**Supplementary Methods**

1. **Tumor cell identification and copy number prediction**

For each sample, the expression profile of all cells was first analyzed by k-means clustering using the Seurat package (https://satijalab.org/seurat/). To reduce the influence of low-frequency genes on copy number prediction, the genes with an expression frequency less than 5% across all cells were removed from subsequent analysis. To eliminate the impact of highly-expressed genes on copy number calling, gene expression values greater than one were replaced with "1".

Next, we used two criteria to identify tumor clusters: 1) the proportion of expressed genes in each cell. We assumed that there should be more activated genes in tumor cells. Therefore, the average ratio of expressed genes in tumor clusters should be higher than in non-tumor clusters, and 2) the variation of average expression of adjacent genes in each cell. We assumed that the genomic instability in tumor clusters should be higher than in non-tumor cells. Thus, we used the frameshift method, as described previously (1,2), to roughly calculate putative copy number scores for each cell. The genome variation was then predicted by the median absolute deviation (MAD) of copy number scores. Using these two criteria, we selected the most tumor-like cluster as a candidate cluster and calculated this cluster’s average copy number scores as a seed. Following this, copy number scores from each cell were subjected to correlation analysis with this seed. The cells with correlation values greater than 0.5 were considered putative tumor cells. Finally, the clusters with a putative tumor cell fraction greater than 50% were labeled as putative tumor clusters. The remaining clusters were defined as non-tumor clusters. Clusters consisting of less than 30 cells were not selected as a candidate cluster. Copy number alteration visualization was performed using inferCNV (3). Finally, the cells in the tumor cluster detected with *PTPRC* expression were also filtered out.

1. **RNA velocity analysis**

The RNA velocity analysis was performed using the method reported by La Manno et al. (4). To further confirm the tendency of tumor cells to transit from each lesion, we defined two transition states and enumerated the number of tumor cells of two lesions for each transition state. Specifically, the transition states included “*leaving from the core of the lesion*” (LC) and “*approaching to the core of the lesion*” (AC). Since the result of RNA velocity analysis provided the coordinate data of transition arrow (i.e., start and end points of the arrow, representing the transition direction of a cell) for each cell, we first identified the core position of tumor cells for each lesion by calculating the average coordinate data of the start points. We defined those tumor cells whose distance between end point and core position of the indicated lesion (*De*) was longer than the distance between start point and core position of the indicated lesion (*Ds*) as LC, otherwise as AC. The distance (*Ds* or *De*) was calculated as following formula:

where denotes the coordinate data (x, y) of tumor lesion *j*. denote the coordinate data (x, y) of start and end points of the arrow of tumor cell *i*, respectively. represents the cell *i* is likely to be an LC state as to lesion *j*. By contrast, represents the cell *i* is likely to be an AC state as to lesion *j*. For each lesion, we count the number of tumor cells either labeled as LC or AC. To exactly determine the association between the transition state and two lesions, we focused on the tumor clusters derived from two lesions that had similar numbers of cells and showed close to each other on the t-SNE plot. Since these tumor clusters could exhibit a coherent transition, facilitating to identify the transition state for each cell. After calculating the exact transition state for each cell, we then performed Fisher’s exact test for evaluating the significance of such a transition process.

1. **NES prediction model**

We first identified DEGs specifically expressed in the tumor cells from the 1st or 2nd lesion in NJ01 and NJ02 since each lesion in these two samples was subjected to a transcriptomic data profiling up to nearly 10,000 cells, which may capture a comprehensive landscape of tumor cells within the tissue. Specifically, we defined the highly expressed genes in the 1st and 2nd lesions as oDEGs and yDEGs, respectively. We required oDEGs whose expression percentage across tumor cells in the 1st lesion was 5% higher than that in the 2nd lesion, vice versa. These genes generally harbored the signatures characterizing distinct features between the older (the 1st) and the younger (the 2nd) GBMs, but may still contain pseudo-features caused by noise. Therefore, we focused on identifying the clusters with the older or younger features and performed differential expression analysis on these clusters. To identify these clusters, we first calculated the putative sequence order (designated as PE) for each tumor cluster using the following formula:

Where and denotes the percentage of oDEGs and yDEGs expressed in cluster , respectively. A high PE value indicates that the cluster may undergo long-term evolution (or be older).

