64Rec_001	64	recurrent	001	immune
64Rec_002	64	recurrent	002	hypoxia
64Rec_003	64	recurrent	003	prolif
67Prim_001	67	primary	001	immune
67Prim_002	67	primary	002	immune
67Prim_003	67	primary	003	unknown
67Rec_001	67	recurrent	001	prolif
67Rec_002	67	recurrent	002	hypoxia
67Rec_003	67	recurrent	003	unknown
71Prim_001	71	primary	001	hypoxia
71Prim_002	71	primary	002	immune
71Prim_003	71	primary	003	prolif
71Rec_001	71	recurrent	001	prolif
71Rec_002	71	recurrent	002	immune
71Rec_003	71	recurrent	003	hypoxia
82Prim_001	82	primary	001	unknown
82Prim_002	82	primary	002	immune
82Prim_003	82	primary	003	prolif
82Rec_001	82	recurrent	001	prolif
82Rec_002	82	recurrent	002	hypoxia
82Rec_003	82	recurrent	003	immune
84Prim_001	84	primary	001	prolif
84Prim_002	84	primary	002	hypoxia
84Prim_003	84	primary	003	immune
84Rec_001	84	recurrent	001	hypoxia
84Rec_002	84	recurrent	002	immune
84Rec_003	84	recurrent	003	prolif

Table S4.3 | IMC marker panel including marker/antibody justification.

The IDs listed in the marker justify field correspond to PubMed identifiers.AC-like; cytometry by time-of-flight (CyTOF); endoplasmic reticulum (ER)); EMT; immunocytochemistry (ICC); IHC; immunofluorescence (IF); knock-out (KO); MES-like; NPC-like; OPC-like; phosphate buffered saline (PBS).

Marker	Category	Type	Cell state	Loc 1	Loc 2	Marker justify	manufacturer (antibody clone)	antibody justify	Concentration (μg/mL)	Isotope	Control tissue
HOPX	cancer	AC-like		cytoplasm		31327527; 31554641; 32641768	abcam (ab230544)		100	Yb171	Brain, Tonsil
SLC1A3_EAAT1	cancer	AC-like		membrane		31327527; 31554641; 32641768	abcam(ab240235)	BSA and azide free	1000	Gd158	Brain
GFAP	normal	astrocyte		cytoplasm		25726916	abcam (ab218309)	PMID: 34174183	100	Sm149	Brain
CD56	normal	neuron		membrane	extracellular	28791027	biolegend (318345)	PMID: 28369679	100	Dy162	Brain
IBA1	immune	macrophage		cytoplasm		32848611	abcam (ab220815)	PMID: 34174183	200	Eu153	Spleen, Tonsil
NeuN_FOX3	normal	neuron		nucleus		20452351	biolegend (834502)	PMID: 34174183	400	Sm147	Brain
ANXA_A1	cancer	MES-like		cytoplasm	membrane	31327527; 31554641; 32641768	abcam (ab222398)	BSA and azide free		Yb172	Tonsil
ANXA_A2	cancer	MES-like		cytoplasm	membrane	31327527; 31554641; 32641768	rndsystems (mab3928)	In PBS with Trehalose	8-25	Er166	Prostate, Tonsil
CHI3L1	cancer	MES-like		cytoplasm	extracellular	31327527; 31554641; 32641768	abcam (ab255864)	BSA and azide free	250	Sm154	Spleen, Brain
SOD2	cancer	MES-like		mitochondria		31327527; 31554641; 32641768	abcam (ab227846)	BSA and azide free, used in IHC	100	Nd146	Prostate
P2Y12R	immune	microglia		membrane		32848611	abcam (ab274386)	BSA and azide free	1000	Lu175	Brain
TMEM119	immune	microglia		cytoplasm	membrane	32848611	sigmaaldrich (HPA051870)	PMID: 31740814	500-1000	Gd155	Brain, Tonsil
NKP46	immune	NK cell		membrane		31784984	rndsystems (mab1850)	PMID: 36689332	5-25	Nd144	Spleen
BCAN	cancer	NPC-like		extracellular	nucleus	31327527; 31554641; 32641768	thermofisher (MA5-27639)	BSA free, used in ICC	50	Gd160	Brain
DLL3	cancer	NPC-like		membrane		31327527; 31554641; 32641768	abcam (ab255694)	BSA and azide free	100	Nd148	Brain
MOG	normal	oligodendrocyte		membrane		2649509	rndsystems (mab1850)		5-25	Gd157	Brain
OLIG1	cancer	OPC-like		nucleus		31327527; 31554641; 32641768	rndsystems (mab2417)	carrier free, used in IHC	8-25	Yb174	Skin
SCD5	cancer	OPC-like		ER		31327527; 31554641; 32641768	thermofisher (PA5-59963)	used in IHC	50	Tm169	Brain
CD3	immune	T cell		membrane		29768164	fluidigm (3170019D)	PMID: 36689332	75-200	Er170	Spleen, Tonsil
CD8	immune	T cell		membrane		29768164	biolegend (344727)	PMID: 28369679	200	Ho165	Spleen, Tonsil
DNA1	DNA intercalator			nucleus			fluidigm (201192B)	Preconugated to 191Ir		Ir191	
DNA2	DNA intercalator			nucleus			fluidigm (201192B)	Preconugated to 193Ir		Ir193	
CD45	immune			membrane		12414720	fluidigm (91H029152)	Preconugated to 152Sm	300	Sm152	Spleen, Tonsil
CD31	vasculature			membrane		27055047	fluidigm (3151025D)	Preconugated to 151Eu		Eu151	Skin, Tonsil, Prostate
SMA	vasculature			cytoplasm		19929197	rndsystems (mab1420)	used in Cytof	8-25	Dy164	Prostate, Skin, Tonsil
EZH2			transcript respressive	nucleus		23720055	abcam (ab231165)	BSA and azide free	250	Nd145	Tonsil
HIF1A			hypoxia	cytoplasm	nucleus	11606368	thermofisher (700505)	PMID: 32868913	400	Dy161	Bone marrow
JARID2_C_Terminus			active	nucleus		30573669	developed in house			Nd143	Brain
JARID2_N_Terminus			repressed	nucleus		30573669	abcam (ab251123)	BSA free version, validated (by KO)		Yb173	Brain
Ki67			proliferating	nucleus		29322240	fluidigm (3168001B)	Preconugated to 168Er		Er168	Skin, Tonsil
SNAI1			EMT	nucleus	cytoplasm	33806868	rndsystems (af3639)	BSA and azide free, used in IHC	5-15	Tb159	Ubiquitous
SOX2			proliferating	nucleus		30952620	fluidigm (3150019B)	Preconugated to 150Nd		Nd150	Brain, Tonsil
TGFBeta			quiescent	extracellular		30952620	fluidigm (3163010B)	Pre-conjugated to 163Dy		Dy163	Spleen, Bone marrow, Prostate
TNC			quiescent	extracellular		30952620	rndsystems (mab2138)	used in IF	8-25	Gd156	Uterus

Table S4.4 | Gating criteria used to annotate cell types and cell states.

