Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

An integrative single-cell and bulk transcriptomic framework prioritizes a glioblastoma stemness program associated with immune suppression and adverse outcome

Simin Yang
Southern Medical University
Sihan Ju
Fudan University
Yunjie Liao
Southern Medical University
Shengxuan Ding
Southern Medical University
Hongbo Guo
Southern Medical University
Chengwei Chen
chengv88@smu.edu.cn
Southern Medical University

Research Article

Keywords: glioblastoma, stemness, single-cell RNA sequencing, tumor microenvironment, WGCNA, immune suppression

Posted Date: May 21st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9643632/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

An integrative single-cell and bulk transcriptomic framework prioritizes a glioblastoma stemness program associated with immune suppression and adverse outcome

Research article

Authors

Simin Yang[1],*; Sihan Ju[2],*; Yunjie Liao[1],*; Shengxuan Ding[1],*; Hongbo Guo[1],†; Chengwei Chen[1],†

* These authors contributed equally to this work.

† Corresponding authors.

Affiliations

1 Neurosurgery Center, The National Key Clinical Specialty, The Engineering Technology Research Center of Education Ministry of China on Diagnosis and Treatment of Cerebrovascular Disease, Guangdong Provincial Key Laboratory on Brain Function Repair and Regeneration, The Neurosurgery Institute of Guangdong Province, Zhujiang Hospital, Southern Medical University, Guangzhou 510282, China

2 Neurosurgical Department of Huashan Hospital and MOE Frontiers Center for Brain Science, Fudan University, Shanghai 200040, China

Correspondence

Hongbo Guo (corresponding author)

Neurosurgery Center of Zhujiang Hospital, Southern Medical University, Guangzhou 510282, China

Email: guohongbo911@126.com

Chengwei Chen (corresponding author)

Neurosurgery Center of Zhujiang Hospital, Southern Medical University, Guangzhou 510282, China

Email: chengv88@smu.edu.cn

ORCID: https://orcid.org/0000-0001-6293-5387

Running title: GBM stemness program and immune suppression

Abstract

Purpose: Clinically useful glioblastoma (GBM)-specific stemness biomarker programs remain incompletely defined. We aimed to prioritize a GBM-focused stemness program and candidate markers by integrating single-cell, single-nucleus, and bulk transcriptomic data.

Methods: We integrated scRNA-seq data from 55 GBM samples and snRNA-seq data from 10 GBM samples. Malignant-cell stemness was quantified with CytoTRACE. Stem.Sig was derived from six public GBM single-cell datasets by intersecting genes positively associated with CytoTRACE with genes upregulated in malignant cells. Bulk cohorts were examined for immune associations, tumor-type summary trends with intratumoral heterogeneity (ITH) and tumor mutational burden (TMB), and small-cohort anti-PD-1 outcome modeling. Weighted gene co-expression network analysis (WGCNA) was used to prioritize hub genes, followed by preliminary qPCR and western blot expression support.

Results: Malignant compartments associated with non-response showed higher CytoTRACE scores than treatment-naive compartments in both cellular and nuclear datasets (both $P<0.001$). Stem.Sig was associated with lower immune-related gene expression and lower inferred immune-cell abundance. At the tumor-type summary level, higher median Stem.Sig tended to align with higher ITH and TMB; these comparisons were interpreted descriptively. A Naive Bayes exploratory model showed modest discrimination in pretreatment anti-PD-1 GBM cohorts (validation and test AUCs both 0.65). WGCNA overlap prioritized nine hub genes; SOX6, BCHE, and LINC00511 showed preliminary qPCR support, and SOX6 and BCHE were also supported at the protein level.

Conclusion: This integrative framework prioritizes a GBM stemness program linked to immune-cold biology, heterogeneity, and adverse outcome, and nominates SOX6, BCHE, and LINC00511 as candidate markers for mechanistic follow-up.

Keywords

glioblastoma; stemness; single-cell RNA sequencing; tumor microenvironment; WGCNA; immune suppression

Introduction

Glioblastoma (GBM) is best understood in the contemporary WHO 2021 framework as an adult-type diffuse astrocytic glioma, IDH-wildtype, CNS WHO grade 4, although legacy public datasets do not always include complete molecular annotation [1]. Despite maximal safe resection, radiotherapy, and temozolomide, outcome remains poor [2]. A major biological explanation for treatment failure is the persistence of stem-like tumor states capable of self-renewal, plasticity, invasion, and therapy tolerance [3-7].

Single-cell transcriptomics has substantially refined the current view of GBM by resolving malignant cellular states and the surrounding immune and stromal compartments [8-10]. CytoTRACE provides a practical framework for estimating differentiation status from transcriptional diversity at the single-cell level [11]. Across cancers, transcriptomic stemness programs have been linked to oncogenic dedifferentiation, intratumoral heterogeneity, immune suppression, and immunotherapy-response contexts [12-14]. In GBM, this type of malignant-cell state analysis is especially relevant because a bulk expression profile can confound tumor-intrinsic stemness, tumor purity, and immune microenvironment composition.

We therefore integrated GBM scRNA-seq, snRNA-seq, and bulk transcriptomic datasets to address four questions: whether malignant-cell stemness differs across clinically annotated GBM samples; whether a GBM-focused stemness program can be derived from recurrent single-cell signals; whether the resulting program aligns with immune-cold biology, ITH, and TMB; and whether WGCNA can reduce a broad biological program into a smaller candidate biomarker set supported by preliminary qPCR and western blot assays.

Materials and methods

Data sources and preprocessing

Public datasets are summarized in Online Resource 1, Supplementary Table S1. We integrated scRNA-seq data from 55 GBM samples and snRNA-seq data from 10 GBM samples derived from six GEO datasets (GSE138794, GSE139448, GSE102130, GSE103224, GSE131928a, and GSE131928b). GSE131928a and GSE131928b were handled as separate processed single-cell/single-nucleus resources in the recurrence-based derivation workflow rather than pooled as duplicate patient-level observations. Standard quality-control filters, dimensionality reduction, clustering, and marker-based annotation were applied within each dataset.

Where molecular annotation was available, analyses were restricted to adult IDH-wildtype glioblastoma; when IDH status was unavailable in legacy datasets, this limitation was explicitly considered in interpretation. Malignant-cell annotations were supported by lineage-marker expression and dataset-level malignant labels where available. Ambiguous clusters lacking canonical tumor or lineage-marker patterns were excluded before CytoTRACE scoring. scRNA-seq and snRNA-seq datasets were analyzed separately to avoid platform-driven stemness artifacts. Dataset-level Stem.Sig derivation required recurrence across independent cohorts rather than reliance on a single integrated embedding.

Stemness signature derivation

To derive a GBM-focused stemness program, each of the six single-cell datasets was analyzed independently. Genes positively correlated with CytoTRACE in malignant cells (Spearman R>0, false discovery rate <1×10^-5) were intersected with genes upregulated in malignant versus non-malignant cells (log fold-change >=0.25, false discovery rate <1×10^-5). Genes recurrently satisfying both criteria across datasets were de-duplicated to form the final 235-unique-gene Stem.Sig. The complete 235-gene list and a standard GMT file are supplied in the Online_Resources folder.

