Discovery of peculiar cells in the peri-tumoral regions and peripheral blood of glioblastoma patients by single cell RNA-sequencing

Runwei Yang[1,2*], Guanglong Huang[1,2,3*], Jinglin Guo[1,2*], Yaomin Li[1,2], Haimin Song[1,2], Kaishu Li[1,2], Guozhong Yi[1,2,3], Zhiying Lin[1,2], Xiran Wang[1,2], Sidi Xie[1,2], Zhifeng Liu[4], Xiaofeng Shi[5], Ke Li[2], Songtao Qi[1,2,3#], Yawei Liu[1,2,5#]

1. Department of Neurosurgery, Nanfang Hospital, Southern Medical University, Guangzhou 510515, Guangdong, China;

2. The Laboratory for Precision Neurosurgery, Nanfang Hospital, Southern Medical University, Guangzhou 510515, Guangdong, China;

3. Nanfang Glioma Center, Guangzhou 510515, Guangdong, China;

4. Department of Critical Care Medicine, General Hospital of Guangzhou Military Command, 510010, Guangzhou, China;

5. Department of Neurosurgery, Longgang Central Hospital of Shenzhen, Shenzhen 518116, Guangdong, China.

*Runwei Yang, Guanglong Huang, and Jinglin Guo contributed equally to this work.

#Corresponding authors:

Yawei Liu, The Laboratory for Precision Neurosurgery, New laboratory building, Nanfang Hospital, No. 1838 Guangzhou Avenue North, Guangzhou, Guangdong, China, phone: (86)62786450, email: liuyawei@smu.edu.cn.

Songtao Qi, Department of Neurosurgery, Nanfang Hospital, No. 1838 Guangzhou Avenue North, Guangzhou, Guangdong, China, phone: (86)61641801, email: qisongtaonfyy@126.com.

Abstract

Recurrence is almost inevitable in glioblastoma (GBM) patients even after aggressive multimodal therapies. Most recurrences occur at the resection margin that is outside GBM's contrast-enhancing region (C region) of magnetic resonance imaging (MRI), and always left behind after surgery. Current knowledge about GBM cells outside C region is limited and thus impedes the post-operative therapy. In this study, single cell RNA-sequencing (scRNA-seq) was performed to dissect twelve GBM samples of C regions and outside C regions from four patients. We found a cluster of peculiar GBM cells outside C region different from the GBM cells with typical copy number variations (CNVs), which had only fragmental CNVs in chromosomes 1, 10, 11, 12 and 19, and were termed CNV 4 in this study. Furthermore, transcription factor-cofactor (TF-TCF) interaction network was constructed to uncover the transcriptional regulatory mechanism of GBM cells. The core signature of TF-TCF interaction network of CNV 4 GBM cells shows prognostic significance in GBM patients. Besides, sub-cluster of these peculiar GBM cells possess stem cell-like properties and have potential of tumorigenesis as revealed by in silico and in vitro analyses. To explore whether these peculiar GBM cells exist in the circulation of GBM patients, additional peripheral blood mononuclear cells (PBMC) samples from four GBM patients, three epilepsy patients and three healthy volunteers were analyzed by scRNA-seq. GBM cells were not found in PBMC, but peculiar proliferative immune cells were discovered in the PBMC of GBM patients but epilepsy patients and healthy volunteers. Collectively, these results indicated that peculiar GBM cells

outside C region are responsible for recurrence, and proliferative immune cells are specific to the PBMC of GBM patients. This study sheds light on the hidden and neglected parts of GBM, providing a new perspective to understand the nature of GBM recurrence, to develop monitoring strategy or promising immune therapy against GBM.

Key words

Glioblastoma, Recurrence, Single cell RNA-sequencing, Copy number variation, Transcriptional regulation, Proliferative immune cells

Introduction

Glioblastoma (GBM), the most common and lethal brain malignancy in adults [1, 2], is highly heterogeneous and infiltrative. These features lead to the failure of multimodal treatments and subsequent tumor recurrences. GBM is characterized by contrast enhancement in T1-gadolinium enhanced magnetic resonance imaging (MRI) and the surrounding non-enhancing region of abnormal T2-weighted/fluid-attenuated inversion recovery (FLAIR) signal [3-5]. The former region represents the dense cellular core of GBM with neovascularization and blood-brain barrier (BBB) disruption, which is always resected as much as possible during surgery [5]. The latter region is regarded as edematous tissue with infiltrating GBM cells, part of which is often left behind after surgery, becoming the target of post-operative treatment [5]. Most molecular characterizations of GBM were identified based on the studies of contrast-enhancing samples, and researchers tried efforts to avoid contamination of

other brain tissue [6-9]. However, 90% of GBM recurrences occurred at the resection margin even after total resection of the contrast-enhancing regions [10]. Besides, studies have demonstrated that recurrent GBM cells differed greatly from primary GBM cells in the aspects of genome, transcriptome, and microenvironment [7, 11-13]. Therefore, knowledge accumulated from the studies of GBM contrast-enhancing regions and therapies developed from this knowledge are insufficient to understand and overcome GBM completely.

On the other hand, more and more studies confirmed that the extent of resection (EOR) is positively associated with superior outcomes of GBM patients. For instance, Yan Michael Li and colleagues showed that additional resection of the surrounding T2/FLAIR abnormality beyond total contrast-enhancing resection prolonged median survival time of GBM patients [14]. In addition, lobectomy of GBM in non-eloquent areas benefited patients in both overall survival (OS) and progression-free survival (PFS) [15]. Furthermore, Hanna Mendes Levitin and colleagues identified GBM cells from the non-enhancing GBM margin, and revealed that the expression signature biased toward infiltrated GBM cells was negatively correlated with patients' survival [16]. These results suggest that the less GBM cells outside the contrast-enhancing region, the better patients will turn out. Meanwhile, these results emphasize the prognostic and therapeutic values to understand the residual GBM cells in more details. Although recent studies are beginning to focus on peri-tumoral region [3, 4, 16-18], there are many characteristics of these GBM cells and the underlying regulatory mechanisms remained to be uncovered.

In this study, we performed single cell RNA-sequencing (scRNA-seq) on isocitrate dehydrogenase 1-wild type (IDH1-wt) primary GBM. Samples from contrast-enhancing central region (C region), peri-enhancement edematous region (P region), and peri-edema normal region (N region) of four GBM patients were collected separately under the guidance of MRI. For each of these twelve samples, we carried out pathological examination and primary GBM cells culture. Remarkably, we found peculiar GBM cells in P and N regions that might be responsible for recurrence after surgery. To explore whether GBM cells exist in the circulation of GBM patients, additional peripheral blood mononuclear cells (PBMC) samples from four GBM patients, three epilepsy patients and three healthy volunteers were analyzed by scRNA-seq. To our surprise, proliferative immune cells were discovered in the PBMC of GBM patients but epilepsy patients and healthy volunteers.

Results

GBM cells show remarkable heterogeneity at the aspect of CNV

To comprehensively investigate the characteristics of primary IDH-wt GBM, we separately collected samples of C regions, P regions, and N regions from four GBM patients under the guidance of MRI (Figure 1A, Supplementary Table S1). Hematoxylin and eosin (HE) staining and immunohistochemistry (IHC) staining showed that C regions contained the histological hallmarks of GBM, including microvasculature proliferation and tissue necrosis, as well as high ratio of KI67-positive cells, while P and N regions had no typical GBM features but

resembled normal brain tissue in some extent (Figure S1A, B). After preparation of single-cell suspension, we profiled these 12 GBM samples and 4 corresponding PBMC samples by using 10×Chromium 3' Single Cell Platform. Totally, 72, 620 cells passed our filters of quality control, including 48, 770 cells from GBM samples and 23, 850 cells from PBMC samples (Figure 1B, C). To avoid the disturbance of tissue heterogeneity, GBM samples and PBMC samples were analyzed separately.

All of these cells were classified into malignant cells and non-malignant cells by the following approaches. First, cells were divided into different clusters according to Seurat "FindClusters" function, each of which represented a transcriptome-distinct cell group (Figure 1B, C). Second, differentially expressed genes (DEGs) of each cluster were identified through comparison to the rest of other clusters. Cell type specific genes in the DEGs were used to inferred the cell identities, and gene ontology (GO) enrichment analysis was performed for further confirmation. According to these approaches, we identified oligodendrocytes, macrophages, T cells and putative GBM cells in GBM samples (Figure 1B). T cells, B cells, and monocytes were identified in PBMC samples (Figure 1C). For the putative GBM cells that could not be classified as any normal cell types, CNVs were inferred from transcriptomic information by moving average method, which uses 100 genes in each chromosomal region as the moving window to smooth gene-specific expression patterns [8, 19-21]. This analysis algorithm enabled identification of large-scale amplifications and deletions in putative GBM cells, such as chromosome 7 gain and chromosome 10 loss, which are genetic hallmarks of most GBM cells (Figure 1D).

There was remarkable heterogeneity at the aspect of CNV among GBM cells, which was consistent with previous discoveries [8, 19, 20]. According to CNV profile, GBM cells were classified into four CNV patterns (Figure 1D). GBM cells of CNV pattern 1 had typical hallmarks of chromosome 7 amplification and chromosome 10 deletion, as well as apparent variations in chromosomes 1, 6, 11, 12 and 19. GBM cells of CNV pattern 2 were featured with regional loss of chromosomes 6, 10, 13, 14, 16, 19 and 22, while those of CNV pattern 3 showed amplifications of chromosomes 7, 12, 15 and 18. GBM cells of CNV pattern 4 had fragmentary variations in chromosomes 1, 6, 10, 11, 12, 13 and 19, but they were not as typical as the former three patterns. Notably, most GBM cells of CNV patterns 1, 2 and 3 (77.44%, 95.73%, and 98.90%, respectively), whose CNVs were similar with classical GBM hallmarks, came from C regions, whereas GBM cells of CNV pattern 4 were found mainly in P and N regions (59.43% and 35.05%, respectively) (Chi-square test, $P < 0.0001$) (Figure 1E).

CNV patterns contribute to transcriptomic status of GBM cells in limited extent and show potential transitional relationships

Classification of GBM cells according to CNV profiles was based on transcriptomic information, so we next analyzed the relationship between CNV patterns and transcriptomic sub-clusters. Totally, 6, 715 GBM cells were divided into 11 sub-clusters (Figure 2A). Interestingly, sub-cluster 3 accounted for 93.93% of CNV pattern 3, and sub-cluster 5 accounted for 94.31% of CNV pattern 2, while CNV pattern 1 and 4 were made up of diverse sub-clusters in relatively balanced

proportions (Figure 2B). From another perspective, we found that most transcriptomic sub-clusters were derived from multiple CNV patterns (Figure S2A, B). In other words, there was no strict correspondence between CNV patterns and transcriptomic sub-clusters, and GBM cells of different CNV patterns resembled with each other at the level of transcriptome in certain extent. We then analyzed the developmental trajectory of these GBM cells to explore their underlying relationship. As shown in Figure 2C, GBM cells of CNV pattern 1 and 4 were located across three states, while CNV patterns 2 and 3 GBM cells were mainly isolated in state 2. Interestingly, the results of pseudo-time suggested potential transitions from CNV pattern 4 to CNV patterns 1, 2, and 3, or from CNV pattern 1 to CNV patterns 2 and 3 (Figure 2C, Figure S2C).

Pathway enrichment analysis of DEGs among different states revealed disturbances in many pathways involved in the regulation of gene expression, such as AP1 pathway, glial cell differentiation, regulation of mitotic nuclear division, and canonical Wnt signaling pathway (Figure 2D). Notably, 23% of top 100 DEGs belong to transcription factors and transcription cofactors, including SOX2, SOX4, SOX11, ETV1, and CBX3, which played important roles in GBM as demonstrated by previous studies [22-28] (Figure S2D and Supplemental Table S2). Besides, intersection analysis of the DEGs of four CNV patterns showed that only 11 genes were commonly upregulated. Most of the other DEGs were CNV pattern-specific (Figure 2E). Consistently, pathway enrichment analysis of CNV pattern 1-specific DEGs and CNV pattern 4-specific DEGs indicated differences in the transcriptional regulation of gene

expression (Figure S2E, F). Thus, we inferred that CNV patterns and transcriptional regulation cooperated to shape GBM transcriptomic states.

Transcriptional regulatory networks orchestrate transcriptome of GBM cells and correlate with patient outcomes

To investigate the transcriptional regulation of GBM cells, we focused on the transcription factors (TFs) and transcription cofactors (TCFs). Figure 3A described the analysis strategy for TFs-TCFs interaction network construction. Briefly, we firstly picked out TFs and TCFs in DEGs of four CNV patterns respectively. TFs and TCFs cooperate to exert their functions, so we figured out their interaction relationships according to STRING database. To obtain confident interactions, only those derived from curated databases and experiments and with combined scores larger than 0.9 were considered (Figure S3, S4, S5, and S6). Sorted by the interaction degree of each TF, top three TFs were chosen as hub TFs for further analysis of differentially expressed target genes and the construction of TF-TCF-Target regulatory networks (Figure 3B, Figure S7, S8, S9, S10). We found that JUN played role in four CNV patterns, and MYC in CNV patterns 2, 3 and 4, while TP53 and E2F1 were specific to CNV pattern 1, SMAD2 specific to CNV patterns 3, and FOS specific to CNV pattern 4 (Figure 3B, C). Similarly, Hong Wang and colleagues also identified JUN and MYC as centered TFs in high-grade glioma (HGG) mouse models by combinatory profiling of proteome, phosphoproteome, and transcriptome [29], which proved the effectiveness of our analysis strategy. On the other hand, we identified

differentially expressed TFs and TCFs with fold changes larger than two compared to normal cells. Combining the TFs and TCFs of high interaction degrees and variations, we defined the core TFs and TCFs gene sets for four CNV patterns (Figure 3C, Figure S11A). The sub-networks of core gene sets made up 61.54%, 41.41%, 56.95%, and 58.56% of the whole TF-TCF interaction networks of four CNV patterns, respectively (Figure S3, S4, S5, and S6).