We ranked the clusters from each lesion by PE score and selected the top and bottom 20% as the older and younger clusters. We performed differential expression analysis on expression profiles comprised of the older clusters in the 1st lesion and the younger clusters in the 2nd lesion based on these clusters. The differential analysis used criteria including fold-change (FC) of differentially expressed genes>1.3, adjusted p-value<10-3 (Wilcoxon’s rank-sum test), and the percentage of genes expressed in each group ≥ 5%. We also applied a similar strategy to analyze DEGs for the multifocal GBMs of TT01 and TT02. We included all of tumor clusters into differentially expression analysis on TT01 and TT02 due to the small size of tumor clusters in a single lesion. The significant DEGs were identified with the same criteria: FC>1.3, adjusted p-value<, and the percentage of genes expressed in each group≥5%.

The DEGs derived from four GBMs (NJ01, NJ02, TT01, and TT02) were subjected to intersection analysis, remaining 28 genes. These genes were further filtered for the tumor-intrinsic expressed genes identified in our previous study (5) and subjected to protein-protein interaction network analysis (STRING, https://string-db.org). Finally, twelve genes were retained as a core interaction module and defined as GBM NES (**Supplementary Table 1**). These genes were subjected to the NES prediction model using the ssGSEA algorithm (6).

1. **Trajectory analysis**

The trajectory analysis was performed using Monocle2 (7). The analysis only focused on the tumor cells derived from two paired lesions for each case, of which the expression profile was integrated through the Seurat package (version 2.3.4) with default parameters (8). The NES score for each tumor cell was calculated by NES prediction model.

1. **Public data collection**

TCGA GBM datasets were obtained from the UCSC website (https://xenabrowser.net), including mRNA expression profiles (Affymetrix U133a microarray and Illumina HiSeq 2000 RNA-seq), mutation profiles, and RABIT transcription factor regulatory impact data. CGGA GBM datasets were downloaded from <http://www.cgga.org.cn>. The datasets GSE193884 (**for Fig. 2D**), GSE76184(**for Fig. 2E**), GSE117891 (**for Supplementary Fig. S3I-K**), GSE141946 (**for Supplementary Fig. S4A**), GSE27523 and GSE73556 (**for Fig. 3D**), GSE7835 (**for Fig. 3E, F**), GSE59612 (**for Fig. 4C**), GSE138794 (**for Supplementary Fig. S6, Supplementary Fig. S9A**), GSE126725 (**for Fig. 6F**), GSE77043 (**for Fig. 6K**), GSE147147 (**for Supplementary Fig. S11A-D**) used in this study were obtained from the GEO database (Gene Expression Omnibus (GEO), RRID:SCR\_005012). Specifically, the samples (11 samples) in GSE193884 were collected from the first-, second-, and third-generation patient-derived glioblastoma explants and subjected to single-cell RNA sequencing. Two (i.e., JK142, first- and second-generation) out of 11 samples were excluded from the analysis due to the few numbers of tumor cells (N < 200 cells). Datasets downloaded from GlioVis (9) (<http://gliovis.bioinfo.cnio.es>) included Rembrandt (10) and IVY GAP (11)**.** Association analysis between TFs and tissue development was based on the E-MTAB-6814 dataset obtained from the EMBL-EBI database (European Bioinformatics Institute, RRID:SCR\_004727). The mRNA profile of primary and recurrent GBMs was available from The Glioma Longitudinal AnalySiS (GLASS, <http://www.synapse.org/glass>) (12,13).

1. **Collection of cell molecular signatures**

The BMDM and MG molecular signatures were provided by Müller et al. (14). The T cell gene set consisted of *CD3D*, and *CD3E*. The molecular signature of GBM cellular states and subtypes was provided by Neftel et al. (15) and Wang et al. (5). The metastasis-like signature was collected from a review summarized by Scott Valastyan, and Robert A Weinberg (16) (**Supplementary Table 3**). All signatures were subjected to a construction model using ssGSEA (17). To identify the exact molecular signature in each cell analyzed by scRNA-seq, the signature score calculated by ssGSEA was evaluated using the permutation analysis devised in our previous study (5). Only one signature whose permutation p-value was less than 0.05 (permutation p-values for other signatures were all greater than 0.2) was taken as the main feature for the indicated cells.

