Neuro-Oncology Advances

3(1), 1–15, 2021 | doi:10.1093/noajnl/vdab049 | Advance Access date 16 April 2021

An integrative analysis of genome-wide 5-hydroxymethylcytosines in circulating cell-free DNA detects noninvasive diagnostic markers for gliomas

Jiajun Cai, Chang Zeng, Wei Hua, Zengxin Qi, Yanqun Song, Xingyu Lu, Dongdong Li, Zhou Zhang, Xiaolong Cui, Xin Zhang, Zixiao Yang, Jinsen Zhang, Kai Quan, Wei Zhu, Jiabin Cai, Chuan He, Shi-Yuan Cheng, Wei Zhang, and Ying Mao

Department of Neurosurgery, Huashan Hospital, Fudan University, Shanghai, China (J-J.C., W.H., Z.Q., X.Z., Z.Y., J.Z., K.Q., W.Zhu, Y.M.); Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA (C.Z., Z.Z., W.Z.); Shanghai Epican Genetech Co., Ltd., Shanghai, China (Y.S., X.L., D.L.); Department of Chemistry, The University of Chicago, Chicago, Illinois, USA (X.C., C.H.); Department of Liver Surgery and Transplantation, Liver Cancer Institute, Zhongshan Hospital, Fudan University, Shanghai, China (J-B.C.); The Ken and Ruth Davee Department of Neurology, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA (S.Y.C.); State Key Laboratory of Medical Neurobiology, School of Basic Medical Sciences, and The Collaborative Innovation Centre for Brain Science, Fudan University, Shanghai, China (Y.M.)

Corresponding authors: Ying Mao, Department of Neurosurgery, Huashan Hospital, Fudan University, Shanghai 200040, China (yingmao168@hotmail.com); Wei Zhang, Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois 60611, USA (wei.zhang1@northwestern.edu).

J-J.C., C.Z., W.H., and Z.Q. share the first authorship on this publication.

Y.M., W.Z, S.Y.C., and C.H. share the senior authorship on this publication.

Abstract

Background. Gliomas, especially the high-grade glioblastomas (GBM), are highly aggressive tumors in the central nervous system (CNS) with dismal clinical outcomes. Effective biomarkers, which are not currently available, may improve clinical outcomes through early detection. We sought to develop a noninvasive diagnostic approach for gliomas based on 5-hydroxymethylcytosines (5hmC) in circulating cell-free DNA (cfDNA).

Methods. We obtained genome-wide 5hmC profiles using the 5hmC-Seal technique in cfDNA samples from 111 prospectively enrolled patients with gliomas and 111 age-, gender-matched healthy individuals, which were split into a training set and a validation set. Integrated models comprised 5hmC levels summarized for gene bodies, long noncoding RNAs (lncRNAs), *cis*-regulatory elements, and repetitive elements were developed using the elastic net regularization under a case–control design.

Results. The integrated 5hmC-based models differentiated healthy individuals from gliomas (area under the curve [AUC] = 84%; 95% confidence interval [CI], 74–93%), GBM patients (AUC = 84%; 95% CI, 74–94%), WHO II-III glioma patients (AUC = 86%; 95% CI, 76–96%), regardless of *IDH1* (encoding isocitrate dehydrogenase) mutation status or other glioma-related pathological features such as TERT, TP53 in the validation set. Furthermore, the 5hmC biomarkers in cfDNA showed the potential as an independent indicator from *IDH1* mutation status and worked in synergy with *IDH1* mutation to distinguish GBM from WHO II-III gliomas. Exploration of the 5hmC biomarkers for gliomas revealed relevance to glioma biology.

Conclusions. The 5hmC-Seal in cfDNA offers the promise as a noninvasive approach for effective detection of gliomas in a screening program.

© The Author(s) 2021. Published by Oxford University Press, the Society for Neuro-Oncology and the European Association of Neuro-Oncology. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Key Points

- Genome-wide 5hmC features demonstrated gene regulatory relevance and tissue origin.
- Integrated 5hmC-based models can distinguish gliomas from healthy individuals.
- The 5hmC biomarkers can independently separate GBM from WHO grade II-III gliomas.

Importance of the Study

The clinical outcomes of malignant gliomas especially glioblastoma (GBM) remain dismal. For example, only ~5% of GBM patients survive 5 years after diagnosis. However, early detection of gliomas and GBM has been challenging. Because of the blood–brain barrier, protein biomarkers in serum that are diagnostic in other human cancers are not reliable for gliomas. The 5hmC (5-hydroxymethylcytosine), an emerging cytosine modification with gene regulatory roles and tissue specificity, has been recognized as a potential cancer biomarker. Importantly, recent technical advancement has allowed robust profiling of 5hmC in convenient liquid biopsies, such as the circulating cell-free DNA (cfDNA) from plasma. In this study, the 5hmC in cfDNA was used to develop integrated diagnostic models for gliomas. The 5hmC-Seal technique in cfDNA offers a clinically feasible approach for effective diagnosis of gliomas and GBM, providing an attractive alternative for early detection of this deadly cancer.

Glioma is a common primary intracranial tumor with dismal prognosis. The 5-year survival rate for glioblastoma (GBM, WHO grade IV) is only ~5%, and ~25% for grade III glioma.[1] Although there have been advances in diagnosis and personalized medicine, malignant gliomas (i.e., WHO grade III-IV) are still characterized by high degree of anaplasia, aggressiveness, and poor clinical outcomes.[2,3] Therefore, effective biomarkers for early detection of asymptomatic gliomas among healthy individuals are urgently needed for improving clinical outcomes.[4]

High-tech imaging approaches such as MR spectroscopy (MRS) and Positron Emission Tomography-Computed Tomography (PET-CT) have enabled the detection and diagnosis of a wide range of diseases in the brain and the spine. Radiomics or deep learning technique might predict some molecular information, such as *IDH* (encoding isocitrate dehydrogenase) mutations,[5] chromosome 1p/19q co-deletion,[5,6] and *MGMT* (encoding O-6-methylguanine-DNA methyltransferase) promoter methylation[5] that may stratify patients for treatments, but these modalities still have limited clinical utility.[5,7] Liquid biopsies, such as the blood, CSF (cerebrospinal fluid) have been exploited to determine *IDH* mutations and other circulating biomarkers such as microRNAs in gliomas.[4,8–11] Despite the promising results, the clinical utility of these studies might be compromised by limited sample size, unsatisfactory reproducibility across different profiling platforms and lack of proper study design.[8,12]

Because the epigenome is inherently more stable compared to mRNA, it is feasible to assess epigenetic patterns in a variety of clinical specimens, from genomic DNA to circulating cell-free DNA (cfDNA). The latter may be released from gliomas through apoptosis and necrosis into the peripheral blood or CSF, thus providing genetic and epigenetic information of the tumor.[13–15] As a noninvasive approach, an epigenetics-based test in cfDNA would be a convenient and practical tool for the clinical workup of brain tumors, such as in a screening program targeting general population and outpatient diagnosis of asymptomatic patients among healthy individuals.

The 5-hydroxymethylcytosines (5hmC) are emerging epigenetic markers with distinct gene regulatory functions and genomic distributions from the more abundant 5-methylcytosines (5mC). In the cell, 5mC can be oxidized into 5hmC catalyzed by the TET (ten-eleven translocation) family of enzymes in an active demethylation process.[16] Unlike 5mC that represses not only protein-coding genes but also a vast amount of transposons in the human genome, 5hmC, a stable modification generated in an active demethylation process, could better reflect specific gene activation changes.[17] A recent study on a 5hmC map of human tissues also showed that the distribution of 5hmC is particularly enriched in tissue-specific enhancers, distinct from 5mC.[18] In addition, reduced global levels of 5hmC were found in various cancers including gliomas, indicating its relevance in cancer pathobiology. Notably, epigenetic dysregulation of *TET2* was shown to repress

the mRNA expression of *TET2*, which further affected tumor growth in GBM.[19] The levels of 5hmC in tumor tissues were also found to be associated with survival in a study of 30 GBM patients.[20] Specifically, the cluster of patients with lower 5hmC levels are associated with older age at diagnosis and shorter median survival.[19] Importantly, unlike conventional bisulfite conversion-based approaches, which is unable to differentiate 5hmC from 5mC, recent technical advancement has enabled robust and direct profiling of 5hmC in clinical specimens including cfDNA from peripheral blood, demonstrating the promise of exploiting 5hmC in cfDNA as noninvasive biomarkers for cancer diagnosis.

