Article

Programs, origins and immunomodulatory functions of myeloid cells in glioma

https://doi.org/10.1038/s41586-025-08633-8

Received: 9 October 2023

Accepted: 10 January 2025

Published online: 26 February 2025

Open access

Check for updates

Tyler E. Miller[1,2,3,4,5,15,17], Chadi A. El Farran[1,2,4,5,17], Charles P. Couturier[1,2,6,7,8,9,16,17], Zeyu Chen[1,2,4], Joshua P. D'Antonio[1,4], Julia Verga[1], Martin A. Villanueva[2,6,7,9], L. Nicolas Gonzalez Castro[2,10], Yuzhou Evelyn Tong[2,6,7,9], Tariq Al Saadi[11], Andrew N. Chiocca[1], Yuanyuan Zhang[1,2], David S. Fischer[2], Dieter Henrik Heiland[12,13], Jennifer L. Guerriero[5,14], Kevin Petrecca[11], Mario L. Suva[2,3], Alex K. Shalek[2,6,7,9] & Bradley E. Bernstein[1,2,4,5] ✉

Gliomas are incurable malignancies notable for having an immunosuppressive microenvironment with abundant myeloid cells, the immunomodulatory phenotypes of which remain poorly defined[1]. Here we systematically investigate these phenotypes by integrating single-cell RNA sequencing, chromatin accessibility, spatial transcriptomics and glioma organoid explant systems. We discovered four immunomodulatory expression programs: microglial inflammatory and scavenger immunosuppressive programs, which are both unique to primary brain tumours, and systemic inflammatory and complement immunosuppressive programs, which are also expressed by non-brain tumours. The programs are not contingent on myeloid cell type, developmental origin or tumour mutational state, but instead are driven by microenvironmental cues, including tumour hypoxia, interleukin-1β, TGFβ and standard-of-care dexamethasone treatment. Their relative expression can predict immunotherapy response and overall survival. By associating the respective programs with mediating genomic elements, transcription factors and signalling pathways, we uncover strategies for manipulating myeloid-cell phenotypes. Our study provides a framework to understand immunomodulation by myeloid cells in glioma and a foundation for the development of more-effective immunotherapies.

Diffuse gliomas are the most common primary malignant brain tumours and are ultimately fatal[1,2]. Although immunotherapy has revolutionized the treatment of many cancers, gliomas represent a challenge for immunotherapy owing to the unique immune microenvironment of the brain, restricted access of systemic therapies, and the need to balance therapeutic immune responses with potentially fatal inflammation-induced oedema[3].

In many solid tumours, including glioma, myeloid cells can create an immunosuppressive microenvironment and are associated with poor prognosis[2–4]. In gliomas, myeloid cells are the most prevalent non-malignant cell type, comprising up to 50% of cells in a tumour[5]. Tumour-associated myeloid cells can influence the molecular state of malignant cells[6] as well as tumour-infiltrating T cells[6–8], which are the main effector cells of checkpoint blockade, vaccine and adoptive cell therapies. They can also recruit and suppress other myeloid cells[1]. However, the specific myeloid cell types and expression programs that orchestrate these functions remain to be determined.

Single-cell RNA sequencing (scRNA-seq) studies have documented a spectrum of malignant cell states in gliomas[9–13]. Previous studies have also investigated myeloid cells in human and mouse gliomas[5–7,11,14–18,19]. Even so, many questions remain. First, we lack consensus on the definition of myeloid-cell states and expression programs, or how they inform the clinical and biological features of gliomas. Myeloid cells have conventionally been classified and studied according to cell type: microglia, macrophage, monocyte, conventional dendritic cell (cDC) or neutrophil[5,7,11,15–17,19]. However, evaluating their functional activities independent of cell type has been difficult and is poorly addressed by standard single-cell computational tools. Second, the developmental origins of myeloid cells are typically inferred to be microglia derived or bone-marrow derived on the basis of marker genes identified

[1]Department of Cancer Biology, Dana Farber Cancer Institute, Boston, MA, USA. [2]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [3]Department of Pathology and Krantz Family Center for Cancer Research, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. [4]Departments of Cell Biology and Pathology, Harvard Medical School, Boston, MA, USA. [5]Ludwig Center at Harvard Medical School, Boston, MA, USA. [6]Institute for Medical Engineering and Sciences and Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA, USA. [7]Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA, USA. [8]Department of Neurosurgery, Brigham and Women's Hospital, Boston, MA, USA. [9]Ragon Institute of MGH, Massachusetts Institute of Technology and Harvard University, Cambridge, MA, USA. [10]Center for Neuro-Oncology, Dana-Farber Cancer Institute and Department of Neurology, Brigham and Women's Hospital, Boston, MA, USA. [11]Department of Neurology and Neurosurgery, Montreal Neurological Institute-Hospital, McGill University, Montreal, Quebec, Canada. [12]Microenvironment and Immunology Research Laboratory, Medical Center - University of Freiburg, Freiburg, Germany. [13]Department of Neurological Surgery, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. [14]Division of Breast Surgery, Department of Surgery, Brigham and Women's Hospital, Boston, MA, USA. [15]Present address: Department of Pathology, Case Western Reserve University School of Medicine and Case Comprehensive Cancer Center, Cleveland, OH, USA. [16]Present address: Brain Tumour Research Center, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada. [17]These authors contributed equally: Tyler E. Miller, Chadi A. El Farran, Charles P. Couturier. ✉e-mail: bradley_bernstein@dfci.harvard.edu

from lineage-tracing experiments in mice[20]. However, the origins of myeloid cells in human gliomas remain uncertain[21]. Third, we lack an understanding of how interactions between myeloid cells and other malignant and non-malignant cells create niches and immune microenvironments. A final impediment pertains to experimental modelling. Tumour-associated macrophages change state quickly in monolayer culture, and mouse models fail to completely recapitulate microglia biology and macrophage programs in human tumours[22].

Here we sought to overcome these limitations through a comprehensive study of myeloid cells in human gliomas. We leveraged scRNA-seq data for 85 diverse gliomas and emerging computational methods to identify 4 dominant immunomodulatory programs shared across myeloid cell types. We then integrated lineage tracing, spatial transcriptomics, chromatin accessibility and ex vivo models to discover the origins, niches and drivers of these programs. Our analyses portray a dynamic and interconnected myeloid compartment that becomes highly immunosuppressive in response to microenvironmental cues and therapies. They provide a foundation for advancing diagnostic and immunotherapeutic strategies for gliomas.

Consensus myeloid expression programs

We used scRNA-seq to characterize immune and non-immune cell types in freshly resected human adult diffuse gliomas. We included a wide array of tumours, spanning isocitrate dehydrogenase (IDH)-mutant and wild-type (WT) tumours, primary and recurrent, and tumours exposed to different therapies. We combined 44 prospectively collected tumour profiles with 41 consolidated from previous publications[7,13,23]. We divided these 85 profiles (Supplementary Table 1) into a discovery and a validation dataset, annotated cells based on marker-gene expression, and called copy number alterations to distinguish malignant cells (Fig. 1a, Extended Data Fig. 1a–c and Methods).

To discover the consensus gene-expression programs in myeloid cells, we used an unbiased method, consensus non-negative matrix factorization[24] (cNMF), to identify sets of genes, which we call programs, that were coordinately regulated across myeloid cells in each cohort of our discovery dataset. Hierarchical clustering of these programs identified recurrent expression programs found in all 3 cohorts, from which we derived 14 consensus gene programs. These consensus programs were found across our representative set of 85 gliomas and were recapitulated in our validation cohort (Fig. 1a,b, Extended Data Fig. 1 and Supplementary Table 2).

The consensus programs included cell-identity programs and cell-activity programs (Fig. 1c). The identity programs contain classical marker genes for myeloid cell types, including microglia, macrophages, monocytes, dendritic cells and neutrophils. To compare the tumour-associated monocyte program with peripheral monocytes, we performed scRNA-seq on myeloid cells from 17 matched blood samples, and identified 3 peripheral monocyte programs ($CD14^+$, $CD16^+$ and $CD163^+$). The tumour-associated monocyte program shared features with the $CD14^+$ and $CD163^+$ programs in peripheral monocytes but not with the $CD16^+$ program (Extended Data Fig. 2a). It also included genes involved in cell adhesion, migration, differentiation and initial inflammatory response (for example, *VCAN*, *FCN1*, *LYZ*, *CD44* and *CCR2*), which indicates that the monocytes are undergoing differentiation in the tumour tissue.

Four principal immunomodulatory programs

The cNMF cell-type programs were complemented by nine activity programs. Five of these correspond to known cellular programs composed of classical marker genes for responses to hypoxia, interferon, unfolded proteins or the cell cycle. The other four programs contain genes with immunomodulatory functions and could be split into two inflammatory and two immunosuppressive programs. One inflammatory program, which we named the 'systemic inflammatory' program, includes genes encoding cytokines and chemokines that have established roles in myeloid-cell recruitment and systemic inflammatory responses (*IL1B*, *IL1A*, *CCL2*, *TNF*, *OSM* and *CXCL8*). A 'microglial inflammatory' program includes genes involved in lymphocyte and monocyte recruitment (*CXCR4*, *CXCL12*, *CCL3*, *CCL4* and *CX3CR1*), stress responses (*RHOB*, *JUN*, *KLF2* and *EGR1*) and neural-interaction genes (*PDK4* and *P2RY13*). On the immunosuppressive side, the 'complement immunosuppressive' program includes genes that encode complement factors implicated in wound resolution and anti-inflammatory cytokine responses[25] and associated with immunosuppression in other tumours[26] (*C1QA*, *C1QB*, *C1QC* and *C3*), as well as other established immunosuppressive markers (*VSIG4* and *CD163*)[27]. Finally, the 'scavenger immunosuppressive' program includes genes encoding scavenger receptors (*MRC1* (also known as *CD206*), *MSR1* (also known as *CD204*), *CD163*, *LYVE1*, *COLEC12* and *STAB1*) and other potentially immunosuppressive genes (*NRP1*, *RNASE1* and *CTSB*). Many of these genes have been shown to suppress T cell function, including *CD163* (ref. 28) and *VSIG4* (ref. 27), which inhibit T cell proliferation, and *MSR1*, which suppresses STAT1 signalling[29].

Immunomodulatory programs are shared

The immunomodulatory programs were the most widely utilized cNMF programs, with 91% of glioma-associated myeloid cells expressing one of the four programs (Fig. 1d). All four programs were expressed across multiple cell types (Extended Data Fig. 2b–d and Supplementary Note 1), and the systemic inflammatory program was found in subsets of all myeloid-cell types. Each myeloid cell type can use more than one program. For example, macrophages can express any of the four immunomodulatory programs but are enriched for the two immunosuppressive programs. Microglia can express both inflammatory programs and the complement immunosuppressive program, but they rarely express the scavenger immunosuppressive program.

This widespread but differential immunomodulatory-program usage across the myeloid cell types was a prominent feature in our systematic cNMF analysis of the 183,062 myeloid cells from 85 tumours (Fig. 1c–f, Extended Data Fig. 2e–h and Supplementary Note 2). However, when we reanalysed these single-cell data using standard Louvain clustering and uniform manifold approximation and projection (UMAP) visualization, the immunomodulatory programs were not recapitulated. These conventional approaches, which treat cells as singular units and cluster them by similarity to other cells, organized the data primarily by cell type and failed to capture the activity programs (Extended Data Fig. 3).

Consistently, exhaustive comparisons with previously published literature confirmed that the cNMF immunomodulatory programs are distinct from reported myeloid cell states and gene sets, which were derived through whole-cell clustering methods. For example, a reported state for suppressive macrophage-1 (ref. 7) reflects a composite of the cNMF macrophage program with our two immunosuppressive programs, and a reported state for inflammatory microglia[7] aggregates the cNMF microglia with our two inflammatory programs[7] (Extended Data Fig. 3d,e). A simple overlap of gene sets from previous studies shows a similar pattern of composite program usage, with prior studies each deriving different gene sets based on how they partitioned closely related cells (Extended Data Fig. 4a–d).

These analyses indicate that our cNMF programs represent fundamental components of myeloid cells that cannot be distinguished by conventional clustering. The power of cNMF to isolate activity programs from cell type enabled us to take an activity-centric approach to our investigations of tumour-associated myeloid cells, and we focused our further investigations on the functional and physiological consequences of the four dominant immunomodulatory programs.

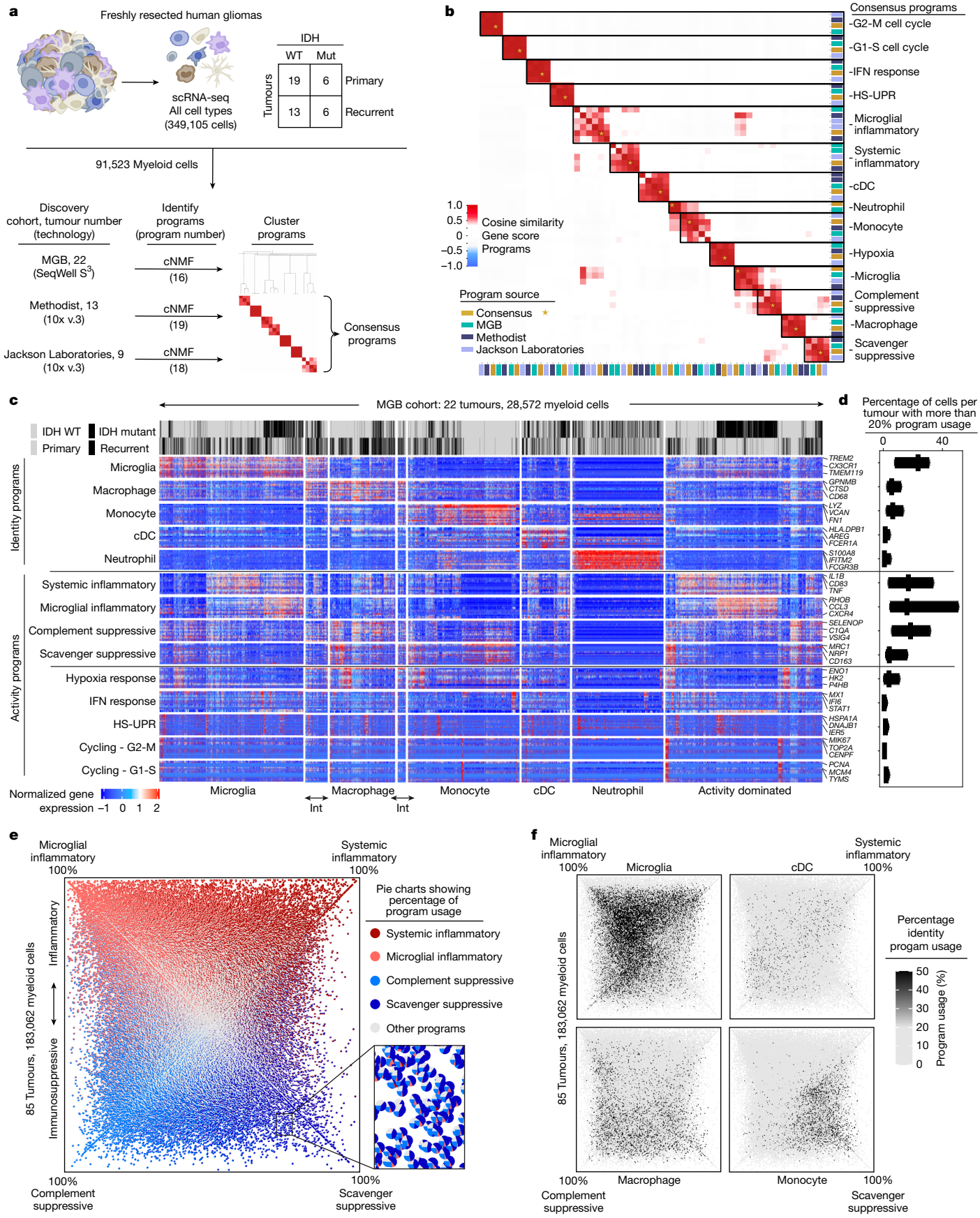


Fig. 1 | See next page for caption.

Fig. 1 | Identification of consensus superimposable myeloid-cell identity and activity programs. a, Analysis pipeline for identifying recurrent myeloid programs across three discovery glioma cohorts. Mut, mutant. **b**, Heatmap depicting similarity of the gene spectra scores of each program in the three discovery cohorts. Consensus programs created from their average scores are also included. HS-UPR, heat shock–unfolded protein response; IFN, interferon. **c**, Heatmap depicting expression of genes in recurrent myeloid programs (rows) in single cells (columns) grouped by cell type, as defined by myeloid-identity program usage. Intermediate (Int) cells express both adjacent identity programs. Right, selected marker genes for each program. **d**, Box plots depicting the percentage of cells per tumour that express the corresponding myeloid program (far left in **c**) across the three discovery cohorts. The line represents the median and boxes represent the first and third quartiles; $n = 44$ tumours, 91,523 myeloid cells. **e**, Quadrant plot positioning myeloid cells from the discovery and validation cohorts according to their relative expression of four prominent immunomodulatory programs. Diagonal axes correspond to the differential between microglial inflammatory and scavenger immunosuppressive program usage or between systemic inflammatory and complement immunosuppressive program usage. Programs at opposite ends of the diagonal axes are largely exclusive in individual cells. Each dot is a small pie chart depicting the prevalence of each immunomodulatory program (colours) and the combination of all other programs (grey) in that cell. **f**, Quadrant plots depicting myeloid cell identity usage for cells positioned as in **e**. Four prominent immunomodulatory activity programs are each used across multiple myeloid cell types.

Immunomodulatory states span CNS tumours

All 4 immunomodulatory programs were detectable in subsets of myeloid cells in all 85 gliomas, which included tumours that were IDH-mutant, IDH-WT, primary, recurrent, low-grade, high-grade, representative of varied mutations and treated with different therapies (Fig. 1c,d and Extended Data Figs. 1d and 2g,h). To examine the programs in other central nervous system (CNS) and non-CNS tumours, we collated published scRNA-seq datasets and calculated program usages in the myeloid cells (Extended Data Fig. 4e). In other primary CNS tumours, including paediatric gliomas and ependymomas, the full set of immunomodulatory programs were represented. However, examination of myeloid cells from non-CNS tumours yielded a different result. The systemic inflammatory and complement immunosuppressive programs were present in essentially all tumour types, but the microglial inflammatory and scavenger immunosuppressive programs were largely specific to CNS tumours. Notably, the myeloid compartments of brain metastases of non-CNS tumours recapitulated the originating tumours and were largely devoid of the microglial inflammatory and scavenger immunosuppressive programs. Finally, myeloid cells from non-neoplastic brains expressed three of the immunomodulatory programs, with only the scavenger immunosuppressive program being absent, further supporting its specificity to primary CNS tumours. These results indicate that gliomas have a unique myeloid environment with a prominent and specific immunosuppressive phenotype that is largely absent from non-CNS tumours and their brain metastases.

Origins and plasticity of myeloid states

We next considered the inter-relationships between developmental origin, myeloid cell type and immunomodulatory phenotypes. We began by inferring the lineage relationships and origins of individual myeloid cells from their mitochondrial DNA mutations. We used MAESTER[30] to call mitochondrial DNA mutations in tumour-associated myeloid cells and matched peripheral blood monocytes from four patients. We presumed that myeloid cells in the tumour that had peripheral blood variants were blood derived. By contrast, those with distinct variants were likely to represent resident microglia-derived myeloid cells in the brain (Extended Data Fig. 5a–c). Consistent with our expectations, we found that cells expressing a microglia program were most likely to harbour resident cell-specific variants, whereas other myeloid cell types were more likely to have peripheral blood variants. Intermediate cells that co-express microglia and macrophage programs were also enriched for peripheral variants. These findings indicate that bone-marrow-derived myeloid cells can activate a microglia-like phenotype in tumours. The immunomodulatory programs also manifested across the different cell identities and presumed origins, a further indication that myeloid cells maintain substantial epigenetic plasticity.

These results prompted us to test the peripheral blood myeloid cells from patients with glioma for expression of the activity programs. Of the four immunomodulatory programs, only the systemic inflammatory program was detected, strongly suggesting that myeloid activity states must shift on tumour infiltration (Extended Data Fig. 2a).

To evaluate the capacity of bone-marrow-derived peripheral cells to acquire glioma-associated myeloid phenotypes, we leveraged a glioma organoid (GBO) system[31]. GBOs were established from resected tumours and passaged to deplete all the immune cells. Matched peripheral blood mononuclear cells (PBMCs) were then applied to GBOs. After one week of co-culture, the GBOs were extensively infiltrated by myeloid cells (Extended Data Fig. 5d–h and Methods). Although many of the myeloid cells differentiated towards a macrophage phenotype, subsets of infiltrating myeloid cells upregulated canonical microglia markers (TMEM119 and P2RY12), confirming that peripheral monocytes can acquire features of tissue-resident microglia. By contrast, myeloid cells that remained outside the organoids were much less likely to express these markers. Immunohistochemistry confirmed that there was robust infiltration of immune cells, including myeloid cells expressing both microglia markers.

Together, these data indicate that blood-derived monocytes can activate all the glioma-associated myeloid programs, including the microglia program. They can also engage four distinct immunomodulatory programs, despite our observation that peripheral blood myeloid cells from patients with glioma can activate only one. They demonstrate that expression of immunomodulatory programs is not contingent on developmental origin or cell type, and highlight the plasticity of myeloid cells in the glioma microenvironment.

Myeloid states linked to tumour niches

To investigate potential microenvironmental drivers of the myeloid programs, we integrated our scRNA-seq data with Visium spatial transcriptomic data from glioblastomas[32]. We first identified a set of spatial gene programs that captured the transcriptomic heterogeneity across 68,830 50 μm spots in 23 sections (Extended Data Fig. 6a,b and Methods). This distinguished six prominent 'niche programs' that included grey- and white-matter structural niches, hypoxic and vascular metabolic niches, a niche composed of proliferative cancer cells, and an inflammatory niche composed of immune cells and reactive astrocytic genes. In parallel, we estimated the cellular content of each 50 μm spot by quantifying the expression of our cNMF programs[33]. These spatial analyses revealed clear niche-specific patterns of myeloid programs, cancer cell programs and other cell types in the tumours (Fig. 2a and Extended Data Fig. 6a,c).

To collate recurrent spatial relationships, we computed spatial correlations between cellular and niche programs across all Visium tumour sections. This revealed robust niche–niche, niche–cell and cell–cell associations (Fig. 2b,c and Extended Data Fig. 6d–f). The niche–niche relationships revealed a recurrent architecture in which a hypoxic niche is flanked by an inflammatory niche, which in turn is adjacent to a proliferative cancer niche that then runs into white matter, which is consistent with the clinical observation that most gliomas are present in white matter[34]. The hypoxic and inflammatory niches were also flanked by vascular niches, indicative of vascular proliferation in response

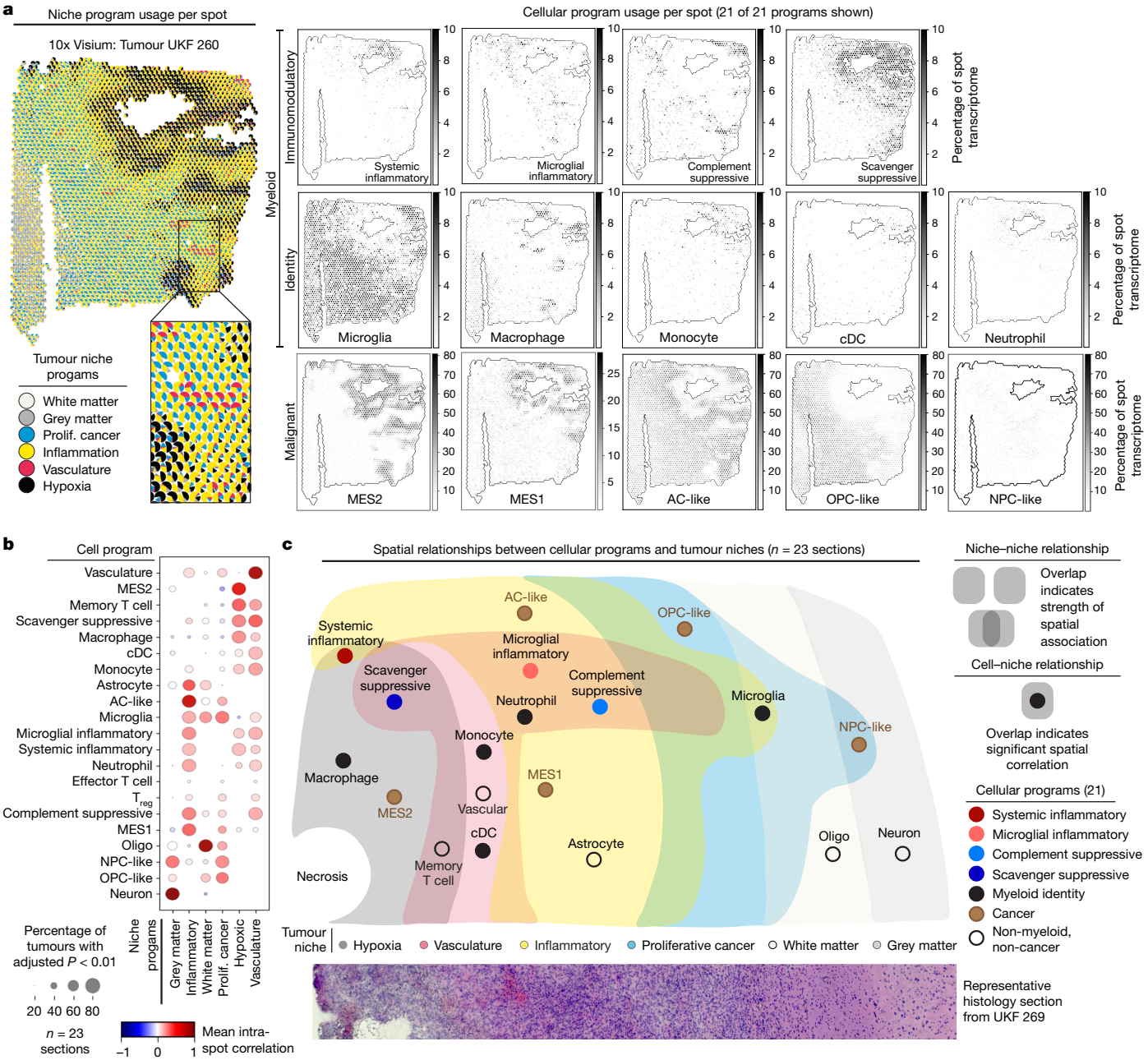


Fig. 2 | Immunomodulatory myeloid programs associated with spatial tumour niches. a, Scatter pie plot (left) of a representative 10× Visium section of a glioma specimen from ref. 32. Each pie chart depicts expression of the indicated tumour niche programs (colours) in a spot. The scatter plots (right) show the same section with spots shaded by expression of the indicated cNMF myeloid cell or malignant cell program. AC, astrocyte; MES, mesenchymal; NPC, neural progenitor cell; OPC, oligodendrocyte progenitor cell; prolif., proliferative. **b**, Dot plot depicting intraspot correlation between niche and cell program scores, calculated independently for each sample. The dot size shows the proportion of samples with a significant correlation (adjusted $P < 0.01$, two-sided, from Pearson correlation probability density function) and the colour indicates a positive (red) or negative (blue) correlation. **c**, Cell–niche map illustrating the conserved spatial relationships of myeloid, malignant and neural cell types (circles) and transcriptomic niches (shaded areas) across spatial samples. These spatial analyses show that the immunomodulatory myeloid programs are enriched in distinct niches in a conserved tumour architecture.

to hypoxia and potentially representing an entry point for immune infiltration. These patterns are largely consistent with recent reports[35].

