1 **Comprehensive molecular characterization of *PRKCG* for glioma diagnosis,**

2 **prognosis and treatment prediction**

3 Lin Liu[1,2,3,4,7], Guangyu Wang[1,2,3,4,5,7], Liguo Wang[6], Chunlei Yu[1,2,3,4], Mengwei Li[1,2,3,4],

4 Shuhui Song[1,2,3,4], Lili Hao[1,2,3,4], Lina Ma[1,2,3,4,*], Zhang Zhang[1,2,3,4,8,*]

5 [1] National Genomics Data Center, Beijing 100101, China

6 [2] BIG Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing

7 100101, China

8 [3] CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of

9 Genomics, Chinese Academy of Sciences, Beijing 100101, China

10 [4] University of Chinese Academy of Sciences, Beijing 100101, China

11 [5] Present Address: The Methodist Hospital Research Institute, 6670 Bertner Ave,

12 Houston, Texas 77030, United States

13 [6] Division of Biomedical Statistics and Informatics, Mayo Clinic College of Medicine,

14 Rochester, MN 55905, United States

15 [7] These authors contributed equally.

16 [8] Lead Contact

17

18 *Correspondence: malina@big.ac.cn (M.L.N.), zhangzhang@big.ac.cn (Z.Z.)

19

1 **Summary**

2 To explore glioma biomarkers with high specificity, enable non-invasive detection and

3 possess clinical significance, we perform large-scale integrative analyses on multi-omics

4 multi-cohort datasets collected from public resources. We identify *PRKCG* as a brain-

5 specific gene that is highly expressed in cerebrospinal fluid, thus achieving higher

6 specificity and detectability in the periphery. Importantly, *PRKCG* bears great potential in

7 clinical application for glioma diagnosis, prognosis, and treatment prediction, which have

8 been consistently testified on a number of independent discovery and validation datasets

9 consisting of multi-omics molecular profiles and different ethnic populations/countries.

10 Based on the comprehensive characterization of multi-omics molecular profiles, our

11 findings suggest the reliability and potential application of *PRKCG* as a biomarker for

12 glioma.

13 **Keywords:** glioma, multi-omics, *PRKCG*, biomarker, cerebrospinal fluid

14

15 **Significance**

16 Glioma is one of the most lethal human malignancies and exhibits low resection rate and

17 high recurrence risk. Powered by high-throughput sequencing technologies, the public

18 availability of vast amounts of multi-omics molecular datasets provides unprecedented

19 opportunities to identify effective biomarkers for glioma. Here, we assemble a large-scale

20 collection of multi-omics multi-cohort glioma datasets from public resources. Our

21 integrative molecular analyses reveal that *PRKCG* achieves higher specificity and

22 detectability in the periphery and bears the significant potential in glioma diagnosis,

1   prognosis and treatment prediction as testified on different datasets. Our study features

2   comprehensive molecular characterization of *PRKCG* in glioma and highlights the value

3   of integrative multi-omics data analysis toward accurate therapeutic strategies and

4   precision healthcare in the era of big data.

5

6

7

8

9

10

11

12

13

14

15

16

17

1 **Introduction**

2 Glioma, one of the serious central nervous system (CNS) tumors, represents 80% of

3 malignant brain tumors (A.Maher. et al., 2019; Schwartzbaum et al., 2006) and features

4 low resection rate and high recurrence risk (Stewart, 2002). Since tumor classification

5 benefits accurate diagnosis and facilitates precise treatment, gliomas can be classified,

6 according to the 2007 World Health Organization (WHO) grading scheme (Louis et al.,

7 2007), into low-grade gliomas (LGG: astrocytoma, oligodendroglioma,

8 oligoastrocytoma) and high-grade gliomas (GBM: glioblastoma). Therefore,

9 identification of effective biomarkers for precise classification of different-grade gliomas

10 is crucial to aid tumor diagnosis, establish appropriate therapies, recognize prognostic