These CNV-specific core gene sets were derived from scRNA-seq of only four GBM patients, so we asked whether they were applicable in larger cohort. We found that these gene sets were also highly expressed in GBM tissue of The Cancer Genome Atlas (TCGA) database compared to normal brain tissue in Genotype-Tissue Expression (GTEx) database (Figure 3D, Figure S11B, C, D, E, F, G, H). Interestingly, CNV pattern 4-specific TFs JUNB and FOS showed opposite expression with comparison to other CNV patterns, and they played different roles in different types of cancers (Figure 3C, Figure S12A, B, C, D). JUNB and FOS were both upregulated in GBM, acute myelocytic leukemia (LAML), and pancreatic adenocarcinoma (PAAD), but downregulated in other malignancies, such as bladder urothelial carcinoma (BLCA), lung adenocarcinoma (LUAD), and ovarian serous cystadenocarcinoma (OV) (Figure S12A, B, C, D). Besides, we applied these gene sets to survival analysis of TCGA GBM cohort. To our surprise, only the core gene set of CNV pattern 4 was negatively correlated with both overall survival (OS) and disease free survival (DFS) (Figure 3E, Figure S13A, B). We then compared the expression of CNV pattern 4 core gene set between primary GBM and recurrent GBM. The results showed that JUNB,

ATF3, ACTN1, CEBPB, EGR1, and NFKBIA were significantly highly expressed in recurrent GBM, and they showed positively associated with each other (Figure S13C, D). As expected, these recurrence-related genes (RRGs) were negatively associated with DFS (Figure 3F).

GBM cells lacking typical CNVs possess properties of stem cell

Many published studies revealed that recurrence of GBM after surgery were often located at the peri-tumoral regions (P and N regions). Therefore, we wondered whether CNV 4 GBM cells in P and N regions have the ability of gliomagenesis. Analysis of CNVs compositions of each GBM sample showed that more than 94% GBM cells in the P and N regions of GBM-3 and GBM-4 belonged to CNV pattern 4 (Figure 4A, Figure S14A). In particular, 100% GBM cells in the N region of GBM-3 were clustered into CNV pattern 4 (Figure 4A). We then performed sphere culture of primary GBM cells derived from these samples, the results of which showed that GBM cells from C region, P region as well as N region could form GBM sphere and express stem cell markers CD133 (Figure 4B), indicating that these cells had the potential of tumorigenesis [22, 30-32].

To characterize the properties of CNV 4 GBM cells in more details, we furtherly analyzed their sub-clusters. As shown in figure 4C, GBM cells of CNV 4 were divided into 9 sub-clusters. Interestingly, GO enrichment analysis of sub-clusters' DEGs revealed that sub-cluster 3 cells were involved in neurogenesis, gliogenesis, and stem cell differentiation (Figure 4D, Figure S14B), and they were mainly located

in P regions (166/173) (Figure 4E). We noticed that GBM stem like-cell marker SOX2 was highly expressed in sub-cluster 3 (Figure 4D), and the results of IHC demonstrated that SOX2 positive and Nestin positive cells existed in both P and N regions (Figure S15A). On the other hand, some of the other sub-clusters' DEGs were enriched in the biological processes associated with angiogenesis and muscle structure development. These results brought to mind the GBM stem cell (GCS) and vasculogenic mimicry that can be formed by CD133 positive GCS [33, 34]. We then evaluated the resemblance of sub-clusters to endothelial cell, pericyte, and other central nervous system cell types, including astrocyte, oligodendrocyte, oligodendrocyte progenitor cell (OPC), neuron, and neural progenitor cell (NPC) (Cell type specific gene sets were listed in Supplementary Table S3) [35, 36]. As expected, the sub-cluster 3 expressed markers of astrocyte, oligodendrocyte, OPC and NPC simultaneously (11.97%, 5.36%, 24.24%, 8.03%, respectively), while the other sub-clusters' DEGs overlapped with endothelial cell or pericyte marker genes in different extents (from 0% to 21.57% and from 0% to 23.44%, respectively) (Figure 4F). In addition, we established stemness score, endothelial cell score and pericyte score according to the average expression of corresponding maker genes to estimate the properties of CNV pattern 4 GBM cells. These three scores revealed continuums rather than separate distribution of stemness and differentiations towards endothelial cell or pericyte (Figure S14C). Furthermore, pseudotime analysis by Monocle algorithm drew the developmental trajectories of CNV pattern 4 GBM cells based on their transcriptomic similarities [37, 38]. Consistent with the former result, sub-cluster 3

cells were located at initial state and the other sub-clusters were distributed along later states (Figure 4G, Figure S14D). Collectively, these results revealed that some GBM cells lacking typical CNVs possessed properties of stem cell and might differentiate into endothelial cell or pericyte.

Sub-clusters of intracranial immune cells and PBMC from GBM patients show abnormal phenotypes

Finally, we turned to the analysis of intracranial immune cells, the largest proportion in GBM tissue samples. CNV analysis demonstrated that these cells had uniform expression pattern but no hallmarks of GBM cell (Figure 5A). Notably, sub-clusters 12 and 17 showed special phenotypes (Figure 5B, Figure S16A). Both sub-clusters expressed the marker genes of microglia/macrophage, CSF1R, CD14, CD68, AIF1, and HLA-DRB1 (Figure 5B). Meanwhile, sub-cluster 12 simultaneously showed high expression of EGFR, PTPRZ1, SOX2, SLC35E3, and MDM2, the known GBM-associated genes; while sub-cluster 17 were featured by the expression of proliferation markers, KIAA0101, BIRC5, MKI67, NUSAP1, and TOP2A (Figure 5B). Gene ontology (GO) enrichment analysis of their DEGs indicated that sub-cluster 12 were associated with nervous system development (Figure 5C), and sub-cluster 17 were enriched in the biological processes of cell cycle (Figure 5D). However, when comparing to GBM cells, these cells exhibited the characteristics of immune cells (Figure 5E, Figure S16B).

Leukocyte-associated processes were matched in both sub-clusters 12 and 17 (Figure

5E, Figure S16B), so we wondered whether these cells came from peripheral blood. Besides, to explore whether peculiar GBM cells outside C regions also exist in the circulation of GBM patients, PBMC samples from these four GBM patients before surgery were analyzed by scRNA-seq. Interestingly, a cluster of proliferative cell was discovered in the PBMC of GBM patients (Figure 5F, G). However, we did not found cells expressing both monocyte markers and GBM-associated markers (Figure S16D, E). These proliferative cells were composed of T cells and B cells, but not monocytes (Figure 5F, Figure S16E). When comparing to GBM cells, these cells also showed involvement in immune system-related processes (Figure S16F). To investigate whether these cells were specific to the PBMC of GBM patients, we additionally examined PBMC samples collected from three healthy volunteers and three non-tumoral patients (epilepsy patients) by scRNA-seq. Clusters of cells highly expressing proliferation markers were not found in the PBMC of healthy volunteers and epilepsy patients (Figure S16G, H).

Proliferation is regarded as a hallmark of malignant cells. However, by comparing the DEGs of proliferative GBM cells and immune cells, as well as previously reported proliferative glioma cells [16, 19, 20], we found that different kinds of cells shared a common signature of proliferation, BIRC5, RRM2, AURKB, NUSAP1, and TOP2A (Figure 5H, Figure S16I, Supplementary Table S4).

Discussion

Recurrence is almost inevitable in GBM patients even after aggressive multimodal

therapies including surgery, radiotherapy, and chemotherapy [10]. Recently, targeted therapy and immunotherapy have been developed to eliminate the refractory GBM cells, yet the effects of which were far away for satisfaction [39-43]. It is worth noting that most characterizations about GBM focused on C region that will be removed during surgery, but recurrences often occur outside the C region, indicating the existence of GBM cells that diffusely invade P and N regions. Extensive investigations of primary GBM cells in C region have dissected GBM from genetic and epigenetic aspects thoroughly [6-8, 11, 13, 20, 44]. However, conclusions derived from these cells could be modestly generalized to GBM cells outside C regions as well as recurrent GBM cells, as many studies have revealed great differences between them, such as responses to irradiation, temozolomide, and lomustine [3, 4, 7, 11, 45]. On the other hand, many studies have showed that additional resection beyond total contrast-enhancing resection could favor patients' outcomes [14, 15]. Therefore, it is of vital prognostic and therapeutic significance to investigate the GBM cells in P and N regions directly to develop novel post-operative treatment for GBM patients.

Since the first isolation and characterization of GBM cells outside C region by Daniel L. Silbergeld and colleagues in 1997 [46], more and more investigators focused on this field and tried efforts to elucidate the features specific to these cells. However, previous studies reported diverse properties of GBM cells outside C regions compared to those in C region. For instance, Daniel L. Silbergeld found no difference in motility between GBM cells isolated from C region and P/N regions [46], while some other studies demonstrated more invasive features of GBM cells in P/N regions than those

in C region [45, 47]. Similarly, some researchers showed that GBM cells derived from P/N regions proliferated faster than those in C region did [45, 46], whereas some others observed opposite results [17, 48]. These seemingly contradictory results might be partially explained by the heterogeneity of GBM cells, which had been extensively revealed by multiomics studies [6-9, 13, 20]. Meanwhile, these results point out the fact that current understanding of GBM cells outside C region is incomplete. In this study, we applied scRNA-seq to GBM samples of C, P and N regions. We discovered peculiar GBM cells in P and N regions distinct from those in C region, characterized by CNVs. The characteristics of peculiar GBM cells in P and N regions have hardly been discussed before.

The origin of GBM cells remains unclear. Some studies proposed that GBM was derived from the malignant transformation of neural progenitor cells (NPCs)/neural stem cells (NSCs) [49-51]. No matter what the precursor cells of GBM are, at the aspect of genome, they are normal diploid cells, whereas GBM cells are characterized by trisomy 7 and monosomy 10, as well as some other remarkable CNVs (amplification of EGFR, PDGFRA, MET, and deletion of CNKN2A, PTEN, etc.) [52]. From normal diploid karyotype to diverse CNVs currently detected, from normal precursor cells to full-fledged malignant cells, there might be complicated courses of evolution. Cyril Neftel and colleagues had inferred heterogeneous CNV patterns of GBM cells based on transcriptome (both inter-tumoral and intra-tumoral) [8]. Notably, the GBM hallmarks of chromosome 7 amplification and chromosome 10 deletion were not detected in the GBM cells of some adult and most pediatric patients [8, 19]. Distinct from

most typical GBM cells, less CNVs were found in these peculiar GBM cells. Another study revealed that aneuploidy cells were found in only 68% and 32% of samples of C and P regions, respectively, and they made up 3 to 84% of all cells [4]. In this study, we identified four CNV patterns in twelve GBM samples. The CNV pattern 4 showed lower degree of CNVs than the other three patterns. Interestingly, more than 90% of CNV pattern 4 GBM cells were distributed in P and N regions of four GBM patients, while the other CNV patterns GBM cells were found mainly in the C region of particular patients, which was consistent with previous studies [17, 20]. Furthermore, pseudotime analysis revealed transitional trend from CNV pattern 4 to other CNV patterns. The homogeneous distribution and moderate CNV state of CNV pattern 4 GBM cells suggested the common mechanism of infiltration among highly heterogeneous GBM cells, and the possibility of CNV pattern 4 as the prodromal and transitional status of fully transformed malignant cells. It has been documented that recurrent GBM cells developed novel genetic variations compared to primary GBM cells, under the selective pressure of radiotherapy and chemotherapy [7, 11, 12, 53]. However, the dynamic changes of CNVs occurrence and progression, as well as the underlying regulatory mechanism remain largely unknown. Further experiments are needed to shed light on the processes of CNVs development.

Although CNV is a pivotal mechanism employed by cancer cell to regulate gene expression, transcriptome is orchestrated by multiple factors. Only 40 to 60% of the amplified genes in cancer cells were founded parallel to corresponding RNA transcripts [54]. Therefore, the inferred CNVs based on RNA-sequencing contained

information of both CNVs and other regulatory mechanisms of transcriptome, and meanwhile omitted some CNVs inconsistent with their RNA transcripts. Our data revealed that GBM cells of different CNV patterns were also controlled by different transcriptional regulation networks. In particular, the core TFs-TCFs gene set of CNV pattern 4 was negatively correlated with both OS and PFS of GBM patients. Despite prominent heterogeneity, we figured out some common TFs among four CNV patterns, such as JUN and MYC, which were also identified as central TFs by recently published study using multiomics to construct kinase-TF centered network in HGG [29]. These common TFs might play essential role in GBM cells. Typically, GBM cells can be classified into three molecular subtypes according to bulk transcriptomic profiles, the classical, mesenchymal, and proneural subtypes [6, 7]. To simplify this model, Li Wang and colleagues focused on the glioma stem-like cells (GSCs) of mesenchymal, proneural, and classical subtypes (mGSC, pGSC, and cGSC, respectively) [55]. They discovered that cellular states of GSCs ranged from mesenchymal root to proneural terminal, between which were intermediate population. Interestingly, three TFs specific to this intermediate population, ATF3, JUNB, and JUN, were also found in the core TFs-TCFs gene set of CNV pattern 4, which further supports our hypothesis that CNV pattern 4 might be shared transitional state of heterogeneous GBM cells. We also noticed that ATF3 is a bidirectional switch regulating the treatment responses of GBM cells. On the one hand, ATF3 is involved in resistance to temozolomide (TMZ), cisplatin, and ultraviolet light [56, 57], but on the other, it serves as a key mediator in proteasome inhibitor-induced GSC-selective apoptosis [58]. Besides, another two CNV

pattern 4-specific core TFs, FOS and JUNB, were proved to confer GBM cells with TMZ resistance [59]. These reports underscore the intrinsic resistance to chemotherapy of CNV pattern 4 GBM cells, as well as the necessity of more researches to overcome this obstacle.

Remarkably, sub-cluster of CNV pattern 4 GBM cells expressed CD133, and showed ability to form sphere, indicating the stem cell-like properties and potential of tumorigenesis. Consistently, SGM Piccirillo and colleagues showed that GBM cells isolated from P/N regions could form neoplasm in mouse model, although they had lower expression level of stem cell markers Nestin, SOX2, and Notch2 compared to those in C region [48]. Besides, CD133-positive and Nestin-positive invasive GBM cells were founded in murine models and they kept the ability of self-renewal [60, 61]. It has been demonstrated that CD133-positive glioblastoma stem cells (GSC) is the source of vasculogenic mimicry (VM) [33, 34] that leads to compromised effect of anti-angiogenesis therapy in GBM [62]. In this study, we found that sub-clusters of CNV pattern 4 GBM cells resembled endothelial cell and pericyte in different extents, which implied the possibility of CNV pattern 4 GBM cells participating in VM.