1. **Analysis of multi-sector GBMs**

We analyzed the single-cell transcriptome and matched whole-genome sequencing data derived from 46 sectors (or regions) of GBMs across eight patients (i.e., GS1-3, GS5-6, GS11-13, samples were available from Li et al. (18)). The identification of tumor and non-tumor cells was inferred from the pipeline mentioned above. For each patient, we chose two sectors according to these criteria: **(a)** Given the few tumor cells in the most of samples, each sector should harbor more than 20 tumor cells to reduce the bias caused by an insufficient number of tumor cells on the analysis. **(b)** Two sectors are located at maximally distant sites from each other. We next inferred the potential genomic evolutionary order between the two sectors for each patient by performing copy number alterations (CNAs) analysis. Seven of eight cases showed consistent CNAs between two sectors (Spearman’s correlation test, *rho≥0.9, P<0.001*). We also performed the ssGSEA algorithm combined with permutation analysis to dissect the number of NES-high tumor cells for each sector in eight cases.

1. **snATAC-seq analysis**

The snATAC-seq and scRNA-seq matrices from the same patients diagnosed with IDH wild-type GBMs (e.g., SF11979 and SF11956) were downloaded from GSE138794 (19). The peak matrix for snATAC-seq was converted to a gene activity matrix using Seurat (20). Then, latent semantic indexing was performed to reduce the dimensionality of the snATAC-seq data. The expression matrix was also imported into Seurat for quality control and downstream analysis. Low-quality cells (>10% mitochondrial genes, <600 or >6000 genes/cell) were excluded. Gene expression matrices were normalized and scaled using default parameters. The top 2000 highly variable genes were selected and used for PCA, and the top 20 principal components were used for graph-based clustering (res=0.6). To classify malignant cells, genomic instability calculations were applied to the expression matrices. According to a normalized NES score, malignant cells were divided into three groups: hNES, intermediate NES, and lNES. To obtain predicted cell types, we identified anchors between the peak and expression matrices. The anchors were used to transfer the cell types identified from the scRNA-seq data to the snATAC-seq cells. Finally, motif enrichment for the open regions was analyzed using Homer (21), with fragment sizes ranging from -200 to 250 bp.

1. **Co-location analysis**

To explore the molecular signature distribution across various regions, we used ssGSEA to calculate signature scores based on the mRNA expression profile provided by the IVY dataset (11).

1. **Re-clustering of myeloid cells**

We integrated and re-clustering myeloid cells from the four multifocal GBMs (16,051 cells in total) using Seurat (8), resulting in 17 subclusters. The molecular markers for each subcluster were identified by the FindMarkers function embedded in the Seurat algorithm (top 10 markers) (8) and further reviewed and validated according to the existing myeloid subtypes reported by literature (**Table 1**). The mac\_GADD45G and mac\_TF subclusters were excluded from the subsequent analysis due to the limited number of cells (<100 cells).

**Table 1.** The list of markers associated with macrophages

|  |  |  |
| --- | --- | --- |
| **Clusters** | **Markers** | **References (PubMed ID)** |
| mac\_CD69 | *CD69/CCL3/IL1B* | 29020624; 33782623; 35690521 |
| mac\_CD74 | *CD74/C1QB* | 28487994; 30471926; 35690521 |
| mac\_BAG3 | *BAG3* | 26522614 |
| mac\_APOE | *APOE/SPP1* | 33782623; 30471926; 35690521 |
| mac\_PLIN2 | *PLIN2/NUPR1* | 35372038; 35690521 |
| mac\_FCGBP | *FCGBP* | 32840654 |
| mac\_F13A1 | *F13A1* | 22192929; 29777108; 35690521 |
| mac\_GINS2 | *GINS2* | 35069907 |
| mac\_SEMA5A | *SEMA5A* | - |
| mac\_CXCL10 | *CXCL10* | 33879239; 31373067; 30471926 |
| mac\_CCNB2 | *CCNB2* | 30471926 |
| mac\_FCN1 | *FCN1* | 33782623 |
| mac\_IL32 | *IL32* | 18296636 |
| mac\_TPM2 | *TPM2* | - |
| mac\_MANF | *MANF* | 31226359; 30515104 |
| mac\_TF\* | *TF* | - |
| mac\_GADD45G\* | *GADD45G* | - |

\*: the number of cells in the indicated subcluster is less than 100 (<1% of all cells)

1. **Cell-cell communication analysis**

We mapped receptor-ligand pairs using CellPhoneDB (RRID:SCR\_017054) (22) (www.cellphonedb.org) onto our cell subsets to identify cell-cell interactions between different cell types. For this analysis, we filtered the raw cell-cell interaction matrices by p-value. Significant interactions (p<0.01) were identified between cell subsets and visualized by CellChat (23). Only the ligand-receptor pairs expressed in more than 10% of the specific cell types were considered while analyzing interactions between malignant cells and BMDMs or MGs.