11 outcome and predict therapeutic response (Cancer Genome Atlas Research et al., 2015;

12 Wesseling et al., 2011).

13 Powered by high-throughput sequencing technologies, a set of molecular biomarkers

14 have been discovered from different omics levels to assist glioma diagnosis and treatment

15 (Kim et al., 2013; Wiestler et al., 2013). Among them, isocitrate dehydrogenase (*IDH*)

16 mutation and 1p/19q co-deletion (1p/19q codel) are two most important molecular

17 signatures for glioma grading (Cohen et al., 2013; Eckel-Passow et al., 2015; Waitkus et

18 al., 2018). Patients with *IDH* mutation (*IDH*-mut) have longer survival than those with

19 *IDH* wild-type (*IDH*-WT) (Cohen et al., 2013; Guo et al., 2011; Ichimura et al., 2009;

20 Kloosterhof et al., 2011; Turcan et al., 2012). And the 1p/19q codel is a distinctive feature

21 of oligodendroglioma (Cairncross et al., 2013; Eckel-Passow et al., 2015; Jenkins et al.,

22 2006; van den Bent et al., 2013). Furthermore, based on these two signatures,

23 accumulated evidence suggested that gliomas can be divided into three subtypes (*IDH*-

1   mut & 1p/19q codel, *IDH*-mut & 1p/19q non-codel, and *IDH*-WT & 1p/19q non-codel),

2   which are associated with diverse clinical outcomes (Mur et al., 2015). Accordingly, in

3   2016, the WHO combined histology and genetic signatures to divide gliomas into five

4   categories (Molinaro et al., 2019), including three low-grade gliomas (diffuse

5   astrocytoma, *IDH*-mut & 1p/19q non-codel; oligodendroglioma, *IDH*-mut & 1p/19q

6   codel; diffuse astrocytoma, *IDH*-WT & 1p/19q non-codel) and two high-grade gliomas

7   (GBM, *IDH*-mut; GBM, *IDH*-WT) (see a review in (Louis et al., 2016)). Meanwhile,

8   biomarkers at the transcriptome level have also been identified (Flynn et al., 2008; Wang

9   et al., 2016); for example, an increased expression of Epidermal Growth Factor Receptor

10   (*EGFR*) has been reported to associate with malignant progression of gliomas (Fan et al.,

11   2009; van den Bent et al., 2015; Verhaak et al., 2010). In addition, epigenetic

12   modifications are also implicated in glioma (Esteller, 2007; Park et al., 2014; Zhang et

13   al., 2017). One classical biomarker is O6-methylguanine-DNA-methyltransferase

14   (*MGMT*) (Binabaj et al., 2018; Donson et al., 2007); patients with methylated *MGMT*

15   promoter have better clinical outcomes and are more sensitive to the alkylating

16   chemotherapy than those without methylated *MGMT* promoter (Binabaj et al., 2018; Hegi

17   et al., 2005; Rivera et al., 2010; Wick et al., 2014).

18   Although tremendous efforts have been devoted for better understanding of glioma

19   tumorigenesis and identification of molecular biomarkers, existing glioma biomarkers

20   have several drawbacks. First, they are lack of glioma specificity. For instance, the *IDH*

21   mutation has also been identified as a biomarker in other tumor types including acute

22   myeloid leukemia (AML) (Dang et al., 2010; Dang et al., 2016), chondrosarcoma and

23   intrahepatic cholangiocarcinoma (Kerr et al., 2013). Second, existing biomarkers are

1   unable to differentiate glioma subtypes with high accuracy, which would be beneficial to

2   glioma grading and personalized medicine. Third, although biomarkers have been

3   discovered to aid the prognosis and prediction of GBM (Hu et al., 2018; McNamara et al.,

4   2013) and LGG (Cancer Genome Atlas Research et al., 2015; Wesseling et al., 2015),

5   they are incapable to reflect the progression from low- to high-grade gliomas. Last but

6   foremost, considering the practical significance of detectability in periphery, genetic

7   tumor profiling based on extant biomarkers yet involves brain surgery (Miller et al.,

8   2019), limiting its clinical applicability for glioma diagnosis and prognosis.