Bulk cohort analyses

Stem.Sig was scored in bulk expression cohorts with gene set variation analysis (GSVA) [15]. Public and processed expression matrices analyzed here included TCGA-GBM, GSE68848, GSE4290, and comparative TCGA tumor types (LUAD, LUSC, PAAD, OV, and UVM). Immune-related genes were assessed using curated immune-gene categories, and immune-cell abundance was inferred using MCP-counter [16, 17]. ITH and TMB were analyzed at the tumor-type summary level for descriptive cross-tumor comparisons. Complete tumor-purity estimates were not uniformly available across all processed matrices, so purity-adjusted immune analyses were treated as an important sensitivity analysis for future work rather than as confirmatory evidence in the present study.

Exploratory anti-PD-1 outcome modeling

For small-cohort outcome modeling, pretreatment GBM cohorts were assembled from published anti-PD-1 glioblastoma studies, including Cloughesy et al. 2019/GSE121810, Schalper et al. 2019, and Zhao et al. 2019. The processed modeling set was divided into training (n=29), validation (n=14), and test (n=16) sets; cohort

source, treatment context, pretreatment status, endpoint definition, and split assignment are summarized in Online Resource 1, Supplementary Table S1. Six algorithms (Naive Bayes, support vector machine, random forest, k-nearest neighbors, AdaBoost, and LogitBoost) were evaluated, and the retained exploratory model was selected by validation-set AUC. The anti-PD-1 component was treated as hypothesis-generating because the sample size was limited and the study design did not support deployment, calibration, or patient-level prediction claims.

WGCNA and preliminary expression support

WGCNA was performed on 178 GBM and control bulk samples from GSE68848 and GSE4290 using a soft-thresholding power of 13 [18]. Modules associated with the GBM phenotype were examined by module membership and gene significance metrics, using $|MM|>0.8$ and $|GS|>0.2$ as hub-gene filters. Because WGCNA was performed across GBM and control samples, some module structure may reflect tumor-normal contrast; therefore, hub genes were interpreted as prioritization candidates rather than definitive stemness drivers.

Candidate hub genes were intersected with Stem.Sig to obtain a smaller biomarker shortlist. Preliminary qPCR expression support was performed for SOX6, BCHE, and LINC00511 using de-identified clinical residual/previously archived human tissue specimens from patients at Zhujiang Hospital, Southern Medical University, including three GBM and three non-tumor control specimens in the available assay set. Total RNA was extracted using RNAiso Plus, reverse-transcribed with the PrimeScript RT reagent Kit, and quantified using SYBR Green chemistry on an ABI ViiA7 instrument with GAPDH as the reference gene; reaction conditions and assay/specimen details are summarized in Online Resource 1, Supplementary Table S4. Western blot expression support was available for SOX6 and BCHE only and is shown with molecular-weight labels and source images in Fig. 5f and Online Resource 1, Supplementary Fig. S4. The use of these human specimens was approved by the Ethics Committee of Zhujiang Hospital, Southern Medical University (approval no. 2024BA0040-GC), with waiver of informed consent because the specimens were clinical residual/previously archived materials and all patient information was anonymized. LINC00511 is a long non-coding RNA and was therefore not treated as a western blot target. The expression-support assay values were interpreted as preliminary and nominal because the available sample size was 3 versus 3.

Statistical analysis

Two-sided Wilcoxon tests were used for CytoTRACE comparisons, Spearman correlation was used for association analyses, and Kaplan-Meier plots were used for descriptive survival separation. For the preliminary expression-support assays, qPCR and western blot P values are reported as nominal exploratory values. Summary-level pan-cancer ITH/TMB panels were interpreted descriptively because the displayed unit was tumor-type median rather than patient-level observations.

Results

Single-cell and single-nucleus analyses identify higher stemness in clinically unfavorable samples

Integrated single-cell and single-nucleus analyses resolved the expected major GBM compartments, including malignant, immune, and stromal populations (Fig. 1a,b) [8-10]. Within malignant cells/nuclei, CytoTRACE and phenotype maps showed substantial intratumoral variation, consistent with state plasticity rather than a uniform malignant compartment (Fig. 1c,d). Group comparisons showed higher CytoTRACE scores in non-responder-associated malignant compartments than in treatment-naive malignant compartments (Fig. 1e,f).

Because CytoTRACE comparisons were performed at the cell/nucleus level, these P values should be interpreted as compartment-level evidence; patient-level validation remains necessary. These comparisons should not be overread as complete responder-versus-non-responder validation, but they support the premise that clinically unfavorable GBM compartments are enriched for higher stemness states.

Derivation and functional annotation of the GBM Stem.Sig program

Applying a common derivation strategy across six single-cell cohorts yielded a GBM-focused 235-unique-gene Stem.Sig program. The derivation required both a positive relationship with CytoTRACE and higher expression in malignant cells, thereby excluding genes that merely tracked microenvironmental content. Functional annotation showed enrichment for neural/developmental, NTRK-related, extracellular matrix, migration/invasion, cell-cycle/DNA-damage, and immune/inflammatory processes (Fig. 2a-c). These patterns support interpretation of Stem.Sig as a biological state program rather than an immediately deployable clinical assay.

Stem.Sig is associated with an immune-cold microenvironment and tumor heterogeneity

In TCGA-GBM and comparative TCGA tumor sets, higher Stem.Sig scores were associated with broad reductions in immune-related gene expression and lower inferred abundance of T cells, CD8 T cells, cytotoxic lymphocytes, B-lineage cells, NK cells, monocytes, and myeloid dendritic cells (Fig. 3a,b). The Hallmark-pathway analysis further linked Stem.Sig to pathways compatible with immune exclusion and aggressive tumor biology (Fig. 3c).

At the tumor-type summary level, higher median Stem.Sig tended to align with higher ITH and TMB (Fig. 3e,f). This pan-cancer comparison was interpreted descriptively because only a small number of tumor-type median points were displayed. When bulk GBM samples were stratified jointly by Stem.Sig and TMB, the low Stem.Sig/high TMB group showed the highest inferred immune-cell abundance, whereas the high Stem.Sig/low TMB group showed the lowest abundance (Fig. 3d). These results support an immune-cold interpretation of high Stem.Sig while acknowledging that patient-level purity-adjusted sensitivity analyses remain important.

Exploratory outcome modeling shows modest discrimination

An exploratory Naive Bayes model achieved the highest validation-set discrimination and was therefore retained for small-cohort evaluation (Fig. 4a). The validation-set and test-set ROC curves are shown in Fig. 4c,d, with AUCs of 0.65 in both settings. Comparator-signature benchmarking showed that Stem.Sig ranked among the stronger available signatures in this processed GBM setting, but the absolute discrimination remained modest (Fig. 4b).

Kaplan-Meier curves showed exploratory survival separation between the model-defined groups in the validation and test sets (Fig. 4e,f). Because the cohorts were small and the original modeling did not include a full multivariable framework incorporating steroid exposure, MGMT status, extent of resection, recurrence status, or treatment context, these survival analyses are presented as hypothesis-generating and not as a clinically deployable assay.