Cell Type	Protein Marker(s)	Expression Ranks
Endothelial Cell	SMA^+ $CD31^+$	≥ 15
NK cell	$CD45^+$ $NKP46^+$	≥ 14
T Cell	$CD45^+$ $CD3^+$	≥ 11
Macrophage	$CD45^+$ $IBA1^+$	≥ 15
Microglia	$CD45^+$ $IBA1^+$ $TMEM119^+$	≥ 12
Neuron	$NeuN^+$ $NKP46^+$ $CD3^+$	≥ 16
Astrocyte	$GFAP^+$	≥ 17
Oligodendrocyte	MOG^+	≥ 18
AC-like	$SLC1A3\_EAAT1^+$ $HOPX^+$	≥ 17
MES-like	$SOD2^+$ $CHI3L1^+$ $ANEXIN\_A2^+$ $ANXA1^+$	≥ 17
NPC-like	$DLL3^+$ $BCAN^+$	≥ 15
OPC-like	$SCD5^+$ $OLIG1^+$	≥ 15

Table S4.5 | ANOVA comparing intra-tumour (across patient ROIs) and inter-tumour (within patient and surgery sample) heterogeneity of cell categories.

The p value significance levels: ****p < 0.0001; ***p < 0.001; **p < 0.01; *p < 0.05; n.s (not significant).

Cell category	Source of variation	Effect size (F statistic)	p value	p significance
Immune	patient & surgery	5.82	7.62×10^{-4}	***
Cancer	patient & surgery	9.47	3.33×10^{-5}	****
Normal	patient & surgery	10.11	2.11×10^{-5}	****
Vasculature	patient & surgery	4.56	3.01×10^{-3}	**
Immune	ROI	0.32	7.29×10^{-1}	n.s
Cancer	ROI	2.63	9.97×10^{-2}	n.s
Normal	ROI	1.23	3.15×10^{-1}	n.s
Vasculature	ROI	1.75	2.02×10^{-1}	n.s

Table S4.6 | Comparison of cell category prevalence between primary and recurrent samples.

Statistical significance was assessed using an unpaired *Wilcoxon test*, with adjusted p-values calculated using the false discovery rate (FDR) method.

Cell category	Comparison groups	n (per comparison group)	p value	Adjusted p value
Immune	Prim vs Rec	15	9.75×10^{-2}	9.75×10^{-2}
Cancer	Prim vs Rec	15	6.13×10^{-2}	8.17×10^{-2}
Normal	Prim vs Rec	15	1.13×10^{-4}	4.52×10^{-4}
Vasculature	Prim vs Rec	15	4.94×10^{-3}	9.88×10^{-3}

Table S4.7 | Comparison of cell type prevalence between primary and recurrent samples.

Statistical significance was assessed using an unpaired *Wilcoxon test*, with adjusted p-values calculated using the FDR method.

Cell tyoe	Comparison groups	n (per comparison group)	p value	Adjusted p value
T cell	Prim vs Rec	15	5.12×10^{-1}	6.83×10^{-1}
NK cell	Prim vs Rec	15	2.02×10^{-1}	3.72×10^{-1}
Macrophage	Prim vs Rec	15	2.17×10^{-1}	3.72×10^{-1}
Microglia	Prim vs Rec	15	2.50×10^{-1}	3.75×10^{-1}
AC-like	Prim vs Rec	15	6.53×10^{-2}	1.57×10^{-1}
MES-like	Prim vs Rec	15	9.17×10^{-1}	9.17×10^{-1}
NPC-like	Prim vs Rec	15	8.33×10^{-1}	9.09×10^{-1}
OPC-like	Prim vs Rec	15	8.14×10^{-1}	9.09×10^{-1}
Neuron	Prim vs Rec	15	4.17×10^{-3}	1.98×10^{-2}
Astrocyte	Prim vs Rec	15	3.02×10^{-3}	1.98×10^{-2}
Oligodendrocyte	Prim vs Rec	15	2.25×10^{-2}	6.75×10^{-2}
Endothelial	Prim vs Rec	15	4.94×10^{-3}	1.98×10^{-2}

Table S4.8 | Comparison of Shannon's entropy (H) between primary and recurrent samples, quantifying intra-tumour cellular heterogeneity.

Patient(s)	Comparison group	n (per comparison group)	p value	Adjusted p value
All	Prim vs Rec	150	3.93×10^{-3}	3.93×10^{-3}
64	Prim vs Rec	30	1.43×10^{-4}	3.58×10^{-4}
67	Prim vs Rec	30	2.63×10^{-2}	2.63×10^{-2}
71	Prim vs Rec	30	1.69×10^{-17}	8.45×10^{-17}
82	Prim vs Rec	30	1.52×10^{-3}	1.90×10^{-3}
84	Prim vs Rec	30	4.26×10^{-4}	7.10×10^{-4}

Table S4.9 | Changes in GBM cancer cell types across hypoxia and EMT cellular states in primary and recurrent samples.

Statistical significance was assessed using the unpaired, *Wilcoxon test*, with adjusted p-values calculated using the FDR method. The p value significance levels are denoted using the following symbols: ****p < 0.0001); ***p < 0.001); **p <0.01); **p < 0.05); n.s (not significant).

Cell type	Cellular State	Comparison	n cells (primary)	n cells (recurrent)	p value	Adjusted p value	p significance
AC-like	hypoxia	Prim vs Rec	5522	686	4.98×10^{-115}	9.96×10^{-115}	****
MES-like	hypoxia	Prim vs Rec	9141	3119	9.62×10^{-125}	3.85×10^{-124}	****
NPC-like	hypoxia	Prim vs Rec	1664	613	5.49×10^{-70}	7.32×10^{-70}	****
OPC-like	hypoxia	Prim vs Rec	1743	616	7.41×10^{-1}	7.41×10^{-1}	n.s
AC-like	EMT	Prim vs Rec	5522	686	5.29×10^{-161}	1.06×10^{-160}	****
MES-like	EMT	Prim vs Rec	9141	3119	1.03×10^{-7}	1.37×10^{-7}	****
NPC-like	EMT	Prim vs Rec	1664	613	2.32×10^{-183}	9.28×10^{-183}	****
OPC-like	EMT	Prim vs Rec	1743	616	4.40×10^{-3}	4.40×10^{-3}	**

Table S4.10 | Summarised pair-wise cell-cell interactions compared to a null model of spatial randomness.