In this study, we aimed to develop integrated diagnostic models based on the 5hmC profiles in cfDNA for distinguishing patients with gliomas from healthy individuals, as well as GBM from WHO II-III gliomas. The 5hmC-Seal technique,[21] a highly sensitive chemical labeling technique was employed to profile 5hmC in cfDNA.[22] The genome-wide 5hmC profiles allowed us to explore the diagnostic biomarkers not only at gene level, but also the possibility to evaluate their synergy with a variety of genomic feature types, such as long noncoding RNAs (lncRNAs), histone marks, and repetitive elements. Moreover, we explored functional relevance of the 5hmC biomarkers for gliomas. Our innovative findings lay the foundation for further development of a robust liquid biopsy-based tool that can be utilized for early detection of gliomas.

Materials and Methods

Study Participants

A total of 111 adult patients (≥18 years) with newly diagnosed primary gliomas (WHO II, n = 32; WHO III, n = 15; WHO IV, n = 64) were prospectively enrolled at Huashan Hospital of Fudan University in Shanghai, China between February 2017 and February 2018 (Table 1). Patient diagnosis and tumor grading were confirmed by a study neuropathologist, following the WHO classification and grading system for CNS tumors.[3,23] Lower grade glioma was defined as WHO grade II-III tumors.[13,24] Peripheral blood was collected before any radical treatment (i.e., surgical resection, chemotherapy, or radiation therapy) or steroid treatment. We obtained baseline clinical, pathological, and treatment data from medical records, as well as *IDH1*, *TERT, TP53,* and *ATRX* mutation status and 1p/19q co-deletion, which were determined using immunohistochemistry or the next-generation sequencing (NGS) (Table 1, Supplementary Methods). In addition, 111 age-, gender-matched healthy participants were recruited from individuals who underwent regular physical examinations at Zhongshan Hospital of Fudan University in Shanghai, China (GSE112679).[25] An additional set of 27 patients with gliomas were recruited at Huashan Hospital in 2019 for independent validation. This study was reviewed and approved by the Ethics Committee at Huashan Hospital. Written informed consent was obtained from each participant.

Sample Preparation, 5hmC-Seal Profiling, and Data Processing

For each study participant, approximately 5 mL of frozen plasma was collected from peripheral blood, followed by cfDNA extraction and the 5hmC-Seal profiling. Details about the 5hmC-Seal library construction, the NGS, and the data processing pipelines are described in Supplementary Methods.[21,22] The raw and processed 5hmC-Seal data have been deposited into the NIH Gene Expression Omnibus (GEO): GSE132118.

Developing Diagnostic Models for Gliomas

The 111 WHO II-IV glioma patients and the 111 age-, gender-matched healthy individuals were randomly grouped into a training set (WHO II-III, n = 33; GBM, n = 40; healthy individuals, n = 75) and an internal validation set (WHO II-III, n = 14; GBM, n = 24; healthy individuals, n = 36) with a balanced distribution of age, gender, and *IDH1* mutation status and other pathological features (Figure 1, Table 1). The brain-derived H3K4me1 and H3K27ac loci (hg19) were obtained from the Roadmap Epigenomics Project.[26] For each genomic feature type, a separate 5hmC-based diagnostic model for gliomas was developed using a 2-step procedure as described in Supplementary Methods. Integrative models were further evaluated by incorporating different combinations of genomic feature types (Supplementary Methods).

Functional Exploration and Co-localization with Gene Regulatory Elements

We explored the underlying biological connections of the detected 5hmC biomarkers that were differentially modified between healthy individuals and patients with gliomas, based on Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG)[27,28] as described in Supplementary Methods.

Results

Demographics and Clinical Characteristics of the Study Participants

As given in Table 1, the median age of the 111 patients with newly diagnosed gliomas (WHO II-IV) was 49.0 years (range, 20–72 years) and 66.7% (n = 74) were males, comparable to the healthy individuals. Among the 64 GBM patients, 90% patients (n = 58) were *IDH1* wild type, 26.6% patients (n = 17) were *TERT* wild type, and among the 40 patients with available 1p/19q co-deletion information, all patients had no chromosome 1p/19q co-deletion. Among the 47 WHO II-III patients, 46.8% (n = 22) were astrocytic gliomas, and 40.4% (n = 19) were oligodendrogliomas. Of note, there were no significant differences in terms of gender, age, and *IDH1, TERT, TP53,* and *ATRX* mutation status between the training set and the validation set. Regarding location, 25.0% (n = 16) had tumors in the

Table 1. Demographics and Clinical Characteristics of the Study Participants

	WHO II/WHO III Glioma					WHO IV Glioma (GBM)					Healthy Subjects				
	Training (*N* = 33)		Validation (*N* = 14)			Training (*N* = 40)		Validation (*N* = 24)			Training (*N* = 75)		Validation (*N* = 36)		
	No./Value	%	No./Value	%	*P*	No./Value	%	No./Value	%	*P*	No./Value	%	No./Value	%	*P*
Age –Years (mean ± sd)	42.5 ± 11.4		41.1 ± 9.6		.67[a]	51.8 ± 12.7		49.5 ± 12.5		.50[b]	47.4 ± 13.0		46.9 ± 13.0		.83[a]
Sex (Male)	23	69.7	11	78.6	.79[b]	26	65	14	58.3	.79[b]	50	66.7	24	66.7	1.00[b]
IDH1[IHC/NGS]					.79[b]					.82[b]					
Mutant	23	69.7	11	78.6		3	7.5	3	12.5						
Wild-type	10	30.3	3	21.4		37	92.5	21	87.5						
TERT[NGS]					1.00[b]					.64[b]					
Mutant	9	27.3	3	33.3		12	30	11	45.8						
Wild-type	13	39.4	6	66.7		11	27.5	6	25						
1p/19q co-deletion[NGS]					.97[b]										
Yes	10	30.3	3	33.3		/	/	/	/						
No	14	42.4	6	66.7		23	57.5	17	70.8						
TP53[IHC]					.80[b]					.49[b]					
Mutant	9	27.3	5	35.7		15	37.5	10	41.7						
Wild-type	22	66.7	8	57.1		19	47.5	7	29.2						
ATRX[IHC]					.51[b]					.96[b]					
Mutant	21	63.6	8	57.1		33	82.5	20	83.3						
Wild-type	8	24.2	6	42.9		5	12.5	2	8.3						
Hemisphere					.44[b]					1.00[b]					
Left	17	51.5	6	42.9		12	30	7	29.2						
Right	11	33.3	8	57.1		25	62.5	14	58.3						
Location															
Frontal	14	42.4	10	71.4		15	37.5	11	45.8						
Temporal	11	33.3	5	35.7		14	35	6	25						
Parietal	3	9.1	2	14.3		2	5	3	12.5						
Occipital	1	3				5	12.5	2	8.3						
Thalamus	1	3				2	5	1	4.2						
Ultimate Treatment															
Radiation therapy	32	97	13	92.9		36	87.5	20	83.3						
Chemotherapy	32	97	12	85.7		34	85	19	79.2						

Table 1. Continued

	WHO II/WHO III Glioma					WHO IV Glioma (GBM)					Healthy Subjects			
	Training ($N = 33$)		Validation ($N = 14$)			Training ($N = 40$)		Validation ($N = 24$)			Training ($N = 75$)		Validation ($N = 36$)	
	No./Value	%	No./Value	%	*P*	No./Value	%	No./Value	%	*P*	No./Value	%	No./Value	%	*P*
Event					.91[b]					.55[b]					
Event	4	12.1	1	7.1		21	52.5	6	25						
Event Time - Months	11.0 ± 6.8		17.0 ± 0.0			10.0 ± 5.1		7.7 ± 2.9							
Event Free	15	45.5	8	57.1		11	27.5	6	25						
Event Time - Months	15.3 ± 3.0		17.9 ± 4.3			14.6 ± 2.3		19.2 ± 6.4							

ATRX, ATRX chromatin remodeler; IDH, isocitrate dehydrogenase; TERT, telomerase reverse transcriptase; TP53, tumor protein p53; WHO, World Health Organization.
p: *P*-values obtained from [a] the 2-tailed *t*-test and [b] the Pearson's chi-squared test for 2-sample proportions.
IHC: the results were from immunohistochemical assay.
NGS: the results were from next-generation sequencing.
No./Value: the continuous variables will be displayed as (mean ± sd); the categorical variables will be displayed as the number of patients in each category.
%: percentage out of the total number.
/: data not available.

temporal lobes, 37.5% (n = 24) in the frontal lobes, and 29.7% (n = 19) in the right hemispheres. The majority of the patients (91.0%) ultimately received standardized treatments after blood collection, including surgery and adjuvant radiochemotherapy. In addition, the additional 27 independent glioma patients (WHO II-III, n = 10; GBM, n = 17) had a median age of 48.0 years and 70.4% were males.