WGCNA prioritizes candidate hub genes with preliminary expression support

WGCNA of 178 GBM and control samples identified brown and yellow modules as the two modules most strongly linked to the GBM phenotype. Intersecting genes from GBM-associated modules with Stem.Sig yielded nine candidate hub genes: BCAN, BCHE, DSEL, ETV1, IGF2BP3, LINC00511, RBP1, SOX6, and TRAF4 (Fig. 5a-c; Online Resource 1, Supplementary Table S2). Single-cell localization and enrichment

analyses supported malignant-cell relevance for these candidates (Fig. 5d; Online Resource 1, Supplementary Fig. S3 and S6).

Based on biological prioritization and assay availability, SOX6, BCHE, and LINC00511 were selected for preliminary expression support. qPCR supported higher expression of all three genes in GBM versus non-tumor control specimens, while western blot supported increased SOX6 and BCHE protein abundance normalized to Actin (Fig. 5e,f; Online Resource 1, Supplementary Table S3). These data should be interpreted as preliminary expression support rather than functional validation.

Discussion

This study integrates GBM single-cell, single-nucleus, and bulk transcriptomic resources to prioritize a GBM-specific stemness program and candidate hub genes. The main contribution is a layered framework: first defining malignant-cell stemness states, then translating recurrent state-associated genes into a bulk-scored program, and finally shrinking the program to a smaller WGCNA-overlap biomarker shortlist supported by preliminary expression assays.

Several features of Stem.Sig suggest that it captures more than generic proliferation. GBM single-cell studies have emphasized that malignant states distribute along developmental and mesenchymal axes rather than fitting fixed histologic bins [8-10]. In the present analysis, high-stemness tumors were enriched for immune-cold and adverse biological contexts. This supports the view that a stemness program may contribute to immune exclusion, therapy tolerance, and tumor heterogeneity, while remaining cautious about causal interpretation.

The immune findings place stemness within the broader architecture of GBM immune resistance. Immune checkpoint blockade has shown limited activity in unselected GBM, and response likely depends on factors beyond PD-1/PD-L1 expression alone [19-24]. Stem.Sig was inversely associated with immune-gene expression and inferred immune-cell abundance, consistent with an immune-cold phenotype. However, tumor purity is a potential confounder in bulk immune deconvolution, and future work should include tumor-purity-adjusted analyses and orthogonal deconvolution methods.

The tumor-type summary trends with ITH and TMB also require nuanced interpretation. Cancer stemness has been associated with intratumoral heterogeneity across tumor types, and GBM-specific studies have linked stemness and evolutionary heterogeneity to clinically relevant phenotypes [13, 25-27]. TMB is often treated as a proxy for immunogenicity, but in GBM the relationship between mutational burden and effective antitumor immunity is context dependent [21, 24, 28, 29]. Our summary-level panels suggest that high Stem.Sig may coexist with genomic complexity, but patient-level analyses will be needed to define whether Stem.Sig modifies the relationship between TMB, ITH, and immunotherapy response.

These considerations explain why the anti-PD-1 modeling was deliberately framed as exploratory. Broader immune-response signatures and tertiary lymphoid structure biology provide important context for interpreting checkpoint-blockade responsiveness, but the present analysis was not designed to replace these established immune biomarkers [30, 31]. AUC values around 0.65 are compatible with modest signal detection, not clinical implementation. The model should not be interpreted as a validated immunotherapy-response assay. Instead, it suggests that the GBM stemness program may carry clinically relevant information that should be retested in larger, better annotated pretreatment cohorts with nested validation, calibration, and clinical covariate adjustment.

The 235-gene Stem.Sig is best viewed as a biological state program rather than an immediately deployable clinical assay. The WGCNA-overlap strategy provides a pragmatic route to a smaller candidate biomarker set. SOX6, BCHE, and LINC00511 were prioritized because they lie at the intersection of malignant-cell stemness, GBM-associated coexpression structure, and preliminary expression support. Prior literature has linked SOX6 to glioma expression and antigenicity, LINC00511 to GBM biology, and cholinesterase-related proteins to glioma tissue biology [32-35]. These genes should be advanced to mechanistic perturbation studies rather than treated as proven drivers.

The study has limitations. The clinically annotated single-cell comparison was non-responder-associated versus treatment-naive rather than a complete responder-versus-non-responder design, and cell/nucleus-level P values may overstate evidence if interpreted as patient-level inference. Legacy public datasets did not uniformly include modern IDH and treatment annotations. WGCNA across GBM and control samples may partly capture tumor-normal contrast. The anti-PD-1 cohorts were small, and the preliminary expression support was limited to 3 versus 3 samples. Finally, qPCR/western blot assays support expression differences, not function. These constraints justify a prioritization rather than validation claim. Future validation in patient-derived GBM organoids and microglia/macrophage-aware models will be needed to test functional relevance [36, 37].

References

1. Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, et al. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. Neuro-Oncology. 2021;23:1231-1251. https://doi.org/10.1093/neuonc/noab106

2. Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJB, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. New England Journal of Medicine. 2005;352:987-996. https://doi.org/10.1056/NEJMoa043330

3. Lathia JD, Mack SC, Mulkearns-Hubert EE, Valentim CLL, Rich JN. Cancer stem cells in glioblastoma. Genes & Development. 2015;29:1203-1217. https://doi.org/10.1101/gad.261982.115

4. Suva ML, Rheinbay E, Gillespie SM, Patel AP, Wakimoto H, Rabkin SD, et al. Reconstructing and reprogramming the tumor-propagating potential of glioblastoma stem-like cells. Cell. 2014;157:580-594. https://doi.org/10.1016/j.cell.2014.02.030

5. Bao S, Wu Q, McLendon RE, Hao Y, Shi Q, Hjelmeland AB, et al. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. Nature. 2006;444:756-760. https://doi.org/10.1038/nature05236

6. Chen P, Hsu WH, Han J, Xia Y, DePinho RA. Cancer stemness meets immunity: from mechanism to therapy. Cell Reports. 2021;34:108597. https://doi.org/10.1016/j.celrep.2020.108597

7. Lim M, Xia Y, Bettegowda C, Weller M. Current state of immunotherapy for glioblastoma. Nature Reviews Clinical Oncology. 2018;15:422-442. https://doi.org/10.1038/s41571-018-0003-5

8. 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;178:835-849.e21. https://doi.org/10.1016/j.cell.2019.06.024

9. 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;344:1396-1401. https://doi.org/10.1126/science.1254257

10. Darmanis S, Sloan SA, Croote D, Mignardi M, Chernikova S, Samghababi P, et al. Single-cell RNA-seq analysis of infiltrating neoplastic cells at the migrating front of human glioblastoma. Cell Reports. 2017;21:1399-1410. https://doi.org/10.1016/j.celrep.2017.10.030

11. Gulati GS, Sikandar SS, Wesche DJ, Manjunath A, Bharadwaj A, Berger MJ, et al. Single-cell transcriptional diversity is a hallmark of developmental potential. Science. 2020;367:405-411. https://doi.org/10.1126/science.aax0249

12. Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, et al. Machine learning identifies stemness features associated with oncogenic dedifferentiation. Cell. 2018;173:338-354.e15. https://doi.org/10.1016/j.cell.2018.03.034

13. Miranda A, Hamilton PT, Zhang AW, Pattnaik S, Becht E, Mezheyeuski A, et al. Cancer stemness, intratumoral heterogeneity, and immune response across cancers. Proceedings of the National Academy of Sciences USA. 2019;116:9020-9029. https://doi.org/10.1073/pnas.1818210116

14. Zhang Z, Wang ZX, Chen YX, Wu HX, Yin L, Zhao Q, et al. Integrated analysis of single-cell and bulk RNA sequencing data reveals a pan-cancer stemness signature predicting immunotherapy response. Genome Medicine. 2022;14:45. https://doi.org/10.1186/s13073-022-01050-w

15. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013;14:7. https://doi.org/10.1186/1471-2105-14-7

16. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The immune landscape of cancer. Immunity. 2018;48:812-830.e14. https://doi.org/10.1016/j.immuni.2018.03.023

17. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biology. 2016;17:218. https://doi.org/10.1186/s13059-016-1070-5

18. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559. https://doi.org/10.1186/1471-2105-9-559

19. Verhaak RGW, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, 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. https://doi.org/10.1016/j.ccr.2009.12.020

20. Jackson CM, Choi J, Lim M. Mechanisms of immunotherapy resistance: lessons from glioblastoma. Nature Immunology. 2019;20:1100-1109. https://doi.org/10.1038/s41590-019-0433-y

21. Reardon DA, Brandes AA, Omuro A, Mulholland P, Lim M, Wick A, et al. Effect of nivolumab versus bevacizumab in patients with recurrent glioblastoma: the CheckMate 143 randomized clinical trial. JAMA Oncology. 2020;6:1003-1010. https://doi.org/10.1001/jamaoncol.2020.1024

22. Cloughesy TF, Mochizuki AY, Orpilla JR, Hugo W, Lee AH, Davidson TB, et al. Neoadjuvant anti-PD-1 immunotherapy promotes a survival benefit with intratumoral and systemic immune responses in recurrent glioblastoma. Nature Medicine. 2019;25:477-486. https://doi.org/10.1038/s41591-018-0337-7

23. Schalper KA, Rodriguez-Ruiz ME, Diez-Valle R, Lopez-Janeiro A, Porciuncula A, Idoate MA, et al. Neoadjuvant nivolumab modifies the tumor immune microenvironment in resectable glioblastoma. Nature Medicine. 2019;25:470-476. https://doi.org/10.1038/s41591-018-0339-5

24. Zhao J, Chen AX, Gartrell RD, Silverman AM, Aparicio L, Chu T, et al. Immune and genomic correlates of response to anti-PD-1 immunotherapy in glioblastoma. Nature Medicine. 2019;25:462-469. https://doi.org/10.1038/s41591-019-0349-y

25. Wang Z, Wang Y, Yang T, Xing H, Wang Y, Gao L, et al. Machine learning revealed stemness features and a novel stemness-based classification with appealing implications in discriminating the prognosis, immunotherapy and temozolomide responses of 906 glioblastoma patients. Briefings in Bioinformatics. 2021;22:bbab032. https://doi.org/10.1093/bib/bbab032

26. McGranahan N, Swanton C. Clonal heterogeneity and tumor evolution: past, present, and the future. Cell. 2017;168:613-628. https://doi.org/10.1016/j.cell.2017.01.018

27. Sottoriva A, Spiteri I, Piccirillo SGM, Touloumis A, Collins VP, Marioni JC, et al. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. Proceedings of the National Academy of Sciences USA. 2013;110:4009-4014. https://doi.org/10.1073/pnas.1219747110

28. Yarchoan M, Hopkins A, Jaffee EM. Tumor mutational burden and response rate to PD-1 inhibition. New England Journal of Medicine. 2017;377:2500-2501. https://doi.org/10.1056/NEJMc1713444

29. Cristescu R, Mogg R, Ayers M, Albright A, Murphy E, Yearley J, et al. Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. Science. 2018;362:eaar3593. https://doi.org/10.1126/science.aar3593

30. Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR, et al. IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. Journal of Clinical Investigation. 2017;127:2930-2940. https://doi.org/10.1172/JCI91190

31. Sautès-Fridman C, Petitprez F, Calderaro J, Fridman WH. Tertiary lymphoid structures in the era of cancer immunotherapy. Nature Reviews Cancer. 2019;19:307-325. https://doi.org/10.1038/s41568-019-0144-6

32. Ueda R, Yoshida K, Kawakami Y, Kawase T, Toda M. Expression of a transcriptional factor, SOX6, in human gliomas. Brain Tumor Pathology. 2004;21:35-38. https://doi.org/10.1007/BF02482175

33. Ueda R, Iizuka Y, Yoshida K, Kawase T, Kawakami Y, Toda M. Identification of a human glioma antigen, SOX6, recognized by patients' sera. Oncogene. 2004;23:1420-1427. https://doi.org/10.1038/sj.onc.1207252

34. Du X, Tu Y, Liu S, Zhao P, Bao Z, Li C, et al. LINC00511 contributes to glioblastoma tumorigenesis and epithelial-mesenchymal transition via LINC00511/miR-524-5p/YB1/ZEB1 positive feedback loop. Journal of Cellular and Molecular Medicine. 2020;24:1474-1487. https://doi.org/10.1111/jcmm.14829

35. García-Ayllón MS, Sáez-Valero J, Muñoz-Delgado E, Vidal CJ. Identification of hybrid cholinesterase forms consisting of acetyl- and butyrylcholinesterase subunits in human glioma. Neuroscience. 2001;107:199-208. https://doi.org/10.1016/S0306-4522(01)00355-4

36. Jacob F, Salinas RD, Zhang DY, Nguyen PTT, Schnoll JG, Wong SZH, et al. A patient-derived glioblastoma organoid model and biobank recapitulates inter- and intratumoral heterogeneity. Cell. 2020;180:188-204.e22. https://doi.org/10.1016/j.cell.2019.11.036

37. Hambardzumyan D, Gutmann DH, Kettenmann H. The role of microglia and macrophages in glioma maintenance and progression. Nature Neuroscience. 2016;19:20-27. https://doi.org/10.1038/nn.4185

Figure legends

Fig. 1. Single-cell and single-nucleus analyses indicate higher stemness in treatment non-response-associated GBM malignant compartments. (a,b) UMAPs of the scRNA-seq and snRNA-seq GBM cohorts annotated into malignant, immune, and stromal compartments. (c,d) CytoTRACE and phenotype maps of malignant cells or malignant nuclei in the scRNA-seq and snRNA-seq cohorts. (e,f) CytoTRACE score comparisons between non-responder-associated and treatment-naive malignant compartments in the scRNA-seq and snRNA-seq cohorts. NR-associated/NR, non-responder-associated; TN, treatment-naive.

Fig. 2. Derivation and functional annotation of the GBM-focused stemness program. (a) Workflow used to derive Stem.Sig from six GBM single-cell datasets. (b) Enrichment summary focused on neuro-oncology-relevant biological processes. (c) Gene-concept network illustrating representative genes within enriched processes.