Hypoxia was also organizing for the myeloid programs. The scavenger immunosuppressive program was enriched in vascular and hypoxic niches, whereas the complement immunosuppressive program was excluded from hypoxic niches and associated with surrounding inflammatory and vascular niches. The systemic inflammatory program was associated with hypoxic and inflammatory niches, and the microglial inflammatory program was enriched in inflammatory and vascular niches. Microglia were excluded from hypoxic niches but were found throughout the rest of the tumour field.

These spatial analyses indicated a particularly robust role for the scavenger immunosuppressive program in the glioma microenvironment. A spatial regression model of cell–cell relationships identified multiple interactions between the scavenger immunosuppressive program and nearly every cell program in the hypoxic and vascular niches, including the MES2, MES1 and monocyte programs. We validated these connections orthogonally using the scRNA-seq dataset, which revealed that average usage of these programs was highly correlated with the

scavenger immunosuppressive program across surgical samples (Extended Data Fig. 6g). Overall, spatial analyses highlight specific associations between immunomodulatory myeloid programs and tumour niches, and indicate that the scavenger immunosuppressive program may be a key determinant of the tumour microenvironment.

Clinical correlates of myeloid programs

We next investigated whether the myeloid cell identities and immunomodulatory programs correlate with clinical factors such as mutational status and therapies. Previous studies have reported increased inflammatory phenotypes in IDH-mutant tumours, a finding attributed to the oncometabolite 2-hydroxyglutarate[18,36]. We consistently found that IDH-mutant tumours have a distinct composition of our immunomodulatory programs, characterized by strong enrichment of the microglial inflammatory program and depletion of both immunosuppressive programs (Extended Data Fig. 7a–c). However, this distinction was most strongly associated with tumour grade, rather than IDH mutation, as the myeloid composition of grade 4 IDH-mutant tumours closely approximated that of grade 4 IDH-WT tumours (Fig. 3a and Extended Data Fig. 7d,e). Moreover, examination of a larger cohort revealed that the myeloid program composition of low-grade IDH-WT tumours (as per the 2016 WHO classification) mirrored that of low-grade IDH-mutant tumours (Fig. 3b and Supplementary Note 3). As a result, purported differences in the myeloid compartment of IDH-mutant and IDH-WT tumours are more likely to reflect the tumour grade.

We also investigated whether clinical therapies affect myeloid states. We first compared program composition in primary tumours versus recurrent tumours, all of which had previous tumour-targeting therapy, but we found no significant changes when controlling for grade. However, we did find that the complement immunosuppressive program is specifically and significantly associated with dexamethasone (Dex). Dex is a potent corticosteroid that is routinely administered to patients with glioma to reduce vasogenic oedema in the brain pre- and post-operatively. The complement immunosuppressive program increases with escalating dexamethasone dose and is significantly reduced in untreated tumours when controlling for hypoxia, which prevents the anti-inflammatory effects of glucocorticoids in myeloid cells[37] (Fig. 3c–e and Extended Data Fig. 8a).

The observation that glioma-associated myeloid cells and programs are derived from infiltrating monocytes indicates that Dex phenotypes may initially arise in peripheral monocytes. Indeed, when we examined scRNA-seq data for peripheral monocytes, we identified a potentially immunosuppressive program that was specifically increased in patients treated with Dex. This program includes *CD163* and some other markers found in the complement immunosuppressive program, raising the possibility that this is a precursor program that further develops in the tumour. In support of this, we observed a positive correlation between the Dex-related program in circulating monocytes and the complement immunosuppressive program in tumour-associated monocytes from the same patient (Extended Data Fig. 8b,c).

Myeloid cells and immunotherapy response

We also related our myeloid programs to immunotherapy response, tumour immune state and overall survival. A recent scRNA-seq study of 12 patients with glioma treated with neoadjuvant PD1 blockade identified a population of SIGLEC9-expressing macrophages that were associated with non-responsive tumours[38]. Reanalysis of these data using our cNMF framework revealed that SIGLEC9-positive cells were heterogeneous in their expression of our immunomodulatory programs. Notably, only the SIGLEC9-positive cells expressing the scavenger immunosuppressive program were enriched in non-responders. Moreover, the scavenger immunosuppressive program was more closely associated with non-responding gliomas than SIGLEC9 alone (Fig. 3f,g and Extended Data Fig. 8d–f), indicating that this immunosuppressive program may more accurately explain the refractory phenotype. We did not detect a significant association between immunotherapy response in this cohort and any of our other programs or previously published gene sets of myeloid cell state (Extended Data Fig. 9a–c).

To expand this association to a broader range of tumours, we examined regulatory T (T_{reg}) cells as a proxy for an immunosuppressive environment and poor immune responses[39]. Samples with high T_{reg} cell frequency were enriched for myeloid cells expressing the scavenger and complement immunosuppressive programs but depleted of microglial inflammatory-expressing cells (Fig. 3h). We also detected a spatial association between T_{reg} cells and the complement immunosuppressive program. By contrast, T cells with naive/memory expression signatures were spatially associated with hypoxic niches and the scavenger immunosuppressive program (Fig. 2b and Extended Data Fig. 6e,f). Hence, the respective myeloid programs are associated with distinct T cell enrichments and immune microenvironments.

Finally, we investigated whether our immunomodulatory programs were associated with overall survival. We used the top genes in each program to score IDH-WT glioblastomas in consortium cohorts and adjusted the results on the basis of estimated myeloid content (Methods). We found that the scavenger immunosuppressive program was significantly associated with worse overall survival. Such a significant association was not seen with any of our other myeloid activity programs, nor with previously published myeloid gene sets (Fig. 3i and Extended Data Fig. 9d).

These collective analyses implicate the glioma-specific scavenger immunosuppressive program as a critical determinant of the T cell immune microenvironment, response to checkpoint therapy and overall survival.

Regulators of immunomodulatory programs

We next sought to associate the immunomodulatory expression programs with gene regulatory elements and upstream transcription factors (TFs) (Fig. 4 and Extended Data Fig. 10). We extracted chromatin accessibility profiles for 36,675 individual myeloid cells across 20 human gliomas from a public compendium of single-nucleus assay for transposase-accessible chromatin with sequencing (snATAC-seq) data[40]. We calculated program usage for each myeloid cell on the basis of inferred gene activity scores (Methods). We then created pseudobulk accessibility profiles for each immunomodulatory program by aggregating cells with the highest program usage. Clustering on differentially accessible sites revealed putative program-specific promoters and enhancers (Fig. 4a,b and Extended Data Fig. 10a,b).

A scan of these accessible elements identified TF-binding motifs associated with specific immunomodulatory programs. We complemented these motif scans with single-cell regulatory network inference and clustering (SCENIC)[41], an orthogonal strategy for connecting expression programs to TF networks. These analyses revealed robust associations between the systemic inflammatory program and NF-κB-related TFs. Moreover, the scavenger immunosuppressive program elements were strongly enriched for AP-1 motifs, with order-of-magnitude higher significance scores than systemic inflammatory (Fig. 4b,c). These associations with putative regulatory elements and prominent TF families provided insight into the networks and regulators that drive immunosuppressive myeloid programs in glioma.

Feedback interplay among myeloid states

We also examined signalling pathways and environmental ligands with the potential to drive the immunomodulatory programs and associated TFs. Integrating our scRNA-seq data with the NicheNet database[42] revealed ligands expressed in our tumours that can activate these TFs. Top-ranked ligands associated with NF-κB signalling included interleukin-1β (IL-1β),

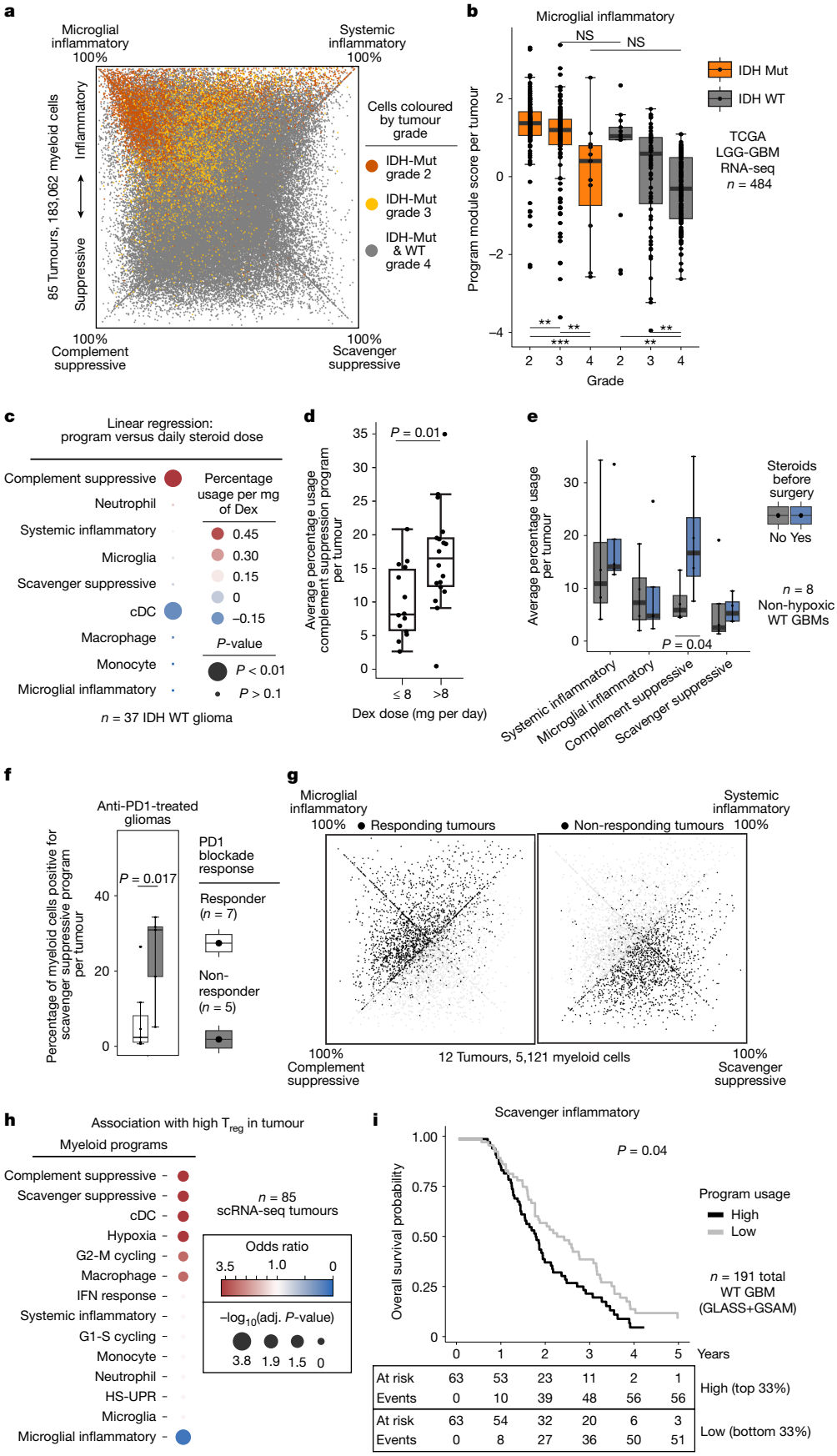


Fig. 3 | See next page for caption.

Fig. 3 | Clinical correlates of immunomodulatory myeloid programs. **a**, Quadrant plot with myeloid cells positioned as in Fig. 1e and coloured by the grade of the originating tumour. **b**, Box plot depicting the distribution of microglial inflammatory program module scores computed for 484 glioma datasets. For all the box plots in this figure, each dot represents a tumour, and lines represent median and quartiles. The whiskers represent 1.5 times the interquartile range. False discovery rate (FDR)-corrected *P*-values were calculated using two-sided Wilcoxon rank sum tests: $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$; NS, not significant. Not all comparisons are shown. GBM, glioblastoma; LGG, low-grade glioma. **c**, Dot plot depicting the linear regression coefficient between the average usage of each myeloid program per glioma sample and the corresponding patient's pre-surgery daily Dex dose. *P*-values are from ordinary least-squares regression model. **d**, Box plot showing average complement immunosuppressive program usage in glioma samples stratified by Dex dose; $n = 37$ IDH-WT glioma; *P*-values calculated using a two-sided Wilcoxon rank sum test. **e**, Box plot showing average usage of programs in glioma samples stratified by Dex use (hypoxic samples excluded). *P*-values were calculated using two-sided Wilcoxon rank sum tests. **f**, Box plot depicting the percentage of cells expressing the scavenger immunosuppressive program across glioma specimens from an anti-PD1 immunotherapy trial[38], stratified by response ($P = 0.017$, FDR when considering all measurements in Extended Data Fig. 8f = 0.088). **g**, Quadrant plots positioning the individual myeloid cells in **f** from radiographic response (left) and non-response (right) samples according to their expression of the immunomodulatory programs. **h**, Dot plot depicting the association between the indicated programs and T_{reg} cell abundance across 85 gliomas. Adj. *P*-value, adjusted *P*-value. **i**, Kaplan–Meyer curve of overall survival for IDH-WT glioblastomas stratified by scavenger immunosuppressive program expression in GLASS[54] and GSAM[55] cohorts. The *P*-value was calculated using a two-sided log-rank test.

TNF, CCL2, IL-1α and IFNγ. The AP-1 module was associated with IL-1β, TNF, PLAU, OSM, IL-10, IL-6 and TGFβ (Fig. 4d,e and Extended Data Fig. 10c). Notably, systemic inflammatory program myeloid cells expressed many ligands that were predicted to activate AP-1 signalling and may therefore potentiate the scavenger immunosuppressive program.

We used the GBOs to test directly whether ligands expressed by the systemic inflammatory program impact the scavenger immunosuppressive program. We applied peripheral monocytes to GBOs distributed across a 96-well plate and cultured them for 5 days in the presence of various ligands. We then collected the organoids and assessed the myeloid programs by flow cytometry and immunohistochemistry of key protein markers (Fig. 5a). IL-1β and TGFβ both induced scavenger immunosuppressive markers, with IL-1β inducing a more than four-fold increase over control. Conversely, IFNγ increased the proportions of cells expressing systemic inflammatory markers and reduced immunosuppressive markers (Fig. 5b).

These predictions and experimental validations identified specific ligands in the glioma microenvironment that had opposite effects on the systemic inflammatory and scavenger immunosuppressive programs. They also indicate that systemic inflammatory ligands may promote the scavenger immunosuppressive program, which is potentially indicative of feedback interplay that restrains inflammation in CNS tumours.

Dex induces durable immunosuppression

In contrast to the systemic and scavenger programs, our analyses did not identify specific TFs or ligands that were strongly associated with the complement immunosuppressive program. However, the close clinical association of the program with Dex treatment prompted us to investigate this connection using GBOs. We found that Dex specifically and significantly induced the complement immunosuppressive program, without affecting the scavenger immunosuppressive program. This effect was evident in fresh GBOs with endogenous myeloid cells and in passaged GBOs infiltrated with peripheral blood myeloid cells ex vivo (Fig. 5b,c and Extended Data Fig. 11a–e).

Dex affects transcriptional activity by driving nuclear translocation and DNA binding of the glucocorticoid receptor[43] (GR). We evaluated direct GR targets using public chromatin immunoprecipitation with sequencing (ChIP–seq) and RNA-seq data for cultured human macrophages treated with Dex for 1–24 h[43]. The RNA-seq data confirmed rapid upregulation of top complement program genes, many of which were direct GR binding targets using ChIP–seq (Fig. 4f–h). By contrast, genes that were downregulated by Dex lacked evidence of GR binding. Instead, they were associated with the NF-κB pathway and, by extension, with our systemic inflammatory program (Fig. 4f and Extended Data Fig. 10d,e). This is consistent with murine studies demonstrating GR-mediated suppression of NF-κB activity through the upregulation of NF-κB inhibitors[44]. Indeed, our analyses identified the NF-κB inhibitor TNFAIP3 as a direct GR target that is rapidly induced by Dex treatment.

Finally, we investigated whether this Dex-induced immunosuppressive phenotype is reversible. This is an important clinical question given that the vast majority of patients with glioma receive corticosteroids perioperatively and that immunotherapy trials are typically initiated after a short Dex cessation period. We applied Dex to myeloid-cell-infiltrated GBOs for two days and then washed out Dex. However, complement immunosuppressive program expression did not reverse even two weeks after drug withdrawal (Fig. 5c and Extended Data Fig. 11f). Although adding IFNγ to the GBOs partly reversed the Dex-induced state change, overall complement immunosuppressive program usage remained high. We also tested whether the effect of Dex on peripheral monocytes was sustained after they entered the tumour. We found that Dex rapidly induced complement markers in peripheral monocytes, and that these changes were sustained as the monocytes differentiated in the GBOs, even after Dex withdrawal (Extended Data Fig. 11g).

Altogether, these data indicate that Dex drives the complement immunosuppressive program through combined induction of hormone-responsive GR targets and suppression of the default NF-κB-associated systemic inflammatory program. The results highlight the plasticity of myeloid cell states during glioma infiltration, differentiation and therapy, but they also demonstrate the remarkable durability of the Dex-induced immunosuppressive state whose irreversibility has major clinical implications.

Targeting of AP-1 and the scavenger program by p300i

Given the prominent roles of the scavenger immunosuppressive program as a suppressor of inflammation and immune responses in glioma, we explored whether it could be reprogrammed to create a more inflammatory microenvironment. Using GBOs from freshly resected tumours, we tested chemical inhibitors of epigenetic regulators for their effects on program usage[45–47] (Extended Data Fig. 12a). We found that GNE-781, a selective p300/CBP bromodomain inhibitor (p300i)[48], had the most striking effect on the immunomodulatory phenotypes. Treatment with GNE-781 for seven days strongly reduced scavenger immunosuppressive program usage, with a compensatory increase in the systemic inflammatory program (Fig. 5d).

To investigate the mechanisms by which GNE-781 moderates the immunosuppressive program, we performed ATAC-seq on myeloid cells isolated from the treated organoids. Loci with reduced accessibility in GNE-781-treated samples were strongly enriched for AP-1 motifs (Fig. 5e). This is consistent with previous studies that implicated p300/CBP in the activation of AP-1 targets[49] and probably explains the striking reduction in the scavenger immunosuppressive program. Conversely, the set of sites whose accessibility increased following GNE-781 treatment were enriched for NF-κB-related motifs, consistent with a previous study that associated CBP/p300 with myeloid-derived suppressor cells[45] and with our observed shift in the myeloid expression phenotypes from immunosuppressive

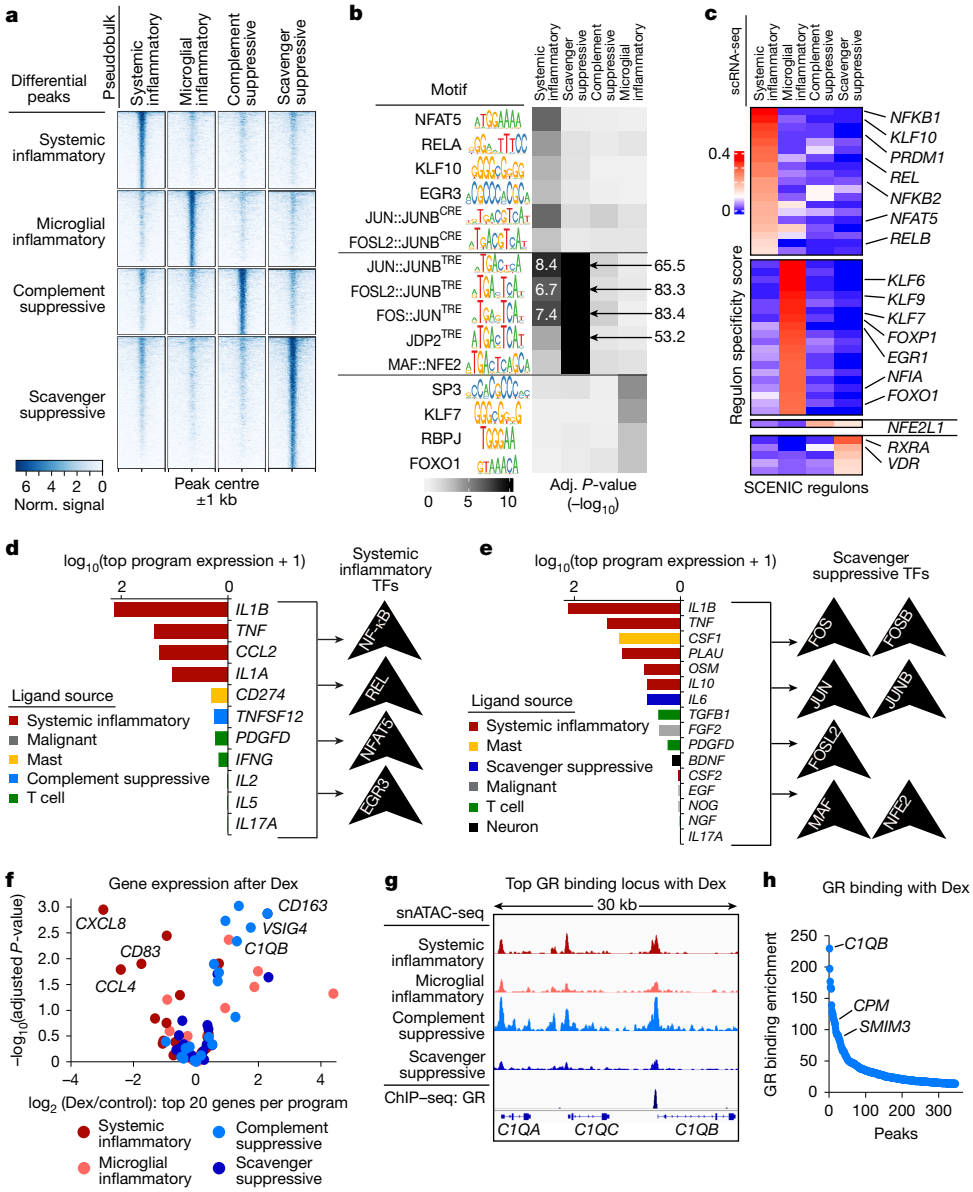


Fig. 4 | Regulatory circuits associated with immunomodulatory myeloid programs. a, Heatmap depicting the accessibility of candidate regulatory loci (rows) associated with each immunomodulatory program in pseudobulk profiles aggregated from snATAC-seq data for glioma-associated myeloid cells with preferential expression of a single program (columns). Norm., normalized. **b**, Heatmap depicting enrichment of the indicated TF motifs (rows) in accessible elements associated with each program (columns). *P*-values were calculated using one-sided Fisher's exact test. **c**, Heatmap displaying the specificity scores for TF regulons (rows) in cells with discrete immunomodulatory program usage (columns). Scores were calculated from scRNA-seq by SCENIC. **d**,**e**, Bar charts showing the expression of ligands predicted to target TFs of the system inflammatory program (**d**) or the scavenger immunosuppressive program (**e**). Values indicate expression in the program with the highest expression of the ligand (indicated by bar colour). **f**, Volcano plot depicting differentially expressed genes in macrophages treated with Dex for 24 h (ref. 43). The top 20 genes in each program are plotted and coloured according to the program. **g**, Genomic view showing snATAC profiles aggregated for each program over the top GR target loci. The GR ChIP–seq data in the bottom row show top GR-bound loci. **h**, Dot plot showing GR binding signal over rank-ordered target sites. Selected complement immunosuppressive program genes near GR target sites are labelled. Together, epigenetic data shed light on the TF regulatory mechanisms underlying immunomodulatory program expression in glioma-associated myeloid cells.

to systemic inflammatory. These data support our hypothesis that high AP-1 signalling is a driver of the scavenger immunosuppressive program, and also nominate a therapeutic strategy with the potential to reprogram the unique immune microenvironment in gliomas (Fig. 5f).