Statistical significance and direction of the interactions were determined using a *permutation test*, with p-values indicating interactions more or less likely than random. The *Signif* field denotes the aggregated strength and direction of the pair-wise interactions for a given *from* and *to* cell pair: ≥ 1 (significantly interacting); ≤ -1 (significantly avoiding interactions); 0 (neutral and/or non-statistically significantly associated). Where interaction *types* are significant across both primary and recurrent surgeries, the surgery-specific p values are listed, otherwise the respective, surgery-specific significant p values are populated in either the *Prim p value* or *Rec p value* fields.

From	To	Type	Across	Signif	Prim p value	Rec p value
AC-like	AC-like	Interacting	Primary	3	9.99×10^{-4}	
AC-like	Astrocyte			0		
AC-like	Endothelial			2		
AC-like	MES-like			-1		
AC-like	Macrophage			1		
AC-like	Microglia			0		
AC-like	NK cell			0		
AC-like	NPC-like			1		
AC-like	Neuron			0		
AC-like	OPC-like			2		
AC-like	Oligodendrocyte			1		
AC-like	T cell			1		
Astrocyte	AC-like			2		
Astrocyte	Astrocyte	Interacting	Both	3	9.99×10^{-4}	9.99×10^{-4}
Astrocyte	Endothelial	Interacting	Recurrent	2		9.99×10^{-4}
Astrocyte	MES-like	Interacting	Recurrent	2		9.99×10^{-4}
Astrocyte	Macrophage	Interacting	Recurrent	1		2.00×10^{-3}
Astrocyte	Microglia			2		
Astrocyte	NK cell	Interacting	Recurrent	0		9.99×10^{-4}
Astrocyte	NPC-like	Interacting	Recurrent	1		9.99×10^{-4}
Astrocyte	Neuron			2		
Astrocyte	OPC-like	Interacting	Recurrent	2		9.99×10^{-4}
Astrocyte	Oligodendrocyte	Interacting	Recurrent	0		9.99×10^{-4}
Astrocyte	T cell			1		
Endothelial	AC-like			2		
Endothelial	Astrocyte	Interacting	Primary	3	9.99×10^{-4}	
Endothelial	Endothelial	Interacting	Both	5	9.99×10^{-4}	9.99×10^{-4}
Endothelial	MES-like	Interacting	Primary	4	9.99×10^{-4}	
Endothelial	Macrophage	Interacting	Recurrent	0		9.99×10^{-4}
Endothelial	Microglia	Interacting	Primary	3	9.99×10^{-4}	
Endothelial	NK cell	Interacting	Both	4	9.99×10^{-4}	9.99×10^{-4}
Endothelial	NPC-like			0		
Endothelial	Neuron			1		
Endothelial	OPC-like			2		
Endothelial	Oligodendrocyte			1		
Endothelial	T cell			2		
MES-like	AC-like			-1		
MES-like	Astrocyte			2		
MES-like	Endothelial	Interacting	Both	3	9.99×10^{-4}	9.99×10^{-4}
MES-like	MES-like	Interacting	Both	5	9.99×10^{-4}	9.99×10^{-4}
MES-like	Macrophage	Interacting	Recurrent	1		9.99×10^{-4}
MES-like	Microglia			2		
MES-like	NK cell	Interacting	Both	3	9.99×10^{-4}	9.99×10^{-4}
MES-like	NPC-like			-2		
MES-like	Neuron			0		
MES-like	OPC-like			-2		
MES-like	Oligodendrocyte	Interacting	Recurrent	1		9.99×10^{-4}
MES-like	T cell	Interacting	Primary	4	9.99×10^{-4}	
Macrophage	AC-like	Interacting	Primary	3	9.99×10^{-4}	
Macrophage	Astrocyte	Interacting	Primary	4	9.99×10^{-4}	
Macrophage	Endothelial	Interacting	Both	5	9.99×10^{-4}	9.99×10^{-4}
Macrophage	MES-like	Interacting	Both	3	9.99×10^{-4}	9.99×10^{-4}
Macrophage	Macrophage	Interacting	Both	3	9.99×10^{-4}	9.99×10^{-4}
Macrophage	Microglia	Interacting	Both	5	9.99×10^{-4}	9.99×10^{-4}
Macrophage	NK cell	Interacting	Both	5	9.99×10^{-4}	9.99×10^{-4}
Macrophage	NPC-like			2		
Macrophage	Neuron	Interacting	Primary	3	9.99×10^{-4}	
Macrophage	OPC-like			1		