To our surprise, proliferative immune cells were discovered in GBM tissue, as well as the PBMC of GBM patients but healthy volunteers and epilepsy patients. Proliferative immune cells in GBM tissue and PBMC were macrophages and lymphocytes (T cells and B cells), respectively. Notably, proliferative T cells simultaneously expressed perforin and granzyme-related genes such as PRF1, GZMA, GZMB, GZMH, and GNLY, while proliferative B cells showed upregulated expression of

immunoglobulin-related genes such as IGKC, IGKV1-12, and JCHAIN, presenting the properties of plasm cell. These transcriptomic profiles suggested that they were highly responsive to GBM. Apart from refractory resistance to standard of care, GBM is notorious for being highly immunosuppressive both locally and systemically [63, 64], which increases challenges to immunotherapy in GBM. Although making up less than 1% of PBMC, these proliferative immune cells represent the seeds of revolution against GBM, and propose a new way to dig out the potential of immune system in GBM treatment. On the other hand, since these peculiar immune cells were specific to the PBMC of GBM patients, they could be suitable targets for recurrence monitoring in addition to MRI that showed limited efficiency in differentiating tumor recurrence and pseudoprogression [65]. Therefore, the biological function and underpinning of these peculiar immune cells remained to be explored in more details.

Taken together, this study discovered peculiar GBM cells in P and N regions distinct from GBM cells in C region. First, these cells have fragmental but not typical CNVs of GBM cells. Second, these cells resemble GBM cells of typical CNVs at transcriptomic level in certain extent and have potential to transform into the cellular states of them. Third, sub-cluster of these cells possesses stem cell-like properties, and may participate in the process of vasculogenic mimicry. Finally, yet importantly, the core signature of transcriptional regulation network of these cells is negatively correlated with GBM patients' outcomes. Further, we characterized both local and systemic immune responses under the circumstance of GBM lesion. We discovered proliferative immune cells that were highly responsive to GBM in both GBM tissue

and PBMC. The peculiar CNV 4 GBM cells and peculiar proliferative immune cells from GBM patients identified by this study can be ideal targets for diagnosis, therapy and prognosis.

Materials and methods

Samples acquisition

This study was approved by the Ethics Committee of Nanfang Hospital. Written informed consents were provided by all patients. Peripheral blood samples were obtained before surgery in Nanfang Hospital, Southern Medical University. GBM samples were obtained under the guidance of MRI during surgery. Normal brain samples were obtained during surgical approaches to ependymoma. Diagnosis of IDH1-wt GBM were confirmed by the Department of Pathology of Nanfang Hospital (Supplementary Table S1: Pathological diagnosis information of four GBM patients).

Hematoxylin and eosin (HE) staining and immunohistochemistry (IHC) staining

HE and IHC staining assays were performed to examine the histological structure and the protein expression in GBM samples and normal brain tissue. These assays were performed as previously reported [66, 67]. Briefly, specimen slices were obtained from paraffin-embedded tissue blocks with a microtome (Leica, EG1150H, Wetlzar, Germany), deparaffinized and rehydrated. For HE staining, slices were immersed in hematoxylin and eosin sequentially. For IHC staining, slices were further processed, including antigen retrieval, blocking of endogenous peroxidase, primary antibody

incubation, secondary antibody incubation, and nuclear staining. Finally, the slices were sealed with mounting medium for imaging. Antibodies used in this study were listed in Supplementary Table S5.

Primary culture of GBM cells and immunocytochemistry (ICC)

GBM samples were stored in ice and sent to the laboratory within one hour. Samples were minced with scalpel, and then enzymatically digested by the Tumor dissociation kit (Miltenyi Biotec, 130-095-929, Germany) on gentleMACS Octo Dissociator with Heaters (Miltenyi Biotec, 130-096-427, Germany) according to the manufacturers' instructions. To isolated stem-like GBM cells, primary GBM cells were cultured in DMEM/F-12 media (Thermo Fisher Scientific, Gibco, #8117165, Shanghai, China) supplemented with B27 (1×, Thermo Fisher Scientific, #17504044, Shanghai, China), EGF (20 ng/ml, Promokine, #C-60170, Guangzhou Jetway Biotechnology Company, Guangzhou, China), and FGF (20 ng/ml, Promokine, #C-60240, Guangzhou Jetway Biotechnology Company, Guangzhou, China), in which some GBM cells without stem cell-like properties would degraded. ICC assay was performed as previously reported [66].

Samples preparation and 10×Genomics single cell-RNA sequencing

GBM tissue samples were washed by precooled Hank's balanced salt solution (HBSS, Thermo Fisher, 88284, USA) twice, during which apparent vessels and pia matter were removed. Clean samples were dissociated mechanistically and enzymatically as

described above (Miltenyi Biotec, 130-095-929; gentleMACS Octo Dissociator with Heaters, 130-096-427, Germany). For peripheral blood samples, they were collected into EDTA tubes from GBM patients, epilepsy patients, and healthy donors. Isolation of PBMC was carried out with Ficoll-Plaque Premium (GE Healthcare, 17544203-1，USA) according to the protocol of manufacturer. Firstly, 4 mL of peripheral blood was added to 3 mL of ficoll solution slowly, which was then centrifuged at 400×g for 30 minutes at 18 Celsius degree. Next, PBMC layer was transferred into another tube for twice washes with PBS at 400×g for 15 minutes at 18 Celsius degree.

Subsequently, cells from GBM tissue samples or peripheral blood samples were resuspended in DMEM (Thermo Fisher, 11320033, USA) containing 10% fetal bovine serum (FBS) (Thermo Fisher, 10099141, USA), and were filtered by 70-μm nylon filter (Corning Falcon, 431751, USA). Filtered cells were subjected to 1×Red Blood Cell Removal Solution (Biogems, 64010-00-100, USA) for 5 minutes to remove red blood cells, and then cell debris were removed by Debris Removal Solution (Miltenyi, 130-109-398, Germany). Following twice washes and resuspension in PBS solution with 0.04% Bull Serum Albumin (BSA) (Thermo Fisher, AM2616, USA), cells were mixed with Trypan Blue (Thermo Fisher, T10282, USA) to assess their viability in a hemocytometer (Thermo Fisher, C10312, USA). Then, the appropriate volume for each sample was calculated for a target capture of 6, 000 cells according to the user guide of Single Cell 3’ Reagent Kits v2 (10×Genomics company, 120237-16, USA). Single cell droplet generation, reverse transcription, and cDNA library preparation were performed according to the protocol of manufacturer. Finally,

the libraries were sequenced on an Illumina Novaseq 6000 with 150bp paired-end sequencing. A median sequencing depth of 50, 000 reads/cell was targeted for each sample.

scRNA-seq data processing and identification of non-malignant cells

Raw sequencing data (bcl files) were converted to fastq files with Illumina bcl2fastq, version 2.19.1 and aligned to the human genome reference sequences (http://cf.10×genomics.com/supp/cell-exp/refdata-cellranger-GRCh38-1.2.0.tar.gz). The CellRanger 2.2.0 (10×Genomics) analysis pipeline was used to generate a digital gene expression matrix from this data according to its guidelines. The raw digital gene expression matrix (UMI counts per gene per cell) was filtered, normalized, and clustered using R 3.5.2 software (https://www.R-project.org/). Cells and genes filtering was performed as followings. For each detected cell: UMIs should be less than the (1-doublet rate) to exclude multiplet. The multiplet rates were determined according to the user guide of 10×Genomics. UMI should be larger than 8% percentile to exclude ambient bias [68]. Mitochondrial RNA should be less than 90% percentile to exclude devitalized cells. For each sample, the number of detected cell was controlled within 6, 000 by adjusting the baseline gene number. Then normalization and centralization of each cells according to their detected genes were performed, and outlier cells with z-score exceeding ±6 were excluded. We utilized “NormalizeData” and “ScaleData” functions in the Seurat package to normalize and scale the single-cell gene expression data. The highly variable genes (HVGs) across

single cells in each samples were figured out by Seurat "FindVariableGenes" function. Canonical correlation analysis (CCA) was applied to correct batch effects observed in twelve GBM tissue samples and four PBMC samples, respectively [69]. For the CCA, we used the union gene sets from top 2000 HVGs of each samples and the combined data were aligned using the first 30 CC dimensions. Aligned cells were clustered by Seurat "FindClusters" function using the first 20 aligned CC dimensions at a resolution of 0.5. Clustering results were visualized using t-distributed Stochastic Neighbor Embedding (t-SNE). DEGs of each cluster were identified by comparison with the rest of other clusters. Markers of cell types and GO enrichment analysis were combined to classify the identity of each cluster. According to these approaches, we identified oligodendrocytes, macrophages, T cells and putative GBM cells in GBM samples. T cells, B cells, and monocytes were identified in PBMC samples. Data analysis were performed using standard packages procedures including Seurat 2.3.4, Monocle 2.6.4, as well as some website platforms such as Gene Ontology (GO: http://geneontology.org/) [70, 71], Metascape (http://metascape.org/gp/index.html#/main/step1) [72], and Gene Expression Profiling Interactive Analysis (GEPIA2: http://gepia2.cancer-pku.cn/#index) [73].

CNV inference from scRNA-seq data

CNVs of putative GBM cells were evaluated according to transcriptomic information by moving average algorithm inferCNV [8, 19-21, 74, 75]. Briefly, all detected genes were sorted by their chromosomal location, and then the expression level of each gene in

each cells was calculated based on additional 100 neighbor genes (50 from upstream and 50 from downstream), which was set as a sliding window for each chromosome. In this way, gene-specific expression patterns could be smoothed in certain extent, and thus the CNVs' effects on transcriptome become prominent. Considering the cell lineages of GBM belong to central nervous system, only oligodendrocytes were used as normal karyotype reference. Hierarchical clustering was performed by "fastcluster 1.1.25" to identify CNV patterns [76].

Construction of TFs-TCFs-Targets regulatory network

Firstly, we obtained the gene lists of TFs and TCFs from Animal Transcription Factors Database (http://bioinfo.life.hust.edu.cn/AnimalTFDB/#!/) [77-79]. Then, gene lists of TFs and TCFs were intersected with DEGs of four CNV patterns respectively to pick out CNV pattern-specific TFs and TCFs. Interaction relationships of selected TFs and TCFs were identified in STRING database (https://string-db.org/) [80]. To find solid interaction relationship, only interactions from experiments and databases with minimum required interaction score of 0.9 were considered. TFs-TCFs interaction networks were constructed by Cytoscape software based on these confident interactions. Then, interaction degree of each TF was calculated to seek out hub TFs for each CNV pattern. The top 3 TFs of highest interaction degrees were chosen for target genes analysis in Cistrome Data Browser (http://cistrome.org/db/#/) [81, 82]. Target gene lists were intersected with DEGs of four CNV patterns respectively to select the differentially expressed target genes (DETGs). Finally, hub TFs, their neighbors in

TFs-TCFs interaction network and DETGs were used for construction of simplified TFs-TCFs-Targets regulatory network for each CNV pattern in Cytoscape software.

Availability of Data

scRNA-seq data of this study are available upon request on corresponding authors.

Statistics

The statistical analyses were performed in SPSS statistical software, version 21.0 (SPSS Inc., Chicago, IL, USA). Survival analysis of TCGA GBM cohort was performed in GEPIA2 platform, which employed Log-rank test for hypothesis test [73]. Rank sum test was performed for recurrence-related genes (RRGs) comparison between primary GBM and recurrent GBM of China Glioma Genome Atlas database (http://www.cgga.org.cn/index.jsp) (Figure S13C). Spearmen correlation coefficient was calculated for RRGs (Figure S13D). Significance cut-off: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; and ****, $P < 0.0001$.

Funding

This study was supported by the National Natural Science Foundation of China (Grant nos. 81773290, 81802505), Key R&D Program of Guangdong Province (2018B090906001), Guangdong Science and Technology Department (2017A030313497), PLA Logistics Research Project of China (18CXZ030, CWH17L020), Science and Technology Planning Project of Guangzhou

(201902020017).

Acknowledgements

We are grateful to Peng Fang, Mei Zheng, Jiaohua Jiang and their colleagues of Guangzhou SALIAI Stemcell Science and Technology Co., Ltd. (Guangzhou, China) for helpful discussion about this study with them.

Contributions

Conception and design of the work: YWL, STQ, RWY, and GLH.

Acquisition, analysis and interpretation of data: RWY, GLH, JLG, YML, HMS, KSL, GZY, ZYL, XRW, SDX, ZFL, XFS, KL, STQ, YWL.

Drafting of the manuscript: RWY and YWL.

Critical revision for important intellectual content: RWY, GLH, JLG, YML, STQ, YWL.

All authors approved the final version of this manuscript.

Competing interests

The authors declare no competing interests.

References

1 Chen W, Sun K, Zheng R *et al.* Cancer incidence and mortality in China, 2014. *Chin J Cancer Res* 2018; **30**:1-12.

2 Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin* 2019; **69**:7-34.

3 Lemee JM, Clavreul A, Menei P. Intratumoral heterogeneity in glioblastoma: don't forget the peritumoral brain zone. *Neuro Oncol* 2015; **17**:1322-1332.

4 Lemee JM, Clavreul A, Aubry M *et al.* Characterizing the peritumoral brain zone in glioblastoma: a multidisciplinary analysis. *J Neurooncol* 2015; **122**:53-61.

5 Administration NHCM. Diagnosis and treatment of brain glioma (2018). *Zhonghua Shen Jing Wai Ke Za Zhi* 2019; **35**:217-239.

6 Verhaak RG, Hoadley KA, Purdom E *et al.* Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell* 2010; **17**:98-110.

7 Wang Q, Hu B, Hu X *et al.* Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. *Cancer Cell* 2017; **32**:42-56 e46.

8 Neftel C, Laffy J, Filbin MG *et al.* An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. *Cell* 2019; **178**:835-849 e821.

9 Yuan J, Levitin HM, Frattini V *et al.* Single-cell transcriptome analysis of lineage diversity in high-grade glioma. *Genome Med* 2018; **10**:57.

10 Petrecca K, Guiot MC, Panet-Raymond V, Souhami L. Failure pattern following complete resection plus radiotherapy and temozolomide is at the resection margin in patients with glioblastoma. *J Neurooncol* 2013; **111**:19-23.

11 Barthel FP, Johnson KC, Varn FS *et al.* Longitudinal molecular trajectories of diffuse glioma in adults. *Nature* 2019; **576**:112-120.

12 Chen X, Wen Q, Stucky A *et al.* Relapse pathway of glioblastoma revealed by single-cell molecular analysis. *Carcinogenesis* 2018; **39**:931-936.

13 Lee JK, Wang J, Sa JK *et al.* Spatiotemporal genomic architecture informs precision oncology in glioblastoma. *Nat Genet* 2017; **49**:594-599.

14 Li YM, Suki D, Hess K, Sawaya R. The influence of maximum safe resection of glioblastoma on survival in 1229 patients: Can we do better than gross-total resection? *J Neurosurg* 2016; **124**:977-988.

15 Roh TH, Kang SG, Moon JH *et al.* Survival benefit of lobectomy over gross-total resection without lobectomy in cases of glioblastoma in the noneloquent area: a retrospective study. *J Neurosurg* 2019:1-7.