Discussion

Our study highlights the plasticity of glioma-associated myeloid cells and the critical impact of the local microenvironment and niche on their phenotypes. By decomposing single-cell transcriptomes into discrete expression programs using cNMF, we disentangled cell identity from cell activity, the fundamental components of myeloid-cell phenotypes. This revealed biological insights that are obscured in gene signatures derived from Louvain/Leiden clusters, which conflate co-expressed programs. Although myeloid cells have typically been classified by cell type or ontogeny, we found neither to be major determinants of their activity in gliomas. Instead, different myeloid cell types can each actuate the same four immunomodulatory programs, which probably determine the overall immune state. Indeed, clinical correlates indicate that these immunomodulatory programs potently modulate immune responses and patient outcomes. Functional investigations reveal TF circuits, signalling pathways and ligands that mediate interplay and

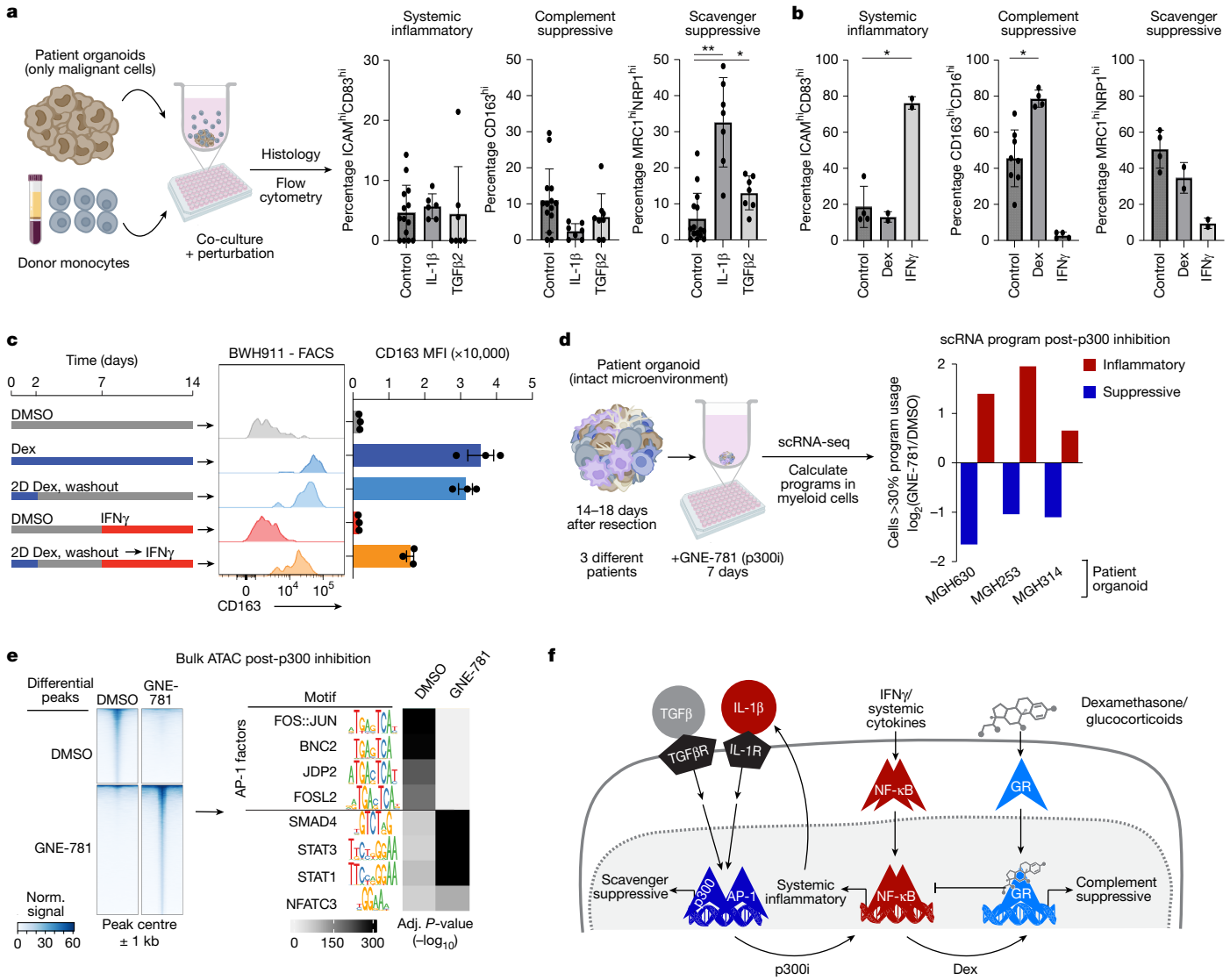


Fig. 5 | Functional induction and reversal of immunosuppressive programs. **a**,**b**, Experimental design and bar plots of flow results from ligand perturbation experiments of IL-1β and TGFβ2, which specifically induce the scavenger immunosuppressive program (**a**) and Dex and IFNγ, which specifically induce the complement immunosuppressive program and systemic inflammatory programs, respectively (**b**). Data points represent individual GBOs. Plots show mean ± s.d. *P*-values for induction over control were calculated using Fisher's exact test; $^{*}P < 0.05$, $^{**}P < 0.01$. **c**, Experimental design and flow-cytometry results from evaluations of the durability of the Dex phenotype. Bar plots show mean signal ± s.d. Data points each represent two pooled GBOs. FACS, fluorescence-activated cell sorting; MFI, mean fluorescence intensity. **d**, Experimental design and results for p300i treatments of GBOs. Bars represent a combination of inflammatory programs (red) or immunosuppressive programs (blue) computed from scRNA-seq. $P = 0.0036$, calculated by two-tailed Student's *t*-test of change in suppressive program versus change in inflammatory program across three samples. **e**, Heatmaps depict differentially accessible candidate regulatory elements between myeloid cells treated with p300i or DMSO control. TF motifs enriched in the differential elements are shown on the right. *P*-values were calculated using one-sided Fisher's exact test. **f**, Schematic depicting the interconnected feedback loops of the immunomodulatory programs and the cytokines, TFs and small molecules that can affect them. 96-well plate (**a** and **d**) and PBMC tube (**a**) created in BioRender. Shalek, A. (2005) https://BioRender.com/b15i535.

balance the respective programs to maintain an immunosuppressive microenvironment.

The four programs have unique associations, niches and drivers but are highly interconnected (Extended Data Fig. 12b). The microglial inflammatory program is emblematic of myeloid plasticity in the brain microenvironment. Despite including classical microglia markers, it is also actuated in subsets of macrophages and dendritic cells derived from infiltrating monocytes. By contrast, the systemic inflammatory program is evident in peripheral monocytes from patients with glioma. In the glioma microenvironment, infiltrating systemic inflammatory cells produce high levels of IL-1β that can induce the glioma-specific scavenger immunosuppressive program, potentially creating a feedback loop that moderates brain inflammation. Notably, a p300/CBP inhibitor downregulates this immunosuppressive program and associated AP-1 signatures, restoring the default inflammatory program. A final program, complement immunosuppressive, actuates different genes and circuits, and occupies distinct glioma niches. It is driven by GR signalling, which can be activated by endogenous glucocorticoids and markedly increased by the synthetic glucocorticoid Dex. The durability of this immunosuppressive phenotype highlights the need for alternatives to Dex, which is widely used for symptom management in glioma.

Our study leaves some critical questions unaddressed. Our data capture the diversity of myeloid cells at the time of tumour resection but do not address their temporal dynamics or rate of turnover. We speculate that microglia and invading monocytes are initially biased towards inflammatory phenotypes but subsequently adopt immunosuppressive

programs in response to microenvironmental cues, as would be consistent with recent mouse studies[50,51] (Supplementary Note 4). The topic is clinically important and timely in light of indications that myeloid-cell lifespan may be extended in brain tumour niches[52]. Further study is also needed to refine our inferences regarding signalling molecules that drive invasion, differentiation and immunomodulatory programs. Our data delineate how glioma niches affect myeloid programs, but they do not address the extent to which feedback from myeloid phenotypes shapes these niches. Nor do we define the causality of the association between tumour grade and myeloid programs, which may further involve interplay between malignant-cell genetics and myeloid phenotypes[6,53]. Nevertheless, our work provides a robust framework for understanding myeloid biology that should catalyse further discovery and ultimately guide new strategies for reprogramming myeloid cells and the immune microenvironment for therapeutic benefit.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-025-08633-8.

1. Gutmann, D. H. & Kettenmann, H. Microglia/brain macrophages as central drivers of brain tumor pathobiology. *Neuron* **104**, 442–449 (2019).
2. Sørensen, M. D., Dahlrot, R. H., Boldt, H. B., Hansen, S. & Kristensen, B. W. Tumour-associated microglia/macrophages predict poor prognosis in high-grade gliomas and correlate with an aggressive tumour subtype. *Neuropathol. Appl. Neurobiol.* **44**, 185–206 (2018).
3. Wang, E. J. et al. Immunotherapy resistance in glioblastoma. *Front. Genet.* **12**, 750675 (2021).
4. Ruffell, B. & Coussens, L. M. Macrophages and therapeutic resistance in cancer. *Cancer Cell* **27**, 462–472 (2015).
5. Klemm, F. et al. Interrogation of the microenvironmental landscape in brain tumors reveals disease-specific alterations of immune cells. *Cell* **181**, 1643–1660 (2020).
6. Hara, T. et al. Interactions between cancer cells and immune cells drive transitions to mesenchymal-like states in glioblastoma. *Cancer Cell* **39**, 779–792 (2021).
7. Abdelfattah, N. et al. Single-cell analysis of human glioma and immune cells identifies S100A4 as an immunotherapy target. *Nat. Commun.* **13**, 767 (2022).
8. Mathewson, N. D. et al. Inhibitory CD161 receptor identified in glioma-infiltrating T cells by single-cell analysis. *Cell* **184**, 1281–1298 (2021).
9. Neftel, C. et al. An integrative model of cellular states, plasticity, and genetics for glioblastoma. *Cell* **178**, 835–849 (2019).
10. Couturier, C. P. et al. Single-cell RNA-seq reveals that glioblastoma recapitulates a normal neurodevelopmental hierarchy. *Nat. Commun.* **11**, 3406 (2020).
11. Venteicher, A. S. et al. Decoupling genetics, lineages, and microenvironment in IDH-mutant gliomas by single-cell RNA-seq. *Science* **355**, eaai8478 (2017).
12. Garofano, L. et al. Pathway-based classification of glioblastoma uncovers a mitochondrial subtype with therapeutic vulnerabilities. *Nat. Cancer* **2**, 141–156 (2021).
13. Johnson, K. C. et al. Single-cell multimodal glioma analyses identify epigenetic regulators of cellular plasticity and environmental stress response. *Nat. Genet.* **53**, 1456–1468 (2021).
14. Alghamri, M. S. et al. G-CSF secreted by mutant IDH1 glioma stem cells abolishes myeloid cell immunosuppression and enhances the efficacy of immunotherapy. *Sci. Adv.* **7**, eabh3243 (2021).
15. Pombo Antunes, A. R. et al. Single-cell profiling of myeloid cells in glioblastoma across species and disease stage reveals macrophage competition and specialization. *Nat. Neurosci.* **24**, 595–610 (2021).
16. Müller, S. et al. Single-cell profiling of human gliomas reveals macrophage ontogeny as a basis for regional differences in macrophage activation in the tumor microenvironment. *Genome Biol.* **18**, 234 (2017).
17. Friebel, E. et al. Single-cell mapping of human brain cancer reveals tumor-specific instruction of tissue-invading leukocytes. *Cell* **181**, 1626–1642 (2020).
18. Friedrich, M. et al. Tryptophan metabolism drives dynamic immunosuppressive myeloid states in IDH-mutant gliomas. *Nat. Cancer* **2**, 723–740 (2021).
19. Lee, A. H. et al. Neoadjuvant PD-1 blockade induces T cell and cDC1 activation but fails to overcome the immunosuppressive tumor associated macrophages in recurrent glioblastoma. *Nat. Commun.* **12**, 6938 (2021).
20. Ginhoux, F. et al. Fate mapping analysis reveals that adult microglia derive from primitive macrophages. *Science* **330**, 841–845 (2010).
21. Hambardzumyan, D., Gutmann, D. H. & Kettenmann, H. The role of microglia and macrophages in glioma maintenance and progression. *Nat. Neurosci.* **19**, 20–27 (2016).
22. Zilionis, R. et al. Single-cell transcriptomics of human and mouse lung cancers reveals conserved myeloid populations across individuals and species. *Immunity* **50**, 1317–1334 (2019).
23. Couturier, C. P. et al. Glioblastoma scRNA-seq shows treatment-induced, immune-dependent increase in mesenchymal cancer cells and structural variants in distal neural stem cells. *Neuro Oncol.* **24**, 1494–1508 (2022).
24. Kotliar, D. et al. Identifying gene expression programs of cell-type identity and cellular activity with single-cell RNA-Seq. *eLife* **8**, e43803 (2019).
25. Bohlson, S. S., O'Conner, S. D., Hulsebus, H. J., Ho, M.-M. & Fraser, D. A. Complement, c1q, and c1q-related molecules regulate macrophage polarization. *Front. Immunol.* **5**, 402 (2014).
26. Zhang, S. et al. C1q$^+$ tumor-associated macrophages contribute to immunosuppression through fatty acid metabolic reprogramming in malignant pleural effusion. *J. Immunother. Cancer* **11**, e007441 (2023).
27. Vogt, L. et al. VSIG4, a B7 family-related protein, is a negative regulator of T cell activation. *J. Clin. Invest.* **116**, 2817–2826 (2006).
28. Timmermann, M., Buck, F., Sorg, C. & Högger, P. Interaction of soluble CD163 with activated T lymphocytes involves its association with non-muscle myosin heavy chain type A. *Immunol. Cell Biol.* **82**, 479–487 (2004).
29. Gudgeon, J., Marín-Rubio, J. L. & Trost, M. The role of macrophage scavenger receptor 1 (MSR1) in inflammatory disorders and cancer. *Front. Immunol.* **13**, 1012002 (2022).
30. Miller, T. E. et al. Mitochondrial variant enrichment from high-throughput single-cell RNA sequencing resolves clonal populations. *Nat. Biotechnol.* **40**, 1030–1034 (2022).
31. Jacob, F. et al. A patient-derived glioblastoma organoid model and biobank recapitulates inter- and intra-tumoral heterogeneity. *Cell* **180**, 188–204 (2020).
32. Ravi, V. M. et al. Spatially resolved multi-omics deciphers bidirectional tumor-host interdependence in glioblastoma. *Cancer Cell* **40**, 639–655 (2022).
33. Cable, D. M. et al. Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat. Biotechnol.* **40**, 517–526 (2022).
34. Scherer, H. J. Structural development in gliomas. *Am. J. Cancer* **34**, 333–351 (1938).
35. Greenwald, A. C. et al. Integrative spatial analysis reveals a multi-layered organization of glioblastoma. *Cell* **187**, 2485–2501 (2024).
36. Amankulor, N. M. et al. Mutant IDH1 regulates the tumor-associated immune system in gliomas. *Genes Dev.* **31**, 774–786 (2017).
37. Auger, J.-P. et al. Metabolic rewiring promotes anti-inflammatory effects of glucocorticoids. *Nature* **629**, 184–192 (2024).
38. Mei, Y. et al. Siglec-9 acts as an immune-checkpoint molecule on macrophages in glioblastoma, restricting T-cell priming and immunotherapy response. *Nat. Cancer* **4**, 1273–1291 (2023).
39. Nishikawa, H. & Koyama, S. Mechanisms of regulatory T cell infiltration in tumors: implications for innovative immune precision therapies. *J. Immunother. Cancer* **9**, e002591 (2021).
40. Terekhanova, N. V. et al. Epigenetic regulation during cancer transitions across 11 tumour types. *Nature* **623**, 432–441 (2023).
41. Aibar, S. et al. SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).
42. Browaeys, R., Saelens, W. & Saeys, Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat. Methods* **17**, 159–162 (2020).
43. Jubb, A. W., Young, R. S., Hume, D. A. & Bickmore, W. A. Enhancer turnover Is associated with a divergent transcriptional response to glucocorticoid in mouse and human macrophages. *J. Immunol.* **196**, 813–822 (2016).
44. Oh, K.-S. et al. Anti-inflammatory chromatinscape suggests alternative mechanisms of glucocorticoid receptor action. *Immunity* **47**, 298–309 (2017).
45. de Almeida Nagata, D. E. et al. Regulation of tumor-associated myeloid cell activity by CBP/EP300 bromodomain modulation of H3K27 acetylation. *Cell Rep.* **27**, 269–281 (2019).
46. Guerriero, J. L. et al. Class IIa HDAC inhibition reduces breast tumours and metastases through anti-tumour macrophages. *Nature* **543**, 428–432 (2017).
47. Pappalardi, M. B. et al. Discovery of a first-in-class reversible DNMT1-selective inhibitor with improved tolerability and efficacy in acute myeloid leukemia. *Nat. Cancer* **2**, 1002–1017 (2021).
48. Romero, F. A. et al. GNE-781, a highly advanced potent and selective bromodomain inhibitor of cyclic adenosine monophosphate response element binding protein, binding protein (CBP). *J. Med. Chem.* **60**, 9162–9183 (2017).
49. Kamei, Y. et al. A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. *Cell* **85**, 403–414 (1996).
50. Kirschenbaum, D. et al. Time-resolved single-cell transcriptomics defines immune trajectories in glioblastoma. *Cell* **187**, 149–165 (2024).
51. Chen, Z. et al. A paracrine circuit of IL-1β/IL-1R1 between myeloid and tumor cells drives genotype-dependent glioblastoma progression. *J. Clin. Invest.* **133**, e163802 (2023).
52. Ballesteros, I. et al. Co-option of neutrophil fates by tissue environments. *Cell* **183**, 1282–1297 (2020).
53. 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 (2023).
54. Varn, F. S. et al. Glioma progression is shaped by genetic evolution and microenvironment interactions. *Cell* **185**, 2184–2199 (2022).
55. Hoogstrate, Y. et al. Transcriptome analysis reveals tumor microenvironment changes in glioblastoma. *Cancer Cell* **41**, 678–692 (2023).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2025

Methods

Human subjects

Adult male and female patients at Massachusetts General Hospital and Brigham and Women's Hospital (MGB) provided preoperative informed consent to take part in the study in all cases under the approved Institutional Review Board Protocol DF/HCC 10-417. Patients' clinical characteristics are summarized in Supplementary Table 1. Patients in other cohorts gave consent according to published methods[7,10,13]. Previously unpublished patient data from the Montreal Neurological Institute was collected as reported with other tumours from McGill University[23].

Primary tumour processing for Seq-Well and GBOs

Fresh tumour samples were collected directly from the operating room at the time of surgery and the presence of glioma was confirmed by frozen section. Samples were dissected into small pieces and mixed. For samples with enough material, we divided up the mixed tumour pieces, with part being used for single-cell dissociation and part for GBO generation.

Single-cell dissociation and Seq-Well. For the MGB cohort, minced tissue pieces were mechanically and enzymatically dissociated using a tumour dissociation kit, according to the manufacturer's instructions, and a GentleMACS Octo Dissociator with Heaters (Miltenyi Biotec) using custom settings. The single-cell suspension was then depleted of dead cells and debris using magnetic-activated cell sorting (Dead Cell Depletion Kit, Miltenyi Biotec). Cells were then processed for Seq-Well as previously described[56]. In brief, cells were resuspended in media and 10,000–15,000 cells were then distributed dropwise onto a Seq-Well microwell array preloaded with mRNA capture beads. The arrays were then sealed with a semipermeable membrane and submerged in lysis buffer to lyse their cellular and nuclear membranes. Lysis buffer was then removed and the array was submerged in a hybridization buffer to allow for attachment of mRNA to the capture beads. After mRNA capture, the seal was removed from the array and the capture beads were manually washed out of the array with a washing buffer. Following removal from the array, the beads were washed to remove any excess lysis reagents and unattached nucleic acids. A reverse-transcription reaction was run on the beads and cDNA was generated by elongation of the capture bead oligos using the captured RNA as a template. An exonuclease reaction was then run on the beads to clean them up. After this, a second strand-synthesis reaction was run to generate a DNA strand complementary to the cDNA strand by polymerase chain reaction. Subsequent to second-strand generation, the now double-stranded cDNA was amplified by polymerase chain reaction in a whole-transcriptome amplification. Sequenceable libraries were generated from the whole-transcriptome amplification using the Nextera Library Prep kit, following the manufacturer's instructions.

MAESTER materials and preparation. From the Seq-Well cDNA, mitochondrial RNA fragments were selectively pulled down, separately amplified and sequenced as previously described[30].

Creation and maintenance of GBOs. Minced tissue pieces were further dissected using two scalpels until pieces were 1–2 mm in diameter. These were washed, further processed and maintained according to a previous protocol[57].

Patient PBMC and monocyte processing

Patient PBMCs were collected at the time of surgery and isolated using SepMate-15 tubes (StemCell Technologies) and Lympholyte-H (Cedarlane) according to the manufacturers' instructions. Cells were either directly processed for Seq-Well as above or enriched for pan-myeloid cells using CD11b beads (Miltenyi Biotec, 130-097-142) on Miltenyi magnet according to the manufacturing protocols, and then processed for Seq-Well. $CD11b^+D45^+$ purity was checked by flow cytometry (purity > 90%).

GBO perturbation and single-cell read-out methods

GBO perturbations. For perturbation experiments, GBOs were pipetted into ultralow-adherence round-bottomed 96-well plates (Corning, 7007) with one GBO per well. GBOs were plated in 100 µl GBO media. Small molecules were then added in an extra 100 µl of media at 2× concentration. Medium was changed every 2–3 days by removing 100 µl and replacing it with 100 µl fresh media with perturbation. Depending on the experiment, each condition had 6–12 GBOs per condition to account for heterogeneity among GBOs. For experiments with flow cytometry or scRNA-seq as a read-out, multiple GBOs were grouped together in replicates per condition and then dissociated to single cells together.

Myeloid–GBO co-culture. Human $CD11b^+CD45^+$ cells isolated from tumour or donor patient PBMCs, as described above, were aliquoted and frozen at 5×10^6–1×10^7 cells per ml per vial. Before co-culturing with GBOs, myeloid cells were gently thawed and washed in warm myeloid cell media (ImmunoCult-SF Macrophage Differentiation Medium, using base medium with only M-CSF (50 ng ml^{-1}; STEMCELL Technologies, 10961), and plated in a 24-well low-attachment plate (Corning) to recover for 30 min in a 37 °C CO_2 incubator. Plates were placed on an orbital rotator at 120 rpm with 2.5×10^6 maximum cells per well to avoid cell attachment and to maintain monocyte morphology. Monocytes (10,000–50,000, depending on the experiment) were then added to each GBO well in 100 µl myeloid cell media with small-molecule perturbation when applicable. Medium was changed every 2–3 days by removing 100 µl and replacing it with 100 µl fresh medium (a 1:1 mix of GBO medium and myeloid cell medium) with perturbation when applicable.

Dissociation of GBOs. In brief, all GBOs in each experimental replicate were grouped together in a 1.7 ml Eppendorf tube, the medium was aspirated and GBOs were washed twice with 1 ml media to remove small molecules and/or cells. GBOs were then dissociated to single cells using dissociation media from a Miltenyi tumour dissociation kit mixed 2:1 with accutase in the 1.7 ml tubes. These were placed at 37 °C and mechanically dissociated every 5–10 min by pipetting up and down until there was a homogeneous single-cell mixture. Cells were passed through a 40 µm filter and then used for downstream assays. Cells were processed for Seq-Well as described above or analysed by flow cytometry as described below.

Flow cytometry. Flow cytometry was based on a previous protocol[58]. In brief, an antibody cocktail was made by a 1:1 ratio of Brilliant Stain Buffer (BD Horizon, 566349) and PBS, and antibodies or dyes were added (antibodies are listed in Supplementary Table 5). Single-cell suspensions were washed by PBS + 0.5% BSA in 1.5-ml Eppendorf tubes or 96-well U-bottomed plates. Cells were pelleted by centrifugation at 300*g* for 5 min at room temperature. After removing the washing buffer, cells were resuspended in 100 µl staining cocktail by pipetting up and down about ten times. Plates or tubes were covered to avoid light and stained in the dark at room temperature for 20–25 min. Cells were washed by PBS + 0.5% BSA and centrifuged at 300*g* at room temperature for 5 min. Cell pellets were resuspended in 200 µl PBS + 0.5% BSA. Flow cytometry was processed on a BD LSRFortessa X-20 according to the manufacturer's procedure. UltraComp eBeads (Invitrogen, 01-3333-42) were used to pre-annotate the compensation. FlowJo v.10 was used for data analysis.