Concentration of cfDNA and Genomic Distributions of 5hmC in cfDNA

Compared with the age- and gender-matched healthy individuals, those patients with gliomas showed higher levels of extracted cfDNA (Wilcoxon text $P < .0001$; Supplementary Figure 1A). Principal components analysis indicated no systemic bias and batch effect in the overall 5hmC-Seal profiling data (Supplementary Figure 1B). Similar to our previous findings in gastrointestinal cancers,[22,25] the detected 5hmC modifications in cfDNA were more abundant in gene bodies and exonic regions relative to the flanking genic regions and depleted at the promoter regions (Figure 2A). Notably, the distribution of 5hmC profiles in cfDNA from patients with gliomas reflected their putative roles in gene activation, significantly co-localized with enhancers markers: H3K4me1 and H3K27ac derived from brain tissues in the Roadmap Epigenomics Project (Figure 2A).[26] Differential analyses identified 653 gene bodies, 146 Alu elements, 255 LINE-1 elements, 918 lncRNAs, 1585 H3K4me1 loci and 907 H3K27ac loci between glioma patients and healthy individuals at a false discovery rate (FDR) < 0.05 and fold change > 20% (Supplementary Figure 1C–G, Supplementary Table S1). The majority of these differential 5hmC features (91–99%) showed higher modification levels in glioma patients compared to healthy individuals (Figure 2B). Hierarchical clustering suggested distinct discriminating capacities in detecting gliomas across different genomic feature types (Figure 2C). For example, differential 5hmC in lncRNAs alone could separate 77.5% healthy individuals and 74.8% into 2 major clusters (Figure 2C), similar to Alu elements, apparently outperforming other features. We then compared the differentially hydroxymethylated (5hmC) genes (FDR < 0.05) in cfDNA with those differentially methylated (5mC) genes between glioma tumors and normal tissue controls from The Cancer Genome Atlas (TCGA), which used the Illumina Infinium 27K microarray.[29] Of the 1270 genes that were differentially methylated from TCGA (FDR < 0.05 and $\Delta\beta > 0.2$),[29] 440 genes showed differential hydroxymethylation in our cfDNA samples (Supplementary Figure 1H).

Tissue Specificity of 5hmC Biomarkers in cfDNA

We evaluated whether the 5hmC modifications in cfDNA between glioma patients and healthy individuals were more likely to be brain relevant. When comparing the 5hmC modifications between brain- and other tissue-derived histone modification marks, the proportion of

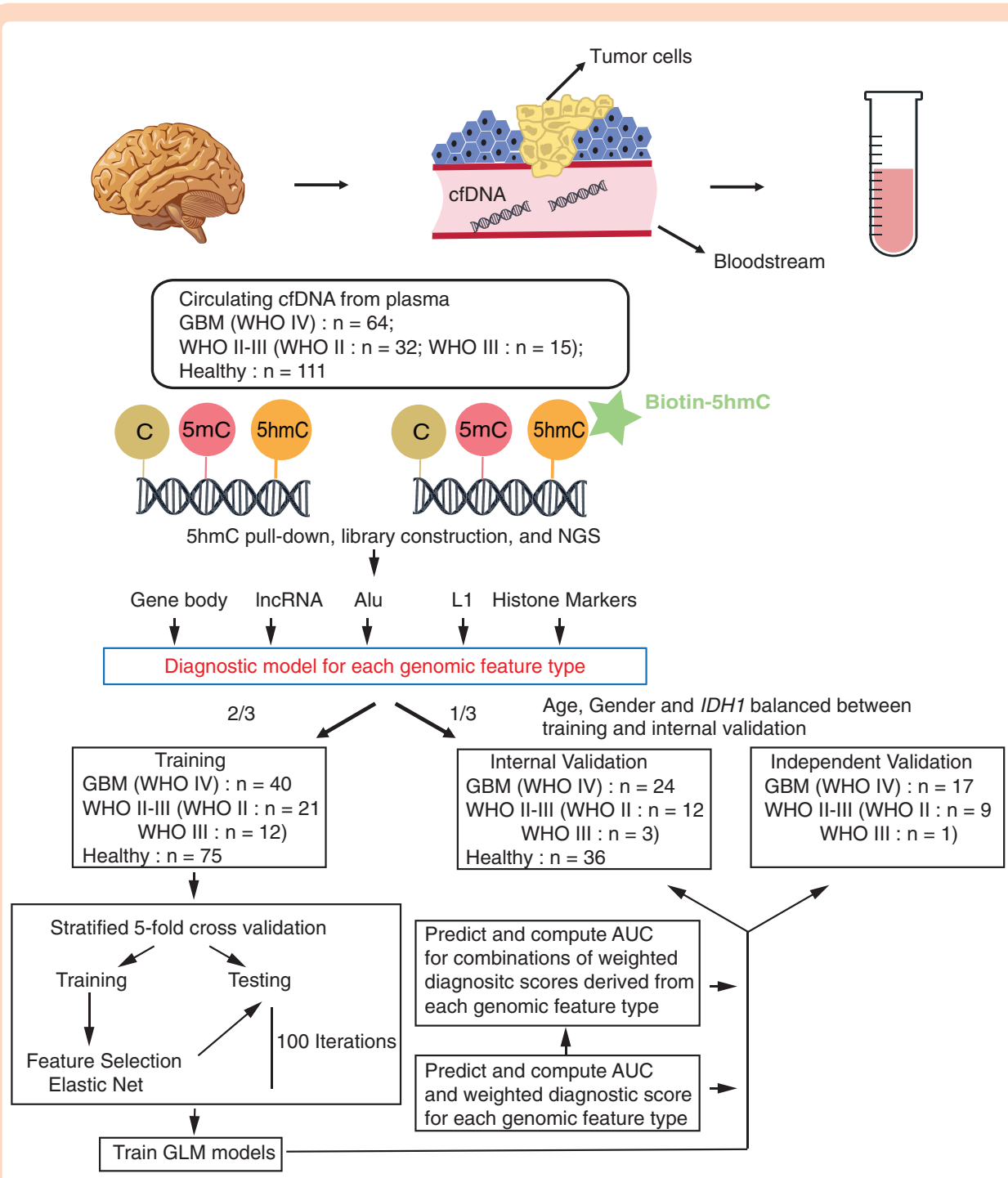

Figure 1. Study design. Study subjects with WHO II-IV gliomas and age- and gender-matched healthy individuals are randomly divided into the training set (2/3) and the internal validation set (1/3) with a balanced distribution of age, gender, and *IDH1* mutation status, followed by modeling. The models are evaluated in an independent validation set. NGS, next-generation sequencing; GBM, WHO IV glioma; lncRNA, long noncoding RNA; GLM, generalized linear model.

differentially hydroxymethylated H3K27ac and H3K4me1 marks between glioma patients and healthy individuals were found to be enriched with brain-derived modification loci (pair-wise 1-tailed 2-proportions z-test $P <$ .01; Figure 2D). Of note, the enrichment of H3K27ac, a histone mark associated with active enhancers, is more prominent in the brain as compared to that of poised enhancer marker H3K4me1. We therefore primarily focus on H3K27ac loci for our future analysis. Furthermore, we took advantage of a dataset comprised of RNA-seq