16 Levitin HM, Yuan J, Cheng YL *et al.* De novo gene signature identification from single-cell RNA-seq with hierarchical Poisson factorization. *Mol Syst Biol* 2019; **15**:e8557.

17 Darmanis S, Sloan SA, Croote D *et al.* Single-Cell RNA-Seq Analysis of Infiltrating Neoplastic Cells at the Migrating Front of Human Glioblastoma. *Cell Rep* 2017; **21**:1399-1410.

18 Gill BJ, Pisapia DJ, Malone HR *et al.* MRI-localized biopsies reveal subtype-specific differences in molecular and cellular composition at the margins of glioblastoma. *Proc Natl Acad Sci U S A* 2014; **111**:12550-12555.

19 Filbin MG, Tirosh I, Hovestadt V *et al.* Developmental and oncogenic programs in H3K27M gliomas dissected by single-cell RNA-seq. *Science* 2018; **360**:331-335.

20 Patel AP, Tirosh I, Trombetta JJ *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 2014; **344**:1396-1401.

21 Venteicher AS, Tirosh I, Hebert C *et al.* Decoupling genetics, lineages, and

microenvironment in IDH-mutant gliomas by single-cell RNA-seq. *Science* 2017; **355**.

22 Tam WL, Ng HH. Sox2: masterminding the root of cancer. *Cancer Cell* 2014; **26**:3-5.

23 Berezovsky AD, Poisson LM, Cherba D *et al.* Sox2 promotes malignancy in glioblastoma by regulating plasticity and astrocytic differentiation. *Neoplasia* 2014; **16**:193-206, 206 e119-125.

24 Ikushima H, Todo T, Ino Y, Takahashi M, Miyazawa K, Miyazono K. Autocrine TGF-beta Signaling Maintains Tumorigenicity of Glioma-Initiating Cells through Sry-Related HMG-Box Factors. *Cell Stem Cell* 2009; **5**:504-514.

25 Hide T, Takezaki T, Nakatani Y, Nakamura H, Kuratsu J, Kondo T. Sox11 prevents tumorigenesis of glioma-initiating cells by inducing neuronal differentiation. *Cancer Res* 2009; **69**:7953-7959.

26 Weigle B, Ebner R, Temme A *et al.* Highly specific overexpression of the transcription factor SOX11 in human malignant gliomas. *Oncol Rep* 2005; **13**:139-144.

27 Baysan M, Woolard K, Cam MC *et al.* Detailed longitudinal sampling of glioma stem cells in situ reveals Chr7 gain and Chr10 loss as repeated events in primary tumor formation and recurrence. *Int J Cancer* 2017; **141**:2002-2013.

28 Wang SQ, Liu F, Wang YH *et al.* Integrated analysis of 34 microarray datasets reveals CBX3 as a diagnostic and prognostic biomarker in glioblastoma. *J Transl Med* 2019; **17**.

29 Wang H, Diaz AK, Shaw TI *et al.* Deep multiomics profiling of brain tumors identifies signaling networks downstream of cancer driver genes. *Nat Commun* 2019; **10**:3718.

30 Najafi M, Farhood B, Mortezaee K. Cancer stem cells (CSCs) in cancer progression and therapy. *J Cell Physiol* 2019; **234**:8381-8395.

31 Manoranjan B, Chokshi C, Venugopal C *et al.* A CD133-AKT-Wnt signaling axis drives

glioblastoma brain tumor-initiating cells. *Oncogene* 2019.

32 Aghajani M, Mansoori B, Mohammadi A, Asadzadeh Z, Baradaran B. New emerging roles of CD133 in cancer stem cell: Signaling pathway and miRNA regulation. *J Cell Physiol* 2019; **234**:21642-21661.

33 El Hallani S, Boisselier B, Peglion F *et al.* A new alternative mechanism in glioblastoma vascularization: tubular vasculogenic mimicry. *Brain* 2010; **133**:973-982.

34 Ricci-Vitiani L, Pallini R, Biffoni M *et al.* Tumour vascularization via endothelial differentiation of glioblastoma stem-like cells. *Nature* 2010; **468**:824-828.

35 Lake BB, Chen S, Sos BC *et al.* Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. *Nat Biotechnol* 2018; **36**:70-80.

36 Darmanis S, Sloan SA, Zhang Y *et al.* A survey of human brain transcriptome diversity at the single cell level. *Proc Natl Acad Sci U S A* 2015; **112**:7285-7290.

37 Trapnell C, Cacchiarelli D, Grimsby J *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* 2014; **32**:381-386.

38 Qiu X, Mao Q, Tang Y *et al.* Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods* 2017; **14**:979-982.

39 Migliorini D, Dietrich PY, Stupp R, Linette GP, Posey AD, Jr., June CH. CAR T-Cell Therapies in Glioblastoma: A First Look. *Clin Cancer Res* 2018; **24**:535-540.

40 Weller M, Butowski N, Tran DD *et al.* Rindopepimut with temozolomide for patients with newly diagnosed, EGFRvIII-expressing glioblastoma (ACT IV): a randomised, double-blind, international phase 3 trial. *Lancet Oncol* 2017; **18**:1373-1385.

41 O'Rourke DM, Nasrallah MP, Desai A *et al.* A single dose of peripherally infused EGFRvIII-directed CAR T cells mediates antigen loss and induces adaptive resistance in patients with recurrent glioblastoma. *Sci Transl Med* 2017; **9**.

42 Gatson NT, Weathers SP, de Groot JF. ReACT Phase II trial: a critical evaluation of the use of rindopepimut plus bevacizumab to treat EGFRvIII-positive recurrent glioblastoma. *CNS Oncol* 2016; **5**:11-26.

43 Brown CE, Alizadeh D, Starr R *et al.* Regression of Glioblastoma after Chimeric Antigen Receptor T-Cell Therapy. *N Engl J Med* 2016; **375**:2561-2569.

44 Louis. DN, Ohgaki. H, Wiestler. OD, Cavenee. WK. WHO Classification of Tumours of the Central Nervous System (Revised 4th edition). *International Agency for Research on Cancer (IARC) 69372 Lyon Cedex 08, France* 2016.

45 Glas M, Rath BH, Simon M *et al.* Residual tumor cells are unique cellular targets in glioblastoma. *Ann Neurol* 2010; **68**:264-269.

46 Silbergeld DL, Chicoine MR. Isolation and characterization of human malignant glioma cells from histologically normal brain. *J Neurosurg* 1997; **86**:525-531.

47 Ruiz-Ontanon P, Orgaz JL, Aldaz B *et al.* Cellular plasticity confers migratory and invasive advantages to a population of glioblastoma-initiating cells that infiltrate peritumoral tissue. *Stem Cells* 2013; **31**:1075-1085.

48 Piccirillo SG, Dietz S, Madhu B *et al.* Fluorescence-guided surgical sampling of glioblastoma identifies phenotypically distinct tumour-initiating cell populations in the tumour mass and margin. *Br J Cancer* 2012; **107**:462-468.

49 Matarredona ER, Pastor AM. Neural Stem Cells of the Subventricular Zone as the Origin of

Human Glioblastoma Stem Cells. Therapeutic Implications. *Front Oncol* 2019; **9**.

50 Alcantara Llaguno S, Chen J, Kwon CH *et al.* Malignant astrocytomas originate from neural stem/progenitor cells in a somatic tumor suppressor mouse model. *Cancer Cell* 2009; **15**:45-56.

51 Sanai N, Tramontin AD, Quinones-Hinojosa A *et al.* Unique astrocyte ribbon in adult human brain contains neural stem cells but lacks chain migration. *Nature* 2004; **427**:740-744.

52 Reifenberger G, Wirsching HG, Knobbe-Thomsen CB, Weller M. Advances in the molecular genetics of gliomas - implications for classification and therapy. *Nat Rev Clin Oncol* 2017; **14**:434-452.

53 Kim J, Lee IH, Cho HJ *et al.* Spatiotemporal Evolution of the Primary Glioblastoma Genome. *Cancer Cell* 2015; **28**:318-328.

54 Weinberg RA, Abbate E, Abkowitz J, al. e. The biology of cancer (Second edition). United States of America: Garland Science 2014.

55 Wang L, Babikir H, Muller S *et al.* The Phenotypes of Proliferating Glioblastoma Cells Reside on a Single Axis of Variation. *Cancer Discov* 2019; **9**:1708-1719.

56 Le Mercier M, Lefranc F, Mijatovic T *et al.* Evidence of galectin-1 involvement in glioma chemoresistance. *Toxicol Appl Pharm* 2008; **229**:172-183.

57 Hamdi M, Popeijus HE, Carlotti F *et al.* ATF3 and Fra1 have opposite functions in JNK- and ERK-dependent DNA damage responses. *DNA Repair* 2008; **7**:487-496.

58 Yoo YD, Lee DH, Cha-Molstad H *et al.* Glioma-derived cancer stem cells are hypersensitive to proteasomal inhibition. *EMBO Rep* 2017; **18**:150-168.

59 Bandey I, Chiou SH, Huang AP, Tsai JC, Tu PH. Progranulin promotes Temozolomide

resistance of glioblastoma by orchestrating DNA repair and tumor stemness. *Oncogene* 2015; **34**:1853-1864.

60 Ye XZ, Xu SL, Xin YH *et al.* Tumor-associated microglia/macrophages enhance the invasion of glioma stem-like cells via TGF-beta1 signaling pathway. *J Immunol* 2012; **189**:444-453.

61 Molina JR, Hayashi Y, Stephens C, Georgescu MM. Invasive glioblastoma cells acquire stemness and increased Akt activation. *Neoplasia* 2010; **12**:453-463.

62 Mao JM, Liu J, Guo G, Mao XG, Li CX. Glioblastoma vasculogenic mimicry: signaling pathways progression and potential anti-angiogenesis targets. *Biomark Res* 2015; **3**:8.

63 Jackson CM, Choi J, Lim M. Mechanisms of immunotherapy resistance: lessons from glioblastoma. *Nat Immunol* 2019; **20**:1100-1109.

64 Chongsathidkiet P, Jackson C, Koyama S *et al.* Sequestration of T cells in bone marrow in the setting of glioblastoma and other intracranial tumors. *Nat Med* 2018; **24**:1459-+.

65 Wen PY, Macdonald DR, Reardon DA *et al.* Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group. *J Clin Oncol* 2010; **28**:1963-1972.

66 Yi GZ, Huang GL, Guo ML *et al.* Acquired temozolomide resistance in MGMT-deficient glioblastoma cells is associated with regulation of DNA repair by DHC2. *Brain* 2019; **142**:2352-2366.

67 Yang RW, Guo JL, Lin ZY *et al.* The combination of two-dimensional and three-dimensional analysis methods contributes to the understanding of glioblastoma spatial heterogeneity. *J Biophotonics* 2019.

68 Lun ATL, Riesenfeld S, Andrews T *et al.* EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol* 2019; **20**.

69 Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 2018; **36**:411-420.

70 The Gene Ontology C. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res* 2019; **47**:D330-D338.

71 Ashburner M, Ball CA, Blake JA *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000; **25**:25-29.

72 Zhou Y, Zhou B, Pache L *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 2019; **10**:1523.

73 Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res* 2017; **45**:W98-W102.

74 Tirosh I, Izar B, Prakadan SM *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 2016; **352**:189-196.

75 Haas B, C G, Brown M, Tickle T, Weitz E, al. e. inferCNV of the Trinity CTAT Project. Available from: https://github.com/broadinstitute/inferCNV.

76 Mullner D. fastcluster: Fast Hierarchical, Agglomerative Clustering Routines for R and Python. *J Stat Softw* 2013; **53**:1-18.

77 Hu H, Miao YR, Jia LH, Yu QY, Zhang Q, Guo AY. AnimalTFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors. *Nucleic Acids Res* 2019; **47**:D33-D38.

78 Zhang HM, Liu T, Liu CJ *et al.* AnimalTFDB 2.0: a resource for expression, prediction and

functional study of animal transcription factors. *Nucleic Acids Res* 2015; **43**:D76-81.

79 Zhang HM, Chen H, Liu W *et al.* AnimalTFDB: a comprehensive animal transcription factor database. *Nucleic Acids Res* 2012; **40**:D144-149.

80 Szklarczyk D, Gable AL, Lyon D *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 2019; **47**:D607-D613.

81 Zheng R, Wan C, Mei S *et al.* Cistrome Data Browser: expanded datasets and new tools for gene regulatory analysis. *Nucleic Acids Res* 2019; **47**:D729-D735.

82 Mei S, Qin Q, Wu Q *et al.* Cistrome Data Browser: a data portal for ChIP-Seq and chromatin accessibility data in human and mouse. *Nucleic Acids Res* 2017; **45**:D658-D662.

Figure legends

Figure 1. A, Workflow of this study. Samples from contrast-enhancing central region (C region), peri-enhancement edematous region (P region), and peri-edema normal region (N region) were collected separately under the guidance of MRI. Then, these samples were used for pathological examination, scRNA-seq, and primary culture of GBM cells. Meanwhile, PBMC of the corresponding patients were collected for scRNA-seq. B, t-SNE plot of 48, 770 cells from twelve tissue samples of four GBM patients. Left, clustering results of these cells. Right, sample origins of different clusters. C, t-SNE plot of 23, 850 cells from four PBMC samples of the corresponding GBM patients. Left, clustering results of these cells. Right, sample origins of different

clusters. D. Classification of CNV patterns inferred from transcriptomic profiles of 6, 715 GBM cells. CNV patterns 1, 2, and 3 were mainly patient-specific, while CNV pattern 4 contained GBM cells from four patients. E. Spatial distribution of GBM cells of different CNV patterns in C, P and N regions. Chi-square test, $P < 0.0001$.

Figure 2. A, t-SNE plot of 6, 715 GBM cells. B, Cumulative percentage chart showing different transcriptomic clusters compositions of four CNV patterns. C, Developmental trajectory of all GBM cells inferred by monocle 2 algorithm. D. Pathway enrichment results of DEGs among three states. Pathways involved in the regulation of gene expression were marked in red rectangles. E, Venn diagram showing the intersection of DEGs of four CNV patterns. Only 11 DEGs were shared by four CNV patterns. A large proportion of DEGs were specific to the corresponding CNV patterns.

Figure 3. A, Overview of the analysis strategy for construction of TFs-TCFs-Targets regulatory network. B, Top 3 TFs of highest interaction degrees (interacting neighbors) of four CNV patterns. C, Heatmap of core gene sets of TFs-TCFs interaction networks of four CNV patterns. The color gradation showed the log(Fold change) of DEGs. Common TFs and TCFs in different CNV patterns and CNV pattern-specific TFs/TCFs were arranged from left to right. D, Expression analysis of four CNV patterns' core gene sets based on The Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx) databases. E, Prognostic value of CNV pattern

4-core gene set in GBM patients of TCGA database. Expression value of gene set was defined as the average value of all genes in the corresponding gene set. Patients of expression values larger than 75th percentile were grouped in high signature group, and those of expression values less than 25th percentile were in low signature group. F, Prognostic value of recurrence-related gene set in GBM patients of TCGA database.