Histological assessment of GBO experiments. GBOs were fixed in 4% formaldehyde for 30 min and then washed with DPBS and left in a 30% sucrose solution overnight to dehydrate the tissue. The organoids

were then embedded in OCT (Sakura Tissue Tek) and frozen using an isopropanol bath. The tissue was then sectioned at a thickness of 8 µm using a cryostat. For staining, slides were dried at room temperature for 10 min and then prewashed with 1× TBST to remove the OCT. The tissue was blocked with a glycine BSA solution for 1 h at room temperature. Tissue sections were then incubated with primary antibodies either at 4 °C overnight or at room temperature for 2 h (antibodies are listed in Supplementary Table 5). Tissue sections were then washed thoroughly with 1× TBST and incubated with fluorophore-conjugated secondary antibodies and DAPI for 1 h at room temperature. The tissue was washed with 1× TBS and incubated with 1× True Black Autofluorescence quencher for 1 min. Tissue sections were washed with 1× TBS and mounted with Prolonged Gold mounting media (Invitrogen) and covered with glass slips. For imaging, a Leica Thunder microscopy system was used with an automated mechanized stage. Images were taken using the scanning features with a ×40 oil-immersion objective lens. Images were stitched together and enhanced with the fast computational clearing programs of Leica LAS X software v.3.8.2.27713.

Histological analysis. All histological images were analysed using Qupath open-source image-analysis software v.0.4.3. Cells were counted with the cell-detection feature using the DAPI channel. The detected cells were then called for positivity of up to three fluorescent markers using the single-object measurement feature with positivity thresholds adjusted according to the experiment. Thresholds were set by comparison of experimental conditions to control and then applied to all images of the experiment through automated scripts.

Single-cell, spatial and bulk RNA-seq analyses

Cohorts. This study used four cohorts, split into two datasets. The discovery dataset contained the MGB, Houston Methodist[7] and Jackson Laboratories[13] cohorts. Cells in these cohorts were assayed with more-advanced scRNA-seq technologies: Seq-Well S^3 (MGB) or 10x Genomics 3′ v.3 (Methodist and Jackson Laboratories). The validation dataset was composed of samples from the McGill cohort, including a set that had been previously published[10,23] and a set that was not previously published. McGill tumours were assayed using 10x Genomics 3′ v.2 kits. The Methodist data were downloaded (GEO series, GSE182109), the Jackson Labs data were downloaded from https://ega-archive.org/datasets/EGAD00001007772 and published McGill data were downloaded from https://ega-archive.org/datasets/EGAD00001006206.

For other tumours, we obtained raw expression matrices from sources indicated in Supplementary Table 4.

Alignment. We used the Cumulus platform[59] to process the large-scale single-cell RNA-seq experiments. Libraries were aligned to the GRCh38 genome using STARsolo v.2.7.10 (ref. 60). Our GitHub page contains the specific settings (https://github.com/BernsteinLab/Myeloid-Glioma).

Data processing and visualization. The raw matrices outputs of STARsolo (no filtration of cells) for each tumour and GBO library were gzipped and used as input for Seurat[61] using the Read10X() function with the default parameters. The pipeline was done for each cohort independently. Tumours in each cohort were merged using the Seurat merge() function to generate a Seurat object for each cohort. The percentage of mitochondrial gene expression was determined using PercentageFeatureSet() with the pattern set to ^MT. Cells with more than 25% of their transcriptome composed of mitochondrial gene expression were filtered out. We also filtered out cells expressing fewer than 500 genes or more than 6,000 genes. We also filtered out cells with less than 1,000 UMIs and genes expressed in less than three cells in the matrix. The filtering process was done using the Seurat subset() function.

For plotting, normalization, scaling and variable gene detection were done using the SCTransform() function, for which we used the percentage of mitochondrial gene expression as a regression factor. We performed principal components analysis (PCA) using RunPCA() with default parameters and generated an elbow plot using the ElbowPlot() function to determine the dimensions for generating UMAPs and for Louvain clustering (MGB, 24; Houston Methodist, 19; Jackson Laboratories, 16).

UMAPs were generated using RunUMAP() with the reduction set to pca. FindNeighbors() and FindClusters() were used for clustering, with the resolution set to 0.3.

Classification of cell types. To classify tumour cells in all cohorts, we identified the main cell programs in the MGB cohort and identified the top program for each cell in all cohorts. This top program was then used as the cell's classification.

We merged all cells from the 22 tumours in the MGB cohort and used this expression matrix as the input for cNMF. We identified the 4,000 most variable genes using SCTransform, regressing out mitochondrial content. We subsetted the matrix for these genes and the resulting matrix was then subjected to consensus cNMF v.1.0.4 (ref. 24).

For the cNMF 'prepare' function, we performed factorization over K ranges from 2–35. We ensured that all the variable genes were considered for the factorization using the parameter --numgenes 4000. We also performed 500 iterations by inputting --n-iter 500 in the cNMF prepare script. K = 18 was the highest value with a silhouette score greater than 0.8 and was thus chosen for the consensus script of cNMF. cNMF was run with a --local-density-threshold value of 0.015.

We annotated each program on the final gene_spectra output of cNMF by comparing the top 100 genes with previously published gene sets and known marker genes. We used gProfiler[62] to determine enrichment scores for a manually curated gene-set matrix with more than 600 gene sets (Supplementary Table 3). Manual integration of enrichment scores and known marker genes enabled us to determine the names of the programs (Extended Data Fig. 1a). Of note, MGH720, a tumour with a histological diagnosis of giant cell glioblastoma, had a cNMF unique malignant program.

We then used the gene spectra output of the cNMF programs to calculate the usage of these programs by cells in the other published cohorts and in the GBO scRNA-seq libraries. We extracted a raw-counts matrix including the intersection between genes detected in the cohort and the top 4,000 variable genes in the MGB cohort. This matrix was then subjected to the cNMF prepare script for normalization. The --numgenes parameter was set to the number of genes in the matrix. We used sklearn.decomposition.non_negative_factorization in which X is the filtered normalized expression matrix and H is the filtered gene-spectra consensus matrix. The following parameters were used: n_components=18, init = 'random', update_H=False, solver='cd', beta_loss='frobenius', tol=0.0001, max_iter=1000, alpha=0.0, alpha_W = 0.0, alpha_H = 'same', l1_ratio=0.0, regularization=None, random_state=None, verbose=0, shuffle=False. The code is available at https://github.com/BernsteinLab/Myeloid-Glioma.

Finally, each cell was annotated as a cell type using the final 'usage matrix output of cNMF or the calculated usage matrices, as discussed above. The usage scores were normalized to 100% for each cell. For each cell, the usage scores for all programs in each category were summed to create a usage score for the cell-type category. For example, the usage scores for four myeloid programs were summed to create the myeloid usage per cell. Cells were then annotated as one of the cell types using the top-scoring usage for cell-type category.

Of note, cycling cells were considered separately. We used inferCNV to annotate cycling cells as malignant or non-,alignant. Non-malignant cells were then further annotated by the next-highest cell type. These secondary annotations were used when separating cell types for further cell-type-specific analysis.

CNA inference from single-cell data. We selected a group of reference cells not annotated as any malignant program from various tumours

(a mix of myeloid cells, T cells, oligos and vasculature cells). We extracted and merged the raw counts of these reference cells into a single matrix. The reference cells used are given in https://github.com/BernsteinLab/Myeloid-Glioma. We then used the inferCNV package (inferCNV v.1.3.3 of the Trinity CTAT Project; https://github.com/broadinstitute/inferCNV). We performed the analysis for each tumour separately. In the annotation file, we included the reference cells and annotated the cells of each tumour, as discussed above. We concatenated the raw matrix of each tumour with the reference cells' raw matrix. We constructed the gene-order file required for inferCNV using the gtf_to_position_file.py script of the inferCNV package. We included the following extra arguments: --denoise --HMM --cluster_by_groups --cutoff 0.1. We also ensured that the --ref_group_names match the names given to the reference cells in the annotation files. The selection of the reference cells was performed separately for each cohort.

Doublet detection. Doublets were determined using integration of cNMF and inferCNV data. Cells were considered doublets by cNMF if they expressed a second program above a specific threshold (Supplementary Table 6). Cell-type-specific thresholds were selected by subsetting by cell type, then plotting the usage of each potential second program. From this plot, we found the value that separated the background usage of a second program from doublets. Cells were also considered doublets if their cNMF annotation was not compatible with the inferCNV profile. Of note, the cycling programs were not considered in doublet analysis. A more detailed description of annotation and doublet-detection methods can be found at https://github.com/BernsteinLab/Myeloid-Glioma/blob/main/Processing%20of%20scRNA-Seq%20Files%20(Related%20to%20Figure%201)/15-%20Annotation%20and%20Doublet%20Detection.pdf.

Integrated definition of malignant cells. If a tumour had CNVs that were detectable using inferCNV, cells from that tumour had to meet the following criteria: non-doublet, positive for CNV and not annotated as a non-malignant cell type by the cNMF program. For those tumours in which CNVs could not be readily detected by inferCNV, we relied on annotations based on cNMF.

Gene-program identifications. For more granular analysis of cell programs for a specific cell type (myeloid cells, T cells or malignant cells), we took cells in each specific category and removed doublets on the basis of the method described in the 'Doublet detection' section above. We then input those cells identified as singlets into another cNMF analysis for each category.

Myeloid cells. We used the MGB, Jackson Laboratories and Methodist cohorts for identifying the cNMF programs in myeloid cells in gliomas. The cNMF was carried out in two rounds for each cohort. The first round was used to identify cells using programs that are not myeloid (that is, have a different cell-type identity) or programs used by fewer than 100 myeloid cells. We removed such cells for subsequent analyses. The second round was used to determine the myeloid programs.

In the first round, raw counts of all cells annotated as myeloid and singlets (non-doublets) from each cohort were used to create a Seurat object independently. We then normalized the Seurat object using NormalizeData() and identified the top 2,000 variable genes with mean expression greater than 0.001 in expressing cells in each cohort using the FindVariableFeatures(). Subsequently, we output the three matrices. These matrices were subjected separately to cNMF with the following parameters in the prepare script: --n-iter 500 --total-workers 1 --seed 14 --numgenes 2000. Then we performed factorization and generated the K-plots using the factorize, combine and k_selection_plot scripts of cNMF. We then chose the following Ks: MGB - 22, Houston Methodist - 23, Jackson's laboratories - 14. We then performed the consensus script with the above Ks and a local-density threshold of 0.02.

In the second round, we removed cells from each cohort as discussed above and created a merged Seurat object from the three cleaned matrices using the Seurat merge() function. Then we normalized the merged Seurat object and detected variable genes using NormalizeData() and FindVariableFeatures(). We then filtered out the genes with a mean expression value of less than 0.01 in expressing cells and standardized variance of less than 1. We then filtered the cleaned myeloid matrix of each cohort to include the variable genes that met the criteria mentioned above. Similar to round 1, these matrices were subjected to cNMF individually with the following parameters in the prepare script: --n-iter 500 --total-workers 1 --seed 14 --numgenes 2276. Then we also ran the factorization and generated the K-plots using the factorize, combine and k_selection_plot scripts of cNMF. We then chose the following Ks in the second round: MGB - 18 (we filtered out programs that are not myeloid), Houston Methodist - 19, Jackson Laboratories - 18. Finally, we performed the consensus script with the above-mentioned Ks and a local-density threshold of 0.02.

To find the consensus programs, we performed a cosine similarity of the gene-spectra 'dt' output of each discovery cohort from round 2 of cNMF. We applied a hierarchical clustering algorithm called Ward's method to the similarity matrix to visualize the relationships between programs in a heatmap. To derive the spectra scores for the consensus programs, we averaged the spectra dt scores of the programs with a cosine similarity score of 0.5 or higher. This resulted in 14 consensus myeloid programs across the 3 cohorts, which we annotated as discussed above.

Malignant and T cells. Malignant cells and T cells were obtained from the MGB data in separate cNMF runs, similar to the two-step cNMF used for myeloid cells. We selected a K-value of 7 for the malignant cells based on the silhouette plot's stability, consistent with previously published glioblastoma signatures represented in our five chosen programs[61]. For the T cells, we found the optimal program count to be 4. We calculated the usage of these programs in the other cohorts in a way similar to the all-cell-type NMF mentioned above.

Processing and NMF for PBMC scRNA-seq libraries. The PBMC libraries were processed for cNMF in a similar way to the primary tumour libraries. We merged the expression matrix of all the PBMC libraries using the Seurat merge function. The Seurat object was then normalized using NormalizeData() and ScaleData(). We then used FindVariableFeatures() to calculate the variance score for every gene. We selected the top 3,000 variable genes after removing genes of less than 0.001 mean expression (in expressing cells) and then subsetted the gene-expression matrix to include only the variable genes. As described above, cNMF was performed with --numgenes 3000 and a value K = 18 for the consensus script of cNMF, annotation was done using gProfiler, and non-doublet cells were identified. We isolated myeloid cells, identified the top 2,000 most-variable genes, and performed two rounds of cNMF (K = 16).

Calculation of the usage of myeloid programs. We extracted a raw counts matrix including the intersection between genes detected in the GBO matrix or other tumour types of matrices and the genes found in the myeloid-cell programs gene-spectra file. This matrix was then subjected to the cNMF prepare script for normalization. The --numgenes parameter was set to the number of genes in the matrix. We used sklearn.decomposition.non_negative_factorization in which X is the filtered normalized expression matrix and H is the filtered gene-spectra consensus matrix. The following parameters were used: n_components=14, init='random', update_H=False, solver='cd', beta_loss='frobenius', tol=0.0001, max_iter=1000, alpha=0.0, alpha_W=0.0, alpha_H='same', l1_ratio=0.0, regularization=None, random_state=None, verbose=0, shuffle=False. The code is available at https://github.com/BernsteinLab/Myeloid-Glioma.

The usage scores were normalized to 100% for each cell.

Comparison of gene programs. To assess the similarity of 2 given gene programs, we took the top 100 genes in those programs and compared their make-up using a Jaccard index. *P*-values were measured by assessing the probability of observed gene matches being obtained by random chance using a binomial test in which k is the number of matches, n is the size of the gene set, and P is probability of randomly drawing matches from all genes scored in the program.

Comparative UMAP and clustering of myeloid cells. We extracted the raw counts of all MGB cells annotated as myeloid and singlets (non-doublets) from each tumour. Then normalization was performed for each tumour separately using NormalizeData() with default settings, followed by FindVariableGenes() with the following settings (selection.method = vst, nfeatures = 2,000). We then ran FindIntegrationAnchors() k.filter set at 30 (to ensure that tumours with few myeloid cells were included. We then used the anchors identified as input to batch correct the objects using IntegrateData(), setting features.to.integrate as the intersection of genes detected in all tumours in and dims to 1:30.

The batch-corrected Seurat object was then subjected to ScaleData(), RunPCA() and ElbowPlot() with default parameters to identify the number of dimensions to use for Louvain clustering and UMAP generation. We generated the UMAP using RunUMAP() with dims set to 1:8 and reductions set to pca. We performed the clustering using FindNeighbors() with dims set to 1:8, followed by FindClusters() with a 0.3 resolution. UMAPs were generated using the DimPlot() function.

Generation of heatmap for gene-expression programs. To generate the gene-expression heatmap of the NMF programs, we assigned the myeloid cells to one of the following categories:

Microglia. This had a minimum 10% usage of the microglia program and other identity programs were all less than the usage value of the microglia program (macrophages must be below 10%).

Microglia-like. This had a minimum 10% usage of microglia and 10% usage of the monocytes or macrophages program. Other identity programs should be below the usage value of these two programs (otherwise it was assigned as a microglia).

Macrophages. This had a minimum 10% usage of the macrophage program and other identity programs were all below the usage value of the macrophage program (monocytes below 10%).

Mono_macro. This had a minimum 10% usage of macrophages and 10% usage of the monocytes program. Other identity programs were below the usage value of these two programs.

Monocytes. This had a minimum 10% usage of the macrophage program and other identity programs were below the usage value of the monocytes program.

cDC. This had a minimum 10% usage of the cDC program and other identity programs were all below the usage value of the cDCs program.

Neutrophils. This had a minimum 10% usage of the neutrophils program and other identity programs were all below the usage value of the neutrophils program.

Activity dominated. All identity programs were below 10% usage.

For selecting the genes included in the heatmaps, we identified the top 100 genes in the averaged gene-spectra output of the myeloid cNMF programs for each program. We counted the number of myeloid cells expressing the top 100 genes in each program. We included the top 20 genes with the highest number of myeloid cells expressing them in each program.

We used the ComplexHeatmap library in R[63] to generate the heatmaps. We *z*-score scaled the log1p normalized gene-expression values across all the myeloid cells (regardless of the categorization). We then set an upper limit of 2 and a lower limit of −1. We generated a heatmap for each category separately. We turned off row clustering (genes) by setting cluster_rows = FALSE in the heatmap function, and we allowed default column clustering in each category (cells).

Generation of quadrant plots. We plotted 183,062 myeloid cells from 85 tumours on the basis of their expression of the four immunomodulatory programs, orienting largely exclusive inflammatory and immunosuppressive programs oppositely on the same axes. The x axis of the quadrant plots (bottom left to top right) was calculated by subtracting the usage of the complement immunosuppressive program from the systemic pro-inflammatory program. The y axis (top left to bottom right) was calculated by subtracting the scavenger immunosuppressive program usage from the microglial inflammatory program for each myeloid cell. The resulting coordinates were transformed using a 45° rotation matrix. For the quadrant plot with scatterpies as dots, we used the scatterpie library (https://github.com/GuangchuangYu/scatterpie). The 'others' category for the pie charts was calculated by summing the usages of all programs excluding the four immunomodulatory programs.

Determining cut-off usage values for myeloid cells. To determine the cut-off threshold for myeloid identity programs in primary tumours, we created an elbow plot to analyse the usage values for each myeloid-identity cNMF program. Based on our findings, we identified the minimum inflection point for an identity program at around 9% and set the cut-off level at 10% usage. Similarly, we observed that the minimum inflection point is approximately 18% for activity programs, leading us to select 20% usage as the cut-off level. A single myeloid cell could be classified using multiple programs. For example, a myeloid cell could be considered to be microglia using the scavenger immunosuppressive program if it had at least 10% usage of the microglia program and 20% usage of the scavenger immunosuppressive program. These cut-offs were used for Fig. 1d and Extended Data Figs. 1e and 8b. For Extended Data Fig. 2d, we calculated the observed/expected ratios of the co-occurrence of a myeloid identity program and a myeloid activity program and used a hypergeometric test to assess significance.

For GBOs, as part of optimizing the model, we determined the inflection points of the activity programs in the myeloid cells. We observed that the inflection points were approximately 27% for activity programs, leading us to select 30% usage as the cut-off level in GBO experiments. These cut-offs were used for all GBO scRNA-seq analyses, as shown in Fig. 5d and Extended Data Fig. 12a.

Definite annotation for myeloid cells and T cells. Specific analyses (MAESTER, Extended Data Fig. 5; and heatmap generation, Fig. 1c) required definite annotation of a myeloid cell instead of only considering whether the myeloid cell uses a particular program or not. The cut-offs for such analyses are mentioned in their respective sections.

T cell program usages were more distinct. We therefore simply defined them by the program of their top usage.

Creating discretized matrix and identifying marker genes. To enable the discovery of specific markers and the spatial cellular demultiplexing described below, we created a discretized matrix of MGB cells with the strongest expression of each tumour-cell program, including myeloid, T cell and cancer programs, thereby excluding intermediate cells. For the outputs of myeloid, malignant and T cell cNMFs, cells with a minimum 2.5-fold higher usage of a particular program over the second-most-used program were annotated with that program as a discrete cell. For oligodendrocytes and vasculature, the usages from all cell types cNMF outputs were used to annotate oligo or vasculature discrete cells. Doublets, cycling programs and cycling cells were excluded from the analysis.

Furthermore, we downloaded scRNA-seq for normal brain tissue from individuals 25 and 40 years old[64]. These libraries were processed using the above-mentioned Seurat processing; we normalized the cells and used 1:14 as dims for generating UMAPs and identifying clusters. We performed FindMarkers and extracted neurons and astrocytes from the published matrix on the basis of differentially expressed genes. We then merged these cells with the discrete cells matrix. We generated a UMAP as described above.

To identify markers for the immunomodulatory programs, we extracted discrete cells annotated as scavenger immunosuppressive, complement immunosuppressive, systemic inflammatory or microglial inflammatory using the Seurat subset() function (discrete myeloid immunomodulatory object). A similar processing pipeline was performed with the option dims set to 1:16 in FindNeighbors() and FindClusters(). The UMAP coordinates for these cells were obtained using Embeddings() with the option reduction set to umap. Then the normalized matrix of these cells was extracted using the function GetAssayData() with the option slot set to data. These files were used as input for COMET[65] to identify the significant markers that distinguish each immunomodulatory program. We also used the Seurat function DotPlot() to generate the expression dot plots shown in Extended Data Fig. 10c.

Regulon analysis. We used the normalized discrete myeloid immunomodulatory expression matrix as an input for SCENIC[41]. We also downloaded the feather files hg38_10kbp_up_10kbp_down_full_tx_v10_clust.genes_vs_motifs.rankings.feather and hg38_500bp_up_100bp_down_full_tx_v10_clust.genes_vs_motifs.rankings.feather from https://resources.aertslab.org/cistarget/databases/homo_sapiens/hg38/refseq_r80/mc_v10_clust/gene_based and used these as input for SCENIC. Genes in the matrix were filtered on the basis of the default settings, and we ran the four steps runSCENIC_1_coexNetwork2modules, runSCENIC_2_createRegulons, runSCENIC_3_scoreCells(scenicOptions, exprMat_filtered) and runSCENIC_4_aucell_binarize. To calculate the regulon specificity score for each regulon in each immunomodulatory annotation, we used the function calculate_rrs() of scFunction (https://github.com/FloWuenne/scFunctions/). The full codes can be found at https://github.com/BernsteinLab/Myeloid-Glioma.

MAESTER analysis for determining myeloid-cell origin. To determine the origins of the various myeloid-cell identities, we processed the MAESTER libraries in three stages: preprocessing, identifying variants of interest, and measuring the enrichment of the identified variants in the various myeloid identities.

Stage 1: preprocessing. The preprocessing was done as previously described[30]. In brief, we trimmed high-quality reads, aligned them using STAR (hg38) and processed them using MAEGATK with the default settings except for - -min-reads, which was set to 3 to ensure a minimum of three reads to cover the base in the cell[30].

Stage 2: low-resolution pseudobulking to identify variants of interests. To determine the origin of the myeloid-cell identities, we identified variants specific to myeloid cells in the tumour microenvironment but not present in the PBMC. We also identified variants present in the myeloid cells in the PBMC but absent in the tumour microenvironment. We pseudobulked the primary tumour libraries to the following categories according to their RNA-seq annotation. For the primary tumour libraries Malignant, Tumor-Associated Myeloid (TAMs), Stromal, Oligo and Tcells, and for PBMC libraries, we considered only Neutrophils and Monocytes (Myeloid_PBMCs). This stage involved multiple steps.

Step 1: we extracted the number of UMIs supporting each possible variant at every possible location from the MAEGATK output using the following script in R:

```
computeAFMutMatrix <- function(SE){
ref_allele <- as.character(rowRanges(SE)$refAllele)
getMutMatrix <- function(letter){
mat <- (assays(SE)[[paste0(letter, "_counts_fw")]] + assays(SE)[[paste0(letter, "_counts_rev")]])
rownames(mat) <- paste0(as.character(1:dim(mat)[1]), "_", toupper(ref_allele), ">", letter)
return(mat[toupper(ref_allele)!= letter,])
}
rbind(getMutMatrix("A"), getMutMatrix("C"), getMutMatrix("G"), getMutMatrix("T"))
}
```

The computeAFMutMatrix generated a matrix in which the rows are every possible variant in the mitochondrial genome, and the columns are barcodes (cells). The values represent the UMIs supporting each variant in the given barcode.

Step 2: we extracted the coverage of the Maester libraries at each base of the MT genome in every cell from the output of MAEGATK, which was stored in assays(maegatk.rse)[[coverage]], whereby maegatk.rse is the R object output of MAEGATK.

Step 3: we annotated the cells in each matrix (obtained from step 1 and step 2) based on the scRNA-seq library (as mentioned in the 'Classification of cell types' section). We created a data frame in which one column has the barcodes and the other has the annotation.

Step 4: we pseudobulked the UMI count matrix (obtained from step 1) using R with the following steps: we subsetted the matrix into each annotation using tibble; we used the sum() function to sum all the rows in each matrix, creating a pseudobulked number for each annotation; and we merged all the summed values in each matrix for each variant possibility into a pseudobulked matrix in which each column represents an annotation.

Step 5: we pseudobulked the coverage matrix (obtained from step 2) similarly to step 4.