9   The public availability of multi-omics datasets for glioma (Cancer Genome Atlas

10   Research et al., 2015; Ceccarelli et al., 2016; Zhao et al., 2019) as well as cerebrospinal

11   fluid (CSF) which is the only accessible source to obtain genes stemmed from human

12   CNS (Miller et al., 2019; Mouliere et al., 2018; Sasayama et al., 2017), provides

13   unprecedented opportunities for identifying biomarkers without craniotomy operation and

14   comprehensively characterizing the pathophysiology of brain tumors. To this end, we

15   collected a large-scale assemble of multi-omics multi-cohort datasets from public

16   resources, established a methodological strategy on integrative identification of

17   biomarkers with higher specificity and feasible detectability from periphery, and revealed

18   that *PRKCG* is specifically expressed in brain and also detectable in CSF. Through

19   comprehensive integrative analyses on a total of five discovery datasets and fourteen

20   validation datasets, we systematically characterized *PRKCG* as a potential biomarker for

21   glioma diagnosis, prognosis and treatment prediction.

1   **STAR Methods**

2   *Data collection*

3   In this study, we collected a comprehensive assemble of multi-omics datasets (including

4   genomics, transcriptomics, DNA methylomics and proteomics) from The Cancer Genome

5   Atlas (TCGA, https://portal.gdc.cancer.gov/) (Ceccarelli et al., 2016), Genotype-Tissue

6   Expression Portal (GTEx, https://gtexportal.org/home/) (John Lonsdale, 2013), Gene

7   Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo), Ivy Glioblastoma Atlas

8   Project (Ivy GAP, http://glioblastoma.alleninstitute.org) (Puchalski. et al., 2018) and

9   Chinese Glioma Genome Atlas (CGGA, http://www.cgga.org.cn) (Bao et al., 2014; Sun et

10   al., 2014). Particularly, discovery datasets were derived from TCGA, GTEx and large

11   cohort studies in GEO (GSE83710, GSE16011 and GSE36278 for protein, expression

12   and methylation, respectively). As a result, a total of five discovery datasets and fourteen

13   validation datasets were obtained. For convenience, each dataset collected in this study is

14   assigned   a   unique   accession   number   with   the   format: [D/V][i]-

15   [TCGA/GTEx/GEO/CGGA/Ivy GAP]-[E/V/P/M], where D/V in the first bracket

16   represents the dataset for discovery or validation, i in the second bracket indicates the

17   dataset number, the third bracket shows the data source (as mentioned above), and the last

18   bracket indicates the data type, namely, E for RNA expression, V for CNV, P for protein

19   expression and M for DNA methylation, respectively. The detailed information about all

20   collected datasets was tabulated in Table 1.

21   *Identification of brain-specific genes*

7

1   To identify brain-specific genes, we used the RNA-Seq dataset from GTEx (2016-01-15;

2   v7) (John Lonsdale, 2013), which contains 11,688 samples across 53 tissue sites of 714

3   donors. Considering that several tissues have multiple different sites, gene expression

4   levels were averaged over sites that are from the same tissue. To reduce background

5   noise, genes with maximal expression levels smaller than 10 TPM (Transcripts Per

6   Million) were removed. Finally, we obtained a total of 15,176 gene expression profiles

7   across 30 tissues (Table S1).

8   Based on the expression levels across 30 tissues, we calculated the tissue specificity

9   index τ (Yanai et al., 2005) for each gene to identify tissue-specific genes. τ is valued

10   between 0 and 1, where 0 represents housekeeping genes that are consistently expressed

11   in different tissues, and 1 indicates tissue-specific genes that are exclusively expressed in

12   only one tissue (Yanai et al., 2005). In this study, brain-specific genes were defined as

13   those genes that are maximally expressed in the brain with τ>0.9. As a consequence, a list

14   of the top 100 brain-specific genes ranked by the τ index were obtained for further

15   analysis (Table S2).