Figure 4. A, CNV patterns compositions of different samples from GBM-3 patient. B, Immunofluorescence of GBM spheres stained with glioma stem cell (GSC) markers, CD133. These GBM cells were derived from C, P and N regions of GBM-3, respectively. C, t-SNE plot of 1, 612 CNV pattern 4 GBM cells. D, Results of GO enrichment analysis of sub-cluster 3 differentially expressed genes (DEGs). DEGs were listed in rounded rectangles. The fold changes (FCs) of DEGs were reflected in color gradation. Larger FC was corresponding to deeper color. Top 10 enriched biological processes (BPs) with the least false discovery rates (FDR) were listed from left to right in circles. Higher enrichment fold was corresponding to larger circle. Black lines depicted connections between DEGs and BPs. E, Bar diagram of the spatial distribution of sub-cluster 3 among twelve GBM samples. F, Spider chart showing the overlap extents between DESs of different sub-clusters and different cell types. G, Developmental trajectory of CNV pattern 4 GBM cells inferred by monocle 2 algorithm.

Figure 5. A, Genomic profiling of 23, 850 immune cells in twelve GBM tissue

samples. B, Heatmaps showing markers of microglia/macrophage, malignant cell, and proliferative cell. C, D, Results of GO enrichment analysis of sub-clusters 12 and 17 DEGs compared to the rest of other sub-clusters, respectively. E, Results of GO enrichment analysis of sub-cluster 12 DEGs compared to GBM cells. F, Violin plot showing the expression of markers of proliferative cell, T cell, and B cell in the PBMC of GBM patients. G, Results of GO enrichment analysis of PBMC sub-cluster 10 DEGs compared to the rest of other sub-clusters. H, Venn diagram of common proliferative genes and another three proliferation-associated gene sets from published studies (Patel et al., Filbin et al., and Levitin et al.).

Supplementary figure legends

Figure S1. A, Hematoxylin and eosin (HE) staining and immunohistochemistry (IHC) staining of twelve samples from four GBM patients. C regions were featured by the histological hallmarks of GBM, including microvasculature proliferation and necrosis, as well as high ratio of KI67-positive cell. P and N regions resembled the normal brain tissue in different extents. B, HE and IHC of normal brain tissue obtained during surgical approach to ependymoma.

Figure S2. A, Cumulative percentage chart showing CNV patterns compositions of different transcriptomic clusters. B, Connections between CNV patterns and transcriptomic clusters. The width of connecting lines are parallel to the number of

cells. C, Developmental trajectory of GBM cells inferred by monocle 2 algorithm. Distributions of different groups along three state axes were showed separately. D, Top 6 DEGs among three states. E, Pathway enrichment results of CNV patterns 1-specific DEGs. F, Pathway enrichment results of CNV patterns 4-specific DEGs. Pathways involved in the regulation of gene expression were marked in red rectangles.

Figure S3, S4, S5, S6. TFs-TCFs interaction networks of four CNV patterns. TFs were listed in circles and TCFs in rounded rectangles. The sizes of TFs and TCFs were determined by the number of their connecting neighbors, which was defined as interaction degree. Upregulation of TFs and TCFs was indicated in red color, while downregulation was in blue color. Larger absolute value of fold change was corresponding to deeper color. The core gene sets of TFs-TCFs interaction networks were listed at the right side and the parameters were listed at the bottom.

Figure S7, S8, S9, S10. Simplified TFs-TCFs-Targets regulatory networks of four CNV patterns. TFs were listed in inner layer (circles), TCFs in second layer (diamond), and targets of TFs in outermost layer (rounded rectangles). Upregulation of expression was indicated in red color, while downregulation was in blue color. Larger absolute value of fold change was corresponding to deeper color. Interactions between TFs and TCFs were depicted by straight line. TFs and their corresponding targets were connected by lines with single arrow pointing to target.

Figure S11. A, Heatmap of core gene sets of TFs-TCFs interaction networks of four CNV patterns. The color gradation showed the interaction degrees of DEGs. Common TFs and TCFs in different CNV patterns and CNV pattern-specific TFs/TCFs were arranged from left to right. B, Expression analysis of CNV pattern-specific TFs based on TCGA and GTEx databases. C, D, E, F, G, H, Expression analysis of CNV pattern-specific TFs among four molecular subtypes of GBM samples in TCGA database.

Figure S12. A, B, C, D Expression analysis of CNV pattern 4-specific TFs, JUNB and FOS, in different types of malignancies in TCGA database. GBM, glioblastoma. LAML, acute myelocytic leukemia. PAAD, pancreatic adenocarcinoma. BLCA, bladder urothelial carcinoma. DLBC, Lymphoid neoplasm diffuse larger B-cell lymphoma. LUAD, lung adenocarcinoma. LUSC, lung squamous cell carcinoma. OV, ovarian serous cystadenocarcinoma. THYM, thymoma. LGG, low-grade glioma. ACC, adrenocortical carcinoma. BRCA, breast invasive carcinoma. HNSC, head and neck squamous cell carcinoma. KICH, kidney chromophobe. KIRP, kidney renal papillary cell carcinoma. LIHC, liver hepatocellular carcinoma. SKCM, skin cutaneous melanoma.

Figure S13. A, B, Prognostic value of CNV patterns' core gene sets in OS (A) and DFS (B) of GBM patients. C. Differential expression of CNV pattern 4-specific genes

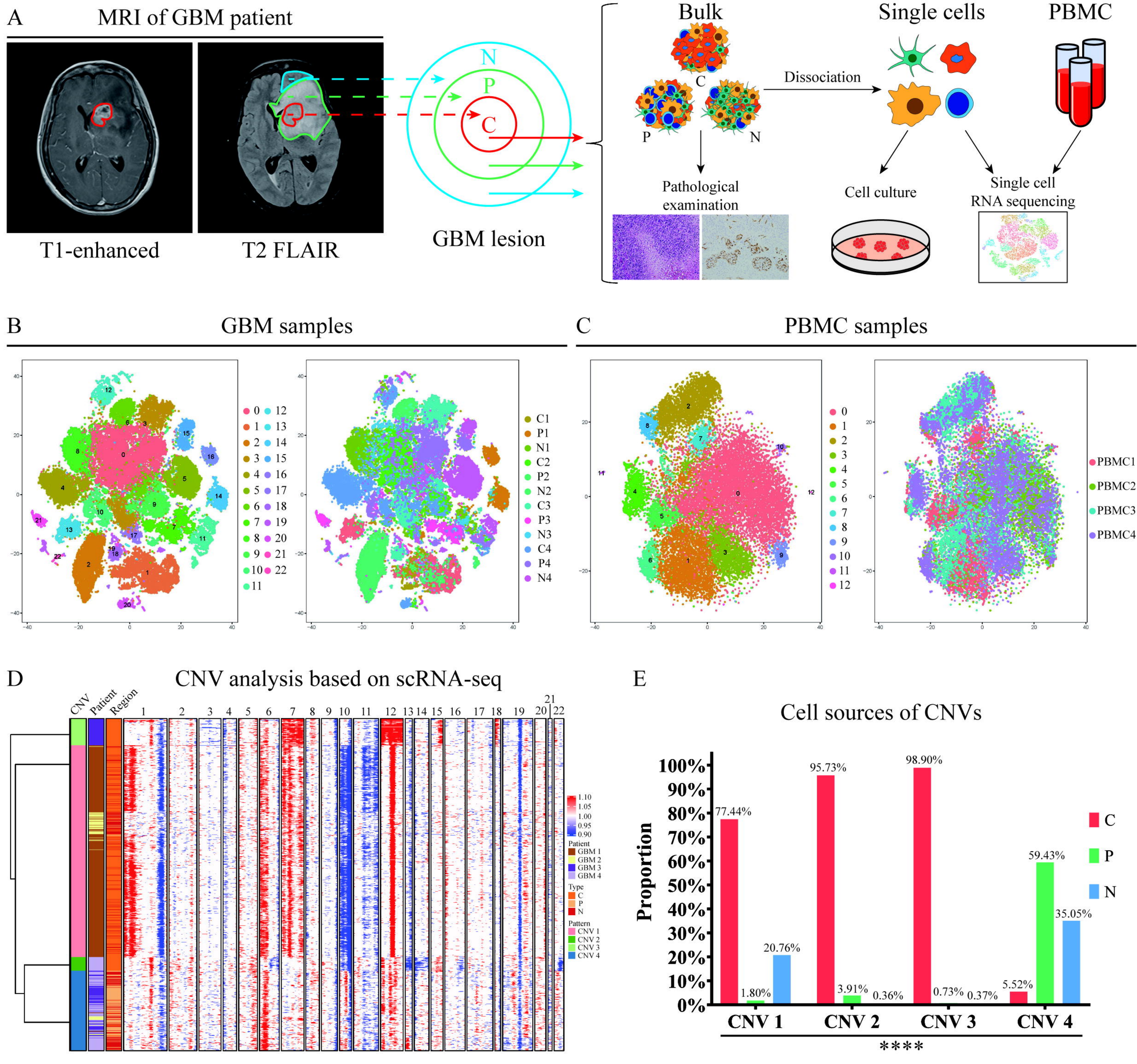
JUNB, ATF3, ACTN1, CEBPB, EGR1, and NFKBIA in primary and recurrent GBM samples from China Glioma Genome Atlas (CGGA) database. D. Spearman correlation analysis of recurrence-related genes of CNV pattern 4.

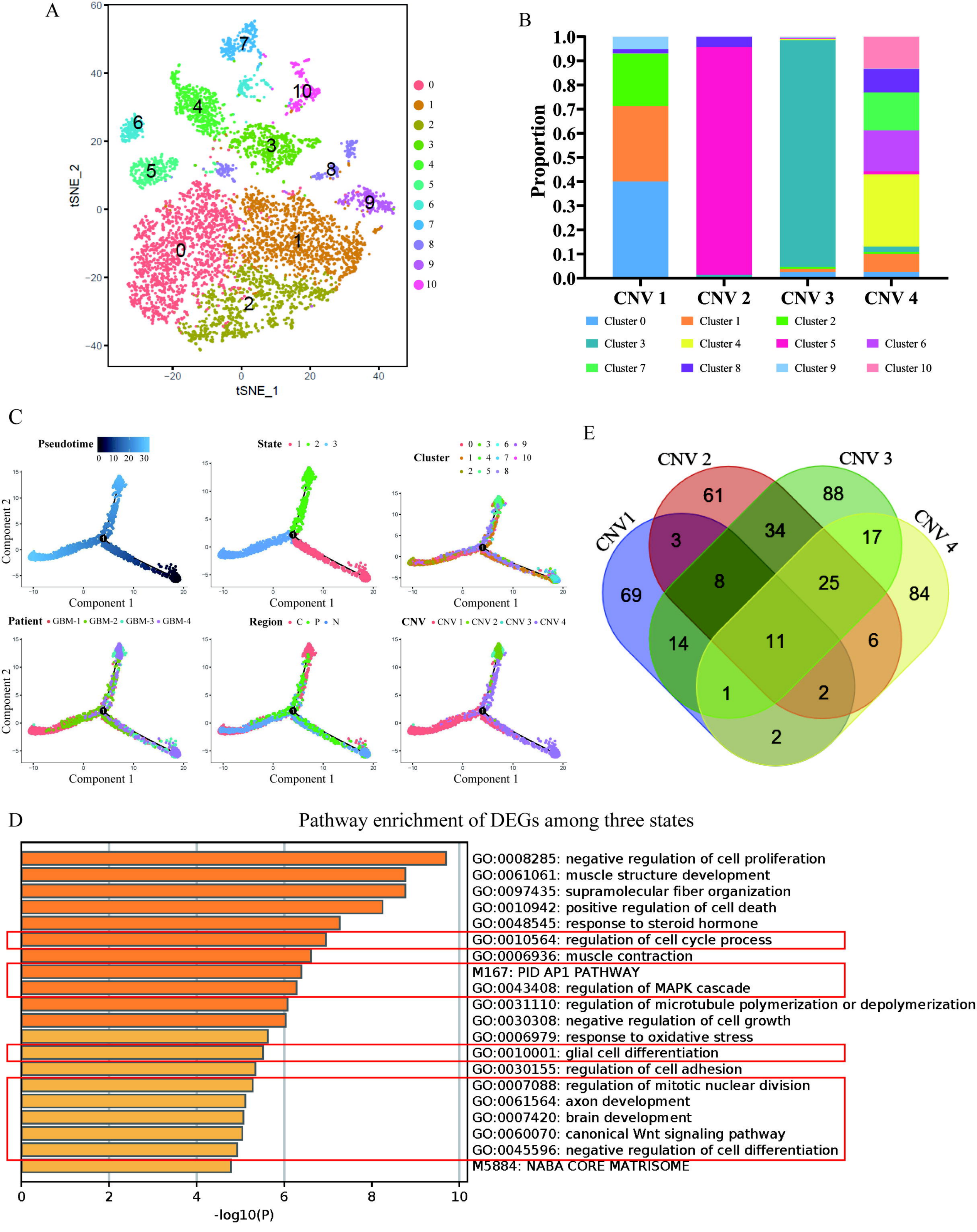
Figure S14. A, Pie charts describing the CNV patterns compositions of different samples from GBM-4 patient. B, Pathway enrichment results of DEGs of sub-cluster 3. C, Three-dimensional coordinate diagram of endothelial cell score (X-axis), pericyte score (Y-axis), and stemness score (Z-axis). GBM cells of CNV pattern 4 were scored in these three dimensions and then plotted in the coordinate frame. D, Developmental trajectory of CNV pattern 4 GBM cells inferred by monocle 2 algorithm. Distributions of sub-clusters along three state axes were presented separately.