Step 6: we calculated the pseudobulked variant allele frequencies (VAFs) using R. We added a pseudo-count 0.000001 to each value in the pseudobulked coverage matrix. We divided each value in the pseudobulked counts matrix by its coverage in the pseudobulked coverage matrix to obtain pseudobulked VAFs for each category.

Step 7: to ensure the specificity of each variant's detection with 99% certainty, we used a binomial model to establish a minimum VAF threshold dependent on coverage.

Step 8: to consider the variant to be specific to the myeloid cells in the tumour microenvironment, the variant had to meet the following criteria: the minimum VAF requirement for TAMs for the coverages in TAMs and Myeloid_PBMC categories; VAF = 0 in the myeloid PBMC category; and VAF > the minimum required in the TAM category.

Step 9: to identify variants specific to the myeloid cells in PBMCs, the variant had to meet the following criteria: first, it had to achieve the minimum VAF requirement for Myeloid_PBMC for the coverages in the Malignant, TAMs and Myeloid_PBMC categories; second, VAF = 0 in the Malignant category (if the tumour library was enriched for malignant cells, this criteria could be replaced with VAF in Myeloid_PBMC, which is 20 times bigger than Malignant); third, VAF > 0 in the TAM category; and fourth, VAF > the minimum required in the Myeloid_PBMC category. The mutations used for subsequent analyses can be found in Supplementary Table 7.

Stage 3: determining the enrichment of variants of interest in various identities of myeloid cells.

We hypothesized that myeloid cells enriched for tumour microenvironment-specific variants were tissue residents, whereas those enriched for PBMC-specific variants were derived from monocytes. Performing the enrichment analysis had several stages. We began by annotating the TAMs with the myeloid identity cNMF programs. We kept cells that met the following criteria to ensure the reliability of the identity and the results:

for Microglia annotation: first, the cell has a minimum of 15% usage of microglia program; second, the microglia program must be at least twice as high as any other identity program; and third, the macrophage program is at least twice as low as any other identity program.

for Microglia_Like annotation: first, the cell has a minimum of 10% usage of the microglia program; second, the cell has a minimum of 10% usage of the macrophage program; third, the microglia program must be at least twice as high as any other identity program (excluding macrophages and monocytes); and fourth, the macrophage program must be at least twice as high as any other identity program (excluding microglia and monocytes).

for Mono_Macro annotation: first, the cell has a minimum of 10% usage of the monocytes program; second, the cell has a minimum of 10% usage of the macrophage program; third, the monocyte program must be at least twice as high as any other identity program (excluding macrophages); and fourth, the macrophage program must be at least twice as high as any other identity program (excluding monocytes).

for Macrophages annotation: first, the cell has a minimum of 15% usage of the macrophages program; second, the macrophage program must be at least twice as high as any other identity program; third, the monocytes program must be at least twice as high as any other identity program; and fourth, the microglia program must be at least twice as high as any other identity program.

For Monocytes annotation: first, the cell has a minimum of 15% usage of the monocytes program; second, the monocyte program must be at least twice as high as any other identity program; and third, the macrophage program must be at least twice as high as any other identity program.

For Neutrophils annotation: first, the cell has a minimum of 15% usage of the neutrophils program; and second, the neutrophils program must be at least twice as high as any other identity program.

For cDC annotation: first, the cell has a minimum of 15% usage of the cDC program; and second, the cDC program must be at least twice as high as any other identity program.

We subsetted the single cells MAESTER UMI matrix obtained in stage 2, step 1 to smaller matrices, with each matrix composed of cells annotated with one of the above identities (stage 3, step 1).

We created a matrix in which the rows are the variants of interest identified in stage 3, and each column represents an identity. The values in this matrix denote the number of cells in each identity in which the variant was detected. We used annotation_matrix[,apply(annotation_matrix,2,function(x) sum(x > 0))] in R for each subsetted matrix. This script keeps columns that have UMI > 0. Then we identified the number of cells for which UMI > 0 for each annotation by using ncol(as.matrix(annotation_matrix), whereby annotation_matrix represents each subsetted matrix.

We then removed any variant with a zero value in every pseudobulked category.

We removed any identity (column) that had less than ten total values to ensure proper enrichment results.

We used GSVA[66] by categorizing the rows of the matrix as PBMC specific or tumour microenvironment specific. Because the values are integers (the number of cells having UMIs supporting the variants), we set the option kcdf to Poisson in the gsva script.

We then subtracted the GSVA enrichment for tumour microenvironment-specific variants from GSVA enrichment for PBMC-specific variants for each identity. Positive values indicate a PBMC origin, whereas negative values indicate a tissue-resident origin.

We calculated the prevalence of each identity in each tumour. We plotted the dot plots with the position of the dots representing the subtracted GSVA enrichment values, and the size of the dot representing the prevalence using ggplot2.

We calculated the average percentage usages of the four immunomodulatory programs in each identity, labelled the remaining percentage as others in each tumour, and plotted them as stacked bar plots using ggplot2.

Sample level association analysis. To measure inter-state association, we determined the average usage of the myeloid, T cell and cancer programs in their respective cells in each sample. We used the Spearman correlation to understand how a rise in one myeloid state corresponded to changes in cancer or T cell states, enabling us to examine how variations in one state could influence the behaviour of others in the same sample.

In our state-clinical/molecular associations analysis, we applied the sample state averages derived above. These were compared across various clinical and molecular categories, such as tumour grade, IDH-mutation status and steroid usage, using the Wilcoxon rank-sum test.

To account for the confounding effect of sample hypoxia on the influence of dexamethasone on the complement immunosuppressive myeloid program, we limited this analysis to samples for which the average MES2 program usage in cancer cells was below 20%.

We used a linear least-square regression to investigate changes in average myeloid program usage across samples relative to the daily dose of Dex.

To assess the myeloid and cancer program associations with T_{reg} cells, we classified each sample in the top 50% and bottom 50% expressors of T_{reg} cells and cancer/myeloid program usages and used Fisher's exact test to measure this association. *P*-values were adjusted using the FDR.

Cell-cycle analysis. We evaluated the relationship between the cell cycle and different myeloid cell states in each sample independently. Cells were defined as cycling if the combined usage of the G1S and G2M programs exceeded 20%, and as belonging to a specific state if its usage surpassed 20%. We used Fisher's exact test to measure this association and obtained an odds ratio. For each program, we assessed the probability that its sample distribution was higher than 1 using a one-sample *t*-test on the log-transformed values, so they were approximately normally distributed[67].

Deconvolution of TCGA, GLASS and G-SAM. The analysis pipeline consisted of three steps to deconvolve previously published glioma bulk expression datasets. Steps 2 and 3 of the pipeline were performed for each dataset separately.

Step 1: creating gene sets for each program. The top 50 genes were obtained for each myeloid program by ranking them in the merged gene-spectra output. For T cells and malignant cells, we ranked the gene spectra from their respective cNMF outputs to obtain the top 50 genes for each program. For the other cell types, we used the gene-spectra output of the cNMF of all cells in the tumour microenvironment. We ranked and obtained the top 50 genes for the pericytes, endothelial and oligo programs. Then we selected genes that appeared only in the list of a single program.

Step 2: calculating module scores using Seurat. Raw gene counts for the glioma datasets were obtained. The matrices were then CPM-normalized using EdgeR's DGEList(), calcNormFactors() and cpm() functions[68]. The CPM-normalized matrix was then log-transformed using the log1p() function of R. Afterwards, Seurat objects were created using the CreateSeuratObject() function. The Seurat objects were scaled using ScaleData(). Finally, module scores were calculated using Seurat's AddModuleScore() function. The above-mentioned gene sets were used as input in the features option of AddModuleScore().

Step 3: normalizing the module score. To correct for the purity differences in the published bulk glioma mRNA-expression datasets, we imputed the percentages of malignant, oligo, vasculature, myeloid, T cells and other immune cells in the bulk expression matrices using CIBERSORTx[69]. As input, we used the discretized matrix described above, excluding neurons and astrocytes. The matrix was used as the single_cell_reference in the Fractions module of CIBERSORTx. The published bulk gene expression matrices used as 'mixture' input were also CPM-normalized without log transformation. The module scores obtained above were then divided by the imputed value of their respective cell type in the CIBERSORTx results outputs.

These normalized module scores were then used to correlate with sample clinical data, including IDH-mutation status and molecular grade. For correlation with clinical data, the normalized module scores were $\log_{10}$ transformed. We removed libraries with a CIBERSORTx value of less than 0.1 for myeloid cells. To correlate the module scores with the expression levels of the identified ligands in the lower-grade glioma and glioblastoma TCGA cohorts, we performed a Pearson correlation

between the log-transformed normalized module score and the log-transformed normalized expression. We used the Wilcoxon rank sum test to assess the statistical significance, adjusting the *P*-values using FDR.

Spatial transcriptomic analysis. We used Visium spatial transcriptomics to analyse 23 samples[32]. Our methodology encompassed two distinct approaches: unbiased niche discovery using cNMF; and cell-state demultiplexing using RCTD[33] and the single-cell RNA-seq programs.

In the unbiased niche discovery, we used all samples, both healthy and cancer, for cNMF analysis. We selected the 1,500 top spatially variable genes based on Moran's *I* and ran the cNMF algorithm from $k = 2$ to $k = 25$. The optimal k value ($k = 7$) was determined as the highest one with a silhouette score greater than 0.9. We defined the identity of niche programs through the analysis of gene expression patterns and excluded the program whose top genes were mitochondrial and ribosomal. To ensure uniformity, the usage scores were normalized to 1 in each spot.

For cell-state demultiplexing, we used the discretized matrix described above. Using this single-cell reference, RCTD was run individually for each spatial sample. Again, the scores were normalized to sum to 1 in each spot.

To assess the link between niche and cell state, we used the Spearman correlation to evaluate the relationship between normalized RCTD scores and cNMF usage for individual patient samples. For Fig. 2, a cell type was classified as part of a niche if the correlation had statistical significance (FDR < 0.01) in more than half of the samples.

In our analysis of cell state and niche–niche spatial relationships, we designed a spatial regression model to quantify cell-state proximity, analysing one sample at a time. For any pair of cell or niche states across all spots in a sample, we regressed the presence of one state in a central spot against the average expression of the other state in surrounding spots. In this model, we binarized the central spot's state presence using a state-specific threshold determined by a knee plot (KneeLocator[70]) of cell-state expression across spots in all samples. The resulting regression coefficient indicates the association strength between the states, with the *P*-value demonstrating the statistical significance of this association. The constant represents the background signal of the non-central state. The network graph in Extended Data Fig. 6f was plotted using the Python package NetworkX (v.2.0). The size of each edge represents the ratio of the regression coefficient to the constant, providing a measure of the state-association strength relative to the background signal. Edges were statistically significant (FDR < 0.01) in more than 40% of the samples with an association strength greater than 0.1.

Correlation of immunomodulatory programs with responders and non-responders to immunotherapy. We downloaded the published Seurat object of scRNA-seq libraries of patients with glioma responding or not responding to immunotherapy[38]. Based on the published annotations, we extracted the myeloid cells from the Seurat object. We calculated the usage of the myeloid cNMF programs in each myeloid cell, as shown above. We averaged the usage of each program in the myeloid cells per patient and calculated the difference between the average usage in responding versus non-responding patients, using FDR-corrected Wilcoxon rank-sum tests to assess the significance.

We calculated the module scores using gene sets from refs. 7,15,19 in all myeloid cells from ref. 38. Similarly, we averaged the module scores of each gene set for each patient and calculated the difference between responding versus non-responding patients, using FDR-corrected Wilcoxon rank sum tests to assess the significance.

Survival analysis. We merged the survival information (survival time) and events from the cohorts of G-SAM and GLASS. We included only IDH-WT glioblastomas. We used the normalized module score values (without log transformation) for the cNMF programs. We removed duplicate patient entries corresponding to recurrences by keeping the values from the primary tumour only. We removed any library with a CIBERSORTx value of 0 for the myeloid lineage. We then took the samples in the top 33% and considered this group of samples as the high in each program. For the low group in each program, we selected samples in the bottom 33% of libraries.

The library ggsurvfit (https://github.com/pharmaverse/ggsurvfit) was then used to generate the Kaplan–Meier survival curve. A Cox proportional hazard model (https://github.com/therneau/survival) was used to determine differences in survival probabilities.

We performed the survival analysis in a similar way using gene sets from refs. 7,15,19.

Single-cell ATAC-seq analysis. We downloaded 20 snATAC-seq libraries from ref. 40 and the fragments file for each library. Each library was then processed using Signac[71] v.1.12.0 to remove low-quality cells. In brief, a fragment object for each fragment file was created using the CreateFragmentObject() function of Signac. Then, peaks were called using the CallPeaks() function. We removed peaks mapping to contigs by using keepStandardChromosomes() and removed blacklisted regions by using the subsetByOverlaps() function with the options ranges = blacklist_hg38 and invert = TRUE.

We created a counts matrix for each library using FeatureMatrix(), then created a chromatin assay by inputting the created counts matrix and fragments file of the respective library using the CreateChromatinAssay() function of Signac. We filtered cells that did not include reads in more than 200 features and removed peaks that did not include reads from more than 10 cells. Finally, the Seurat object for each snATAC-seq library was created using the CreateSeuratObject() function.

The cells of each library were subject to QC analyses using the NucleosomeSignal(), TSSEnrichment(), FractionCountsInRegion() and FRiP() functions to filter out cells with high nucleosomal noise, those high in reads mapped to blacklisted regions, low enrichment near TSS and/or low in the number of reads contributing to peaks. Thresholds selected for each library can be found at https://github.com/BernsteinLab/Myeloid-Glioma. The Seurat object and peaks for each library were saved after filtering out low-quality cells.

Next, we combined all the Seurat objects into one as follows. First, we merged all the peaks using the Reduce() function and filtered out peaks less than 20 base pairs (bp) or more than 10,000 bp in length. Then we created a new fragment object for each library by inputting the original fragments file into CreateFragmentObject() but used only high-quality cells. We then created a counts matrix for each library by inputting the newly created fragment objects and combined peaks using the FeatureMatrix() function. Then we created new chromatin assays for each library by inputting the newly created counts matrices and fragment objects using the CreateChromatinAssay() function. Next, we created a new Seurat object for each library using the newly created chromatin assays. Finally, we combined the new Seurat objects using the merge function.

The combined Seurat object was then subjected to clustering analysis and UMAP generation using the RunTFIDF(), FindTopFeatures(), RunSVD(), RunUMAP(), FindNeighbors() and FindClusters() functions. We used the lsi reduction and removed PC1 and PC3 because they were associated with library depth, not biology. The function DepthCor() was used to determine the association of each PC with depth.

The gene activities were calculated using the GeneActivity() function of Signac. Then we created a gene activity matrix using CreateAssayObject(). The values of gene activities were normalized using the NormalizeData() function with the options normalization.method set to RC and scale.factor set to median(Object$nCount_RNA).

The gene activity matrix was then subject to calculation of the cell-types cNMF programs, as described above for the scRNA-seq matrices, to annotate the cells in the matrix. Using these calculations,

the non-doublet myeloid cells were identified in the combined Seurat object as mentioned above.

The Seurat object was filtered to include only myeloid cells by using the subset() function. After that, a new set of peaks was called, resulting in the creation of a new feature matrix and a new chromatin assay. Finally, a gene activity matrix was generated for the combined myeloid-cells object, as discussed previously. The usage of the myeloid programs was determined in the myeloid cells to extract the discrete myeloid cells, as mentioned above, for the scRNA-seq libraries. We extracted discrete myeloid cells for monocyte and microglia identities and for systemic, scavenger, complement and microglial inflammatory for immunomodulatory activities. For systemic inflammatory myeloid cells, we selected the top 400 discrete myeloid cells.

To determine differentially accessible peaks in the immunomodulatory programs, we pseudobulked the discrete myeloid populations using the AggregateExpression() function with the following options: return.seurat = TRUE, group.by = Myeloid_Discrete_Status4, normalization.method = LogNormalize, scale.factor = 10,000. Then we extracted the normalized counts matrix using the LayerData() function on the pseudobulked Seurat object with the layer option set to data. We then converted the log1p values to exponential values and considered a peak to be specific to a particular discrete annotation if it had a count at least 2.5 times greater than the average counts of the other annotations. We generated normalized bigwigs for each annotation using an in-house code, which can be found at https://github.com/BernsteinLab/Myeloid-Glioma. To generate a reference matrix of coverage for the regions spanning 1,000 bps upstream and downstream of the centres of the specific peaks, we utilized the deeptools computeMatrix function. We inputted the normalized bigwigs and the specific peak files while enabling the reference-point and –skipZeros options, and setting –referencePoint to centre and –b and –a to 1,000. Finally, we plotted the heatmap using the deeptools plotHeatmap function[72].

We performed motif analysis on each specific peak set using monaLISA[73] v.1.11.4 to identify significantly enriched motifs. First, we combined the specific peaks and resized them to the median width of the combined peaks with a centre fixation. Subsequently, we divided the combined peak files into bins based on the annotation of the peak. We used the getSeq() function to obtain the sequences of the peaks. We used the PWM core JASPAR2024 (ref. 74) as features and determined the motifs using the calcBinnedMotifEnrR() function of monaLISA. To create a large background for calculating enrichments, we set the background to hg38 genome and the genome.oversample = 500. Finally, the motifs heatmap was generated using the plotMotifHeatmaps() function.

Ligand–TF analysis. To understand which ligands are involved in driving specific programs and regulatory circuitry, we used the NicheNet[42] database (https://zenodo.org/record/7074291/files/gr_network_human_21122021.rds) of file names gr_Network_Nichenet_Database.txt or Signaling_Network_Nichenet_Database) that connects ligands to the TFs they can activate through known ligand–receptor interactions.

We first identified a list of potentially active TFs for each immunomodulatory program through motif enrichment analysis of specific ATAC-seq loci (Fig. 4a,b) and SCENIC analysis (Fig. 4c). Then we used the NicheNet database to identify all ligands that were connected to each TF. We kept only ligands that were expressed in cells in the tumours (any cell type in the tumour microenvironment) at more than 0.3.

Bulk ATAC-seq analysis. The Nextera transposase adaptors were removed from the fastq files by trimming them with the --phred33 --paired --fastqc options using trim_galore[75] (https://github.com/FelixKrueger/TrimGalore). The STAR aligner[60] was then used to map the fastqs with the following settings: -outSAMtype BAM SortedByCoordinate --outFilterMultimapNmax 1000 --outFilterScoreMinOverLread 0.25 --alignIntronMax 1 --alignEndsType EndToEnd. We then used the PICARD MarkDuplicates module (https://broadinstitute.github.io/picard/) with the option REMOVE_DUPLICATES set to TRUE. Afterwards, samtools[76] was used to remove reads mapped to chrM or contigs. The processed bam files were then used to generate bigwigs using the bamCoverage function of deeptools with the RPGC normalization method and an effective genome size of 2,913,022,398. To identify accessible regions, we converted the processed bam files to bed format using bedtools[77] and used awk to add four bases to the start site if the read was on the positive strand and remove five bases from the end site if the read was on the negative strand to correct for tn5 biases. MACS2 v.2.2.9.1 was used to call peaks using the following settings: -g hs -f BED -q 0.01 --nomodel --shift -75 --extsize 150 --keep-dup all. To identify differential accessible sites, we created tag directories using the processed mapped files as input for the makeTagDirectory function of HOMER[78]. The accessible sites for DMSO and p300i libraries were merged using the mergePeaks function of HOMER, and differential accessible sites were determined using the getDifferentialPeaks function of HOMER and setting -F to 2 to keep regions with two-fold higher normalized tag counts in one of the libraries. Deeptools heatmaps were generated as mentioned above, and significant motifs enriched in the differential accessible sites were identified using monaLISA v.1.11.4, similar to the way mentioned above, except for setting genome.oversample to 50, given the number of differentially accessible sites.

GR ChIP–seq analysis. The binding sites of GR were obtained from ref. 43. The binding sites were annotated using the findPeaks module of HOMER. The maximum allowable distance to a gene was 100,000 bp.

Temporal correlation analysis. We obtained gene sets for genes correlated with recent infiltration and remote infiltration into the tumour in a mouse model of glioblastoma[50]. We used gprofiler (https://biit.cs.ut.ee/gprofiler/gost) to determine the enrichment of the myeloid cNMF programs in these gene sets by submitting the top 100 genes for each cNMF program as a query against each gene set as a background.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The raw counts and processed dataset for the discovery and validation cohort are available at the single-cell portal with study ID SCP2389 at https://singlecell.broadinstitute.org/single_cell/study/SCP2389/programs-origins-and-niches-of-immunomodulatory-myeloid-cells-in-human-gliomas. For all scRNA-seq data generated for this manuscript (MGB and the IDH-mutant cohort from McGill University), individual-level raw data were deposited at dbGaP with study accession code phs003756.v1.p1 and can be viewed at the dbGaP study report page at https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs003756.v1.p1. For other data reanalysed in this study, the Methodist data were downloaded from GEO series GSE182109; the Jackson Laboratories data were downloaded from https://ega-archive.org/datasets/EGAD00001007772; and the published McGill University data (IDH-WT samples) were downloaded from https://ega-archive.org/datasets/EGAD00001006206. All other data are available from the corresponding author on reasonable request.

Code availability

The scripts and codes used to generate all the data in the study are available at https://github.com/BernsteinLab/Myeloid-Glioma. A shiny tool to calculate usage of the presented consensus myeloid programs for glioma-associated myeloid cells from other experiments can be found at https://github.com/BernsteinLab/Calculate_Myeloid_cNMF_Usage/.

An online tool to calculate usages of the presented consensus myeloid programs for glioma-associated myeloid cells from other experiments can be found at https://consensus-myeloid-program-calculator.shinyapps.io/shinyapp/. This tool enables users to upload their own gene-expression matrix from scRNA-seq data and output consensus program usages for each cell.

56. Hughes, T. K. et al. Second-strand synthesis-based massively parallel scRNA-Seq reveals cellular states and molecular features of human inflammatory skin pathologies. *Immunity* **53**, 878–894 (2020).
57. Jacob, F., Ming, G.-L. & Song, H. Generation and biobanking of patient-derived glioblastoma organoids and their application in CAR T cell testing. *Nat. Protoc.* **15**, 4000–4033 (2020).
58. Chen, Z. et al. Integrative dissection of gene regulatory elements at base resolution. *Cell Genom.* **3**, 100318 (2023).
59. Li, B. et al. Cumulus provides cloud-based data analysis for large-scale single-cell and single-nucleus RNA-seq. *Nat. Methods* **17**, 793–798 (2020).
60. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
61. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 (2021).
62. Raudvere, U. et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **47**, W191–W198 (2019).
63. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).
64. Herring, C. A. et al. Human prefrontal cortex gene regulatory dynamics from gestation to adulthood at single-cell resolution. *Cell* **185**, 4428–4447 (2022).
65. Delaney, C. et al. Combinatorial prediction of marker panels from single-cell transcriptomic data. *Mol. Syst. Biol.* **15**, e9005 (2019).
66. Hänzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* **14**, 7 (2013).
67. Bland, J. M. & Altman, D. G. Statistics notes. The odds ratio. *BMJ* **320**, 1468 (2000).
68. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
69. Newman, A. M. et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* **37**, 773–782 (2019).
70. Arvai, K. et al. Arvkevi/kneed: Documentation! Zenodo https://doi.org/10.5281/zenodo.6496267 (2020).
71. Stuart, T., Srivastava, A., Madad, S., Lareau, C. A. & Satija, R. Single-cell chromatin state analysis with Signac. *Nat. Methods* **18**, 1333–1341 (2021).
72. Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
73. Machlab, D. et al. monaLisa: an R/Bioconductor package for identifying regulatory motifs. *Bioinformatics* **38**, 2624–2625 (2022).
74. Rauluseviciute, I. et al. JASPAR 2024: 20th anniversary of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **52**, D174–D182 (2024).
75. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
76. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
77. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
78. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
79. Kolberg, L. et al. g:Profiler-interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Res.* **51**, W207–W212 (2023).

Acknowledgements We thank the patients and their families for donating tumour tissue for this study; B. Li for technical help with the Cumulus pipeline; E. Estrada for help determining the specificity of our programs across different tumour types; I. Tirosh and all members of the Bernstein lab for feedback; and J. M. Larson, M. A. De Sauvage, E. Summers, staff at the MGH Pathology Tissue Bank and MGH Neurosurgery, M. C. Prabhu, C. C. Bossi, K. L. Ligon and staff at the Neuro-oncology Tissue and Data Bank for help acquiring fresh tissue resections. T.E.M. and this study were supported by the UK Brain Tumour Charity's Future Leaders Awards (GN-000701 and GN-000722) in partnership with the Jake McCarthy Foundation and the Ria Melvin Fund, the American Brain Tumor Association basic research fellowship in honour of Joel A. Gingras Jr, the SITC-Bristol Myers Squibb Postdoctoral Cancer Immunotherapy Translational Fellowship, the NCI K08 award K08CA276819 and the NIH T32 training grant CA9216. C.P.C. was supported by postdoctoral fellowships from the Ludwig Center at MIT's Koch Institute, the Canadian Institutes of Health Research (181907), the Fond de Recherche du Québec - Santé and a grant from the J. H. and E. V Wade Fund at MIT. L.N.G.C. was supported by the Robert Wood Johnson Foundation. A.K.S. was supported by NIH grants NCI P30 CA14051, NCI 1U2C CA23319501 and NCI U54 CA217377, the J. H. and E. V Wade Fund at MIT and the Pew-Stewart Scholars Program for Cancer Research. Z.C. was supported by NCI-CA-234842. D.S.F. was supported by the Eric and Wendy Schmidt Center at the Broad Institute of MIT and Harvard. J.L.G. was supported by R37CA269499. This project was supported by funds from the Brilliant Night Foundation (to K.P.), the NCI/NIH Director's Fund (DP1CA216873 to B.E.B.), the Ludwig Center at Harvard and the Emerson Collective. B.E.B. is the Richard and Nancy Lubin Family chair at the Dana Farber Cancer Institute, an American Cancer Society research professor and an investigator in the Ludwig Center at Harvard.