Fig. 3. Stem.Sig is associated with immune suppression and adverse genomic context in bulk cohorts. (a) Summary of associations between Stem.Sig and immune-related genes. (b) Associations between Stem.Sig and inferred immune-cell abundance. (c) Hallmark pathway associations. (d) Immune-cell abundance across combined Stem.Sig/TMB strata. HSHT, high Stem.Sig/high TMB; HSLT, high Stem.Sig/low TMB; LSHT, low Stem.Sig/high TMB; LSLT, low Stem.Sig/low TMB. (e,f) Tumor-type median trends between Stem.Sig and ITH or TMB, interpreted descriptively because the displayed unit is tumor-type median.

Fig. 4. Exploratory association of Stem.Sig with pretreatment anti-PD-1 GBM outcomes. (a) Validation-set AUC ranking across candidate algorithms. (b) Comparator-signature benchmark. (c,d) Validation and test ROC curves for the retained exploratory model. (e,f) Kaplan-Meier plots showing exploratory survival separation between model-defined groups. These analyses are hypothesis-generating and not a clinically deployable predictive assay.

Fig. 5. WGCNA-based prioritization and preliminary expression support for candidate hub genes. (a) Brown-module membership versus gene significance. (b) Yellow-module membership versus gene significance. (c) Overlap between Stem.Sig genes and genes from GBM-associated WGCNA modules. (d) Dot plot showing average expression and compartment frequency for candidate hub genes. (e) qPCR expression support for SOX6, BCHE, and LINC00511. (f) Representative western blot and quantification for SOX6 and BCHE normalized to Actin. LINC00511 was assessed by qPCR only.

Statements and declarations

Funding: No funding was received for this study.

Competing interests: The authors declare no relevant financial or non-financial interests.

Author contributions: Hongbo Guo and Chengwei Chen conceived and supervised the study. Simin Yang, Sihan Ju, Yunjie Liao, and Shengxuan Ding contributed equally to data curation, bioinformatic analysis, figure preparation, preliminary expression-support assays, and manuscript drafting. Hongbo Guo and Chengwei Chen interpreted the data and revised the manuscript. All authors reviewed and approved the final manuscript.

Ethics approval: The public transcriptomic components of this study used publicly available, previously published, de-identified datasets. The local expression-support experiments using de-identified clinical residual/previously archived human tissue specimens from patients at Zhujiang Hospital were approved by the Ethics Committee of Zhujiang Hospital, Southern Medical University (approval no. 2024BA0040-GC). The Ethics Committee waived the requirement for written informed consent because the study used clinical residual/previously archived specimens, all patient information was anonymized, and no identifiable personal information was involved. The study was conducted in accordance with the ethical principles of the Declaration of Helsinki.

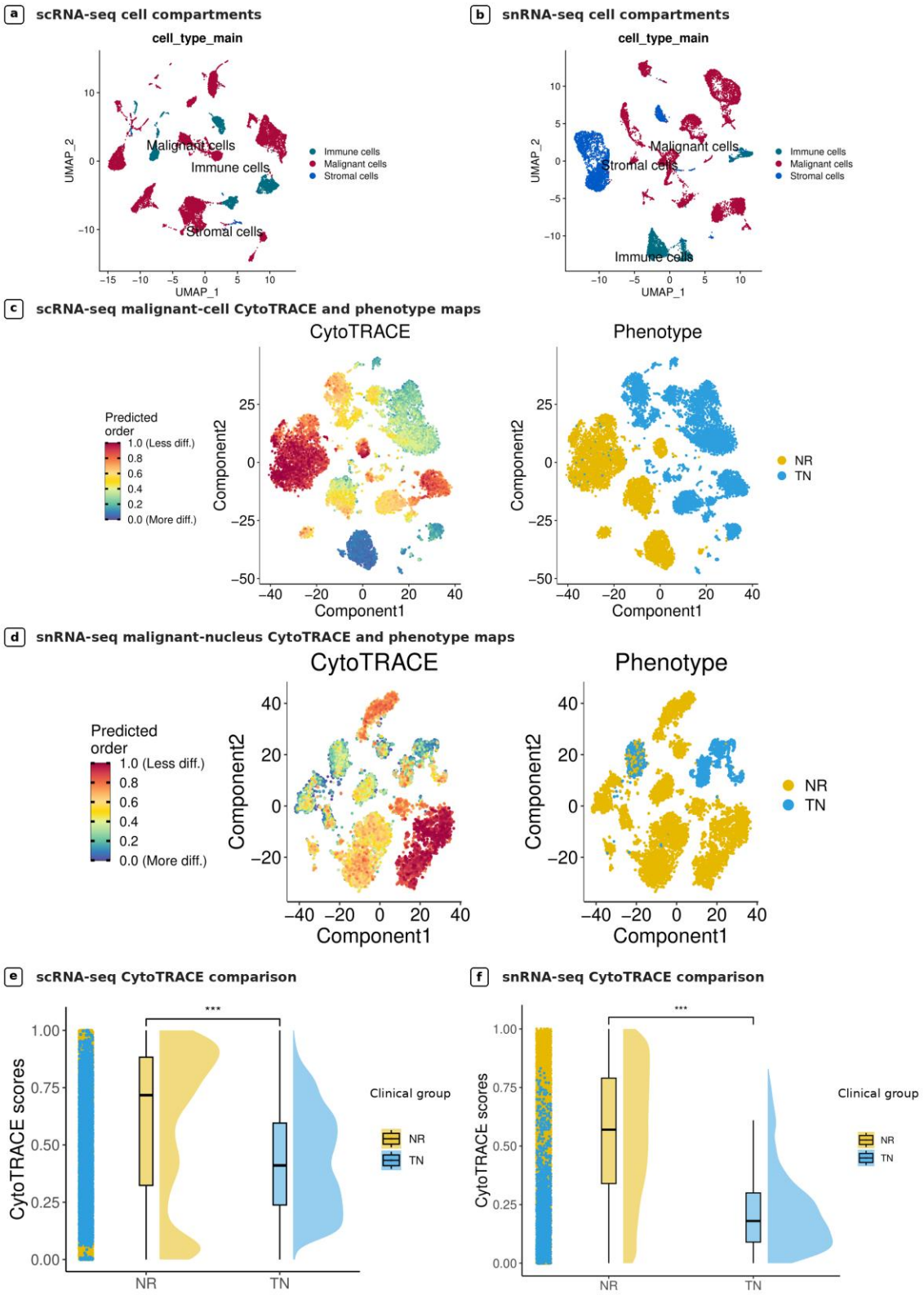
Consent to participate: The requirement for written informed consent was waived by the Ethics Committee of Zhujiang Hospital, Southern Medical University (approval no. 2024BA0040-GC) for the local de-identified clinical residual/previously archived specimen component. The public transcriptomic datasets analyzed in this study were previously published and de-identified.

Consent for publication: Not applicable; this manuscript contains no identifiable individual data or images.