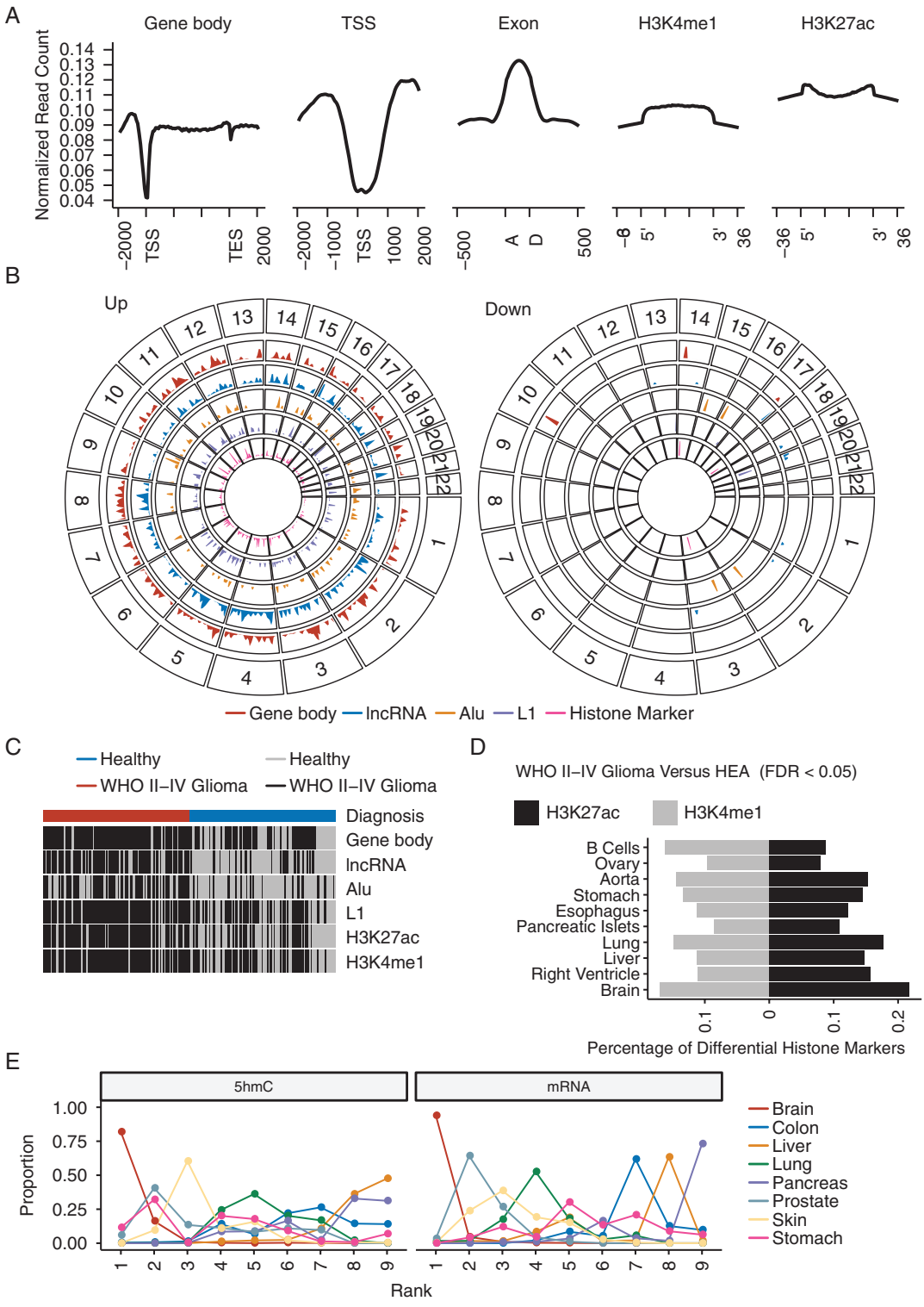

Figure 2. Genomic distributions of 5hmC in cfDNA. The 5hmC-Seal data from the 111 glioma patients (WHO II-IV) and 111 healthy individuals are used to characterize genomic distributions. (A) The 5hmC profiles are distinctly distributed across various genomic feature types. The read counts are normalized to per million counts. (B) Genomic distributions of differentially hydroxymethylated features (FDR < 0.05 and fold change > 20%) between glioma patients and healthy individuals. Up: up-modified; Down: down-modified. (C) Hierarchical clustering of samples based on differentially hydroxymethylated features (FDR < 0.05 and fold-change > 20%) between glioma patients and healthy individuals. (D) Comparison across tissues shows tissue specificity regarding the proportions of differentially hydroxymethylated H3K4me1 and H3K27ac sites (FDR < 0.05). (E) Comparison of the ranks of differentially hydroxymethylated genes (FDR < 0.01) from cfDNA in normal tissues (5hmC and mRNA) shows enrichment in the brain. TSS: transcription start site; TES: transcription end site; D: splice donor site; A: splice acceptor site; FDR: false discovery rate.

and 5hmC-Seal data profiled in normal adult tissues, including brain, colon, liver, lung, pancreas, prostate, skin, and stomach,[18] to evaluate tissue relevance of the glioma-associated gene bodies in cfDNA. Interestingly, these glioma-associated gene bodies were more likely to be ranked as brain-derived in terms of probability for both 5hmC profiles and mRNA expression across various tissues, further supporting their tumor relevance (Figure 2E).

Integrative Diagnostic Models for Gliomas

In the training set, following feature selection using the elastic net, we identified a total of 4 gene bodies, 8 lncRNAs, 3 Alu elements, 9 LINE-1 elements, and 18 H3K27ac histone marks (Figure 3A, Table 2) to distinguish gliomas from healthy individuals as well as GBM from WHO II-III gliomas (Figure 1). Initially, the diagnostic models based on individual genomic feature types were evaluated separately. The wd-scores computed based on the 8 lncRNAs showed the highest overall capacity over other genomic feature types for distinguishing healthy individuals from patients with gliomas (training: area under the curve [AUC] = 0.87, 95% confidence interval [CI], 0.81–0.93; testing: AUC = 0.83, 95% CI, 0.73–0.92), GBM (training AUC = 0.90, 95% CI, 0.84–0.96; testing: AUC = 0.84, 95% CI, 0.73–0.94), WHO II-III (training: AUC = 0.83, 95% CI, 0.75–0.91; testing: AUC = 0.83, 95% CI, 0.71–0.94), as well as from *IDH1* wild-type gliomas (training: AUC = 0.87, 95% CI, 0.80–0.93; testing: AUC = 0.83, 95% CI, 0.72–0.94), and *IDH1* mutant gliomas (training: AUC = 0.89, 95% CI, 0.83–0.95; testing: AUC = 0.84, 95% CI, 0.73–0.95). In more challenging scenarios, the lncRNA-based model achieved the AUC of 0.66 (95% CI, 0.48–0.83) in distinguishing GBM from WHO II-III gliomas and the AUC of 0.74 (95% CI, 0.57–0.91) in distinguishing *IDH1* wild-type and *IDH1* mutant patients in the testing set of samples (Supplementary Table S2).

Before building an integrative model, we evaluated the concordance of models based on different genomic feature types with the Lin's concordance correlation coefficient (CCC) (Supplementary Figure 2A–J).[30] The highest concordance was between the gene body-based and lncRNA-based models (CCC: 0.68, 95% CI, 0.61–0.74) (Supplementary Figure 2E) and the lowest concordance was between the lncRNA-based and Alu-based models (CCC: 0.32, 95% CI, 0.21–0.42; Supplementary Figure 2H), thus suggesting the relative independence of these models and the feasibility of integrating these features to improve performance. Specifically, an integrated model of 3 Alu elements and 8 lncRNAs achieved the highest predictive accuracy for distinguishing gliomas from healthy individuals in the validation set (AUC = 0.84, 95% CI, 0.74–0.93), regardless of their pathological features. For example, 85.7%, 85.7%, 80.0%, and 82.1% glioma patients with *IDH1*, *TERT*, *TP53*, and *ATRX* mutations were accurately identified using this integrated model in the validation set. In contrast, an integrated model of 4 gene bodies, 3 Alu elements, and 8 lncRNAs showed the highest distinguishing capacity for GBM from healthy individuals in the validation set (AUC = 0.84, 95% CI, 0.74–0.94). In comparison, an integrated model of 9 LINE-1 elements, 18 H3K27ac loci, and 8 lncRNAs outperformed individual genomic feature types regarding distinguishing WHO II-III gliomas from healthy individuals (Figure 3B–G, Supplementary Table S2).

Interestingly, the integrated model of 3 Alu elements and 4 gene bodies showed improved capacity for distinguishing GBM from WHO II-III gliomas with an AUC of 0.76 (95% CI, 0.60–0.93) in the validation set, compared with any other combinations or individual genomic feature type alone (Figure 3H, Supplementary Table S2). Furthermore, a multivariable analysis suggested that when combined with *IDH1* mutation, the integrated model of 3 Alu elements and gene bodies demonstrated improved discriminatory capability for GBM and WHO II-III (AUC = 0.88; 95% CI, 0.77–1.00), outperforming *IDH1* mutation status or the 5hmC model alone in the validation set (Figure 3I).

We then evaluated the predictive performance of the 5hmC models in the 27 independent glioma samples. Using the integrated model for gliomas (ie, 3 Alu elements and 8 lncRNAs) (Figure 3B and Supplementary Table S2), we detected 20 glioma patients with a slightly higher accuracy for WHO II-III (80%) compared to 70% for GBM (Figure 3J).