Author contributions T.E.M, C.A.E.F., C.P.C., M.L.S., A.K.S. and B.E.B. conceived and designed the study. T.E.M., C.P.C., Z.C., J.P.D., J.V., L.N.G.C. and A.N.C. collected primary tumours, created GBOs, did scRNA-seq and performed experiments on GBOs. T.A.S. and K.P. provided the McGill tumour cohort with clinical data and supported the interpretation of results. D.H.H. provided the 10X Visium cohort and supported the analysis and interpretation of results. T.E.M., C.A.E.F., C.P.C., Z.C., M.A.V., Y.E.T., A.N.C., Y.Z., D.S.F. and B.E.B. contributed to analytical approaches and performed statistical analyses. J.L.G., M.L.S., A.K.S. and B.E.B. supervised the project and contributed to the interpretation of the data. T.E.M., C.A.E.F., C.P.C. and B.E.B. interpreted data and wrote the manuscript. All authors reviewed and edited the manuscript.

Competing interests T.E.M. has financial interests in Reify Health, Care Access Research and Telomere Diagnostics. C.P.C. has financial interest in Axoft. L.N.G.C. receives consulting fees from Elsevier, Oakstone Publishing, Prime Education, BMJ Best Practice and Servier, and research funding from Merck & Co (to the Dana-Farber Cancer Institute). J.L.G. is consultant and serves on the scientific advisory board of Array BioPharma, AstraZeneca, BD Biosciences, Carisma, Codagenix, Duke Street Bio, GlaxoSmithKline, Kowa, Kymera, OncoOne and Verseau Therapeutics, and receives research support from Array BioPharma/Pfizer, Eli Lilly, GlaxoSmithKline and Merck. M.L.S. is an equity holder, scientific co-founder and advisory board member of Immunitas Therapeutics. A.K.S. receives compensation for consulting and/or scientific advisory board membership from Honeycomb Biotechnologies, Cellarity, Ochre Bio, Relation Therapeutics, FL86/Quotient Therapeutics, Parabalis Medicines, Passkey Therapeutics, IntrECate Biotherapeutics, Danaher, Senda Biosciences and Dahlia Biosciences that is unrelated to this work. B.E.B. has financial interests in Fulcrum Therapeutics, HiFiBio, Arsenal Biosciences, Chroma Medicine and Cell Signaling Technologies. The other authors declare no competing interests.

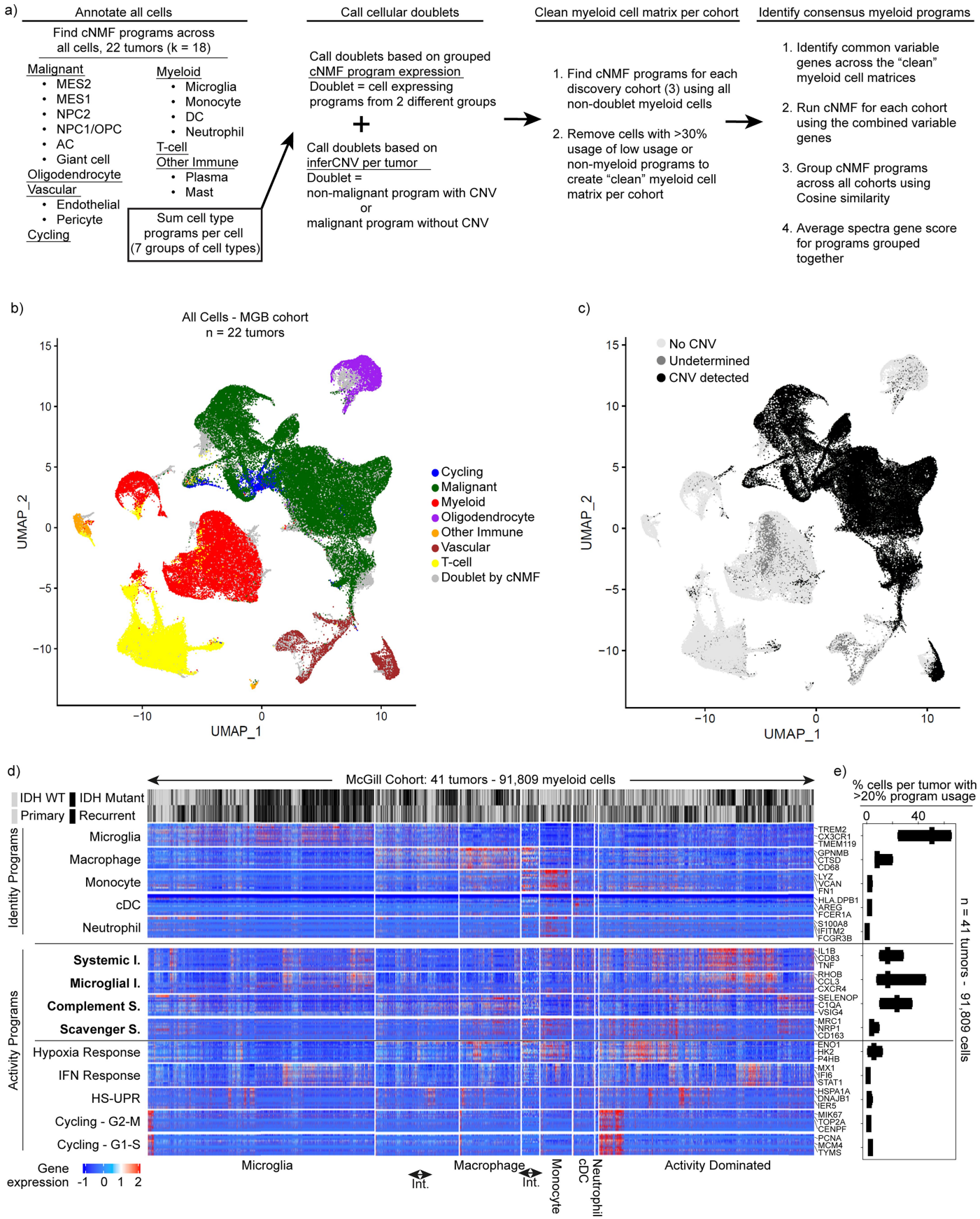
Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-025-08633-8.

Correspondence and requests for materials should be addressed to Bradley E. Bernstein.

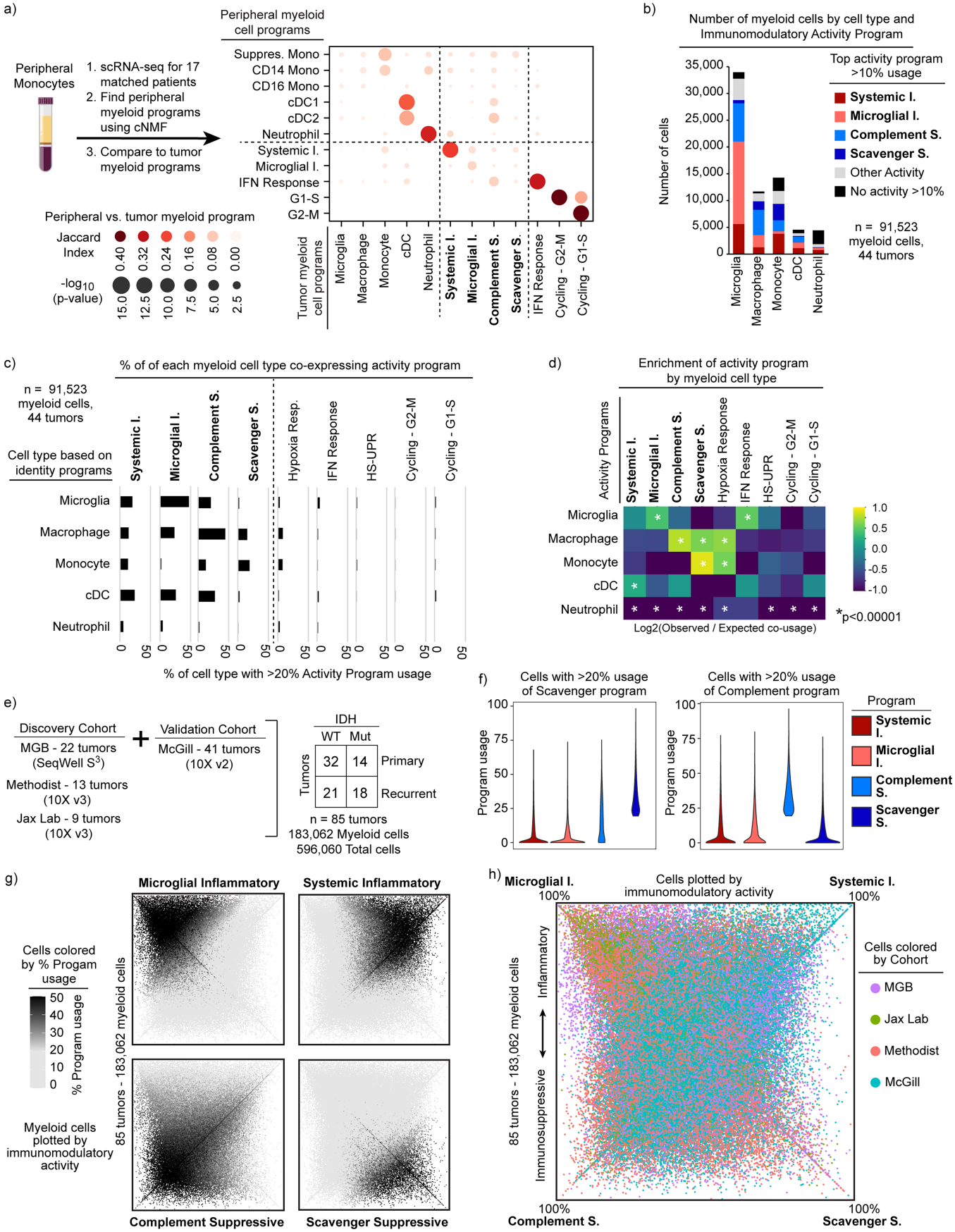
Peer review information *Nature* thanks Leila Akkari and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at http://www.nature.com/reprints.



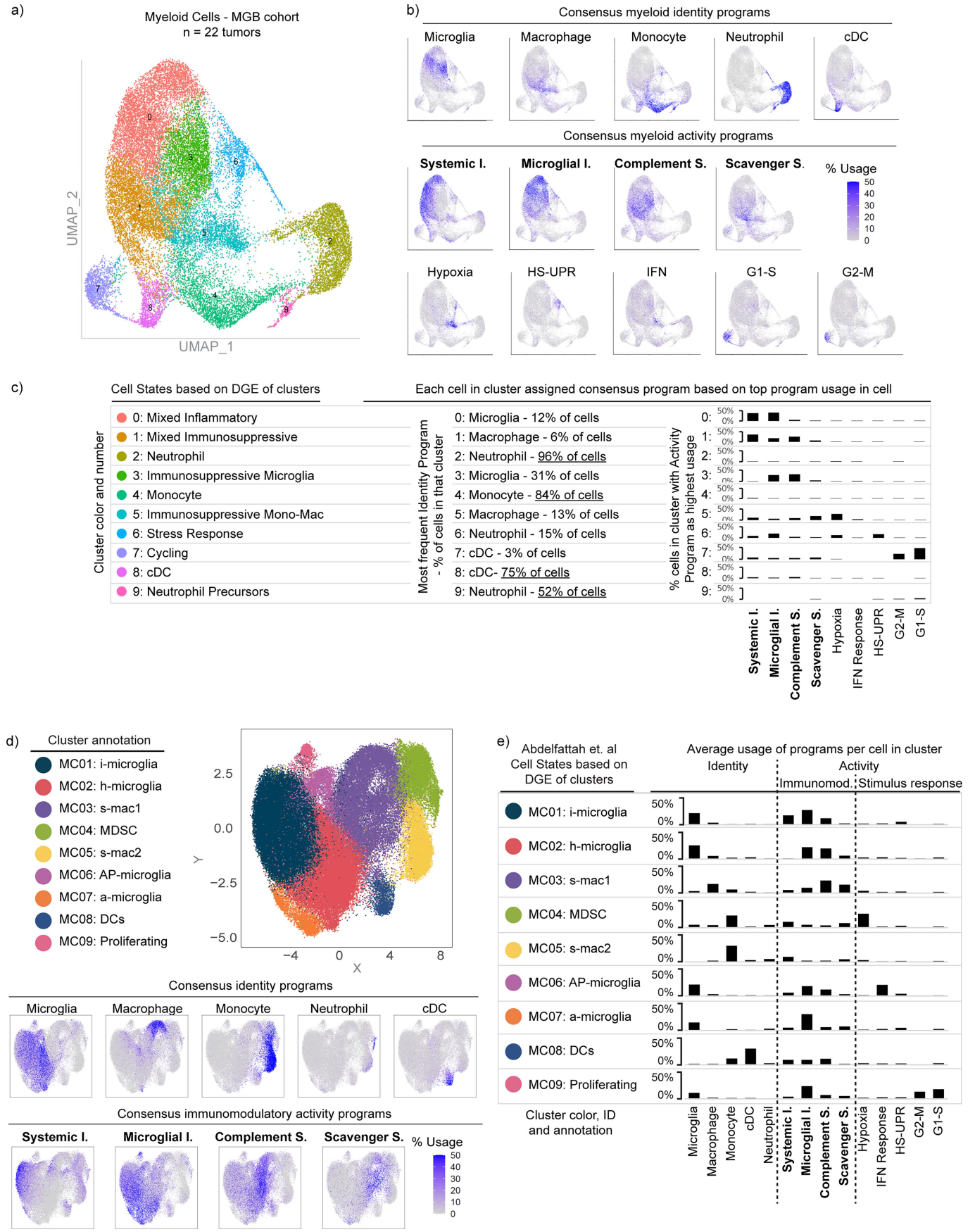
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Identification of consensus myeloid programs in discovery and validation cohorts. a) Computational pipeline used to identify recurrent myeloid programs in scRNA-seq data from three discovery cohorts, and applied separately to validation cohort (McGill). b) UMAP for all cells in the MGB discovery cohort colored by nominal cell type. c) Same UMAP with cells annotated by the presence (black) or absence (gray) of copy number variations. d) Heatmap depicts the expression of genes in recurrent myeloid programs (rows) across cells (column) grouped by cell type, analogous to Fig. 1c, but for the validation cohort (McGill). Cell type determined from the usage of cNMF myeloid identity programs. e) Boxplots depict the percentage of cells per tumor expressing the corresponding myeloid program (far left) across the validation cohort. Line represents median and boxes represent 1st and 3rd quartiles. Consensus myeloid programs were highly consistent between discovery and validation cohorts.



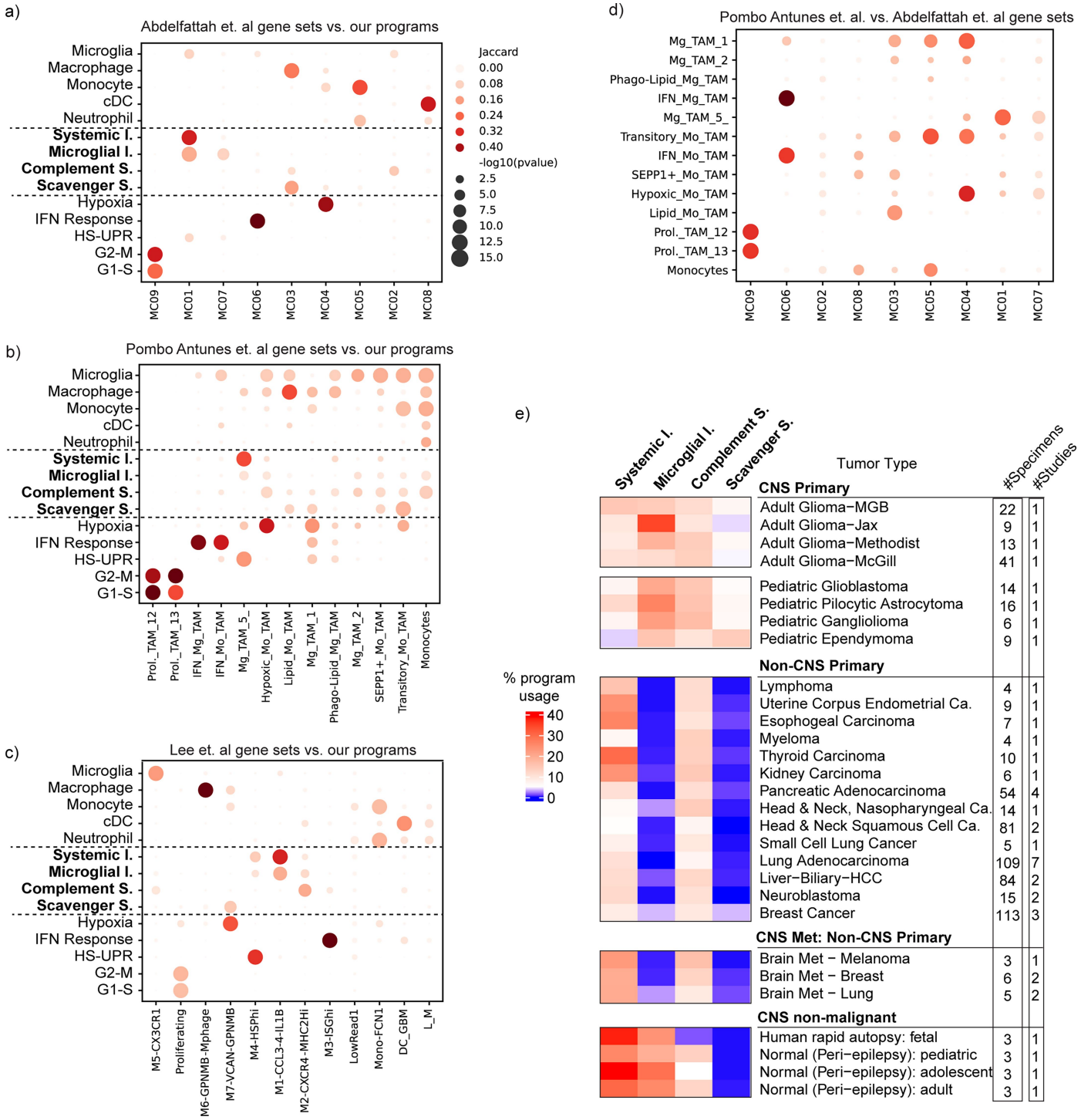
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Identification and comparison of cNMF programs in peripheral myeloid cells to tumor programs. a) cNMF programs identified from scRNA-seq profiles of peripheral blood from glioma patients. Plot compares cNMF programs between peripheral myeloid cells and glioma-associated myeloid cells. Heat and dot size correspond to Jaccard Index between gene spectra. b) For each myeloid cell type, stacked bar shows the absolute number of cells with top usage of indicated activity programs. c) For each myeloid cell type (rows), horizontal bars indicate the percent of cells with >20% usage of the indicated activity program. d) Heatmap shows relative enrichment of the indicated activity programs (columns) across the different cell identities (rows). P value calculated using hypergeometric distribution. e) Schematics depicts expansion of the cohort to include the McGill samples in all subsequent analyses. f) Violin plots shows the usage of each immunomodulatory program in cells with >20% usage of the Scavenger Immunosuppressive program (left) or the Complement Immunosuppressive program (right). g) Quadrant plots position cells (dots) by their expression of the immunomodulatory programs as in Fig. 1e. Grayscale depicts usage of the indicated immunomodulatory program. The position of each dot represents the difference in the usage of immunosuppressive and inflammatory programs by that cell (the upper part of the plot is more inflammatory, while the lower part is more immunosuppressive). h) Quadrant plot as in g but with cells colored by cohort. Overall, the four immunomodulatory programs are shared across myeloid cell types, but are rarely utilized by the same cells. PBMC tube, Shalek, A. (2005) https://BioRender.com/b15i535.

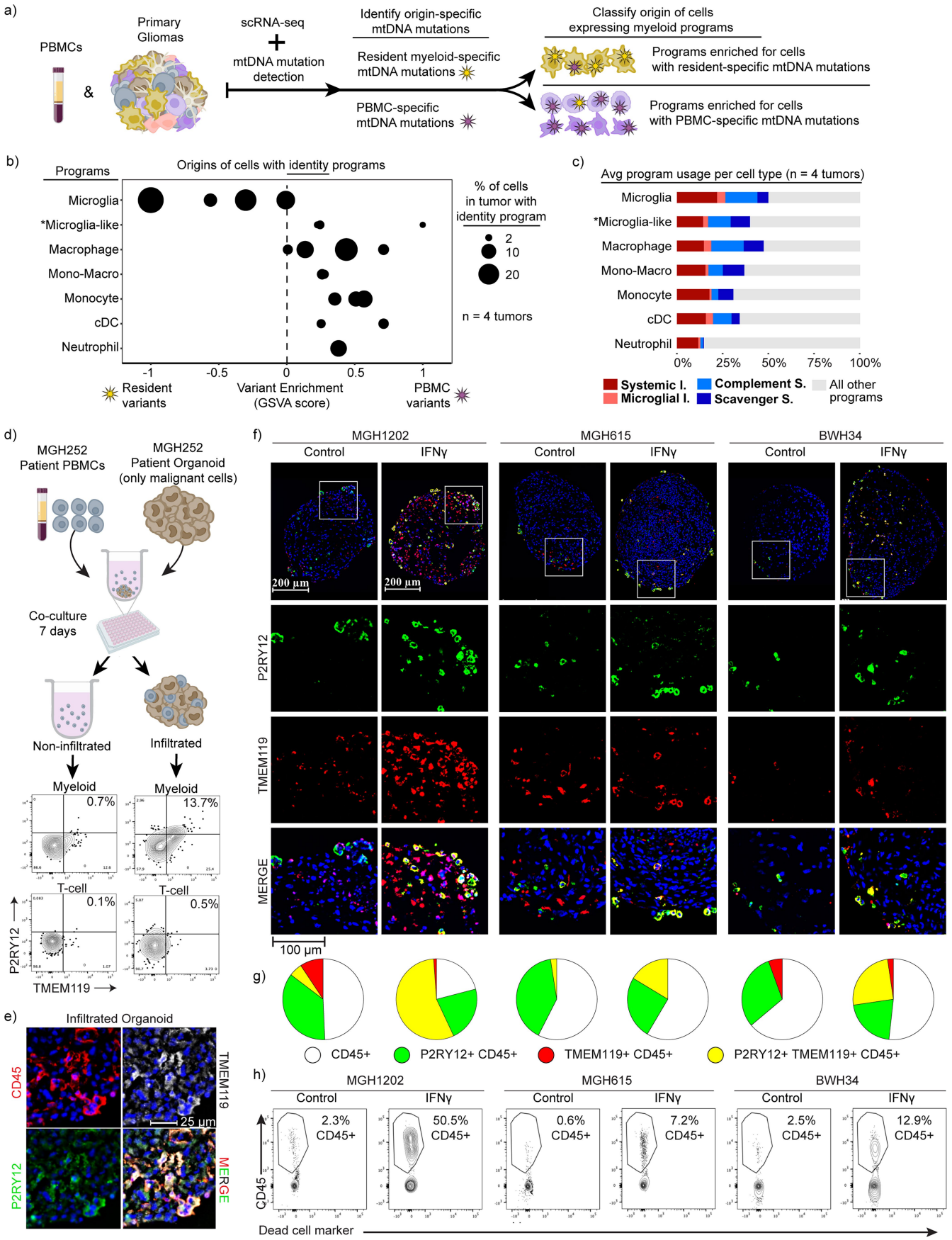


Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Direct comparison of Louvain clustering and cNMF programs. a) UMAP displays the Louvain clusters of batch-corrected singlet myeloid cells of the MGB cohort. b) UMAPs of the same MGB cohort with cells annotated by their usage of the cell identity (top) and cell activity programs (bottom). c) Louvain clusters in panel a were annotated based on differential gene expression analysis (left), or by the name and frequency of most frequent cell type per cNMF identity program usage (middle). For each Louvain cluster (rows), bar chart (right) indicates the percent of cells with high expression of the indicated activity program (columns). d) (Top) UMAP exhibiting the Louvain clusters of batch-corrected singlet myeloid cells of the Abdelfattah et al. cohort. (Bottom) UMAPs of the myeloid cells of the Abdelfattah et al. cohort demonstrating the usage of indicated programs at the top of each UMAP. e) (left) Annotations of Louvain clusters in (d) based on standard differential gene expression analysis of clusters by Abdelfattah et al. (Right) bar chart of the average usage of identity and activity cNMF programs in the cells of the Louvain clusters in (d). The clusters are composed of cells that are heterogeneous combinations of the cNMF programs, with the clusters largely segregating by identity program usage, while the activity programs are shared across clusters and cell types.