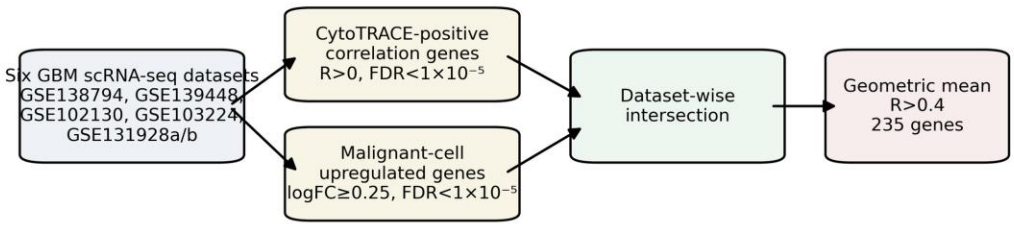
Data availability: Public datasets analyzed in this study include GSE138794, GSE139448, GSE102130, GSE103224, GSE131928a, GSE131928b, TCGA-GBM, GSE68848, GSE4290, and anti-PD-1 GBM cohorts summarized in Online Resource 1, Supplementary Table S1, including GSE121810 where accessioned data were available. The final StemSig_235_unique_genes.csv file and the standard Stem.Sig.gmt file are included in the Online_Resources folder. De-identified summary qPCR and western blot values from the local expression-support experiments are summarized in Online Resource 1, Supplementary Table S3, and assay/material details are summarized in Supplementary Table S4; additional processed values are available from the corresponding authors on reasonable request.

Code availability: Analysis code is available from the corresponding authors on reasonable request.

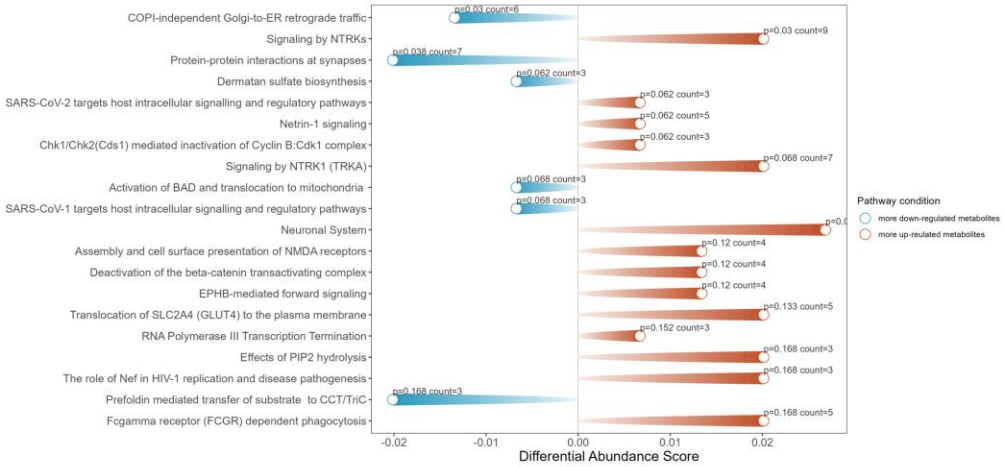
Acknowledgements: The authors thank the investigators and data portals that made the public GBM transcriptomic datasets available for research reuse.


a scRNA-seq cell compartments
cell_type_main
Malignant cells
Immune cells
Stromal cells
UMAP_1
UMAP_2
b snRNA-seq cell compartments
cell_type_main
Malignant cells
Stromal cells
Immune cells
UMAP_1
UMAP_2
c scRNA-seq malignant-cell CytoTRACE and phenotype maps
CytoTRACE
Phenotype
Predicted order
1.0 (Less diff.)
0.8
0.6
0.4
0.2
0.0 (More diff.)
Component1
Component2
NR
TN
d snRNA-seq malignant-nucleus CytoTRACE and phenotype maps
CytoTRACE
Phenotype
Predicted order
1.0 (Less diff.)
0.8
0.6
0.4
0.2
0.0 (More diff.)
Component1
Component2
NR
TN
e scRNA-seq CytoTRACE comparison
CytoTRACE scores
***
Clinical group
NR
TN
f snRNA-seq CytoTRACE comparison
CytoTRACE scores
***
Clinical group
NR
TN