From	To	Type	Across	Signif	Prim p value	Rec p value
Macrophage	Oligodendrocyte	Interacting	Recurrent	2		9.99×10^{-4}
Macrophage	T cell	Interacting	Recurrent	2		9.99×10^{-4}
Microglia	AC-like			2		
Microglia	Astrocyte	Interacting	Primary	4	9.99×10^{-4}	
Microglia	Endothelial	Interacting	Both	5	9.99×10^{-4}	9.99×10^{-4}
Microglia	MES-like	Interacting	Both	4	9.99×10^{-4}	9.99×10^{-4}
Microglia	Macrophage	Interacting	Both	3	9.99×10^{-4}	9.99×10^{-4}
Microglia	Microglia	Interacting	Both	4	9.99×10^{-4}	9.99×10^{-4}
Microglia	NK cell	Interacting	Recurrent	2		9.99×10^{-4}
Microglia	NPC-like	Interacting	Recurrent	1		9.99×10^{-4}
Microglia	Neuron			0		
Microglia	OPC-like			2		
Microglia	Oligodendrocyte	Interacting	Recurrent	2		9.99×10^{-4}
Microglia	T cell	Interacting	Both	4	9.99×10^{-4}	9.99×10^{-4}
NK cell	AC-like			2		
NK cell	Astrocyte			0		
NK cell	Endothelial	Interacting	Both	5	9.99×10^{-4}	9.99×10^{-4}
NK cell	MES-like			1		
NK cell	Macrophage	Interacting	Recurrent	2		9.99×10^{-4}
NK cell	Microglia	Interacting	Both	4	9.99×10^{-4}	9.99×10^{-4}
NK cell	NK cell	Interacting	Both	3	9.99×10^{-4}	9.99×10^{-4}
NK cell	NPC-like			2		
NK cell	Neuron			1		
NK cell	OPC-like			2		
NK cell	Oligodendrocyte	Interacting	Recurrent	0		9.99×10^{-4}
NK cell	T cell	Interacting	Both	3	9.99×10^{-4}	9.99×10^{-4}
NPC-like	AC-like			1		
NPC-like	Astrocyte	Interacting	Recurrent	1		9.99×10^{-4}
NPC-like	Endothelial			2		
NPC-like	MES-like			0		
NPC-like	Macrophage			1		
NPC-like	Microglia			1		
NPC-like	NK cell			-1		
NPC-like	NPC-like	Interacting	Recurrent	1		9.99×10^{-4}
NPC-like	Neuron			2		
NPC-like	OPC-like			2		
NPC-like	Oligodendrocyte			1		
NPC-like	T cell			0		
Neuron	AC-like			2		
Neuron	Astrocyte	Interacting	Recurrent	2	9.99×10^{-4}	9.99×10^{-4}
Neuron	Endothelial	Interacting	Recurrent	1	9.99×10^{-4}	9.99×10^{-4}
Neuron	MES-like	Interacting	Recurrent	2	9.99×10^{-4}	9.99×10^{-4}
Neuron	Macrophage			2		
Neuron	Microglia			1		
Neuron	NK cell			-1		
Neuron	NPC-like			2		
Neuron	Neuron			2		
Neuron	OPC-like			2		
Neuron	Oligodendrocyte			1		
Neuron	T cell			1		
OPC-like	AC-like			1		
OPC-like	Astrocyte			1		
OPC-like	Endothelial			1		
OPC-like	MES-like			1		
OPC-like	Macrophage			1		
OPC-like	Microglia			1		
OPC-like	NK cell			-1		
OPC-like	NPC-like			1		
OPC-like	Neuron			2		
OPC-like	OPC-like			2		
OPC-like	Oligodendrocyte			1		
OPC-like	T cell			1		
Oligodendrocyte	AC-like			1		
Oligodendrocyte	Astrocyte			1		
Oligodendrocyte	Endothelial	Interacting	Recurrent	2		9.99×10^{-4}
Oligodendrocyte	MES-like	Interacting	Primary	3	9.99×10^{-4}	
Oligodendrocyte	Macrophage			2		
Oligodendrocyte	Microglia	Interacting	Recurrent	2		9.99×10^{-4}
Oligodendrocyte	NK cell			-1		
Oligodendrocyte	NPC-like			1		
Oligodendrocyte	Neuron			1		
Oligodendrocyte	OPC-like			1		
Oligodendrocyte	Oligodendrocyte	Interacting	Recurrent	2		9.99×10^{-4}
Oligodendrocyte	T cell	Interacting	Recurrent	2		9.99×10^{-4}
T cell	AC-like			2		
T cell	Astrocyte			1		
T cell	Endothelial	Interacting	Both	5	9.99×10^{-4}	9.99×10^{-4}
T cell	MES-like	Interacting	Both	5	9.99×10^{-4}	9.99×10^{-4}

From	To	Type	Across	Signif	Prim p value	Rec p value
T cell	Macrophage			0		
T cell	Microglia	Interacting	Both	4	9.99×10^{-4}	9.99×10^{-4}
T cell	NK cell	Interacting	Both	3	9.99×10^{-4}	9.99×10^{-4}
T cell	NPC-like			1		
T cell	Neuron			1		
T cell	OPC-like			2		
T cell	Oligodendrocyte	Interacting	Recurrent	2		9.99×10^{-4}
T cell	T cell	Interacting	Both	3	9.99×10^{-4}	9.99×10^{-4}

Table S4.11 | Summary of defined CNs.

The table details how each CN aligns with Greenwald et al.'s meta programs, and their relative proportions in primary and recurrent samples. Mac:macrophages; Inflammatory-Mac: inflammatory macrophages; MES-Ast:MES-like-astrocyte; Oligo:oligodendrocyte; Vasc:vasculature; AC:AC-like;Reactive-Ast:reactive astrocytes;Reactive-Ast/Neuron:reactive astrocytes/neurons.

CN	Greenwald layer	Greenwald label	n cells (Prim)	Prop (Prim)	n cells (Rec)	Prop (Rec)
CN1	3	Mac	6943	12.4%	26	0.1%
CN2	2	MES-Ast	2938	5.2%	6143	16.1%
CN3	5	Oligo	3454	6.2%	3009	7.9%
CN4	3	T-cell	6721	12.0%	4418	11.6%
CN5	3	Vasc	5171	9.2%	1560	4.1%
CN6	1	MES-Hyp	9186	16.4%	2451	6.4%
CN7	5	Neuron	1469	2.6%	5737	15.1%
CN8	4	AC	7961	14.2%	887	2.3%
CN9	5	Reactive-Ast/Neuron	765	1.4%	4931	12.9%
CN10	3	T-cell	3454	6.2%	543	1.4%
CN11	5	Reactive-Ast	2725	4.9%	7641	20.1%
CN12	2	Inflammatory-Mac	5179	9.3%	746	2.0%

Table S4.12 | Comparison of hypoxia and EMT protein marker abundance across surgeries and CNs.

The first column denotes the Greenwald et al meta programs and the corresponding CNs identified in this study. Statistical significance was assessed using an unpaired *Wilcoxon test*, with adjusted p-values calculated using the FDR method. Mac:macrophages; Inflammatory-Mac: inflammatory macrophages; MES-Ast:MES-like-astrocyte; Oligo:oligodendrocyte; Vasc:vasculature; AC:AC-like;Reactive-Ast:reactive astrocytes;Reactive-Ast/Neuron:reactive astrocytes/neurons.The p value significance thresholds are: $^{****}p < 0.0001$); $^{***}p < 0.001$); $^{**}p < 0.01$); $^{*}p < 0.05$); n.s (not significant).