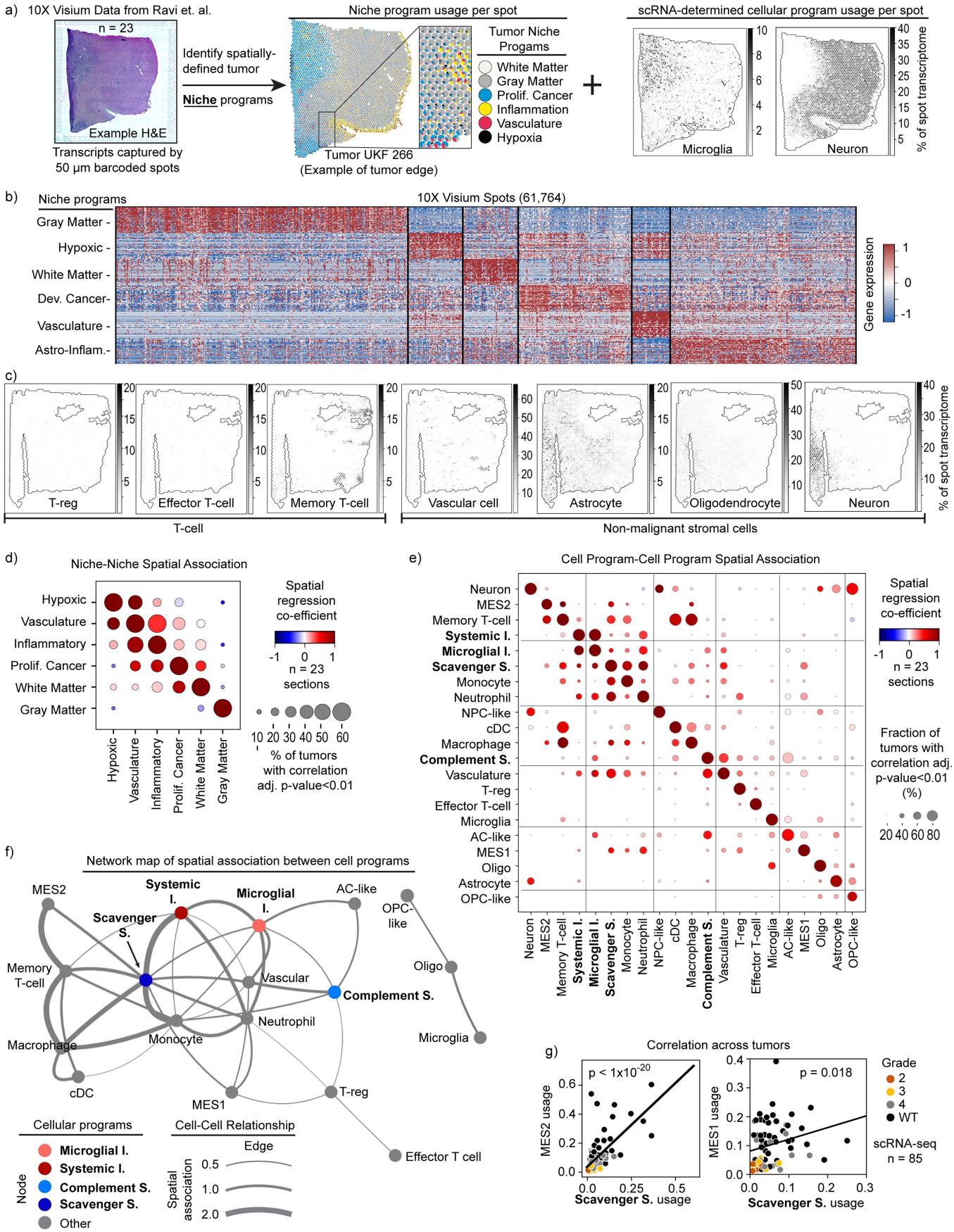


Extended Data Fig. 4 | Myeloid cNMF programs compared to published gene sets and evaluated in other tumor types. a-d) Plots compare top genes in our consensus myeloid cNMF programs to published gene sets for myeloid clusters. Heat and dot size correspond to Jaccard Index and significance between the top genes in our cNMF programs and gene sets derived from the published studies (a-c) or between gene sets from the published studies (d). Gene sets from the other studies were typically derived from differential gene expression of Louvain/Leiden clusters. Our immunomodulatory programs derived from cNMF were mostly distinct from prior gene sets. e) Heatmap of average usage of our consensus cNMF programs in myeloid cells from published scRNA-seq datasets for other tumor types. For cancer types with multiple studies, weighted average was used based on the number of tumors in the cohort. Full list of studies and sources in Supplemental Table 4. The Microglial Inflammatory program was unique to the brain, while the Scavenger Immunosuppressive program was specific to primary brain tumors.



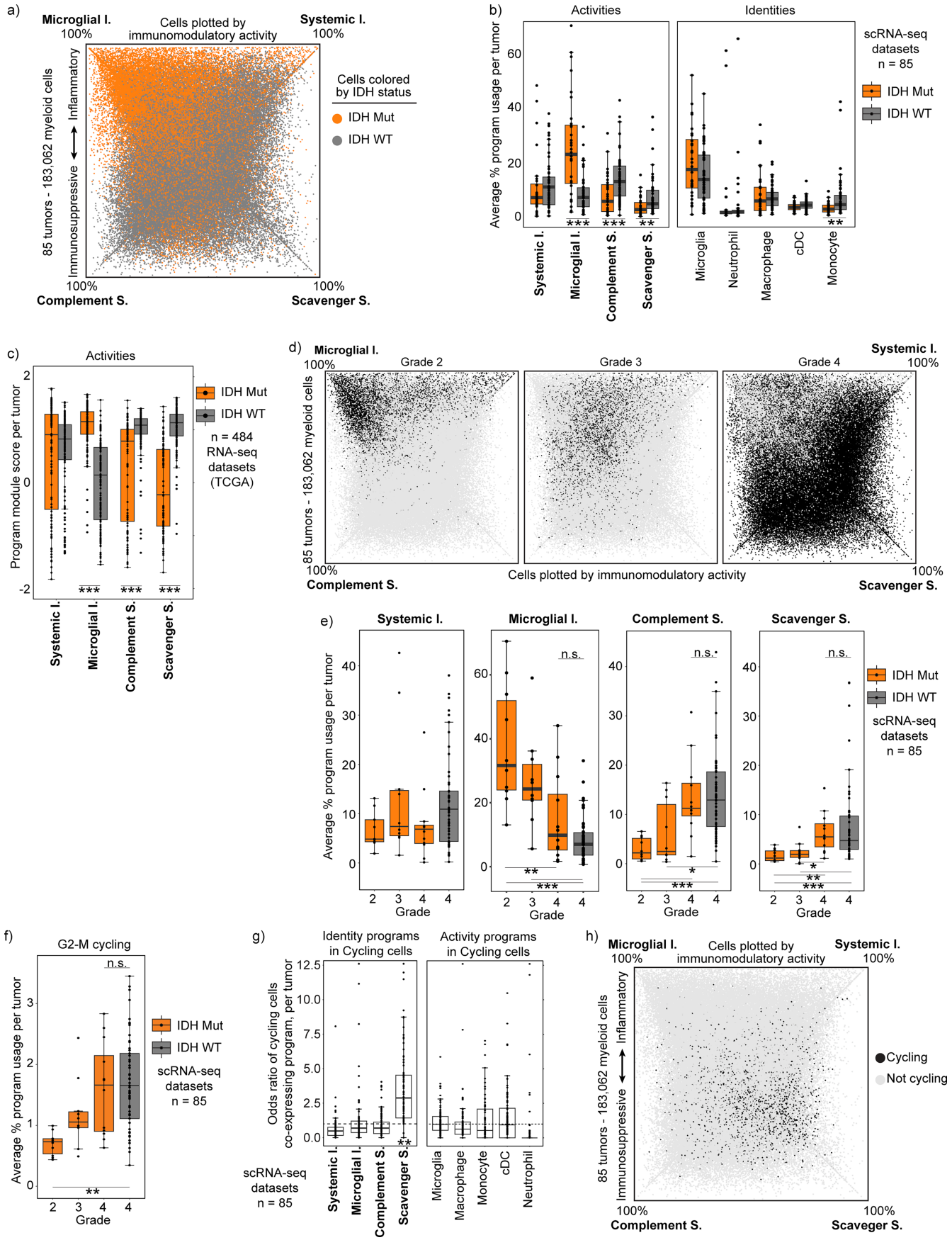
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Convergent phenotypes of microglia- and bone marrow-derived myeloid cells. a) MAESTER analysis pipeline for inferring the origins of myeloid cells in glioma based on detection of mtDNA mutations in scRNA-seq data. b) Plot depicts the enrichment difference between PBMC-specific and Resident-specific variants. Each dot represents the enrichment level for cells with the indicated identity in one patient. X-axis denotes the scaled difference between GSVA enrichment of PBMC-specific variants and Resident variants. c) Stacked bar charts indicate the average usage of the immunomodulatory programs in myeloid cells with the indicated cell identities for the four tumors shown in panel b. The "other programs" category encompasses the other identity and activity programs. d) Experimental design (top) and flow cytometry plots (bottom) for matched patient derived PBMCs applied to immune cell-depleted GBOs. Results are from a combination of 8 GBOs per condition. T cells were used as gating control for P2RY12 and TMEM119. e) Representative immunofluorescence images show PBMCs infiltrated into a GBO. f) Representative immunofluorescence images of organoid sections from different patients using the same experimental setup as in (d) except using normal donor monocytes. g) Pie charts depict quantification of images in (f). Results are from 4 GBOs combined for each condition. h) Flow cytometry plots evaluate the percent of CD45+ cell infiltration into the GBOs. Credit: 96-well plate and PBMC tube, Shalek, A. (2005) https://BioRender.com/b15i535.



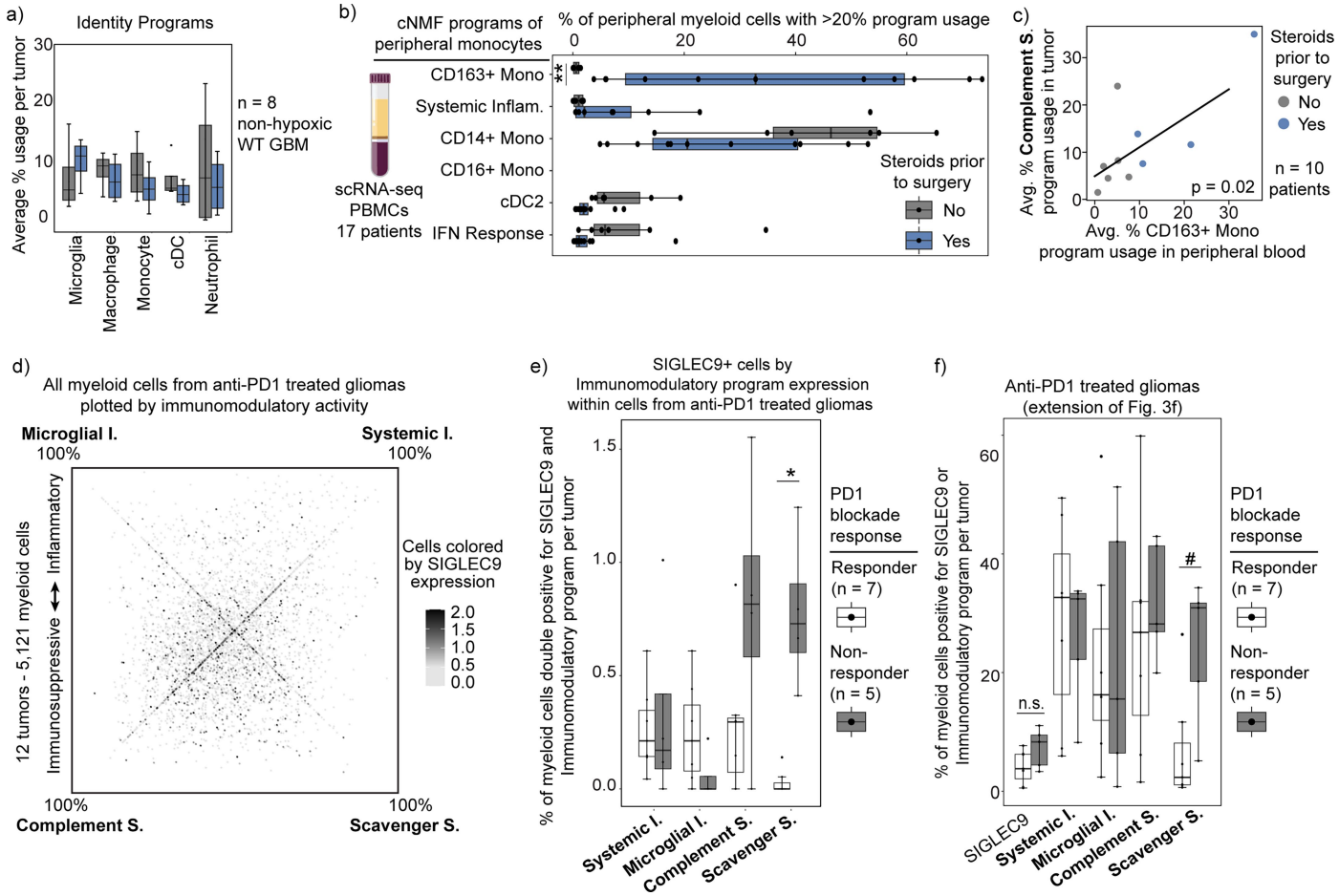
Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Spatial associations of myeloid cells and glioma niches. a) Schematic illustrates the analysis approach for spatial transcriptomics samples: cNMF was applied to spatial transcriptomic data from Ravi et al. to define broad transcriptomic tumor niche programs; in parallel, the cell content of each spot was inferred based on scRNA-seq signatures (RCTD demultiplexing; see Methods). Spatial Pie plot (middle) of a representative 10X Visium section depicts tumor niche programs. Each pie represents a spot and is colored by expression of the indicated tumor niche programs. Spatial plot (right) of the same section depicts relative expression of the indicated cell type programs (gray scale indicates RCTD-predicted spot proportions). b) Heatmap shows expression of genes (rows) grouped by niche program across all spots (columns) in the cohort of spatial transcriptomic samples. Top 40 genes of each niche program are shown. Gene expression data is column normalized, then log normalized and scaled by variance. c) Spatial plots show relative expression of cellular programs in the 10X Visium section shown in Fig. 2a (gray scale indicates RCTD-predicted spot proportions). d) Dotplot displays the spatial proximity enrichment score between niche programs, calculated independently per sample (p-value ordinary least squares, see Methods). Dot size denotes the proportion of samples showing a significant correlation (p-adj <0.01), while color signifies a positive (red) or negative (blue) correlation. e) Dotplot displays the spatial proximity enrichment score between cell programs, calculated independently per sample (p-value ordinary least squares, see Methods). f) Network graph illustrates recurrent spatial relationships of cell types in tumor microenvironment across spatial transcriptomic samples. Nodes denote cell types, with edges marking significantly enriched proximities between cell types, observed in at least 40% of samples with an average enrichment score >0.1. Edge width reflects this average score. g) Scatter plot exhibits the mean Scavenger Immunosuppressive program score (x-axis) versus the MES2 or MES1 cancer program score (y-axis) across glioma samples in the scRNA-seq dataset. Ordinary least square results are shown (line, and p-value).

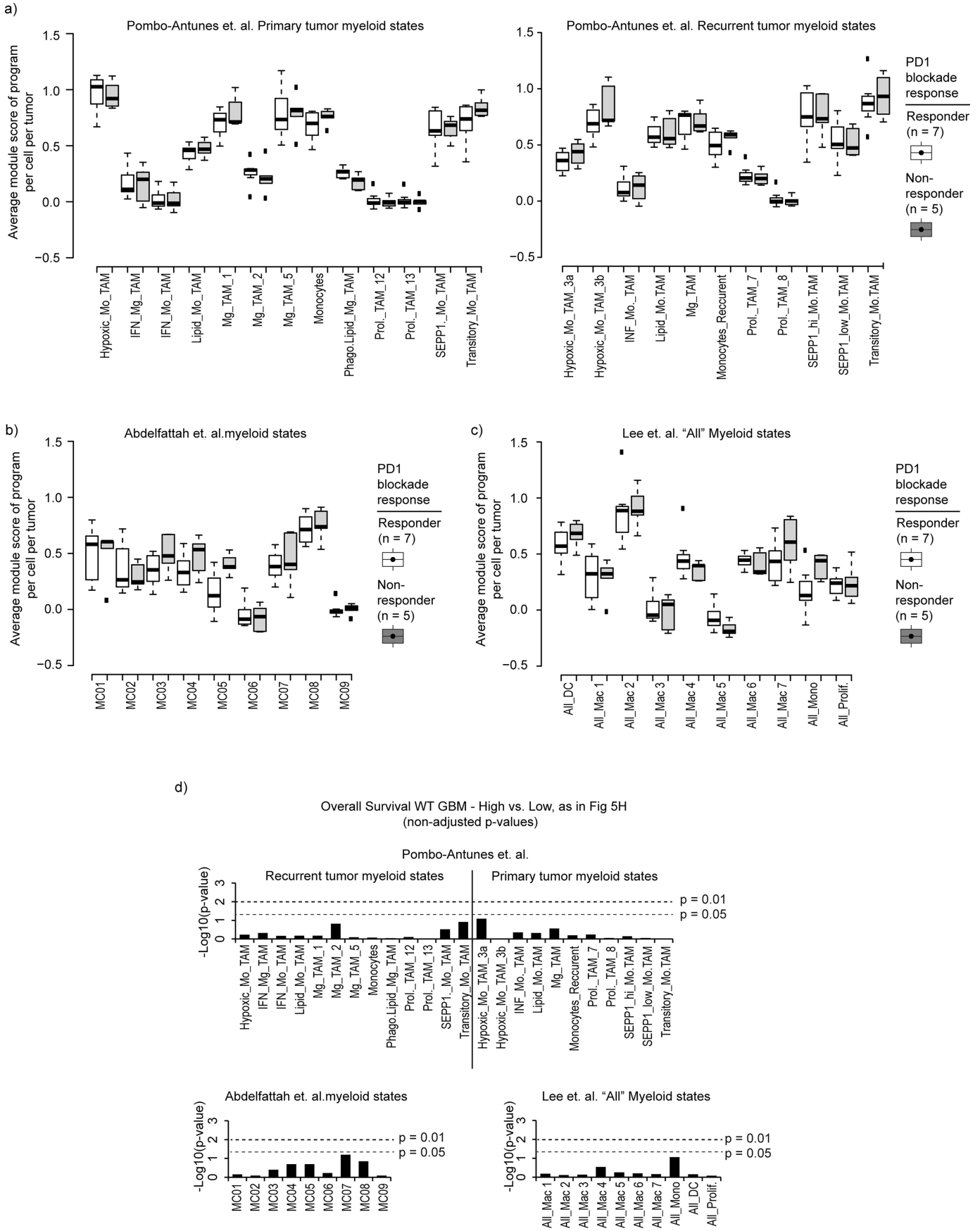


Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Myeloid program composition varies with histopathological tumor grade. a) Quadrant plot with myeloid cells positioned as in Fig. 1e with each cell colored by the *IDH* mutation status of the corresponding tumor. b) Box plots show the average usage of cell activity (left) and identity (right) programs in myeloid cells from the single-cell RNA-Seq cohorts, stratified by *IDH* mutation status of the corresponding tumor. For all box plots in figure, each dot represents a tumor, and lines represent median and quartiles. The whiskers represent 1.5 times the interquartile range. FDR corrected p-value calculated using two-sided Wilcoxon Rank Sum Test: * <0.05, ** <0.01, ***<0.001. c) Box plot exhibiting the program average module scores in tumors of the TCGA cohort stratified by *IDH* mutation status. d) Quadrant plots as in panel (a) in which the black color indicates whether a myeloid cell is from a grade 2 (left), grade 3 (center), or grade 4 (right) glioma. e) Box plots depict the average usage of each immunomodulatory program in myeloid cells from gliomas stratified by grade. f) Box plot shows the average usage of the G2M cycling program in myeloid cells in single-cell RNA-Seq data for the glioma cohorts, stratified by grade and *IDH* mutation status. g) Boxplot shows the odds ratio of cycling for myeloid cells expressing the indicated programs, calculated independently for each tumor. 'Cycling' and program defined by a cell usage >20% of both cycling and indicated program. h) Quadrant plot as in Fig. 3a with indication of cycling cells.

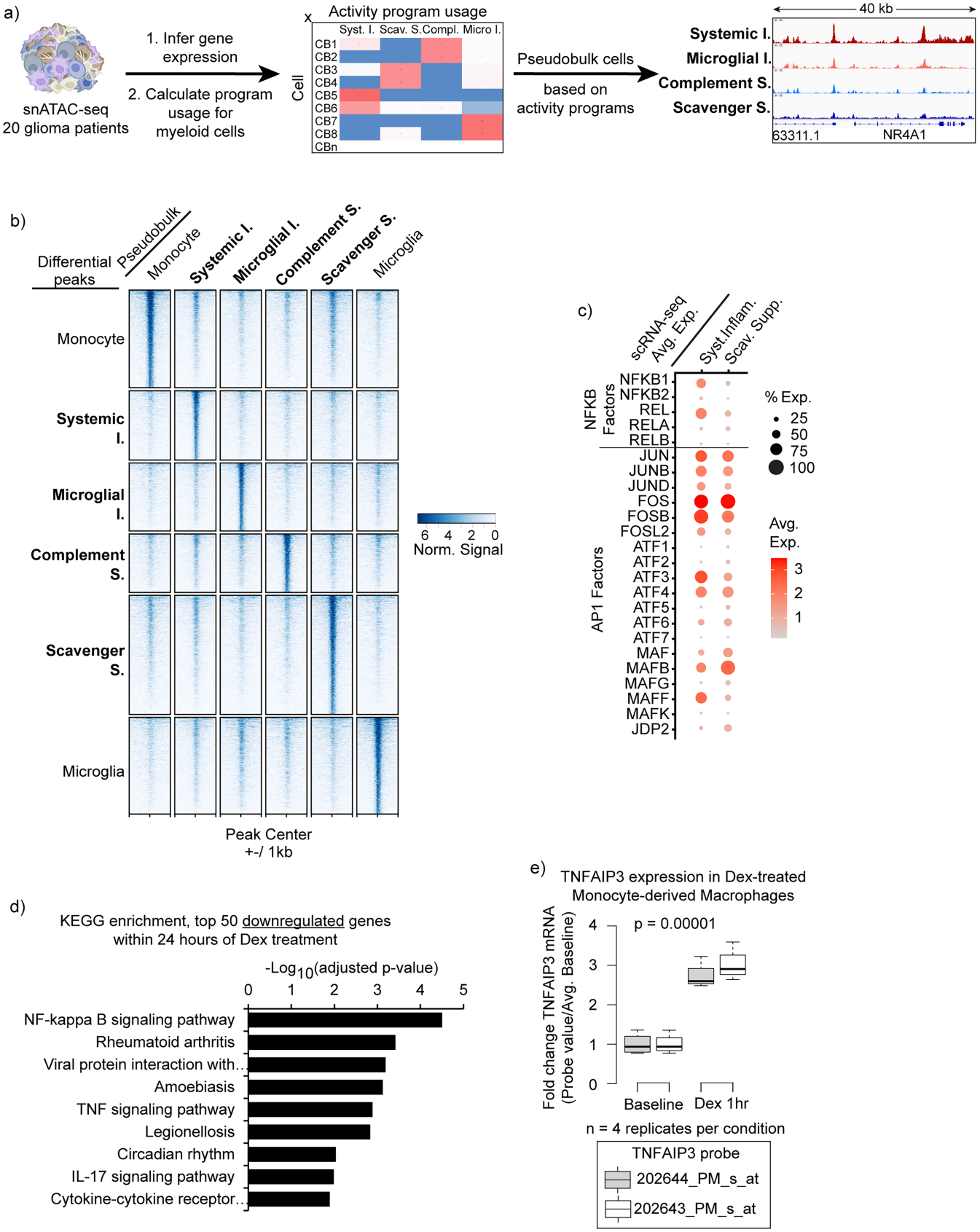


Extended Data Fig. 8 | Immunomodulatory programs associated with immunotherapy resistance. a) Boxplot displays the average usage of the myeloid cell identity programs stratified by use of dexamethasone in IDH-WT tumors with low hypoxic program usage in the MGB cohort. For all box plots in figure, each dot represents a tumor, and lines represent median and quartiles. The whiskers represent 1.5 times the interquartile range. FDR corrected p-value calculated using two-sided Wilcoxon Rank Sum Test: *<0.05, **<0.01, ***<0.001. b) Boxplot shows the percent of peripheral myeloid cells with the indicated peripheral myeloid programs in blood from glioma patients, stratified by administration of Dex treatment. c) Scatterplot depicts average Complement Immunosuppressive program usage in glioma-associated myeloid cells versus average CD163+ Immunosuppressive Monocyte usage in peripheral myeloid cells in matched patients stratified by Dex treatment. Only tumors with low hypoxic program usage are included. P-value from ordinary least-squares regression. d) Quadrant plot positions individual myeloid cells from an anti-PD1 immunotherapy trial[38] (as in Fig. 3g), based on the expression of our immunomodulatory activity programs, highlighting SIGLEC9 expression heterogeneity (grayscale). SIGLEC9 expression was not enriched in any of the immunomodulatory programs. e) Boxplot of SIGLEC9-positive cells from the same immunotherapy trial, grouped by expression of our immunomodulatory programs, then stratified by corresponding tumor response to immunotherapy. Average per tumor plotted. SIGLEC9 expression only stratified patients by response when it was co-expressed with the Scavenger Immunosuppressive program. f) Boxplot of per tumor calculation of cells positive for SIGLEC9 or immunomodulatory program (# p-value = 0.017, FDR = 0.088). When calculated on a per tumor basis, overall SIGLEC9 expression shows no significant difference between responding and non-responding patients. Credit: PBMC tube, Shalek, A. (2005) https://BioRender.com/b15i535.