16   *Sample classification*

17   To comprehensively study the potential of *PRKCG* in glioma diagnosis, we compared its

18   molecular profiles between normal and glioma samples, between LGG and GBM

19   samples, between primary GBM (pGBM) and recurrent GBM (rGBM) samples, and

20   between glioma samples with different anatomic features. We collected 122 GBM

21   samples from the Ivy GAP database (Puchalski. et al., 2018) and grouped them according

22   to their anatomic regions, namely, leading edge (LE, the ratio of tumor/normal cells is

1   about 1–3/100), infiltrating tumor (IT, the ratio of tumor/normal cells is about 10–

2   20/100), cellular tumor (CT, the ratio of tumor/normal cells is about 100/1 to 500/1),

3   pseudo-palisading cells around necrosis (PAN, the narrow boundary of cells along the

4   dead tissue), and microvascular proliferation (MVP, two or more blood vessels sharing a

5   common vessel wall of endothelial).

6   To investigate the prognostic role of *PRKCG*, for each dataset we divided samples into

7   three groups based on its expression/methylation level, where the top 25% are "High",

8   the 25%-75% are "Middle", and the remaining are "Low". When exploring the predictive

9   role of *PRKCG*, we obtained methylation status (methylated and unmethylated) directly

10   from the original study (Ceccarelli et al., 2016), which was defined based on the beta

11   value cutoff 0.3.

Identification of PRKCG-like genes

13   *PRKCG*-like genes were identified by adopting the following criteria (Figure S1): (1)

14   Higher methylation level of at least one CpG site (promoter region; 450K) in glioma

15   samples than normal samples; (2) Higher DNA methylation level in LGG samples than

16   GBM samples; (3) Higher expression level in LGG samples than GBM samples; and (4)

17   Lower expression level in glioma samples than normal samples. As a result, we obtained

18   a total of 720 genes, which were further divided into two groups according to their

19   correlations between gene expression and methylation, namely, 193 genes with positive

20   correlation and 269 genes with negative correlation. These 193 genes were considered as

21   *PRKCG*-like genes.

Statistical analysis

1   All statistical analyses were performed using R version 3.3.2. The Wilcoxon test was

2   used for the analysis of the difference in gene expression/methylation between tumor and

3   normal samples, and between different glioma subtypes. The statistical significance levels

4   were coded by *ns* (not significant) $p > 0.05$, * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$.

5   We performed the survival analysis using the Kaplan-Meier method and estimated the

6   statistical difference using the log-rank test. To identify whether age or gender might

7   exert an influence on survival, the Cox proportional hazards regression model was

8   applied to the univariate and multivariate analyses.

9   *Lead contact and materials availability*

10   This study did not generate new unique reagents. Further information and requests for

11   resources should be directed to and will be fulfilled by the Lead Contact, Z.Z.

12   (zhangzhang@big.ac.cn).

13   *Data availability statement*

14   All datasets integrated in this study were obtained from multiple public database

15   resources   (see   details   in   Table   1),   which   are   freely   available   at

16   ftp://download.big.ac.cn/glioma_data/.

17   **Results**

18   *PRKCG is a brain-specific gene and detectable in cerebrospinal fluid*

19   To identify glioma biomarkers with high brain specificity, we integrated expression data

20   from GTEx (D1-GTEx-E) (John Lonsdale, 2013), explored genes' expression profiles

21   and their tissue-specificity, and identified a list of top 100 brain-specific genes (Table S2).

1   To achieve the detectability in the periphery, we assembled a total of 1,126 CSF-

2   detectable proteins from GEO GSE83710 (D2-GSE83710-P) (Sasayama et al., 2017), due

3   to the critical significance of CSF as the only feasible way to detect genes expressed in

4   human CNS. After integrating brain-specific genes with CSF proteins, we found that

5   there are five brain-specific proteins that can be detected in CSF (Figures 1 and S2), in

6   terms of fluorescence intensity from low to high, namely, *CAMK2A, GFAP*, *OPCML*,

7   *BCAN* and *PRKCG*, which are diversely expressed in different brain regions (Figure S3).