Diagnostic Scores and Clinical Characteristics

We next examined the wd-scores based on each genomic feature type by the WHO grading system and *IDH1* mutation status (Figure 3K). In general, the wd-scores based on different genomic feature types (eg, gene bodies, lncRNAs) showed a similar trend of increasing scores as the WHO grades advanced. As shown in Figure 3K, the gene body-based wd-scores increased in a linear trend as the WHO grades advanced in all glioma patients with available grade information (n = 111), dramatically different from healthy controls (p-trend P-value < .001). Specifically, GBM patients showed a distinct wd-score distribution from that of healthy individuals (Wilcoxon rank sum test P-value < .01; Figure 3K). Glioma patients regardless of *IDH1* mutation status also exhibited significantly higher wd-scores than healthy individuals. Furthermore, glioma patients with *TP53* mutations had significantly higher (Wilcoxon rank sum test P-value = .011) gene body-based wd-scores than patients without *TP53* mutations (Supplementary Figure 2K). In comparison, there were no significant score differences observed between WHO II-III and GBM, or *IDH1* wild-type and *IDH1* mutant, or astrocytic and oligodendroglial gliomas or across different locations of the tumors (eg, left vs right hemispheres, frontal vs temporal lobes). However, distinct 5hmC modification patterns were observed between oligodendroglial and astrocytic gliomas from differential analysis, therefore suggesting the potential association between 5hmC and morphological features (Supplementary Figure 2M). In addition, applying the gene body-based diagnostic model in a set of 1036 patients with hepatocellular carcinoma (HCC) from our previous publication showed distinguishing capability of the wd-scores for glioma and HCC (Student's t-test P-value < .05; Supplementary Figure 2L).[25]

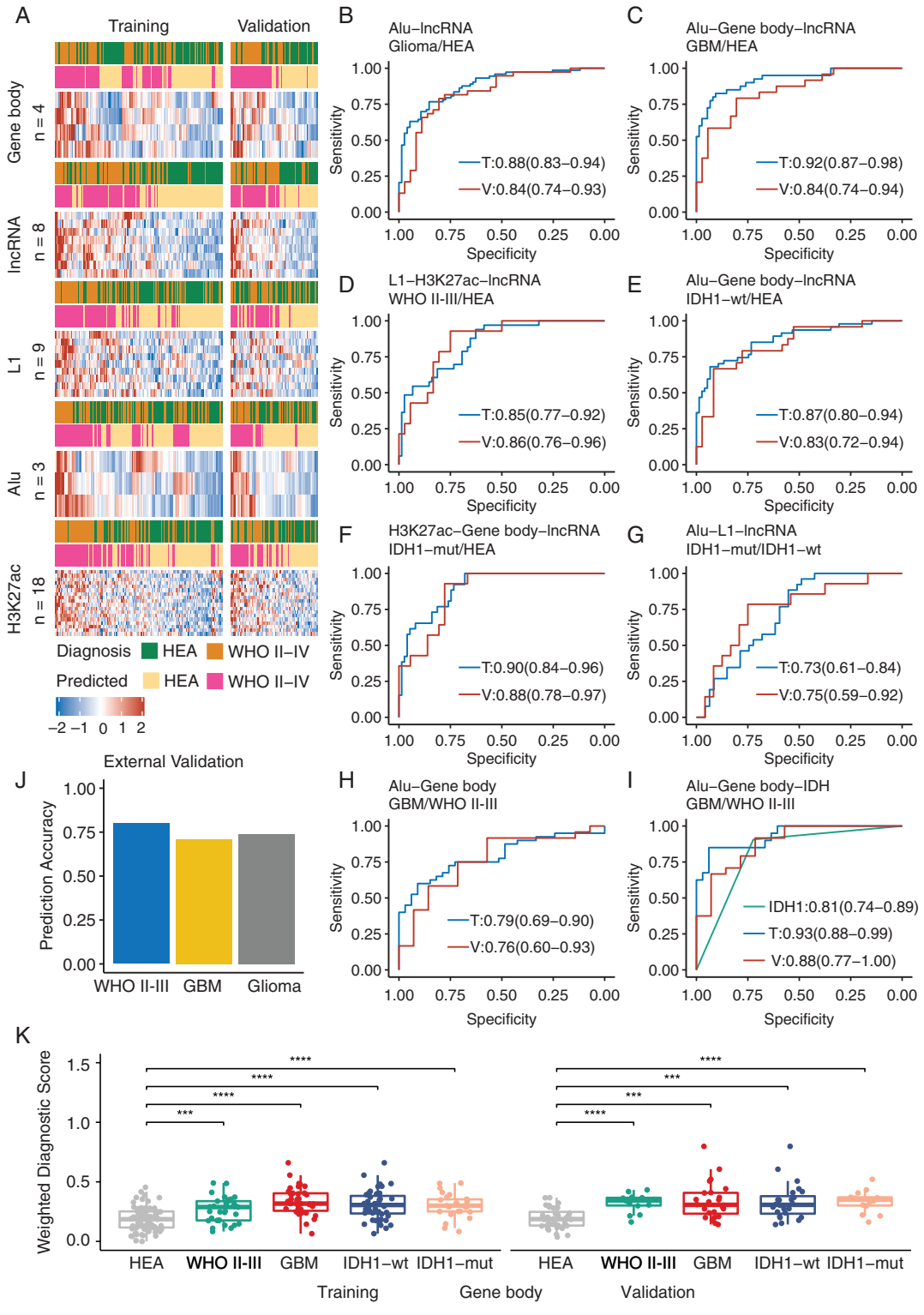

Figure 3. Performance of the integrated diagnostic models for gliomas. (A) The heatmaps show the final features selected in the training set (T) and the internal validating set (V). The AUCs in the training set and internal validation set show the performance of the integrated diagnostic models of (B) Glioma versus HEA: Alu- and lncRNA-based wd-scores; (C) GBM versus HEA: Alu-, gene body-, and lncRNA-based wd-scores; (D) WHO II-III versus HEA: LINE-1-, H3K27ac-, and lncRNA-based wd-scores; (E) *IDH1* wild-type versus HEA: Alu-, gene body-, and lncRNA-based wd-scores for distinguishing; (F) *IDH1* mutant versus HEA: H3K27ac-, gene body-, and lncRNA-based wd-scores; (G) *IDH1* mutant versus wild type: Alu-, LINE-1-, and lncRNA-based wd-scores; (H) GBM versus WHO II-III: Alu- and gene body-based wd-scores; and (I) GBM versus II-III: Alu-, gene body-based

Functional Exploration

Because 5hmC modifications were observed at gene regulatory elements (eg, lncRNAs and histone modification marks; Figure 2A), we further investigated the co-localization patterns between these regulatory features and their host genes. Specifically, genomic features including 146 Alu elements, 255 LINE-1 elements, 918 lncRNAs, and 907 H3K27ac loci with differential 5hmC between gliomas and healthy individuals were scanned for genes within ±2Kb of these regulatory features (Supplementary Table S1). The Venn diagram showed that the host genes of these regulatory features were generally unique and distinct from each other (Figure 4A).

GO enrichment analysis of these host genes together with the 653 differential gene bodies (FDR < 0.05 and fold change > 20%) identified 6, 0, 6, 21, and 6 GO biological processes (gene count ≥ 5 and empirical $P < .01$) for gene bodies, lncRNA-, Alu-, LINE-1-, and H3K27ac- derived host genes, respectively. In comparison, one GO biological process "localization" was shared between Alu- and H3K27ac-derived host genes (Figure 4B, Supplementary Table S3). Several GO biological processes are involved in glioma development such as the "Wnt signaling pathway" and "neuron development and axonogenesis " (Figure 4B, Supplementary Table S3).[31] Interestingly, among the 653 differential gene bodies for gliomas were enriched with KEGG pathways such as "retrograde endocannabinoid signaling" and "ECM-receptor interaction," all of which have been implicated in glioma malignancy or pathogenesis (Figure 4C).[32,33] Additional KEGG enrichment analyses of Alu-, LINE-1-, and lncRNA-derived host genes also highlighted several glioma- or cancer- relevant pathways such as "neurotrophin signaling pathway,"[34] "Rap1 signaling pathway,"[35] "cAMP signaling pathway,"[36] "Ras signaling pathway,"[37] and "Proteoglycans in cancer" (Figure 4D–F, Supplementary Table S3).