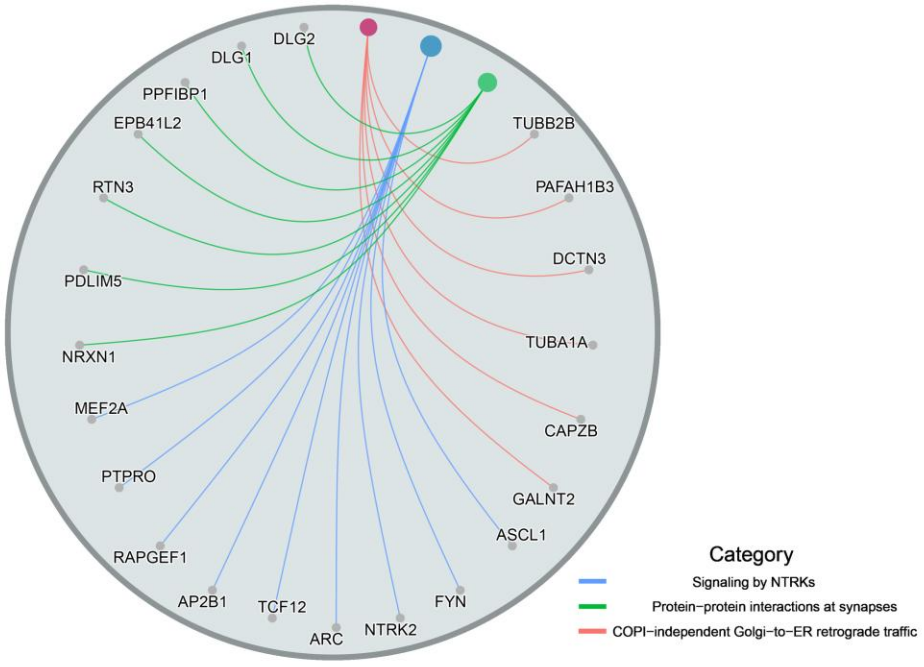
a **Derivation workflow**

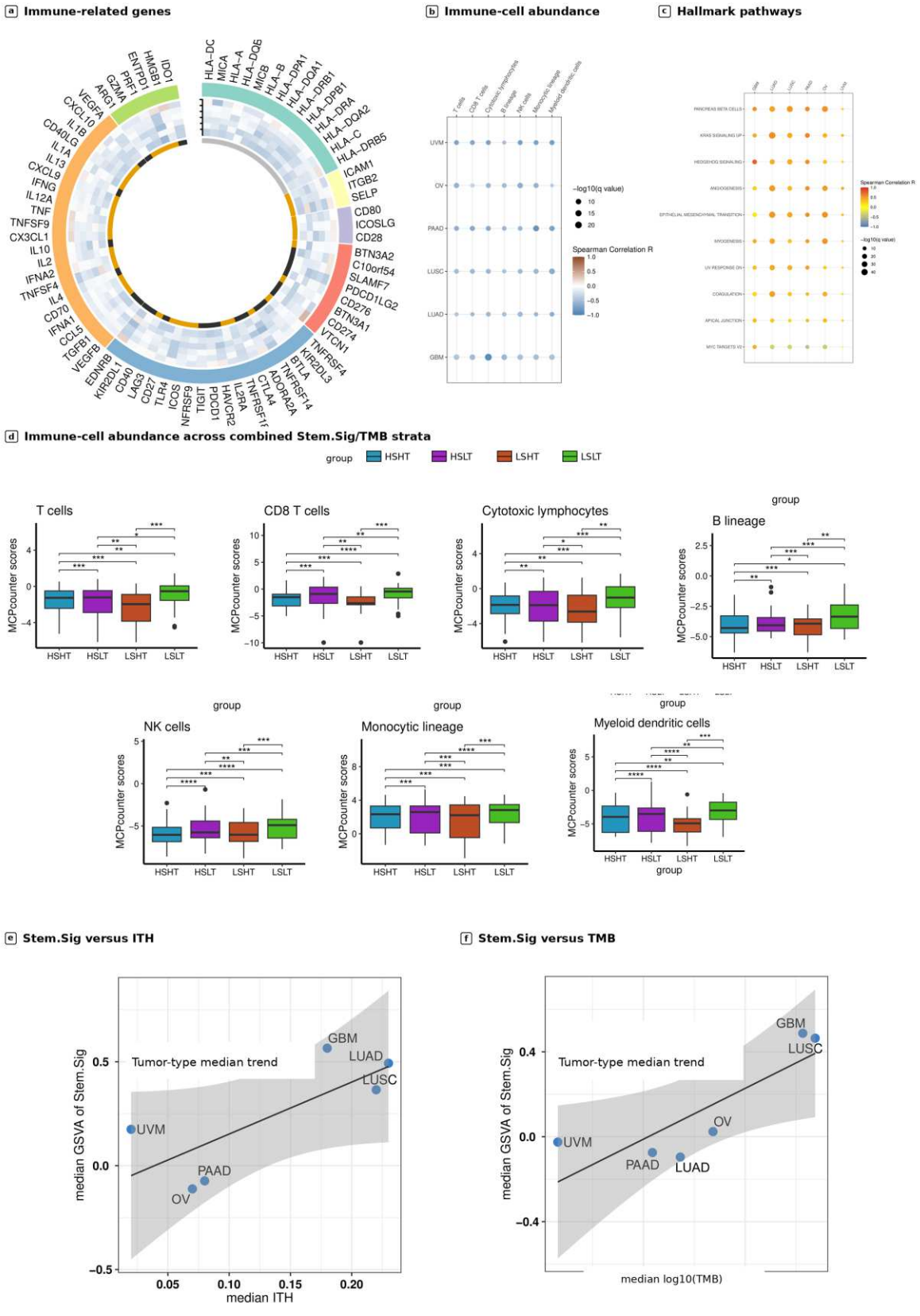


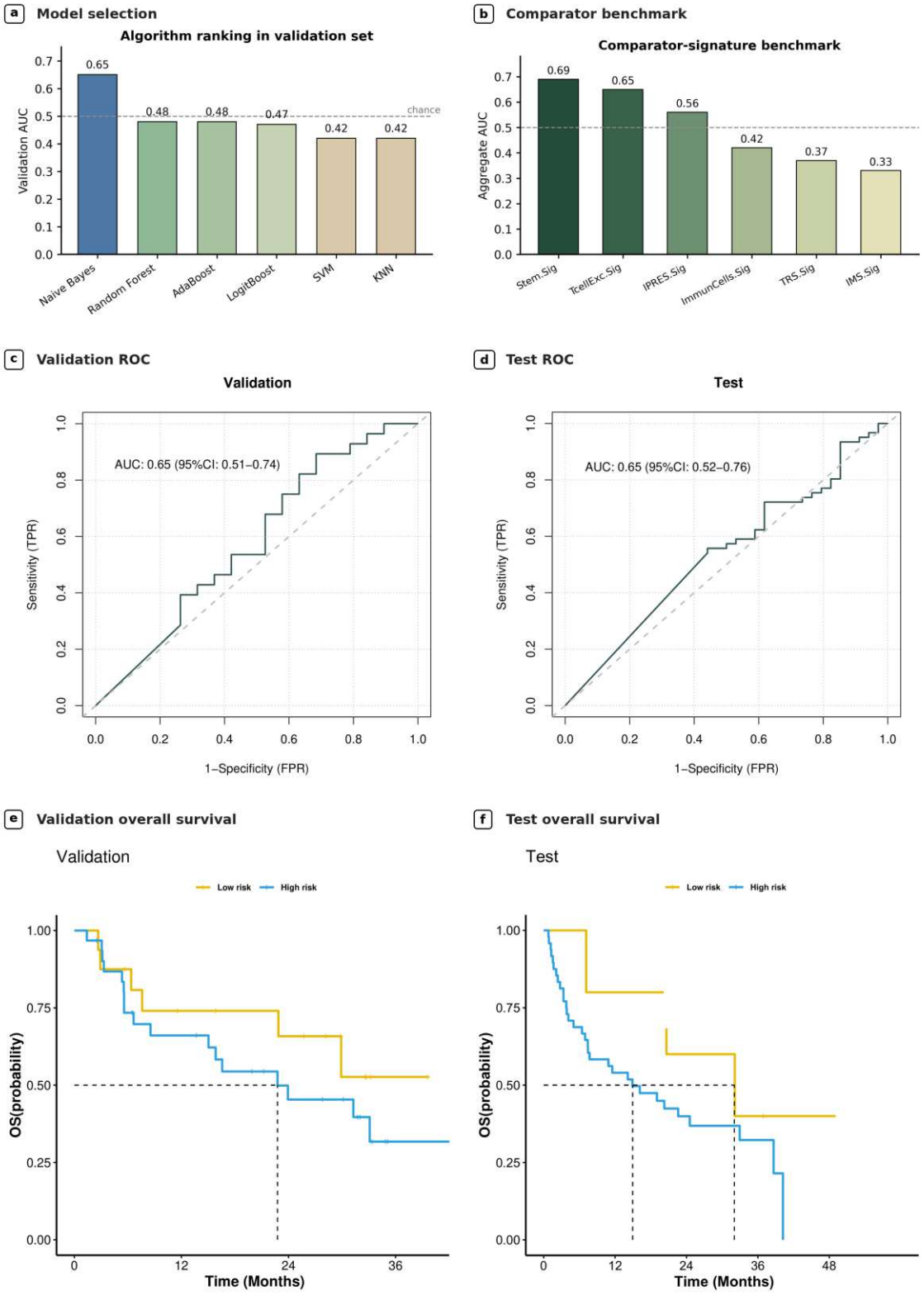
b **Differential pathway abundance**

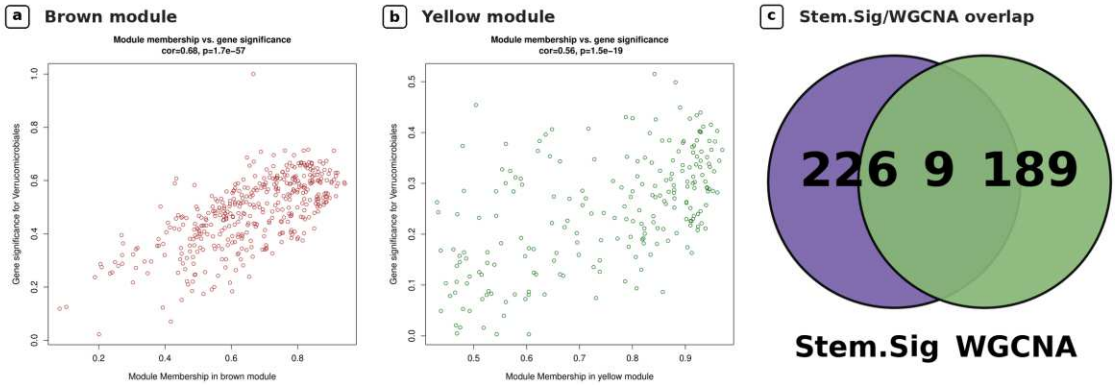


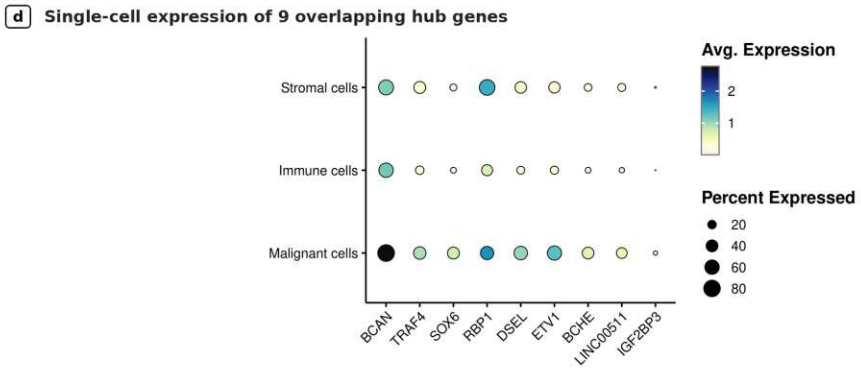
c **Gene-pathway network for representative enriched terms**

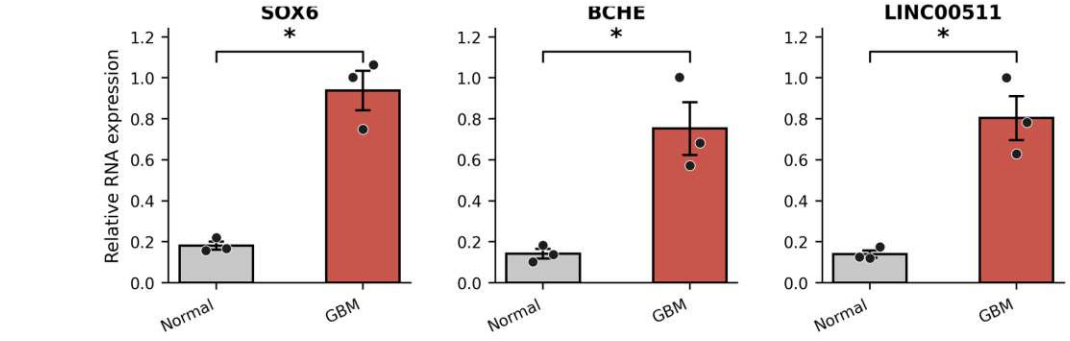


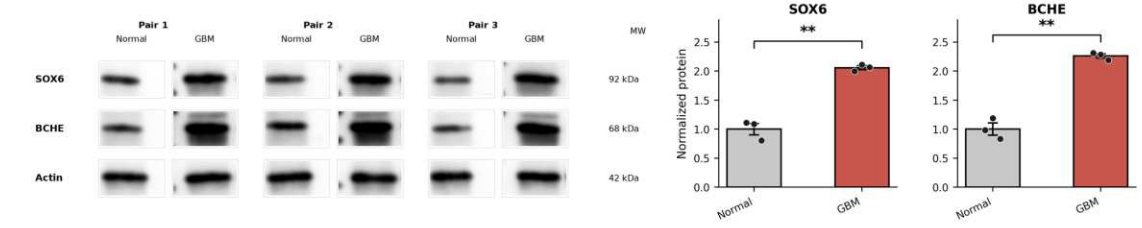

a Immune-related genes
b Immune-cell abundance
c Hallmark pathways
d Immune-cell abundance across combined Stem.Sig/TMB strata
group HSHT HSLT LSHT LSLT
T cells
CD8 T cells
Cytotoxic lymphocytes
B lineage
NK cells
Monocytic lineage
Myeloid dendritic cells
MCPcounter scores
HSHT HSLT LSHT LSLT
e Stem.Sig versus ITH
Tumor-type median trend
median GSVA of Stem.Sig
median ITH
GBM
LUAD
LUSC
UVM
PAAD
OV
f Stem.Sig versus TMB
median log10(TMB)


a Model selection
Algorithm ranking in validation set
Validation AUC
0.65
0.48
0.48
0.47
0.42
0.42
chance
Naive Bayes
Random Forest
AdaBoost
LogitBoost
SVM
KNN
b Comparator benchmark
Comparator-signature benchmark
Aggregate AUC
0.69
0.65
0.56
0.42
0.37
0.33
Stem.Sig
TcellExc.Sig
IPRES.Sig
ImmunCells.Sig
TRS.Sig
IMS.Sig
c Validation ROC
Validation