Greenwald (CN)	State	Comparison	n cells (Prim)	n cells (Rec)	p value	Adjusted p value	Significance
MES-Hyp (CN6)	hypoxia	Prim vs Rec	6527	1407	3.70×10^{-20}	6.78×10^{-20}	****
MES-Ast (CN2)	hypoxia	Prim vs Rec	748	811	1.52×10^{-74}	1.67×10^{-73}	****
Inflammatory-Mac (CN12)	hypoxia	Prim vs Rec	777	71	8.52×10^{-1}	9.13×10^{-1}	n.s
Mac (CN1)	hypoxia	Prim vs Rec	262	2	6.77×10^{-2}	8.27×10^{-2}	n.s
Vasc (CN5)	hypoxia	Prim vs Rec	554	36	8.02×10^{-3}	1.10×10^{-2}	**
T-cell (CN4 & CN10)	hypoxia	Prim vs Rec	2406	621	1.40×10^{-30}	3.08×10^{-30}	****
AC (CN8)	hypoxia	Prim vs Rec	5258	633	6.80×10^{-66}	3.74×10^{-65}	****
Neuron (CN7)	hypoxia	Prim vs Rec	311	310	9.13×10^{-1}	9.13×10^{-1}	n.s
Reactive-Ast/Neuron (CN9)	hypoxia	Prim vs Rec	106	225	1.07×10^{-9}	1.68×10^{-9}	****
Reactive-Ast (CN11)	hypoxia	Prim vs Rec	248	478	1.36×10^{-41}	3.74×10^{-41}	****
Oligo (CN3)	hypoxia	Prim vs Rec	873	440	4.17×10^{-42}	1.53×10^{-41}	****
MES-Hyp (CN6)	EMT	Prim vs Rec	6527	1407	6.11×10^{-1}	6.12×10^{-1}	n.s
MES-Ast (CN2)	EMT	Prim vs Rec	748	811	2.69×10^{-27}	9.86×10^{-27}	****
Inflammatory-Mac (CN12)	EMT	Prim vs Rec	777	71	1.43×10^{-4}	2.25×10^{-4}	***
Mac (CN1)	EMT	Prim vs Rec	262	2	6.12×10^{-1}	6.12×10^{-1}	n.s
Vasc (CN5)	EMT	Prim vs Rec	554	36	4.34×10^{-7}	7.96×10^{-7}	****
T-cell (CN4 & CN10)	EMT	Prim vs Rec	2406	621	2.82×10^{-59}	1.55×10^{-58}	****
AC (CN8)	EMT	Prim vs Rec	5258	633	1.04×10^{-195}	1.14×10^{-194}	****
Neuron (CN7)	EMT	Prim vs Rec	311	310	1.14×10^{-1}	1.39×10^{-1}	n.s
Reactive-Ast/Neuron (CN9)	EMT	Prim vs Rec	106	225	6.90×10^{-4}	9.49×10^{-4}	***
Reactive-Ast (CN11)	EMT	Prim vs Rec	248	478	1.80×10^{-11}	3.96×10^{-11}	****
Oligo (CN3)	EMT	Prim vs Rec	873	440	3.60×10^{-22}	9.90×10^{-22}	****

Chapter 5

Discussion

Glioblastoma (GBM) is arguably the worst cancer diagnosis a person can receive: it is a deadly and incurable brain cancer with a median survival of just ∼15 months despite the standard-of-care treatment[1,2]. Central to this dire prognosis and resistance is the complex heterogeneity which GBM tumours display, comprising genetic, epigenetic, and microenvironmental factors[3–5]. Therefore, to effectively combat this disease, we require comprehensive longitudinal analyses to identify specific mechanisms and processes that better explain why and how GBM cells resist treatment and facilitate regrowth of recurrent tumours. However, conducting such analyses is significantly challenged by patient demographics, where the median age at diagnosis is ∼65 years old; the rapid disease progression; and the relatively low (2.66 in 100,000 people) incidence rate of GBM[6]. Collectively, all these factors limit the opportunities to collect initial and recurrent tumours samples that are vitally needed for the types of longitudinal analysis detailed above. To address this challenge, in our group we have assembled an extensive dataset of paired (primary and recurrent) GBM patient samples, collected both locally and through international collaborations[7–13].

Historically, GBM diagnosis, classification and wider-research has relied extensively on histological techniques, which although very useful, greatly restrict the scope of discovery due to limited sample sizes and arduous collection/processing procedures[14]. More recent advancements in genomic technologies, particularly RNA sequencing (RNA-seq), have transformed this landscape by enabling researchers to comprehensively profile tumours at the molecular level[15]. Such insights have facilitated the identification of distinct GBM molecular subtypes linked with survival outcomes[16], and more recently, GBM malignant cellular states which resemble neurodevelopmental hierarchies and wound-healing programs[5,12,13,17–19].

Crucially, all these studies found that neoplastic cell states are significantly influenced by their tumour microenvironment (TME), comprising immune, normal brain, and vascular cells, with interactions varying distinctly across tumour regions.

Given these recent developments, the overarching aim of this work was to decipher, at scale, how the cellular landscapes of GBM tumours - not just malignant cells, but also immune and other TME cell types - change throughout treatment. Our extensive paired GBM data were profiled using bulk RNA-seq rather than single-cell techniques, due to practical and financial constraints of profiling such large numbers of patient samples. Therefore, to analyse these data, first required the establishment of robust cellular deconvolution methodology. To this end, I developed GBM-specific immune cell signatures and rigorously evaluated them against multiple computational cell deconvolution tools to identify the optimal methods for wider community use, detailed in chapter 2. I then utilised this validated approach, to characterise the cellular changes within our large, paired longitudinal GBM dataset, detailed in chapter 3. Finally, to better understand GBM TME dynamics, I integrated a novel spatial proteomics technique, imaging mass cytometry (IMC), to examine how spatial arrangements of various TME cell types and states evolve through treatment, detailed in chapter 4.

The following sections summarise the key findings of each paper, discusses limitations, and indicates directions for future work.

5.1 Chapter 2 - GBMDeconvoluteR

5.1.1 Summary

In this study, I integrated the data from several publicly available, GBM-specific single-cell RNA sequencing (scRNA-seq) datasets to derive a set of marker genes for immune cell populations: B cells, T cells, natural killer cells (NK cells), microglia, tumour-associated macrophages (TAMs), monocytes, mast cells and dendritic cells (DCs). These were further supplemented with markers corresponding to malignant GBM neoplastic cell states as previously defined by Neftel *et al.*[5]: astrocyte-like (AC-like), neural progenitor-like (NPC-like), oligodendrocyte progenitor-like (OPC-like) and mesenchymal-like (MES-like). These markers were then applied to two prominent computational cell deconvolution methods that were shown to perform well in a comprehensive benchmarking study[20]: the semi-supervised, reference-based MCPcounter[21], and the supervised reference-based CIBERSORTx[22]. The deconvolution cell proportions were experimentally validated using a single-cell resolution, gold standard, based on IMC. This validation demonstrated that MCPcounter, when combined with GBM-specific markers (MCP_{GBM}), provided superior accuracy in quantifying both immune and neoplastic cell populations compared to CIBERSORTx. Further, it highlighted the importance of using tissue- and disease-specific references for optimal performance of reference-based tools.

To facilitate wider community accessibility, the results were packaged into GBMdeconvoluteR, a user-friendly web application[23]. This tool also incorporated an additional single-cell derived marker panel for GBM-specific cell types including vasculature and normal brain cell type markers[24].