Discussion

In this study, we developed a cfDNA-based, noninvasive epigenetic approach that would provide an alternative solution to the early detection of gliomas in a screening context. Specifically, through a systematic and integrative analysis of the genome-wide 5hmC profiles provided by the 5hmC-Seal approach in cfDNA, we identified integrated 5hmC diagnostic models and wd-scores derived from a variety of genomic feature types. The 5hmC models demonstrated robust performance for distinguishing gliomas (or GBM, WHO II-III) from healthy controls. For example, the CSF-derived ctDNA analysis published by Miller et al. identified 42/85 (sensitivity: 49.4%) glioma patients with tumor-derived genetic alteration and 7/7 (specificity: 100%) controls with nonmalignant neurological conditions without the alteration. The integrated model comprised of gene bodies and lncRNAs in our study can achieve a specificity of 94.7% when sensitivity is fixed at 50.0% in the validation set. When adopting the optimal threshold determined by Youden index, the abovementioned model can achieve a sensitivity of 78.9% and specificity of 80.6% in the validation set. However, this comparison needs to be interpreted with caution because of the limited number of controls and lack of validation in the CSF analysis. The 5hmC models also showed excellent capacity for distinguishing GBM from WHO II-III gliomas. Although *IDH* mutation frequencies differ significantly among glioma grades, providing diagnostic values for differentiating WHO II-III gliomas from GBM, it is not always readily available because of the inconvenience of tumor biopsy,[38] or in a general screening program. The 5hmC biomarkers were relatively independent from *IDH1* mutation status, thus offering the possibility of improving differentiating GBM from WHO II-III by combining with *IDH1* mutation.

Although our primary goal was to develop a cfDNA-based diagnostic model, we also explored some basic biology questions regarding tissue origins of cfDNA and comparison of epigenetic and genetic patterns of both tissue and blood samples in glioma patients. Our data suggested that the glioma-associated 5hmC biomarkers showed a genomic distribution that revealed their tumor origin and potential gene regulation relevance. In addition, 34.6% genes that were differentially methylated in TCGA glioma tissue samples were observed to differ in 5hmC levels in our cfDNA data, suggesting the correlation between 5hmC and 5mC. Distinctive sets of genes with differential methylation and hydroxymethylation were also observed. However, these observations should be interpreted with caution because of the different profiling platforms and the fact that the methylation array cannot distinguish 5hmC from 5mC, as well as differences due to sample types (tissue vs blood) and populations.

Furthermore, we evaluated the relevance of the detected 5hmC markers and their host genes with glioma pathobiology. A closer look at the genomic features utilized to develop diagnostic models also shed some light into to the crosstalk between hydroxymethylation and pathobiology in gliomas. For example, Rap1 signaling, a canonical pathway enriched among glioma-associated genomic features, including *RAP1A/B*, which are key components of neurotrophin signaling as well, plays critical roles in mediating cell proliferation in GBM.[39] Among the 4 genes comprising the gene body-based diagnostic model for gliomas are *RGS4* and *NR2F2*. *RGS4* (the regulator of G-protein signaling 4), is a key driver of glioma invasiveness.[40] *NR2F2* (also known as Chicken ovalbumin upstream

wd-scores and *IDH1* mutation status. (J) Performance of the final diagnostic model for gliomas (Panel B) in an independent set of 27 patients. (K) The boxplots show the distributions of wd-score derived from 4 gene bodies in all samples: HEA ($n = 111$), WHO II-III ($n = 47$), GBM ($n = 64$), patients with wild-type *IDH1* ($n = 71$), and patients with mutant *IDH1* ($n = 40$). AUC, area under the curve. HEA: healthy controls; T: training set; V: internal validation set; *IDH1*-wt: *IDH1* wild-type; *IDH1*-mut: *IDH1* mutant.

Table 2. Elastic Net Selected Final Features

Feature	Category	BaseMean	$\log_2$(fold-change)	*P*	FDR	Chr	Start	End	Annotation_2KB	Annotation
RGS4	Gene body	10.76	0.32	.000	0.000	chr1	163038565	163046592		
NPHS2	Gene body	18.17	0.39	.000	0.000	chr1	179519674	179545087		
PRR15	Gene body	10.41	0.28	.000	0.000	chr7	29603427	29606911		
NR2F2	Gene body	17.26	0.39	.000	0.000	chr15	96869167	96883492		
AC092415.1	lncRNA	18.58	0.45	.000	0.000	chr3	28038701	28070095		
RP11-659P15.1	lncRNA	14.73	0.34	.000	0.000	chr11	57811582	57827610		OR9Q1
RP11-643G5.6	lncRNA	27.77	0.47	.000	0.000	chr11	89279805	89322779		
RP11-2L4.1	lncRNA	18.84	0.48	.000	0.000	chr16	82587654	82608815		
RP11-173L6.1	lncRNA	14.79	0.29	.000	0.000	chr18	73556880	73580207		
AC003973.4	lncRNA	10.34	0.64	.000	0.000	chr19	22200761	22219117		
CTC-435M10.6	lncRNA	20.00	0.32	.000	0.000	chr19	41931264	41932142	CTC-435M10.3 B3GNT8	BCKDHA B3GNT8
RP11-87M18.2	lncRNA	40.09	0.40	.000	0.000	chrX	36383741	36458375		
SINE/Alu_AluSx1_197020	Alu	9.48	0.26	.001	0.019	chr1	95143645	95143908		
SINE/Alu_AluSz_2273844	Alu	11.35	0.25	.000	0.007	chr2	26331749	26332078		RAB10
SINE/Alu_AluSc_4223943	Alu	9.59	0.48	.000	0.000	chr7	32707218	32707529		
LINE/L1_L1P2_603915	L1	9.90	0.33	.000	0.002	chr10	97535364	97538037		ENTPD1
LINE/L1_L1MB8_1301438	L1	17.94	0.25	.000	0.002	chr13	101006542	101009464		PCCA
LINE/L1_L1PB1_2691404	L1	17.04	0.39	.000	0.000	chr20	29649078	29651586		
LINE/L1_L1PA4_2945947	L1	11.59	0.29	.000	0.002	chr3	23337809	23343301		UBE2E2
LINE/L1_L1PA5_3107752	L1	10.74	0.19	.003	0.026	chr3	119282028	119288164		
LINE/L1_L1MC4_3200492	L1	9.88	0.27	.001	0.012	chr3	171915141	171916338		FNDC3B
LINE/L1_L1PA2_3230068	L1	11.46	0.27	.000	0.007	chr3	187608009	187614002		
LINE/L1_L1PB_3652941	L1	9.47	0.26	.000	0.005	chr5	53551600	53553176		ARL15
LINE/L1_L1MB3_3745436	L1	9.53	0.23	.001	0.011	chr5	108682005	108682581		PJA2
Rank_2809	H3K27ac	17.33	0.26	.000	0.000	chr2	131625860	131632306		ARHGEF4
Rank_9305	H3K27ac	14.76	0.27	.000	0.000	chr8	145923444	145928208		
Rank_15572	H3K27ac	19.36	0.31	.000	0.000	chr15	100024591	100028923		
Rank_22708	H3K27ac	12.77	0.20	.001	0.007	chr4	15377249	15378595		C1QTNF7
Rank_29494	H3K27ac	13.80	0.22	.000	0.002	chr10	24495172	24499400		KIAA1217
Rank_31171	H3K27ac	13.05	0.42	.000	0.000	chr15	60290939	60300474		
Rank_31379	H3K27ac	10.98	0.24	.001	0.004	chr2	69930813	69932615		ANXA4
Rank_31544	H3K27ac	10.00	0.31	.000	0.000	chr7	29603478	29606668	PRR15	PRR15
Rank_36687	H3K27ac	10.22	0.45	.000	0.000	chr16	10021857	10030768		GRIN2A
Rank_37578	H3K27ac	15.46	0.26	.000	0.000	chr12	92838256	92840403		
Rank_39601	H3K27ac	12.92	0.23	.000	0.001	chr1	203303037	203305772		
Rank_41373	H3K27ac	10.69	0.34	.000	0.000	chr5	172504968	172508599		CREBRF

Table 2. Continued

Feature	Category	BaseMean	$\log_2$(fold-change)	*P*	FDR	Chr	Start	End	Annotation_2KB	Annotation
Rank_41739	H3K27ac	16.92	0.30	.000	0.000	chr10	13862041	13865998		FRMD4A
Rank_42607	H3K27ac	10.29	0.21	.001	0.007	chr12	66537458	66540359		RP11-745O10.4TMBIM4
Rank_47609	H3K27ac	11.08	0.24	.000	0.003	chr8	103134371	103136917	NCALD	NCALD
Rank_49686	H3K27ac	12.59	0.30	.000	0.000	chr5	90421443	90424504		ADGRV1
Rank_50546	H3K27ac	12.70	0.26	.000	0.000	chr10	97455745	97458888		
T2Rank_53694	H3K27ac	26.55	0.30	.000	0.000	chr12	83114658	83128925		TMTC2

BaseMean: mean of normalized counts for all samples.
$\log_2$(fold-change): the logarithmic fold change(ie, $\log_2$(glioma/healthy); $\log_2$(fold-change) > 0, higher 5hmC modification levels in patients with gliomas. $\log_2$(fold-change) < 0, higher 5hmC modification levels in healthy individuals).
P: Wald Statistic *P*-value.
FDR: FDR-adjusted *P*-value.
Chr: chromosome location.
Start: start genomic coordinate.
End: end genomic coordinate.
Annotation: residing gene of the feature.
Annotation 2K: neighboring genes (2KB).

promoter transcription factor II), a member of the nuclear receptor superfamily, is known to play a role in promoting angiogenesis in the tumor environment, and it has also been found to be a prognostic marker in WHO II-III patients with *IDH* mutation and 1p/19q co-deletion.[41] P-Rex1, a core element of the chemokine signaling pathway, is known to regulate dependent responses in neutrophiles. It is also a protein primarily expressed in the immune system and the brain,[42] suggesting the contribution of immune system to glioma pathobiology as well.