**References**

1. Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H*, et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science **2014**;344(6190):1396-401.

2. Puram SV, Tirosh I, Parikh AS, Patel AP, Yizhak K, Gillespie S*, et al.* Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. Cell **2017**;171(7):1611-+.

3. Tickle T, TI GC, Brown M, Haas B. inferCNV of the Trinity CTAT Project. Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA **2019**.

4. La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V*, et al.* RNA velocity of single cells. Nature **2018**;560(7719):494-8 doi 10.1038/s41586-018-0414-6.

5. Wang Q, Hu B, Hu X, Kim H, Squatrito M, Scarpace L*, et al.* Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. Cancer Cell **2018**;33(1):152 doi 10.1016/j.ccell.2017.12.012.

6. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF*, et al.* Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature **2009**;462(7269):108-12 doi 10.1038/nature08460.

7. Qiu X, Hill A, Packer J, Lin D, Ma YA, Trapnell C. Single-cell mRNA quantification and differential analysis with Census. Nat Methods **2017**;14(3):309-15 doi 10.1038/nmeth.4150.

8. 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(5):411-20 doi 10.1038/nbt.4096.

9. Bowman RL, Wang QH, Carro A, Verhaak RGW, Squatrito M. GlioVis data portal for visualization and analysis of brain tumor expression datasets. Neuro-Oncology **2017**;19(1):139-41.

10. Madhavan S, Zenklusen JC, Kotliarov Y, Sahni H, Fine HA, Buetow K. Rembrandt: helping personalized medicine become a reality through integrative translational research. Mol Cancer Res **2009**;7(2):157-67 doi 10.1158/1541-7786.MCR-08-0435.

11. Puchalski RB, Shah N, Miller J, Dalley R, Nomura SR, Yoon JG*, et al.* An anatomic transcriptional atlas of human glioblastoma. Science **2018**;360(6389):660-3 doi 10.1126/science.aaf2666.

12. Barthel FP, Johnson KC, Varn FS, Moskalik AD, Tanner G, Kocakavuk E*, et al.* Longitudinal molecular trajectories of diffuse glioma in adults. Nature **2019**;576(7785):112-+.

13. Varn FS, Johnson KC, Wade TE, Malta TM, Sabedot TS, Barthel FP*, et al.* Longitudinal analysis of diffuse glioma reveals cell state dynamics at recurrence associated with changes in genetics and the microenvironment. bioRxiv **2021**:2021.05.03.442486 doi 10.1101/2021.05.03.442486.

14. Muller S, Kohanbash G, Liu SJ, Alvarado B, Carrera D, Bhaduri A*, 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 **2017**;18.

15. Neftel C, Laffy J, Filbin MG, Hara T, Shore ME, Rahme GJ*, et al.* An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. Cell **2019**;178(4):835-+.

16. Valastyan S, Weinberg RA. Tumor metastasis: molecular insights and evolving paradigms. Cell **2011**;147(2):275-92 doi 10.1016/j.cell.2011.09.024.

17. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF*, et al.* Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature **2009**;462(7269):108-U22.

18. Yu K, Hu Y, Wu F, Guo Q, Qian Z, Hu W*, et al.* Surveying brain tumor heterogeneity by single-cell RNA-sequencing of multi-sector biopsies. Natl Sci Rev **2020**;7(8):1306-18 doi 10.1093/nsr/nwaa099.

19. Wang L, Babikir H, Muller S, Yagnik G, Shamardani K, Catalan F*, et al.* The Phenotypes of Proliferating Glioblastoma Cells Reside on a Single Axis of Variation. Cancer Discov **2019**;9(12):1708-19.

20. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, 3rd*, et al.* Comprehensive Integration of Single-Cell Data. Cell **2019**;177(7):1888-902 e21 doi 10.1016/j.cell.2019.05.031.

21. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P*, et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell **2010**;38(4):576-89 doi 10.1016/j.molcel.2010.05.004.

22. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. Nat Protoc **2020**;15(4):1484-506.

23. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Myung P, Plikus MV*, et al.* Inference and analysis of cell-cell communication using CellChat. bioRxiv **2020**:2020.07.21.214387 doi 10.1101/2020.07.21.214387.