Finally, the utility of GBMDeconvoluteR was demonstrated by analysing data from the cancer genome atlas (TCGA)[16], which confirmed previous findings that mesenchymal GBM cells are strongly associated with immune-rich tumour regions, particularly in patients with aggressive GBM tumours and poorer survival outcomes[25].

5.1.2 Limitations

A key limitation of this work is the restricted inclusion of specific cell types such as activated M2 macrophages and reactive astrocytes, largely due to the inability to reliably define such finer-grain cell sub-populations. This stemmed from the

limited availability of single-cell GBM datasets that met our inclusion criteria at the time the tool was developed.

The IMC validation method was performed on relatively small regions ($3mm^2$ in total for a patient) of tissue, whereas bulk RNA-seq was derived from larger areas. Although we still found strong correlations between this and the computational methods, the possibility of sampling bias due to this mismatch in spatial scale remains a concern, which could have introduced false positives. The difference in modalities (protein vs ribonucleic acid (RNA)) is another such limitation of the chosen validation method used in this study, as transcript-level expression values are more stochastic compared with protein expression and may not always correlate.

Although the marker-panel was specifically designed for isocitrate dehydrogenase wild-type (IDHwt) GBM tumour samples in mind, these markers are not necessarily accurately and/or applicable for other adult-type diffuse gliomas such as isocitrate dehydrogenase mutated (IDHmut) GBM, which does limit broader generalisability.

Finally, GBMdeconvoluteR is based on a semi-supervised, marker-based tool (MCP-Counter) that returns relative cell type proportions. Therefore, comparing abundances across different cell types within the same sample are not directly possible, which restricts analytical depth if users wish to compare proportions between two different cell types.

5.1.3 Perspectives & future work

As part of this work, we sought to identify the most suitable cell deconvolution method that would work well for our specific GBM patient data and subsequently developed GBMdeconvoluteR using this approach. Moreover, cell type deconvolution benchmarking studies have consistently suggested that this is the correct approach, and that the choice of the "best" tool is highly context-dependent[20,26,27]. Given this, one way in which this work could be expanded is by supporting multiple deconvolution tools, allowing users to compare outputs and select the most appropriate method for their specific needs. Similarly, the cell type markers offered in the tool could be further enhanced to include finer immune cell and brain stromal cell sub-populations. A potential approach to facilitate this would be to enable users to upload custom marker panels, rather than selecting one of our pre-built ones. This would definitely improve flexibility, though would have to be balanced against suitability/interpretability of the users supplying poorly defined

markers. We could also expand the existing marker set to include key markers from other adult-type diffuse gliomas, enhancing GBMdeconvoluteR's applicability to more diverse types of GBM samples, whilst maintaining a degree of control over marker validity.

5.2 Chapter 3 - GBM cell changes pre- & post-treatment

5.2.1 Summary

In this study, I applied GBMDeconvoluteR to our longitudinal, paired GBM dataset comprising 219 patient tumour samples. Taking the inferred, relative cell type proportions, I characterised the cellular landscape changes within GBM through treatment by comparing samples, pair-wise across both time-points. The dataset in this study was divided into two distinct cohorts (discovery and validation) to evaluate the consistency of the longitudinal changes. I found notable changes across vasculature and normal brain cell populations, with particularly prominent increases observed in plasma B cells and oligodendrocytes through treatment.

Building on our group's recent publication, which demonstrated that GBM patients can be stratified into Up and Down responder subtypes associated with distinct TME-related treatment resistance mechanisms[7], I also applied this classification to our paired GBM datasets. This confirmed our earlier observation that Up responders exhibit features consistent with a more proneural (PN) phenotype, while Down responders are enriched for mesenchymal (MES) characteristics. Importantly, this study extends those findings by identifying specific cell–cell associations linked to responder status e.g., plasma B cell-oligodendrocyte interactions which were enriched in Up responders[7]. These associations were also further supported using a small validation cohort profiled with both spatial transcriptomics and spatial proteomics technologies from Greenwald *et al.*[28].

The associations of progression-free survival (PFS) and overall survival (OS) survival outcome on specific cell type changes through treatment revealed significant cell type changes, though the clearest patterns aligned with our responder subtype stratification.

Other notable changes include consistent correlations between mast cell populations and both myeloid and lymphoid immune cells. Moreover, associations between radial glial cells and neoplastic GBM cells - particularly in Up responders – suggests a potential treatment-evasion mechanism by which cells transition to a more quiescent state. Finally, predictive modelling of surgery type utilising only the cell type scores underscores that the inferred cell type scores are able to capture biologically meaningful variation, particularly in relation to treatment

response.

5.2.2 Limitations

Bulk RNA-seq inherently lacks spatial and single-cell resolution that is required to confidently assess cell interactions which is why we performed spatial transcriptomics to validate the oligodendrocyte and plasma B cell findings. These validations were performed based on only one and two patients, for the spatial transcriptomics and proteomics, respectively. While these findings were consistent with the associations observed between plasma B and oligodendrocytes, in silico, the small sample size poses a risk of sampling bias. This limitation is particularly relevant given the high level of intra-tumour heterogeneity that already exists within GBM tumours. Similarly, in this study we had to exclude a large number of patient samples when stratifying patients into responder subtypes. This was done because these samples comprised of low tumour purity (<30% across the tumour pair) which we previously showed, affected the malignant GBM cell-driven, responder subtype classification[7].

Another limitation in this study is the age distribution of the patients, which is skewed toward younger patients, when considering that the median age of diagnosis for GBM is ∼65 years old[6]. This likely reflects a greater willingness among younger patients to undergo repeat surgery, but nonetheless, does still introduce potential bias. Similarly, incomplete clinical metadata, particularly for overall patient survival does restrict the statistical power to detect nuanced clinical correlations in this study.

Extending from the GBMDeconvoluteR study, the specificity of marker genes, especially for distinguishing closely related cell populations or activation states, such as M2 macrophages or reactive astrocytes, remains a limitation. While we do still find meaningful associations between cell types, the lack of markers to delineate finer-grain cell populations does restrict the biological resolution and interpretability of the significant findings from this study.

5.2.3 Perspectives & future work

As intimated in the above limitations of this study, the key next step for this work is the comprehensive experimental validation of the major findings, particularly in the role and associations between plasma B cells and oligodendrocytes. In our wider-group we have already started this work by investigating specific antibodies

such as Immunoglobulin G (IgG) and Immunoglobulin E (IgE), which are thought to influence oligodendrocyte myelination. However, at the time of writing, this additional work is still ongoing.