Technically, the 5hmC-Seal approach in cfDNA has showed value for cancer biomarker discovery in various human cancers, such as colorectal cancer and liver cancer.[22,25] Our findings in gliomas further established the potential of this novel approach to be developed into a multi-cancer detection and screening tool in the future, for example using only a few milliliters of plasmas. However, there are several limitations that could be addressed in future studies. Firstly, although tissue or tumor relevance of blood-derived 5hmC biomarkers can be partially supported in the current study, comparison between 5hmC profiles in cfDNA and paired tissue samples from glioma patients would provide direct support for the connection. Secondly, although major clinical variables (eg, gender, age) were well-balanced between the training and validation sets, future independent validation studies with larger sample size in different grades of gliomas and healthy individuals with various epidemiological characteristics (eg, lifestyle) and more comprehensive pathological information will help address problems such as the potential selection bias or suboptimal classification for our samples. Thirdly, this study was conducted in a Chinese patient population. It would be necessary to evaluate the generalizability of the results in other geographical populations. What is more, in the current study, we used a case–control design. Future development phases, including retrospective longitudinal studies, and prospective screening studies, will help validate and establish the ultimate clinical utilities of this approach.[43] Finally, future development needs to consider the distinguishing capacity of the 5hmC models between gliomas and other non-glioma conditions. Nonetheless, our findings from the current study warrant further investigations using this novel approach in brain cancer.

In conclusion, we have developed noninvasive and multi-feature diagnostic models for gliomas through an integrative analysis of genome-wide 5hmC profiled using the highly sensitive 5hmC-Seal technique in cfDNA samples. The 5hmC-based diagnostic approach using cfDNA can be a highly sensitive and specific tool for early detection of gliomas in population screening, especially for those patients with aggressive tumors. Therefore, as a general tool that can be applied in limited amount of specimens (eg, < 5 mL of plasma) that will be ideal for regular screening or disease monitoring, the 5hmC-Seal in cfDNA offers a clinically feasible solution to address the issue of lacking effective biomarkers for gliomas. Given its flexibility, technical robustness, and noninvasiveness, the 5hmC-Seal approach has the potential to be an integrated part of precision medicine tools to improve clinical outcomes of this deadly disease.

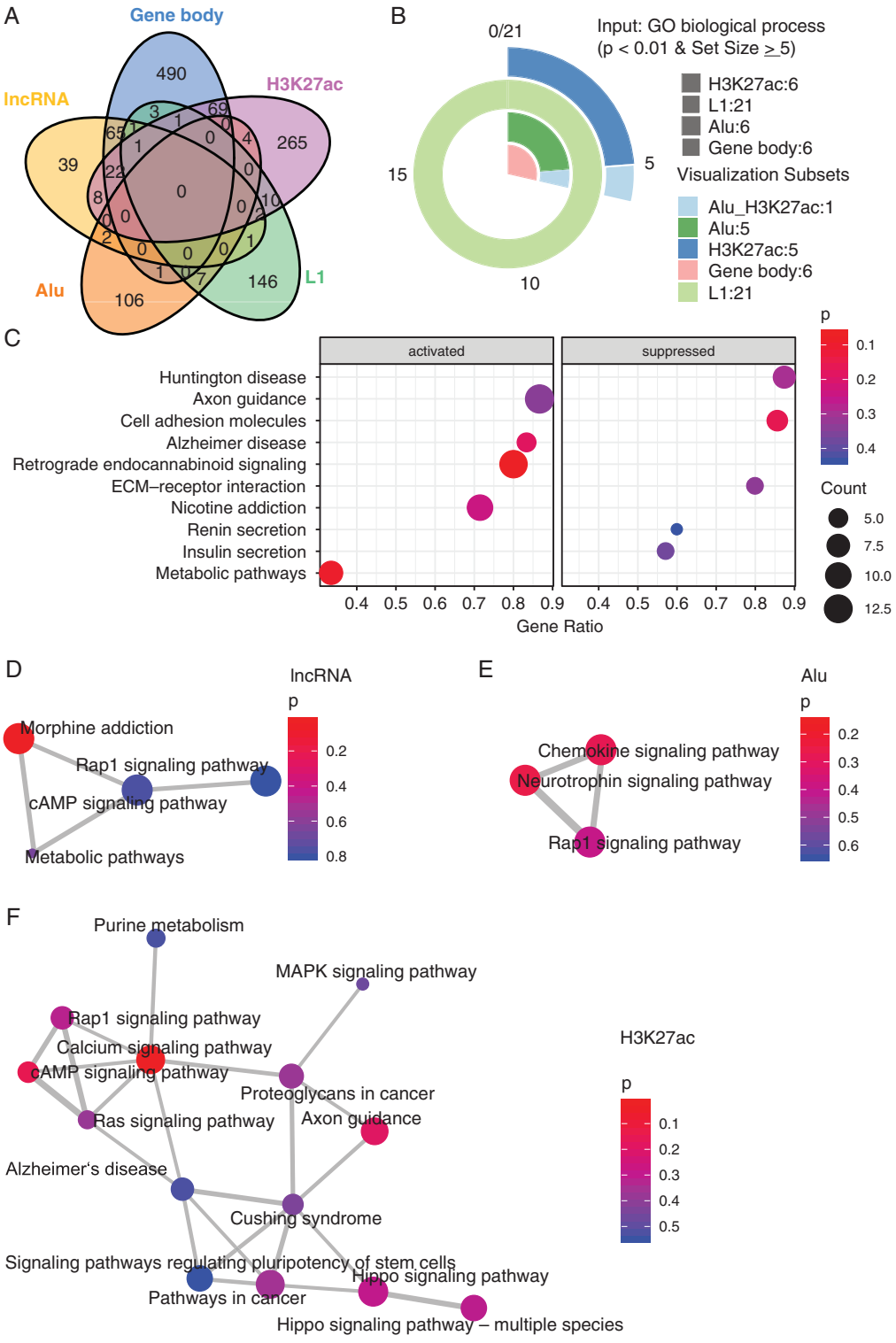

Figure 4. Functional exploration. (A) Venn diagram of residing or neighboring/host genes associated with differentially hydroxymethylated genomic features. (B) The GO enrichment analysis of residing or neighboring/host genes associated with differentially hydroxymethylated genomic features. (C) The KEGG pathway analysis of gene bodies that are differentially modified between gliomas and healthy individuals. (D) The KEGG pathway analysis of lncRNA elements-derived host genes. (E) The KEGG pathway analysis of Alu-derived host genes. (F) The KEGG pathway analysis of H3K27ac loci-derived host genes. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

Supplementary Material

Supplementary material is available at *Neuro-Oncology Advances* online.

Keywords

biomarker | cell-free DNA | diagnosis | glioma | 5-hydroxymethylcytosine

Funding

This study was supported, in part, by grants from the NIH (R21 CA209345 to W.Z. and S.Y.C.), Phi Beta Psi Sorority (to W.Z. and S.Y.C.), the Lou and Jean Malnati Brain Tumor Institute at Northwestern University (to W.Z. and S.Y.C.), the National Natural Science Foundation of China (81572483 and 82072785 to Y.M.; 82072784 to W.H.; 81702461 to Z.Q.; 81502155 to J-B.C.), Shanghai Committee of Science and Technology, China (17430750200 to Y.M.), the International S&T Cooperation Program of China (2014DFA31470 to W.Zhu), and Shanghai Sailing Program (17YF1426600 to Z.Q.). C.H. is a Howard Hughes Medical Institute Investigator.

Conflict of Interest. The 5hmC-Seal technology was invented by C.H. and was licensed by Shanghai Epican Genetech Co., Ltd. for clinical applications in human diseases from the University of Chicago. X.L., Y.S., D.L. are employees and shareholders of Shanghai Epican Genetech Co., Ltd. C.H. and W.Z. are shareholders of Shanghai Epican Genetech Co., Ltd.. C.H. is a scientific founder of Accent Therapeutics, Inc. and a member of its scientific advisory board. All other authors reported no potential conflicts of interest.