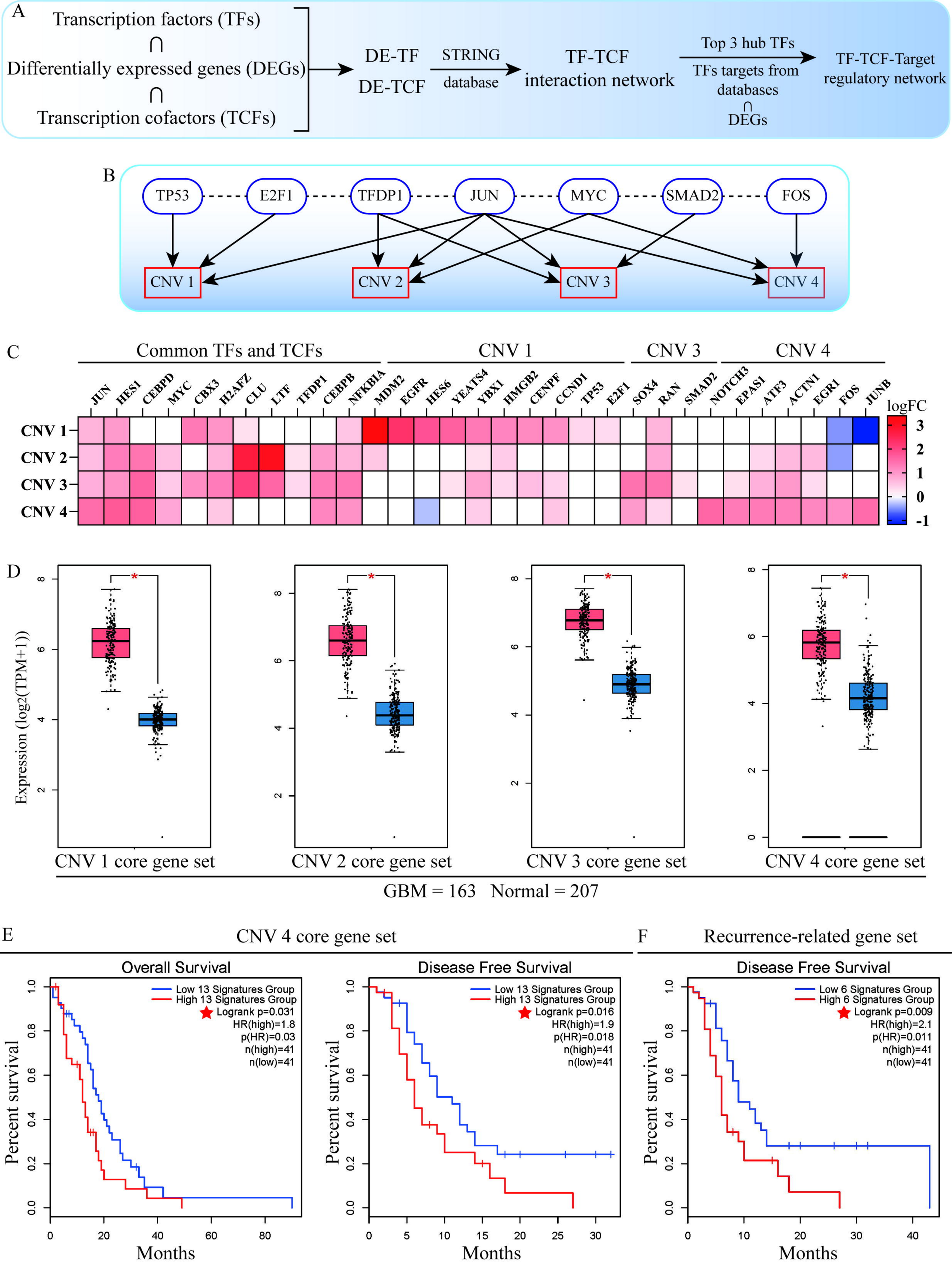
Figure S15. A, IHC staining of SOX2, Nestin, and GFAP in GBM samples from C, P, and N regions. B, IHC staining of SOX2, Nestin, and GFAP in normal brain tissue obtained during surgical approach to ependymoma.

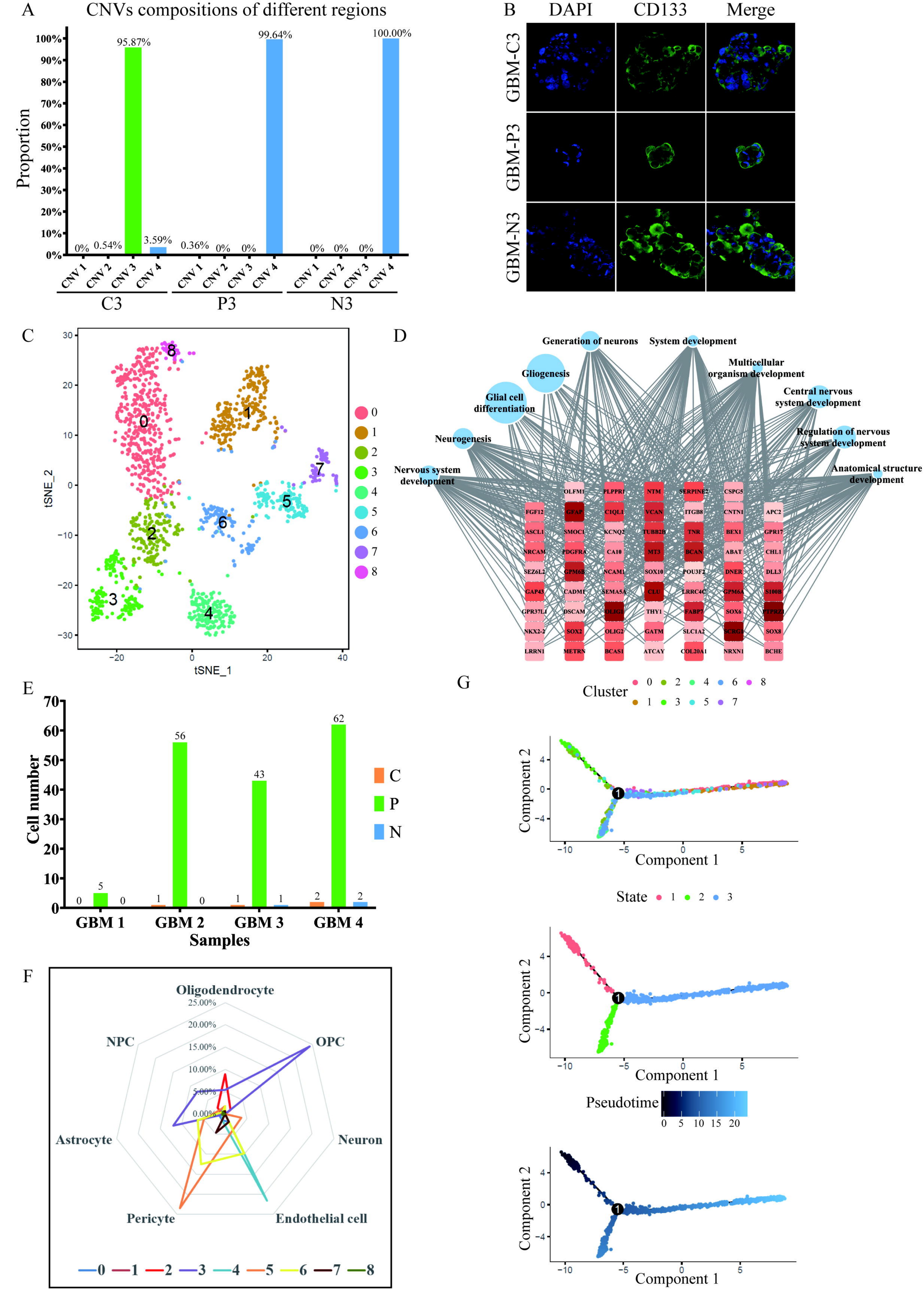
Figure S16. A, t-SNE plot of 23, 850 immune cells in twelve GBM tissue samples. B, Results of GO enrichment analysis of intracranial immune cell sub-cluster 17 DEGs compared to GBM cells. C, D, E, Violin plots showing the expression of some markers of proliferative cell, malignant cell, and monocyte in the PBMC of GBM patients, respectively. F, Results of GO enrichment analysis of PBMC sub-cluster 10

DEGs compared to GBM cells. G, H, Violin plot showing the expression of proliferation-associated genes in the PBMC of healthy volunteers and epilepsy patients, respectively. I, Venn diagram of proliferation-associated genes in GBM cells, PBMC, and intracranial immune cells. Sixteen shared genes were identified, which were termed common gene set of proliferation in this study.

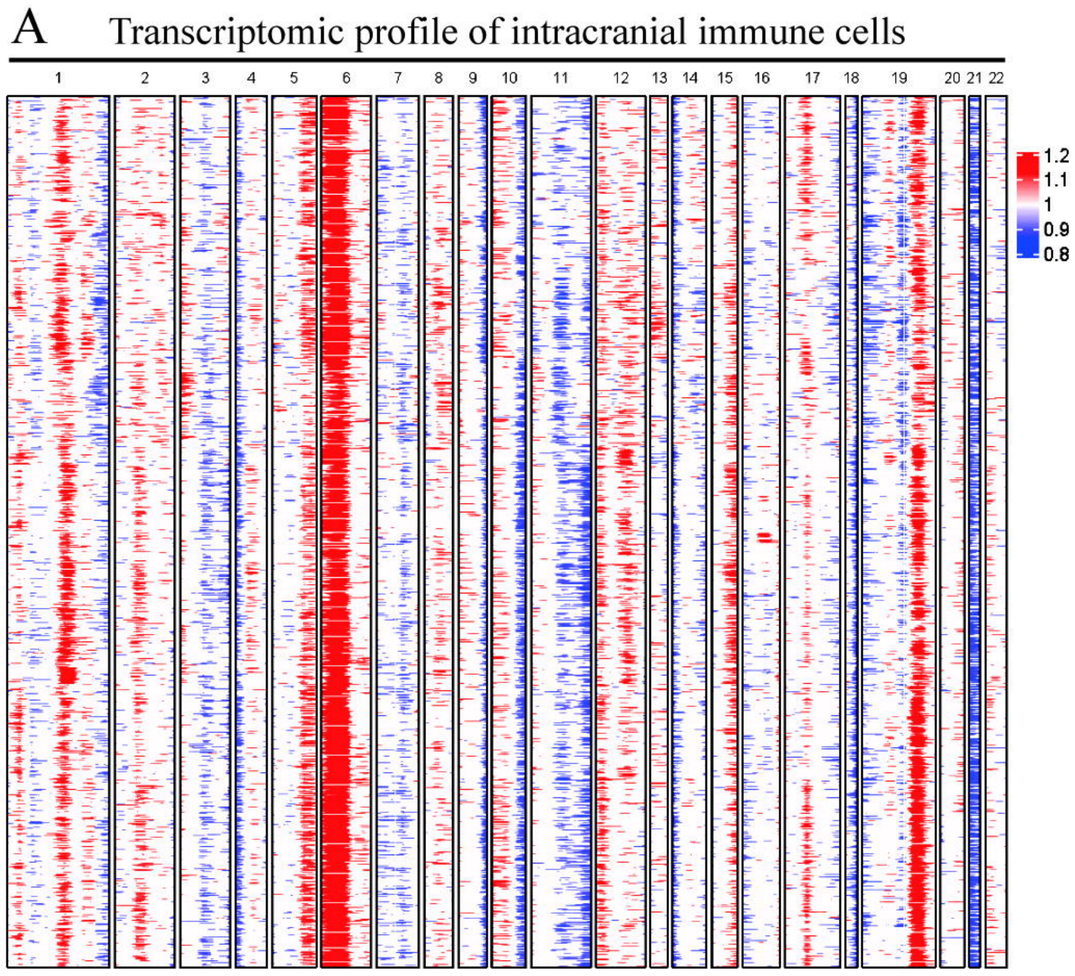

A
MRI of GBM patient
T1-enhanced
T2 FLAIR
N
P
C
GBM lesion
Bulk
Single cells
PBMC
Dissociation
Pathological examination
Cell culture
Single cell RNA sequencing
B
GBM samples
C
PBMC samples
PBMC1
PBMC2
PBMC3
PBMC4
D
CNV analysis based on scRNA-seq
CNV
Patient
Region
Patient
GBM 1
GBM 2
GBM 3
GBM 4
Type
C
P
N
Pattern
CNV 1
CNV 2
CNV 3
CNV 4
E
Cell sources of CNVs
Proportion
77.44%
1.80%
20.76%
95.73%
3.91%
0.36%
98.90%
0.73%
0.37%
5.52%
59.43%
35.05%
CNV 1
CNV 2
CNV 3
CNV 4
****


A
tSNE_1
tSNE_2
0
1
2
3
4
5
6
7
8
9
10
B
Proportion
CNV 1
CNV 2
CNV 3
CNV 4
Cluster 0
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
Cluster 10
C
Pseudotime
State
Cluster
Patient GBM-1 GBM-2 GBM-3 GBM-4
Region C P N
CNV CNV 1 CNV 2 CNV 3 CNV 4
Component 1
Component 2
E
CNV1
CNV 2
CNV 3
CNV 4
61
88
3
34
17
69
8
25
84
14
11
6
1
2
2
D
Pathway enrichment of DEGs among three states
GO:0008285: negative regulation of cell proliferation
GO:0061061: muscle structure development
GO:0097435: supramolecular fiber organization
GO:0010942: positive regulation of cell death
GO:0048545: response to steroid hormone
GO:0010564: regulation of cell cycle process
GO:0006936: muscle contraction
M167: PID AP1 PATHWAY
GO:0043408: regulation of MAPK cascade
GO:0031110: regulation of microtubule polymerization or depolymerization
GO:0030308: negative regulation of cell growth
GO:0006979: response to oxidative stress
GO:0010001: glial cell differentiation
GO:0030155: regulation of cell adhesion
GO:0007088: regulation of mitotic nuclear division
GO:0061564: axon development
GO:0007420: brain development
GO:0060070: canonical Wnt signaling pathway
GO:0045596: negative regulation of cell differentiation
M5884: NABA CORE MATRISOME
-log10(P)