Another aspect of validation this study would greatly benefit from is the integration of additional spatial transcriptomics and proteomics which would strengthen the validation of the inferred cell-type proportions and interactions. Work for this is also already in progress within our broader lab group where we are using the 10X Visium spatial transcriptomics (∼6000 transcripts) platform. This work contributed to the preliminary plasma B cell–oligodendrocyte validation detailed in chapter 3, though by utilising a significantly larger marker panel we should be able to identify more cell types/states and greatly enhance our understanding of the GBM tumour microenvironment (TME).

Further experimental work is also needed to explore other key findings from this study, namely the roles of mast cells and radial glial cells. Specifically, future work would look into the potential immunomodulatory functions of mast cells and the capacity of radial glial cells to support immune evasion or induce a quiescent state in malignant GBM cells, especially in response to temozolomide (TMZ).

5.3 Chapter 4 - Spatially profiling the GBM TME through treatment

5.3.1 Summary

In this paper I applied a novel spatial proteomics technique, IMC to a subset (*n=5*) of the longitudinal paired GBM patient samples used in chapter 3. The aim was to understand how the spatial-organisation of cells within the GBM TME change through treatment. For each patient-specific surgical time-point, three small ($1mm^2$) regions of interests (ROIs) were sampled from different anatomical tumour regions. Following cell-segmentation and phenotyping I assessed broad "cell category" and more-detailed "cell type" changes through treatment, finding there was a significant increase in normal brain cells and a notable reduction in vascular cells. Although the overall cell diversity remained stable through treatment, there was greater ad-mixture of cell types, particularly oligodendrocytes, astrocytes and vasculature cells, suggestive of TME re-organisation. This study builds on and extends the findings of Greenwald *et al.*, who recently used a spot-based, spatial transcriptomics method to show that GBM TME is organised in to layers driven by hypoxia[28]. Using a less stochastic, proteomics-based approach and a paired-sample design, I confirm the finding that hypoxia drives GBM TME organisation. Moreover, I show that this effect is diminished following treatment, giving way to a less structured GBM TME which is shaped more by reactive astrocytes and infiltrating lymphocytes.

5.3.2 Limitations

A clear limitation of this study is the small sample size (*n=5*), which coupled with the significant inter-patient heterogeneity observed in the study, introduces the risk of sampling bias and also severely limits the generalisability of the findings. Similarly, the selection of three small regions per surgical samples may not adequately capture the full spatial heterogeneity that we know is present across GBM tumours.

There are also several technological limitations related to this study. Firstly, cell types were characterised based on a small number of markers (*n=20*) due to the technology being limited to a total of ~40 markers. Although each marker was carefully chosen and based on solid justification of prior work, this limitation

meant that I was not able to delineate all immune cell types. An example of this are B cell populations which were found to be important in chapter 3. Additionally, the restriction to a smaller number of markers also made it difficult to conclusively discern between closely-related cell types such as neoplastic and normal brain cell types and NK cell and T cells. Another key technical limitation of this study relates to the process of cell-segmentation, where I used marker-based pixel intensities to define cell-objects. I was able to accurately capture the nuclear features using the two deoxyribonucleic acid (DNA)-intercalating markers, however, the absence of dedicated cytoplasmic or cell membrane markers meant that cell boundaries had to be approximated using the localisation patterns of available markers. The inclusion of a broad-spectrum cytoskeleton or surface marker would have yielded more accurate cell morphologies and improved segmentation fidelity.

The process for determining cellular neighbourhoods (CNs) and subsequently spatial contexts (SCs), inherently involves user-defined parameters. In this study I used Delaunay triangulation to capture cellular neighbourhoods based on cell interaction graphs as it better accounts for irregular spacing between cells by dividing the image into planes bounded by a convex hull. However, the definition of the interaction distances still remains critical and has the potential to introduce bias, based on the scale chosen. Thus, while our CNs offer valuable insights, they are subject to methodological limitations that should be considered when interpreting the results.

5.3.3 Perspectives & future work

At the time of writing, this study is under review and a recurring recommendation from the reviewers has been the need for orthogonal validation to confirm the reported findings. In response, we are currently in the process of utilising a spot-based spatial transcriptomics approach, 10X Visium, which allows us to profile ∼6000 transcripts, offering a significantly expanded marker-list to use for phenotyping. Therefore, this would allow us to define a more comprehensive marker panel that is more sensitive to immune cell subsets such as M1 and M2 macrophage polarization, and the ratio of $CD8^+$ and regulatory T cells (Tregs). In addition to spatial transcriptomics, this work could be further expanded by incorporating other targeted (guided by current finding) multimodal approaches such as metabolomics. The cell-cell interactions and signalling within the GBM TME are known to be mediated not only by cytokines but also a diverse range of metabolites. For instance, the role of histamine release by malignant GBM cells and immune cells such as mast cells and macrophages is well-documented[29–31].

Therefore, integrating the perspective would allow us to better understand the functional dynamics of the GBM TME.

Given the small sample size of this study, questions around the causality of findings is pertinent. While there are observed changes in cell populations and TME re-organisation following treatment, we need further computational and experimental validation to determine if these changes are directly caused by therapy or reflect natural GBM disease progression. Addressing this distinction is needed not only for understanding underlying biology but, also to translate these findings into clinical application. This may involve the development of therapeutic targets or new stratification tools to better personalise current and/or emerging therapies.

5.4 Conclusion

Overall, this work provides:

1. A validated, benchmarked cellular deconvolution tool – GBMdeconvoluteR, that comprises a custom GBM-specific markers for inferring cell proportions from bulk RNA-seq data.

2. A comprehensive analysis of pre- and post-treatment cellular composition changes present in GBM. This includes novel associations such as those reported between plasma B cells and oligodendrocytes, and which may hold therapeutic potential.

3. An expansion of work previously undertaken within our group showing that GBM patients can be stratified into two distinct responder types, that are linked to specific treatment-resistance mechanisms.

4. New insights into how spatial interactions and organisation within the GBM TME cells changes through treatment.

5. A valuable paired, longitudinal dataset of GBM patient samples that provides a vital resource to the community to further explore how GBM tumours change through stand of care treatment.