AUC: 0.65 (95%CI: 0.51–0.74)
Sensitivity (TPR)
1–Specificity (FPR)
d Test ROC
Test
AUC: 0.65 (95%CI: 0.52–0.76)
Sensitivity (TPR)
1–Specificity (FPR)
e Validation overall survival
Validation
Low risk
High risk
OS(probability)
Time (Months)
f Test overall survival
Test
Low risk
High risk
OS(probability)
Time (Months)


a Brown module
Module membership vs. gene significance
cor=0.68, p=1.7e−57
Gene significance for Verrucomicrobiales
Module Membership in brown module
b Yellow module
Module membership vs. gene significance
cor=0.56, p=1.5e−19
Gene significance for Verrucomicrobiales
Module Membership in yellow module
c Stem.Sig/WGCNA overlap
226 9 189
Stem.Sig WGCNA


d Single-cell expression of 9 overlapping hub genes
Stromal cells
Immune cells
Malignant cells
BCAN
TRAF4
SOX6
RBP1
DSEL
ETV1
BCHE
LINC00511
IGF2BP3
Avg. Expression
Percent Expressed


e qPCR expression support for prioritized candidates
SOX6
BCHE
LINC00511
Relative RNA expression
Normal
GBM


f Western blot expression support with molecular-weight labels
Pair 1
Pair 2
Pair 3
Normal
GBM
SOX6
BCHE
Actin
MW
92 kDa
68 kDa
42 kDa
Normalized protein

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- GBMJNOOnlineResource1Supplementary.docx