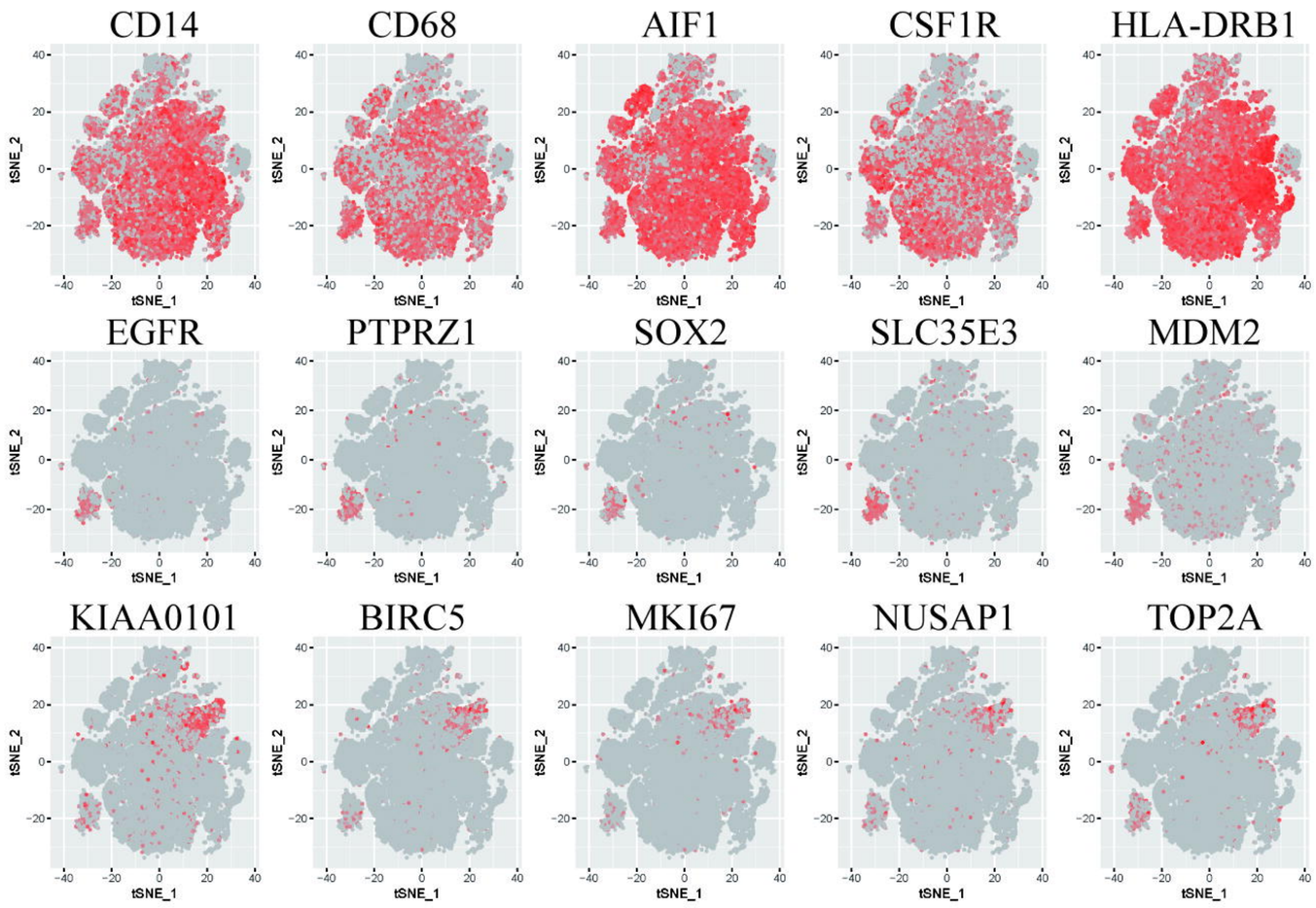

A
Transcription factors (TFs)
∩
Differentially expressed genes (DEGs)
∩
Transcription cofactors (TCFs)
DE-TF
DE-TCF
STRING database
TF-TCF interaction network
Top 3 hub TFs
TFs targets from databases
∩
DEGs
TF-TCF-Target regulatory network
B
TP53
E2F1
TFDP1
JUN
MYC
SMAD2
FOS
CNV 1
CNV 2
CNV 3
CNV 4
C
Common TFs and TCFs
CNV 1
CNV 3
CNV 4
JUN HES1 CEBPD MYC CBX3 H2AFZ CLU LTF TFDP1 CEBPB NFKBIA MDM2 EGFR HES6 YEATS4 YBX1 HMGB2 CENPF CCND1 TP53 E2F1 SOX4 RAN SMAD2 NOTCH3 EPAS1 ATF3 ACTN1 EGR1 FOS JUNB
CNV 1
CNV 2
CNV 3
CNV 4
logFC
3
2
1
0
-1
D
Expression (log2(TPM+1))
CNV 1 core gene set
CNV 2 core gene set
CNV 3 core gene set
CNV 4 core gene set
GBM = 163 Normal = 207
E
CNV 4 core gene set
Overall Survival
Low 13 Signatures Group
High 13 Signatures Group
Logrank p=0.031
HR(high)=1.8
p(HR)=0.03
n(high)=41
n(low)=41
Percent survival
Months
Disease Free Survival
Low 13 Signatures Group
High 13 Signatures Group
Logrank p=0.016
HR(high)=1.9
p(HR)=0.018
n(high)=41
n(low)=41
Percent survival
Months
F
Recurrence-related gene set
Disease Free Survival
Low 6 Signatures Group
High 6 Signatures Group
Logrank p=0.009
HR(high)=2.1
p(HR)=0.011
n(high)=41
n(low)=41
Percent survival
Months


A
CNVs compositions of different regions
Proportion
95.87%
99.64%
100.00%
0%
0.54%
3.59%
0.36%
CNV 1
CNV 2
CNV 3
CNV 4
C3
P3
N3
B
DAPI
CD133
Merge
GBM-C3
GBM-P3
GBM-N3
C
tSNE_1
tSNE_2
D
Generation of neurons
System development
Gliogenesis
Multicellular organism development
Glial cell differentiation
Central nervous system development
Neurogenesis
Regulation of nervous system development
Nervous system development
Anatomical structure development
E
Cell number
Samples
GBM 1
GBM 2
GBM 3
GBM 4
C
P
N
F
Oligodendrocyte
OPC
Neuron
Endothelial cell
Pericyte
Astrocyte
NPC
G
Cluster
Component 1
Component 2
State
Pseudotime

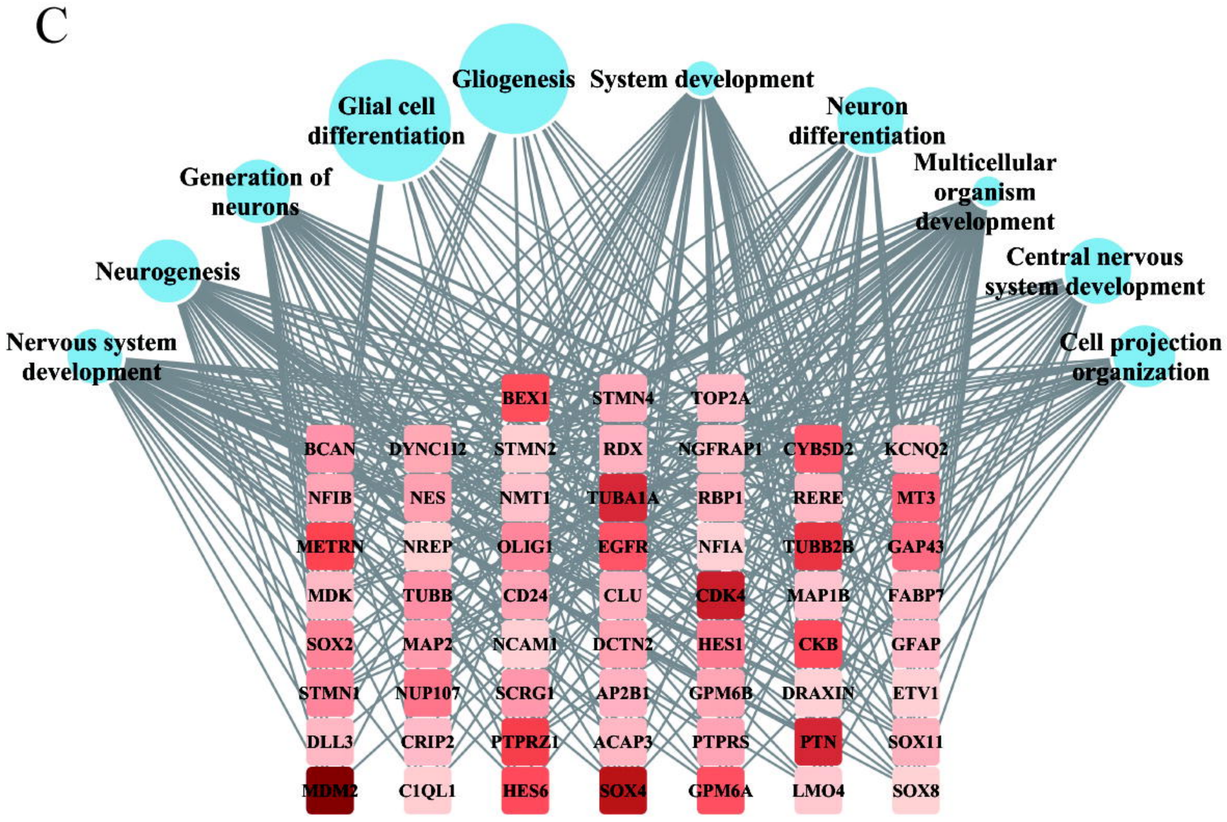
A


Transcriptomic profile of intracranial immune cells
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
1.2
1.1
1
0.9
0.8

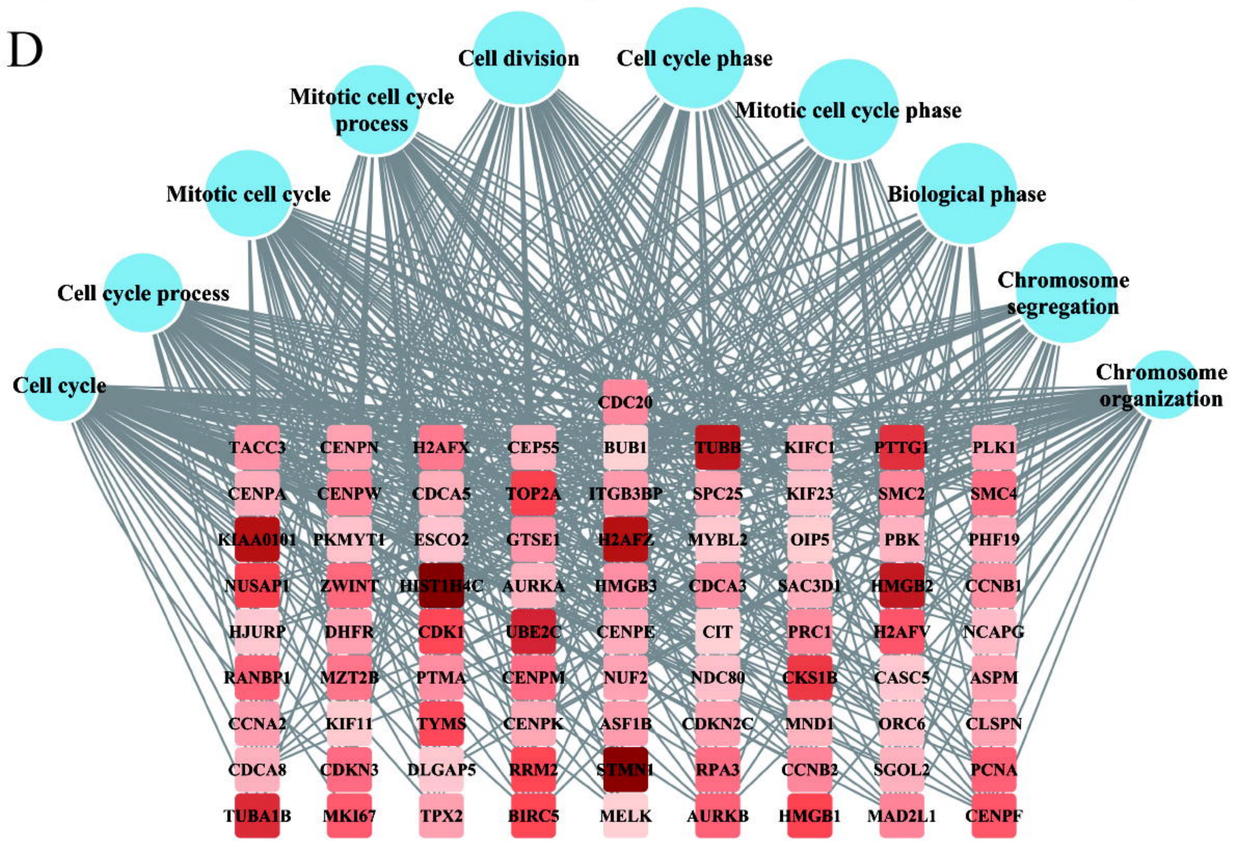
B


CD14
CD68
AIF1
CSF1R
HLA-DRB1
EGFR
PTPRZ1
SOX2
SLC35E3
MDM2
KIAA0101
BIRC5
MKI67
NUSAP1
TOP2A
tSNE_1
tSNE_2

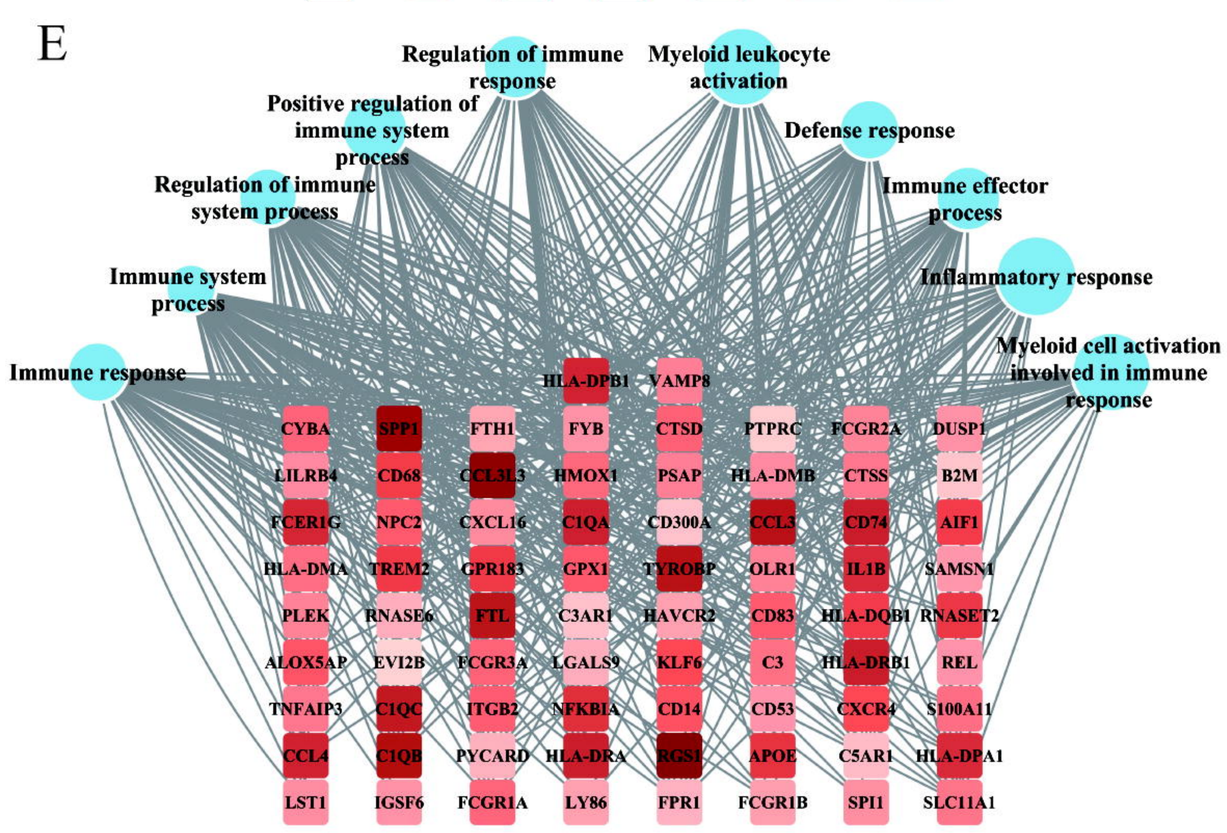
C


Gliogenesis
System development
Glial cell differentiation
Neuron differentiation
Generation of neurons
Multicellular organism development
Neurogenesis
Central nervous system development
Nervous system development
Cell projection organization
BEX1 STMN4 TOP2A
BCAN DYNC1I2 STMN2 RDX NGFRAP1 CYB5D2 KCNQ2
NFIB NES NMT1 TUBA1A RBP1 RERE MT3
METRN NREP OLIG1 EGFR NFIA TUBB2B GAP43
MDK TUBB CD24 CLU CDK4 MAP1B FABP7
SOX2 MAP2 NCAM1 DCTN2 HES1 CKB GFAP
STMN1 NUP107 SCRG1 AP2B1 GPM6B DRAXIN ETV1
DLL3 CRIP2 PTPRZ1 ACAP3 PTPRS PTN SOX11
MDM2 C1QL1 HES6 SOX4 GPM6A LMO4 SOX8