Authorship Statement. Conception and design of this study: Y.M., W.Z., S.Y.C., and C.H.; Writing and editing of this manuscript: J-J.C., C.Z., W.H., X.L., S.Y.C., C.H., W.Z., and Y.M.; Participant recruitment: J-J.C., W.H., Z.Q., X.Z., Z.Y., J.Z., K.Q., W.Zhu, Y.M., and J-B.C.; Technical support for the 5hmC-Seal profiling: Y.S., X.L., and D.L.; Data analysis: C.Z., Z.Z., X.C., and W.Z. All authors have read and approved the final version.

References

1. Omuro A, DeAngelis LM. Glioblastoma and other malignant gliomas: a clinical review. *JAMA*. 2013;310(17):1842–1850.
2. 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(7):434–452.
3. Wesseling P, Capper D. WHO 2016 Classification of gliomas. *Neuropathol Appl Neurobiol*. 2018;44(2):139–150.
4. De Mattos-Arruda L, Mayor R, Ng CKY, et al. Cerebrospinal fluid-derived circulating tumour DNA better represents the genomic alterations of brain tumours than plasma. *Nat Commun*. 2015;6:8839.
5. Chang P, Grinband J, Weinberg BD, et al. Deep-learning convolutional neural networks accurately classify genetic mutations in gliomas. *AJNR AM J Neuroradiol*. 2018;39(7):1201–1207.
6. Akkus Z, Ali I, Sedlář J, et al. Predicting deletion of chromosomal Arms 1p/19q in low-grade gliomas from MR images using machine intelligence. *J Digit Imaging*. 2017;30(4):469–476.
7. Smedley NF, Hsu W. Using deep neural networks for radiogenomic analysis. *Proc IEEE Int Symp Biomed Imaging*. 2018;2018:1529–1533.
8. Kros JM, Mustafa DM, Dekker LJ, Sillevis Smitt PA, Luider TM, Zheng PP. Circulating glioma biomarkers. *Neuro Oncol*. 2015;17(3):343–360.
9. Best MG, Sol N, Zijl S, Reijneveld JC, Wesseling P, Wurdinger T. Liquid biopsies in patients with diffuse glioma. *Acta Neuropathol*. 2015;129(6):849–865.
10. Stoicea N, Du A, Lakis DC, Tipton C, Arias-Morales CE, Bergese SD. The MiRNA journey from theory to practice as a CNS Biomarker. *Front Genet*. 2016;7:11.
11. Martínez-Ricarte F, Mayor R, Martínez-Sáez E, et al. Molecular diagnosis of diffuse gliomas through sequencing of cell-free circulating tumor DNA from cerebrospinal fluid. *Clin Cancer Res*. 2018;24(12):2812–2819.
12. Szopa W, Burley TA, Kramer-Marek G, Kaspera W. Diagnostic and therapeutic biomarkers in glioblastoma: current status and future perspectives. *Biomed Res Int*. 2017;2017:8013575.
13. Miller AM, Shah RH, Pentsova EI, et al. Tracking tumour evolution in glioma through liquid biopsies of cerebrospinal fluid. *Nature*. 2019;565(7741):654–658.
14. Faria G, Silva E, Da Fonseca C, Quirico-Santos T. Circulating cell-free DNA as a prognostic and molecular marker for patients with brain tumors under perillyl alcohol-based therapy. *Int J Mol Sci*. 2018;19(6):1610.
15. Nassiri F, Chakravarthy A, Feng S, et al. Detection and discrimination of intracranial tumors using plasma cell-free DNA methylomes. *Nat Med*. 2020;26(7):1044–1047.
16. Branco MR, Ficz G, Reik W. Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nat Rev Genet*. 2011;13(1):7–13.
17. Greenberg MVC, Bourc'his D. The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol*. 2019;20(10):590–607.
18. Cui XL, Nie J, Ku J, et al. A human tissue map of 5-hydroxymethylcytosines exhibits tissue specificity through gene and enhancer modulation. *Nat Commun*. 2020;11(1):6161.
19. García MG, Carella A, Urdinguio RG, et al. Epigenetic dysregulation of TET2 in human glioblastoma. *Oncotarget*. 2018;9(40):25922–25934.
20. Johnson KC, Houseman EA, King JE, von Herrmann KM, Fadul CE, Christensen BC. 5-Hydroxymethylcytosine localizes to enhancer elements and is associated with survival in glioblastoma patients. *Nat Commun*. 2016;7:13177.
21. Han D, Lu X, Shih AH, et al. A highly sensitive and robust method for genome-wide 5hmC profiling ofrare cell populations. *Mol. Cell*. 2016;63(4):711–719.
22. Li W, Zhang X, Lu X, et al. 5-Hydroxymethylcytosine signatures in circulating cell-free DNA as diagnostic biomarkers for human cancers. *Cell Res*. 2017;27(10):1243–1257.
23. Louis DN, Perry A, Reifenberger G, et al. The 2016 World Health Organization Classification of tumors of the central nervous system: a summary. *Acta Neuropathol*. 2016;131(6):803–820.

24. Ceccarelli M, Barthel FP, Malta TM, et al.; TCGA Research Network. Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. *Cell*. 2016;164(3):550–563.
25. Cai J, Chen L, Zhang Z, et al. Genome-wide mapping of 5-hydroxymethylcytosines in circulating cell-free DNA as a non-invasive approach for early detection of hepatocellular carcinoma. *Gut*. 2019;68(12):2195–2205.
26. Kundaje A, Meuleman W, Ernst J, et al.; Roadmap Epigenomics Consortium. Integrative analysis of 111 reference human epigenomes. *Nature*. 2015;518(7539):317–330.
27. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res*. 2019;47(D1):D419–D426.
28. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27–30.
29. Lai RK, Chen Y, Guan X, et al. Genome-wide methylation analyses in glioblastoma multiforme. *PLoS One*. 2014;9(2):e89376.
30. Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*. 1989;45(1):255–268.
31. Gillespie S, Monje M. The neural regulation of cancer. *Annu Rev Cancer Biol*. 2020;4(1):371–390.
32. Massi P, Valenti M, Solinas M, Parolaro D. Molecular mechanisms involved in the antitumor activity of cannabinoids on gliomas: role for oxidative stress. *Cancers (Basel)*. 2010;2(2):1013–1026.
33. Schulze M, Violonchi C, Swoboda S, et al. RELN signaling modulates glioblastoma growth and substrate-dependent migration. *Brain Pathol*. 2018;28(5):695–709.
34. Lawn S, Krishna N, Pisklakova A, et al. Neurotrophin signaling via TrkB and TrkC receptors promotes the growth of brain tumor-initiating cells. *J Biol Chem*. 2015;290(6):3814–3824.
35. Zhang YL, Wang RC, Cheng K, Ring BZ, Su L. Roles of Rap1 signaling in tumor cell migration and invasion. *Cancer Biol Med*. 2017;14(1):90–99.
36. Daniel PM, Filiz G, Mantamadiotis T. Sensitivity of GBM cells to cAMP agonist-mediated apoptosis correlates with CD44 expression and agonist resistance with MAPK signaling. *Cell Death Dis*. 2016;7(12):e2494.
37. Lo HW. Targeting Ras-RAF-ERK and its interactive pathways as a novel therapy for malignant gliomas. *Curr Cancer Drug Targets*. 2010;10(8):840–848.
38. Cohen AL, Holmen SL, Colman H. IDH1 and IDH2 mutations in gliomas. *Curr Neurol Neurosci Rep*. 2013;13(5):345.
39. Sayyah J, Bartakova A, Nogal N, Quilliam LA, Stupack DG, Brown JH. The Ras-related protein, Rap1A, mediates thrombin-stimulated, integrin-dependent glioblastoma cell proliferation and tumor growth. *J Biol Chem*. 2014;289(25):17689–17698.
40. Weiler M, Pfenning PN, Thiepold AL, et al. Suppression of proinvasive RGS4 by mTOR inhibition optimizes glioma treatment. *Oncogene*. 2013;32(9):1099–1109.
41. Xu Z, Yu S, Hsu CH, Eguchi J, Rosen ED. The orphan nuclear receptor chicken ovalbumin upstream promoter-transcription factor II is a critical regulator of adipogenesis. *Proc Natl Acad Sci USA*. 2008;105(7):2421–2426.
42. Weiner OD. Rac activation: P-Rex1 — a convergence point for PIP3 and Gβγ? *Curr Biol*. 2002;12(12):R429–R431.
43. Drucker E, Krapfenbauer K. Pitfalls and limitations in translation from biomarker discovery to clinical utility in predictive and personalised medicine. *EPMA J*. 2013;4(1):7.