References

1. Price, M. *et al.* CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2017–2021. *Neuro-Oncology* **26**, vi1–vi85 (Oct. 2024). doi:10.1093/neuonc/noae145.
2. Xiao, D. *et al.* National Brain Tumour Registry of China (NBTRC) statistical report of primary brain tumours diagnosed in China in years 2019–2020. *The Lancet Regional Health: Western Pacific* **34**, 100715 (Feb. 2023). doi:10.1016/j.lanwpc.2023.100715.
3. Eyler, C. E. *et al.* Single-cell lineage analysis reveals genetic and epigenetic interplay in glioblastoma drug resistance. *Genome Biology* **21**, 174 (July 2020). doi:10.1186/s13059-020-02085-1.
4. Louis, D. N. *et al.* The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Neuro-Oncology* **23**, 1231 (June 2021). doi:10.1093/neuonc/noab106.
5. Neftel, C. *et al.* An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. *Cell* **178**, 835–849.e21 (Aug. 2019). doi:10.1016/j.cell.2019.06.024.
6. Ostrom, Q. T. *et al.* CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2016—2020. *Neuro-Oncology* **25**, iv1–iv99 (Oct. 2023). doi:10.1093/neuonc/noad149.
7. Tanner, G. *et al.* IDHwt glioblastomas can be stratified by their transcriptional response to standard treatment, with implications for targeted therapy. *Genome Biology* **25**, 45 (Feb. 2024). doi:10.1186/s13059-024-03172-3.
8. Körber, V. *et al.* Evolutionary Trajectories of IDHWT Glioblastomas Reveal a Common Path of Early Tumorigenesis Instigated Years ahead of Initial Diagnosis. *Cancer Cell* **35**, 692–704.e12 (Apr. 2019). doi:10.1016/j.ccell.2019.02.007.
9. Kim, J. *et al.* Spatiotemporal Evolution of the Primary Glioblastoma Genome. *Cancer Cell* **28**, 318–328 (Sept. 2015). doi:10.1016/j.ccell.2015.07.013.
10. Wang, J. *et al.* Clonal Evolution of Glioblastoma under Therapy. *Nature genetics* **48**, 768–776 (July 2016). doi:10.1038/ng.3590.

11. Kim, E. L. *et al.* Intratumoral Heterogeneity and Longitudinal Changes in Gene Expression Predict Differential Drug Sensitivity in Newly Diagnosed and Recurrent Glioblastoma. *Cancers* **12**, 520 (Feb. 2020). doi:10.3390/cancers12020520.

12. Varn, F. S. *et al.* Glioma progression is shaped by genetic evolution and microenvironment interactions. *Cell* **185**, 2184–2199.e16 (June 2022). doi:10.1016/j.cell.2022.04.038.

13. Hoogstrate, Y. *et al.* Transcriptome analysis reveals tumor microenvironment changes in glioblastoma. *Cancer Cell* **41**, 678–692.e7 (Apr. 2023). doi:10.1016/j.ccell.2023.02.019.

14. McCutcheon, I. E. & Preul, M. C. Historical Perspective on Surgery and Survival with Glioblastoma: How Far Have We Come?. *World Neurosurgery* **149**, 148–168 (May 2021). doi:10.1016/j.wneu.2021.02.047.

15. Tzec-Interián, J. A., González-Padilla, D. & Góngora-Castillo, E. B. Bioinformatics perspectives on transcriptomics: A comprehensive review of bulk and single-cell RNA sequencing analyses. *Quantitative Biology* **13**, e78 (2025). doi:10.1002/qub2.78.

16. McLendon, R. *et al.* Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455**, 1061–1068 (Oct. 2008). doi:10.1038/nature07385.

17. Wang, L. *et al.* The Phenotypes of Proliferating Glioblastoma Cells Reside on a Single Axis of Variation. *Cancer Discovery* **9**, 1708–1719 (Dec. 2019). doi:10.1158/2159-8290.CD-19-0329.

18. Couturier, C. P. *et al.* Single-cell RNA-seq reveals that glioblastoma recapitulates a normal neurodevelopmental hierarchy. *Nature Communications* **11**, 3406 (July 2020). doi:10.1038/s41467-020-17186-5.

19. Wang, Q. *et al.* Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. *Cancer Cell* **33**, 152 (Jan. 2018). doi:10.1016/j.ccell.2017.12.012.

20. Sturm, G. *et al.* Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. *Bioinformatics* **35**, i436–i445 (July 2019). doi:10.1093/bioinformatics/btz363.

21. Becht, E. *et al.* Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biology* **17**, 218 (Oct. 2016). doi:10.1186/s13059-016-1070-5.

22. Steen, C. B., Liu, C. L., Alizadeh, A. A. & Newman, A. M. Profiling cell type abundance and expression in bulk tissues with CIBERSORTx. *Methods in molecular biology (Clifton, N.J.)* **2117**, 135–157 (2020). doi:10.1007/978-1-0716-0301-7_7.

23. Ajaib, S. *et al.* GBMdeconvoluteR accurately infers proportions of neoplastic and immune cell populations from bulk glioblastoma transcriptomics data. *Neuro-Oncology* **25**, 1236–1248 (July 2023). doi:10.1093/neuonc/noad021.

24. Ruiz-Moreno, C. *et al. Harmonized single-cell landscape, intercellular crosstalk and tumor architecture of glioblastoma* preprint (Cancer Biology, Aug. 2022). doi:10.1101/2022.08.27.505439.

25. Ravi, V. M. *et al.* Spatially resolved multi-omics deciphers bidirectional tumor-host interdependence in glioblastoma. *Cancer Cell* **40**, 639–655.e13 (June 2022). doi:10.1016/j.ccell.2022.05.009.

26. Avila Cobos, F., Alquicira-Hernandez, J., Powell, J. E., Mestdagh, P. & De Preter, K. Benchmarking of cell type deconvolution pipelines for transcriptomics data. *Nature Communications* **11**, 5650 (Nov. 2020). doi:10.1038/s41467-020-19015-1.

27. Nguyen, H., Nguyen, H., Tran, D., Draghici, S. & Nguyen, T. Fourteen years of cellular deconvolution: methodology, applications, technical evaluation and outstanding challenges. *Nucleic Acids Research* **52**, 4761–4783 (May 2024). doi:10.1093/nar/gkae267.

28. Greenwald, A. C. *et al.* Integrative spatial analysis reveals a multi-layered organization of glioblastoma. *Cell* **187**, 2485–2501.e26 (May 2024). doi:10.1016/j.cell.2024.03.029.

29. Wang, X. *et al.* Mast cells: a double-edged sword in inflammation and fibrosis. *Frontiers in Cell and Developmental Biology* **12**, (Sept. 2024). doi:10.3389/fcell.2024.1466491.

30. Derakhshani, A. *et al.* Mast cells: A double-edged sword in cancer. *Immunology Letters* **209**, 28–35 (May 2019). doi:10.1016/j.imlet.2019.03.011.

31. Chen, J. *et al.* Glioblastoma stem cell-specific histamine secretion drives pro-angiogenic tumor microenvironment remodeling. *Cell Stem Cell* **29**, 1531–1546.e7 (Nov. 2022). doi:10.1016/j.stem.2022.09.009.