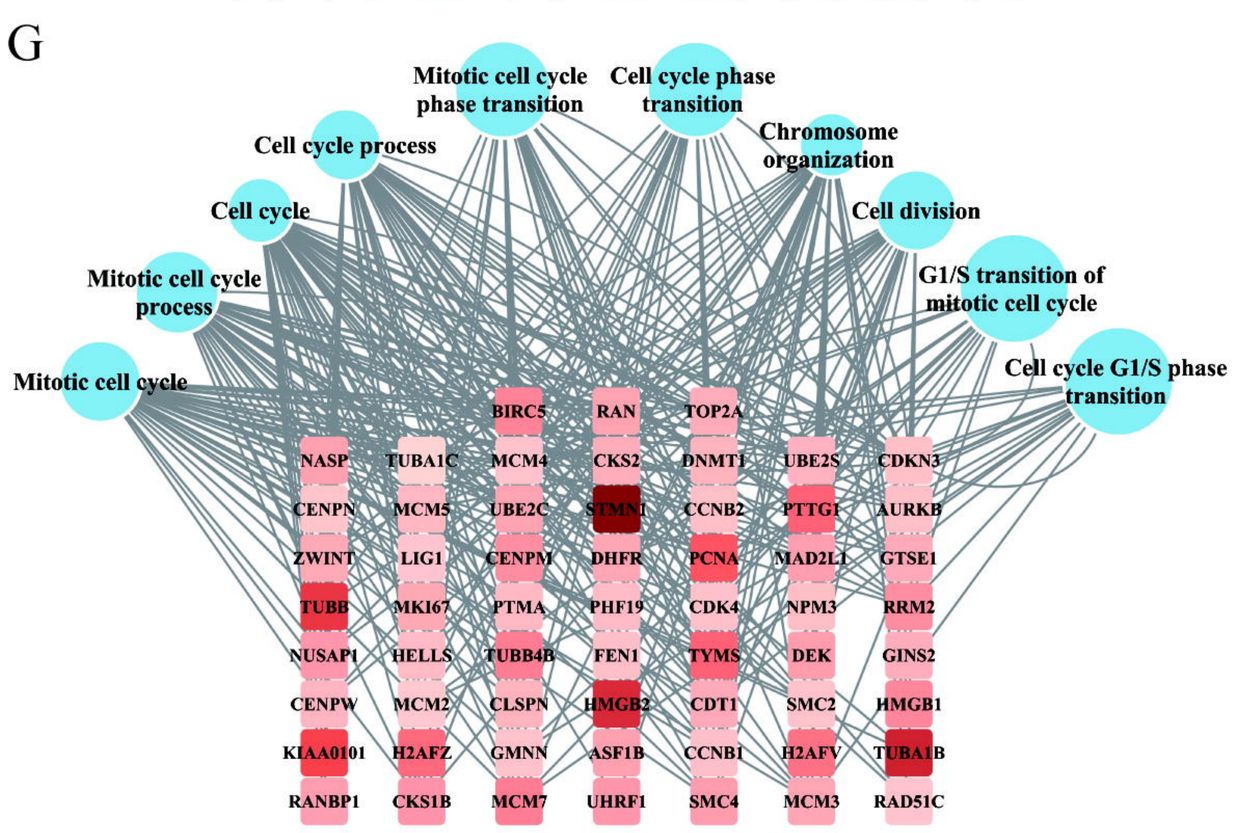
D


Cell division
Cell cycle phase
Mitotic cell cycle process
Mitotic cell cycle phase
Mitotic cell cycle
Biological phase
Cell cycle process
Chromosome segregation
Cell cycle
Chromosome organization
CDC20
TACC3 CENPN H2AFX CEP55 BUB1 TUBB KIFC1 PTTG1 PLK1
CENPA CENPW CDCA5 TOP2A ITGB3BP SPC25 KIF23 SMC2 SMC4
KIAA0101 PKMYT1 ESCO2 GTSE1 H2AFZ MYBL2 OIP5 PBK PHF19
NUSAP1 ZWINT HIST1H4C AURKA HMGB3 CDCA3 SAC3D1 HMGB2 CCNB1
HJURP DHFR CDK1 UBE2C CENPE CIT PRC1 H2AFV NCAPG
RANBP1 MZT2B PTMA CENPM NUF2 NDC80 CKS1B CASC5 ASPM
CCNA2 KIF11 TYMS CENPK ASF1B CDKN2C MND1 ORC6 CLSPN
CDCA8 CDKN3 DLGAP5 RRM2 STMN1 RPA3 CCNB2 SGOL2 PCNA
TUBA1B MKI67 TPX2 BIRC5 MELK AURKB HMGB1 MAD2L1 CENPF

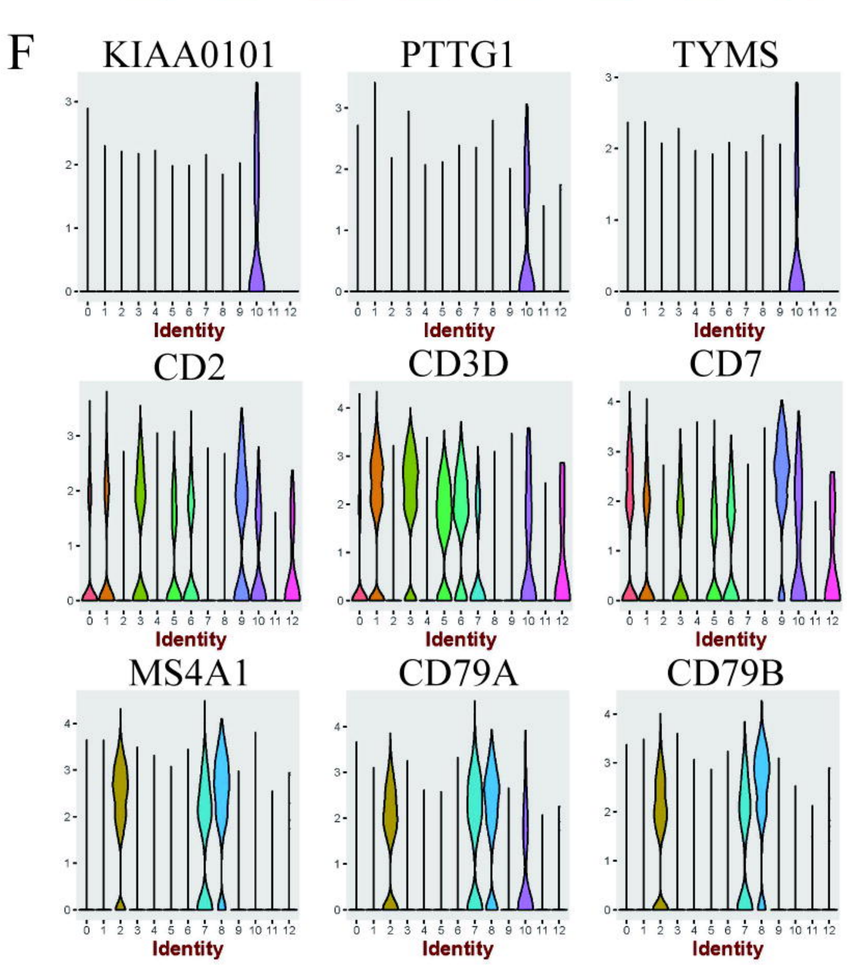
E


Regulation of immune response
Myeloid leukocyte activation
Positive regulation of immune system process
Defense response
Regulation of immune system process
Immune effector process
Immune system process
Inflammatory response
Immune response
Myeloid cell activation involved in immune response
HLA-DPB1 VAMP8
CYBA SPP1 FTH1 FYB CTSD PTPRC FCGR2A DUSP1
LILRB4 CD68 CCL3L3 HMOX1 PSAP HLA-DMB CTSS B2M
FCER1G NPC2 CXCL16 C1QA CD300A CCL3 CD74 AIF1
HLA-DMA TREM2 GPR183 GPX1 TYROBP OLR1 IL1B SAMSN1
PLEK RNASE6 FTL C3AR1 HAVCR2 CD83 HLA-DQB1 RNASET2
ALOX5AP EVI2B FCGR3A LGALS9 KLF6 C3 HLA-DRB1 REL
TNFAIP3 C1QC ITGB2 NFKBIA CD14 CD53 CXCR4 S100A11
CCL4 C1QB PYCARD HLA-DRA RGS1 APOE C5AR1 HLA-DPA1
LST1 IGSF6 FCGR1A LY86 FPR1 FCGR1B SPI1 SLC11A1

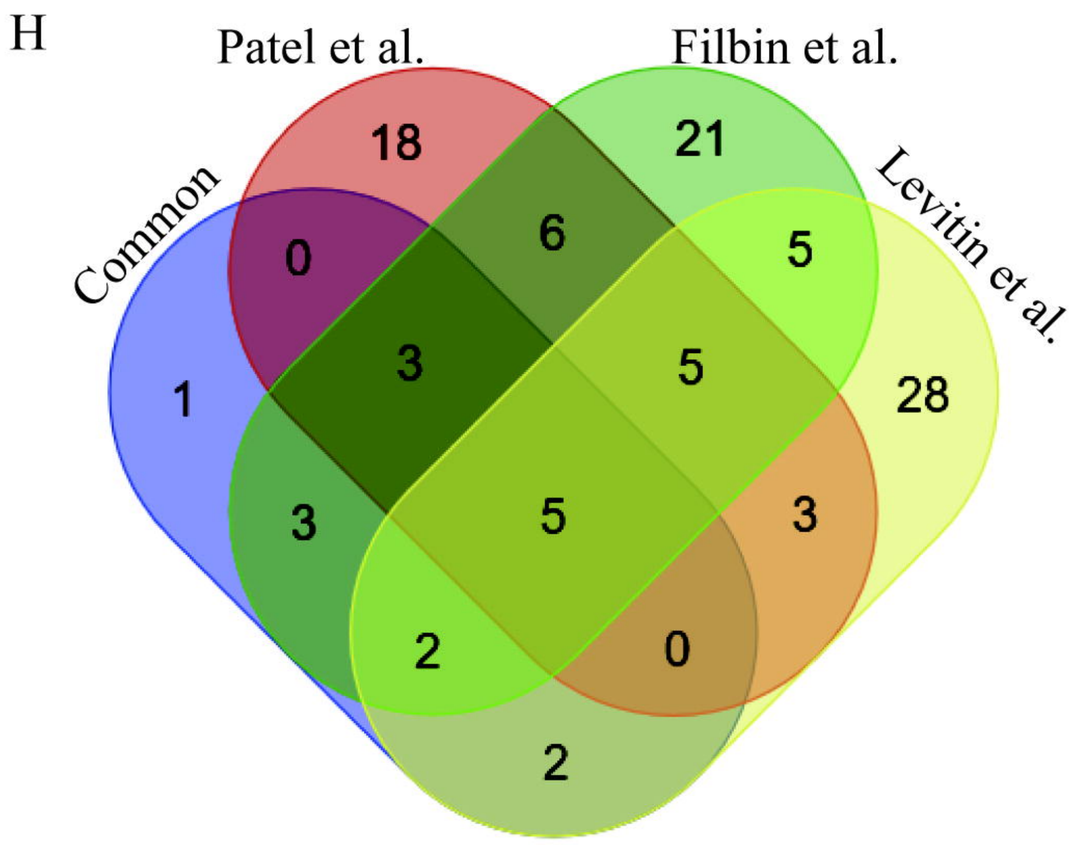
G


Mitotic cell cycle phase transition
Cell cycle phase transition
Cell cycle process
Chromosome organization
Cell cycle
Cell division
Mitotic cell cycle process
G1/S transition of mitotic cell cycle
Mitotic cell cycle
Cell cycle G1/S phase transition
BIRC5 RAN TOP2A
NASP TUBA1C MCM4 CKS2 DNMT1 UBE2S CDKN3
CENPN MCM5 UBE2C STMN1 CCNB2 PTTG1 AURKB
ZWINT LIG1 CENPM DHFR PCNA MAD2L1 GTSE1
TUBB MKI67 PTMA PHF19 CDK4 NPM3 RRM2
NUSAP1 HELLS TUBB4B FEN1 TYMS DEK GINS2
CENPW MCM2 CLSPN HMGB2 CDT1 SMC2 HMGB1
KIAA0101 H2AFZ GMNN ASF1B CCNB1 H2AFV TUBA1B
RANBP1 CKS1B MCM7 UHRF1 SMC4 MCM3 RAD51C

F


KIAA0101
PTTG1
TYMS
CD2
CD3D
CD7
MS4A1
CD79A
CD79B
Identity

H


Patel et al.
Filbin et al.
Levitin et al.
Common
18
21
0
6
5
1
3
5
28
3
5
3
2
0
2