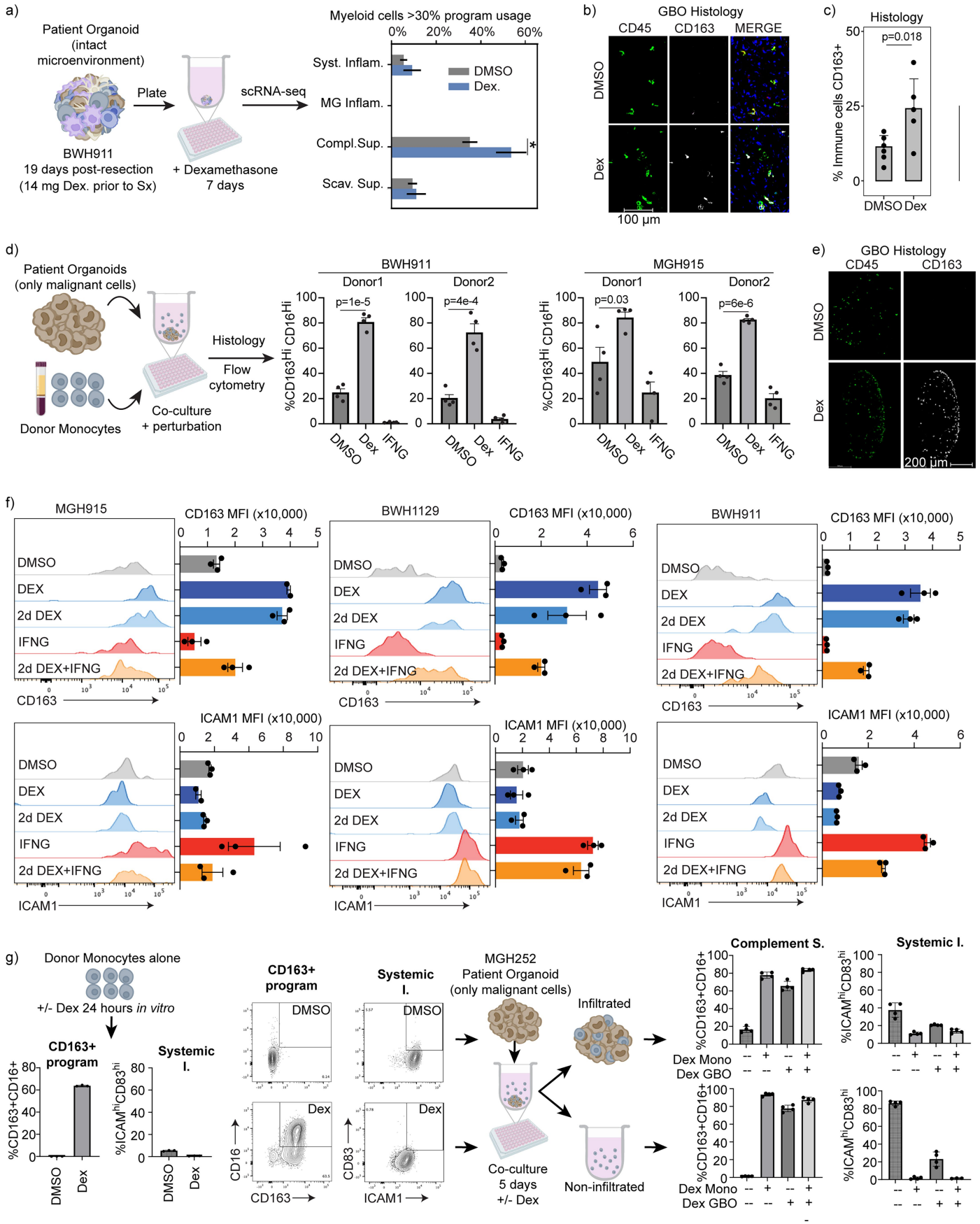
Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Comparisons of clinical outcomes to previously published gene sets. a-c) Boxplot shows the relative module score of published myeloid gene sets from Pombo-Antunes et al. (a), Abdelfattah et al. (b) and Lee et al. (c) in gliomas from Mei et al. stratified by response to immunotherapy. Boxplots represent median and quartiles, with whiskers representing 1.5 times the interquartile range. Significance calculated by FDR-corrected two-sided Wilcoxon Rank Sum Test. All were >0.25. d) Barplots show P-values calculated using a two-sided log-rank test (y-axis) to determine whether the indicated gene sets (x-axis) from Pombo-Antunes et al., Abdelfattah et al., and Lee et al. are associated with survival differences. None met significance of $p < 0.05$.



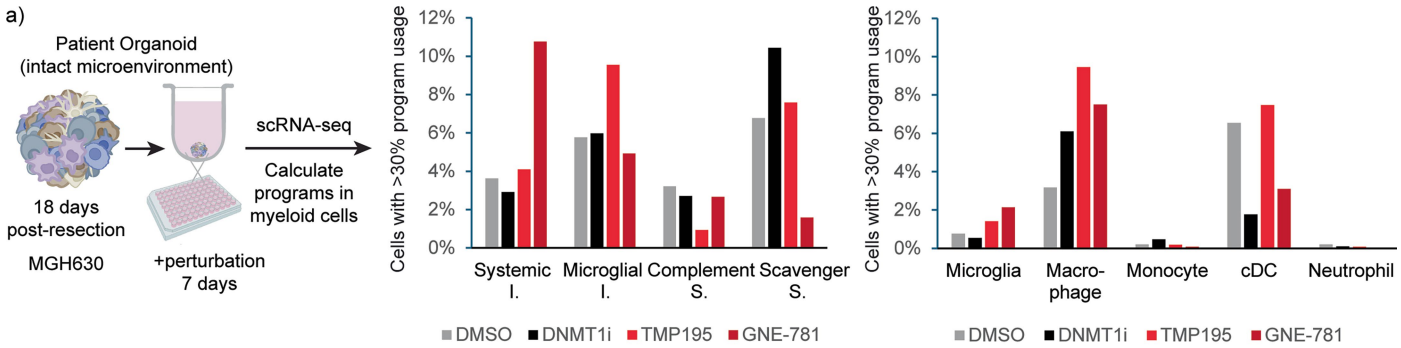
Extended Data Fig. 10 | See next page for caption.

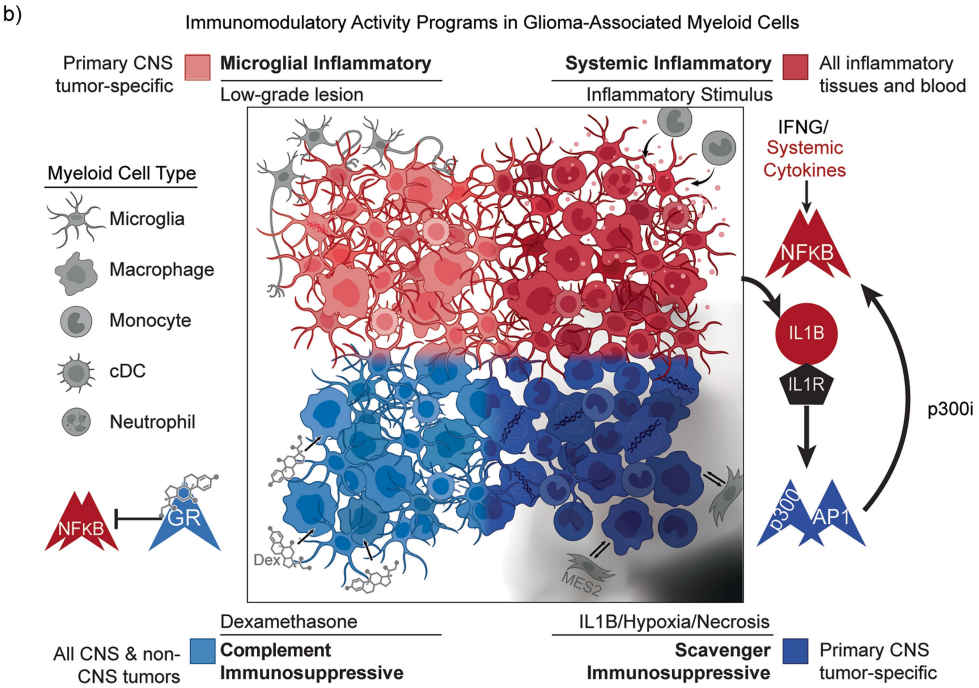
Extended Data Fig. 10 | TF circuits associated with the immunomodulatory programs. a) Schematic of the snATAC-seq analysis pipeline for deriving pseudo-bulk accessibility profiles for the respective immunomodulatory programs. b) Heatmap depicts relative accessibility over candidate regulatory sites for each immunomodulatory program, as in Fig. 4a but including Monocyte and Microglial populations. c) Heatmap shows the average TF expression and percent of cells expressing the TF in myeloid cells with discrete Systemic or Scavenger program expression. d) Barchart shows enrichments of 50 most downregulated genes within 24 h of Dex treatment in human macrophages (KEGG gene sets). Regulated genes derived from Table S1B from Jubb et al. P-value generated by gProfiler[79] using default settings. e) Boxplot of TNFAIP3 expression in macrophages after 1 h of Dex treatment (Jubb et al.). Boxplots represent median and quartiles, with whiskers representing 1.5 times the interquartile range. Significance is adjusted p value (FDR) taken from Table S1A of Jubb et al.

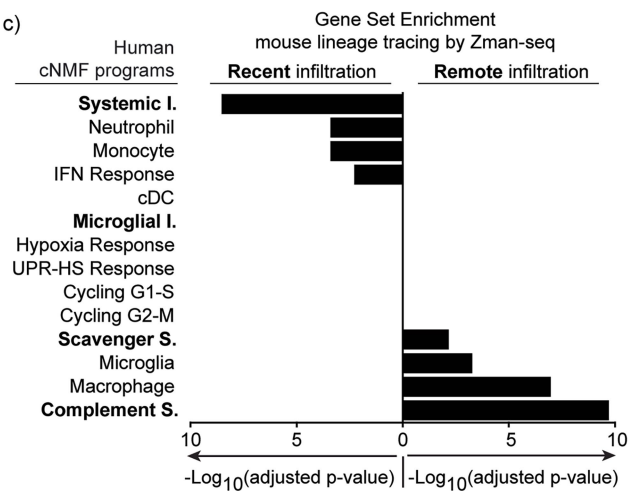


Extended Data Fig. 11 | See next page for caption.

Extended Data Fig. 11 | Dexamethasone rapidly and durably induces the Complement Immunosuppressive program. a) Experimental design and bar graph shows the percentage of myeloid cells expressing the immunomodulatory programs in scRNA-seq data for GBOs treated with Dex for 7 days or control. 16 GBOs from BWH11 were pooled per condition and sequenced; error bars were calculated from the binomial standard error. p-value = 0.018 by Fisher's Exact test, all other comparisons not significant. b) Immunofluorescence images of a GBO with intact endogenous TME cultured for 7 days with 100 nM Dex or DMSO control. c) Barplot shows mean quantification of marker positive cells in sectioned organoids. Each dot represents an organoid. P-value calculated with 2-tailed student's T-test. Error bars represent S.D. d) Experimental design and flow cytometry results for peripheral monocytes applied to immune cell-depleted GBOs in the presence of Dex or IFN-gamma. Data points represent 3 pooled GBOs. Error bars represent S.D. e) Representative histology section of a GBO in the experiment shown in (d). f) Flow cytometry results of GBOs from multiple patients show consistent phenotypes. Experimental design as in Fig. 5c. Top row shows CD163, a marker of the Complement Immunosuppressive program. Bottom row shows ICAM1, a marker of the Systemic Inflammatory program. BWH911 CD163 plot is the same as shown in Fig. 5c. Barplots show mean signal +/− S.D. Data points (4) each represent 2 pooled GBOs. g) Experimental design and results from Dex treatment of infiltrating GBO model. Flow cytometry results after 24 h of treatment, prior to adding cells to GBOs (bottom left). Barcharts show the mean of 3 wells of monocytes treated from the same patient. Error bars +/− S.D. Representative flow plot is shown. Middle schematic shows experimental design after pretreated cells were applied to GBOs. Barchart on the right shows mean flow cytometry results for infiltrated and non-infiltrated myeloid cells, stratified by treatment. Data points (4) each represent 3 pooled wells (for non-infiltrated) or GBOs (for infiltrated). Error bars +/− S.D. Peripheral monocytes exposed to Dex prior to being cultured with GBOs turned on Complement markers and maintained program marker expression after GBO infiltration and through differentiation. The Complement Immunosuppressive marker expression in the GBO was equivalent regardless of exposure timing or duration, indicating monocytes have a memory of Dex exposure even when put in different microenvironments. Credit: 96-well plate and PBMC tube, Shalek, A. (2005) https://BioRender.com/b15i535.







Extended Data Fig. 12 | See next page for caption.

Extended Data Fig. 12 | Impact of microenvironment and perturbations on immunomodulatory programs. a) Experimental design and barplot for evaluations of the impact of epigenetic inhibitors on glioma-associated myeloid cell activity and identity programs. 25 GBOs from MGH630 were pooled per condition and sequenced b) Schematic of the four immunomodulatory activity programs in glioma-associated myeloid cells, along with the microenvironmental associations, TF enrichments, and perturbations shown to induce or reverse program expression. Myeloid cell types depicted in each program quadrant are an approximation of the distribution of cell types expressing the respective program. Program distribution across CNS and non-CNS malignancies also noted. c) Bar plots of gene set enrichments from mouse lineage tracing data (Kirschenbaum et al.). Gene sets derived from Top 100 genes in myeloid cells most correlated with recent infiltration (closer to 12 h) or Top 100 genes most correlated with remote infiltration (being in the tumor for up to 48 h). Gene list taken from Kirschenbaum et al. Supplemental Table 2, sheet "Table_S3_Isotype_Control_Time". P-value generated by gProfiler[79] using default settings with a custom gene set matrix consisting of our cNMF programs. Credit: 96-well plate and PBMC tube, Shalek, A. (2005) https://BioRender.com/b15i535.

nature portfolio

Corresponding author(s): Bradley E. Bernstein

Last updated by author(s): 12/31/2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

Flow cytometry data were collected at: BD LSRFortessa X-20 (BD Bioscience) using DIVA software (BD Bioscience).
Immunofluorescence images were acquired on a Leica Thunder microscope using LAS X software Version 3.8.2.27713
Bulk and single-cell RNA-Sequencing libraries were sequenced on NextSeq500 instruments (Illumina).

Data analysis

Scripts and codes used to generate all the data in the study, including software versions, are available at:
https://github.com/BernsteinLab/Myeloid-Glioma

A shiny tool to calculate usage of the presented consensus myeloid programs for glioma-associated myeloid cells from other experiments can be found at:
https://github.com/BernsteinLab/Calculate_Myeloid_cNMF_Usage/

Genomic analysis tools utilized:
STARsolo V2.7.10b for scRNA-seq alignment
Seurat v4 for scRNA-seq analysis
inferCNV V1.3.3 to identify CNVs in single cells
cNMF V1.0.4 was used for program identification
Signac V1.12.0 for scATAC-seq analysis
MACS2 V2.2.9.1 was used for bulk ATAC-Seq peak calling
monaLISA V1.11.4 to identify significantly enriched TF motifs

Other software used for data analysis:
Flow cytometry data were analyzed with FlowJo (Treestar) version v10
All histological images were analyzed using the Qupath open source image analysis software v0.4.3.

Detailed description of computation methods is provided in Methods section.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The raw counts and processed dataset for the discovery and validation cohort are available at the single-cell portal with study ID: SCP2389 at: https://singlecell.broadinstitute.org/single_cell/study/SCP2389/programs-origins-and-niches-of-immunomodulatory-myeloid-cells-in-human-gliomas

For all scRNA-seq data generated for this manuscript (MGB and IDH-mutant cohort from McGill), individual-level raw data are deposited at dbGaP with Study Accession: phs003756.v1.p1 and can be browsed at the dbGaP study report page at: https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs003756.v1.p1

For other data reanalyzed in this study, Methodist data was downloaded from (GEO Series: GSE182109), Jax Labs data were downloaded from (https://ega-archive.org/datasets/EGAD00001007772), and published McGill data (IDH WT samples) were downloaded from (https://ega-archive.org/datasets/EGAD00001006206).

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Biological sex information is collected and reported, but no analysis was reported in this manuscript. However, we report the data as this dataset will be a resource for the field and others may want to run sex-specific analyses.
Reporting on race, ethnicity, or other socially relevant groupings	We do not collect nor report race, ethnicity, or other socially relevant groupings
Population characteristics	We report the following population characteristics, when available, for the patients included in the human tumor datasets analyzed in this study: Diagnosis, Grade, Recurrence_Status, Age, Sex, Tumor, Location, IDH Status, IDH_mutation, MGMT Methylation, Steroid Use, Steroid Daily Dose. All patients with recurrent tumors were treated with standard of care chemotherapy and radiation. This clinical information is reported in Supplemental Table 1.
Recruitment	Adult male and female patients undergoing neurosurgery for a brain tumor at Massachusetts General Hospital or Brigham and Women's Hospital were offered to consent to biobanking of excess surgical tissue under an IRB-approved biobanking protocol. There were no exclusionary criteria used other than patient's or patient's proxy's willingness to consent to participate. We do not anticipate any biases introduced through this recruitment strategy that would impact the results in this manuscript.
Ethics oversight	Adult male and female patients at Massachusetts General Hospital or Brigham and Women's Hospital (MGB) provided preoperative informed consent to take part in the study in all cases under the approved Institutional Review Board Protocol DF/HCC 10-417. Patients in other cohorts were consented according to their published methods. Previously unpublished patient data from the Montreal Neurological Institute was collected as reported with other tumors from McGill University (see methods).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No clinical trial or animal studies were conducted as part of this manuscript. Sample size of the main tumor cohort analyzed by scRNA-seq was chosen without power calculations, but was chosen based on experience and preliminary analyses with a smaller sample size. With less than 20 tumors, our program results became impacted by adding or subtracting individual tumors. To reach consensus for our cNMF gene expression programs, we also added additional cohorts and then found recurrent programs across cohorts to ensure these were not tumor specific. For other experiments, we sought to achieve at least 3 biological replicates.
Data exclusions	In general, all data were included. In analyses where certain data was excluded, this is explicitly stated in the text and/or methods.
Replication	The number of replicates for each experiment is indicated in the figures or figure legends. All experiments were repeated multiple times with different patients or different cohorts as stated in the figure legends and methods. Many key findings were confirmed using orth
Randomization	No clinical trial or animal studies were conducted as part of this manuscript. For experiments conducted using glioblastoma organoids (GBOs), individual GBOs were randomized into different conditions through mixing in a single plate and then distributing into a 96-well plate. For genomics experiments, all samples were used and therefore no randomization was needed nor used.
Blinding	We did not run a clinical trial or animal experiments that would have required specific blinding strategies. For the other experiments, for instance flow cytometry, immunofluorescence imaging, and sequencing experiments that are based on objective measurements, blinding was not required as results were not subjectively scored.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Anti-SOX2 antibody [20G5] Abcam; Cat# Ab171380, RRID:AB_2732072
Anti Iba1, Rabbit (for Immunocytochemistry) Fujifilm Wako; Cat# 019-19741, RRID:AB_839504
CD3 epsilon antibody [CD3-12] Genetex; Cat# GTX11089, RRID:AB_369097
Cleaved Caspase-3 (Asp175) Antibody #9661 Cell Signaling; Cat# 9661, RRID:AB_2341188
P2Y12 (extracellular) Polyclonal Antibody Thermo Fisher scientific; Cat# PA5-111827, RRID:AB_2857236
CD45 Monoclonal Antibody (YAML501.4) Thermo Fisher scientific; Cat# MA5-17687, RRID:AB_2539077
TMEM119 Polyclonal antibody Proteintech; Cat# 27585-1-AP, RRID:AB_2880915
Anti-EGFR (Ab-1) Mouse mAb (528) Millipore Sigma
Recombinant Anti-VSIG4 antibody [EPR22576-70] Abcam; Cat# Ab252933
CD163 mouse monoclonal antibody,clone 2G12 Origene; Cat# 50-167-6602, RRID:AB_2623740
Anti-Ki67 antibody Abcam; Cat# Ab15580, RRID:AB_443209
CD68 Monoclonal Antibody (KP1) Thermo; Cat# 14-0688-82, RRID:AB_11151139
MRC1(CD206) antibody Biolegend; Cat# 321102, RRID:AB_571923
NRP1 (CD304) Antibody Thermo; Cat# MA5-32179, RRID:AB_2809467
Nurr1 antibody Thermo; Cat# PA5-78097, RRID:AB_2736269
CD83 Antibody Biolegend; Cat# 305302, RRID:AB_314510
RHOB antibody Thermo; Cat# 711274, RRID:AB_2633147
CXCR4(CD184) Antibody Biolegend; Cat# 306502, RRID:AB_314608
P2RY12 Antibody-BV421 Biolegend; Cat# 392106
NRP1(CD304) Antibody- PerCp-Cy5.5 Biolegend; Cat# 304509
CD8 Antibody-BV711 Biolegend; Cat# 344374
TMEM119 Antibody-AF488 Abcam; Cat# AB225497
CD163 Antibody-PE Biolegend; Cat# 333606
CD11b Antibody-PE-Cy5 Biolegend; Cat# 301308

CD68 Antibody-PE-Cy7 Biolegend; Cat# 333816
TREM2 Antibody-APC R&D; Cat# FAB17291A
CD45 Antibody-AF700 Biolegend; Cat# 304012
ICAM1(CD54) Antibody-PacBlue Biolegend; Cat# 332716
CD69 Antibody-BV510 Biolegend; Cat# 310936
CXCR4 Antibody-BV605 Biolegend; Cat# 306522
CD16 Antibody-BV711 Biolegend; Cat# 302044
CD68 Antibody-BV785 Biolegend; Cat# 333826
CD206 Antibody-PE-Cy7 Biolegend; Cat# 321124
CD83 Antibody-APC Biolegend; Cat# 305312

Validation

All antibodies for flow cytometry used in this study were purchased from commercial sources and fully validated antibodies, including KO validated antibodies, were used when available. Well-referenced antibodies validated by other labs were also chosen when manufacturer-validated antibodies were not available. All antibodies used for flow cytometry were purchased from Biolegend, with the exception of TREM2 from R&D (validated with known positive and negative cell populations) and TMEM119 from Abcam (KO validated by Abcam). They are suitable for flow-cytometric analyses and were characterized and validated by Biolegend. These indications are available on the manufacturer's websites (https://www.biolegend.com/en-us/quality/quality-control); For IF and IHC staining, negative controls (isotype and secondary antibody staining) were performed.

Plants

Seed stocks

N/A

Novel plant genotypes

N/A

Authentication

N/A

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Flow cytometry was done based on a prior protocol7. In brief, an antibody cocktail was made by 1:1 ratio of Brilliant Stain Buffer(BD Horizon,566349) and PBS, and then antibodies/dyes were added. Single cell suspensions were washed by PBS +0.5% BSA in 1.5ml Eppendorf tubes or 96-well U bottom plates. Cells were pelleted by centrifugation at room temperature, 300 x g for 5 mins. After removing the washing buffer, cells were resuspended in 100ul staining cocktails via pipetting up and down ~10 times. Plates or tubes were covered to avoid light and stained in a dark at room temperature for 20-25 mins. Cells were washed by PBS+0.5%BSA, centrifuged at room temperature, 300 x g for 5 mins. Cell pellets were resuspended in 200 ul of PBS+0.5%BSA.

Instrument

Flow cytometry was processed on BD LSRFortessa X-20 according to the manufacturing procedure. UltraComp eBeads (Invitrogen, 01-3333-42) are used to pre-annotate the compensation.

Software

FlowJo V10 is used to process data analysis

Cell population abundance

The purity of sorted samples was analyzed by flow cytometry on the same instruments used for sorting.

Gating strategy

All immune cells were gated on FSC/SSC, using PBMC controls for each experiment. Single cells were gated on FSC-a/FSC-h. Zombie NIR Live/Dead dye was used to gate on live cells.

For all GBO perturbation experiments with myeloid cells, myeloid cells were defined CD45+/CD11b+. Untreated monocytes or PBMCs were used as gating controls for myeloid selection and binary populations to define the marker-positive populations.

For Extended Data Fig. 5, myeloid cells were defined as CD45+, and negative/low for CD3 and CD8, using T cells in PBMCs as the gating control for TMEM119 and P2RY12.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.