8   Specifically, *CAMK2A* is a calcium calmodulin-dependent protein kinase and reduced

9   expression of *CAMK2A* is associated with better survival in GBM (John et al., 2017;

10   Long et al., 2017). *GFAP*, encoding one of the major intermediate filament proteins of

11   mature astrocytes (Horst et al., 2007), can be used to assess the differentiation state of

12   astrocytoma (van Bodegraven et al., 2019). *OPCML* encodes a member of the IgLON

13   subfamily in the immunoglobulin proteins and is down-regulated in gliomas and other

14   brain tumors (Carminati. et al., 2010; Reed et al., 2007). *BCAN*, a member of the lectican

15   family of chondroitin sulfate proteoglycans, is highly expressed in glioma and may

16   promote cell motility of brain tumor cells (Phillips et al., 2006; Sydney C. Gary, 2000). In

17   addition, the fusion event between *BCAN* and *NTRK1* (*BCAN-NTRK1)* is a potential

18   glioma driver and therapeutic target (Cook et al., 2018).

19   Remarkably, *PRKCG* (Protein Kinase C Gamma), a member of protein kinase C (PKC)

20   family, exhibits higher fluorescence intensity than the other four genes (Figures 1 and

21   S2). Previous studies have documented that unlike other PKC family members that are

22   expressed in many tissues aside from brain, *PRKCG* is brain-specifically expressed and

23   its localization is restricted to neurons (Saito and Shirai, 2002) and that mutations in

1   *PRKCG* are associated with spinocerebellar ataxia (Klebe et al., 2005; Yabe et al., 2003).

2   Additionally, it has been reported that PKC signaling pathways contribute to the

3   aggressive behavior of glioma cells (do Carmo et al., 2013) and atypical PKC isozymes

4   are fundamental regulators of tumorigenesis (Reina-Campos et al., 2019). Despite this,

5   the potential role of *PRKCG* in glioma remains unknown, and therefore, comprehensive

6   molecular characterization of *PRKCG* across multi-omics glioma datasets is highly

7   desirable.

8   ***PRKCG is a potential biomarker for glioma diagnosis and prognosis***

9   First, we explored the expression profile of *PRKCG* across multiple brain developmental

10   stages, and revealed that its expression is extremely lower in the prenatal stages, but

11   dramatically increased in the infancy stages and then stabilized in the latter stages using

12   the data of GenTree (Shao et al., 2019) (Figure S4). Moreover, using a total of 607

13   TCGA glioma samples, we identified its co-expressed genes (Table S3; *p*-value < 0.01,

14   r^2>0.8) and performed the gene ontology analysis, showing that *PRKCG* co-expressed

15   genes are significantly associated with "chemical synaptic transmission", "axon", and

16   "GABA-A receptor activity".

17   Next, we compared the expression profiles of *PRKCG* between normal and glioma

18   samples. We found that *PRKCG* expression is significantly reduced in gliomas by

19   contrast to normal samples, and this reduced expression is consistently observed in

20   multiple datasets (Figure 2A to 2F; *p*-value < 0.01, Wilcoxon test). Then, using data

21   collected from the Ivy GAP database (Puchalski. et al., 2018) (see Methods), we explored

22   the expression of *PRKCG* among GBMs from different anatomic regions. As expected,

1   *PRKCG* shows significantly different expression patterns among these anatomic regions;

2   its expression is highest in LE (the outermost boundary of the tumor), significantly

3   decreased in IT (the intermediate zone between the LE and the serious CT regions), and

4   lowest in the serious regions (CT, PAN and MVP) (Figure 2G; *p*-value < 0.01, Wilcoxon

5   test). We further examined *PRKCG* expression in gliomas with different grades, and

6   found that its expression level is significantly lower in GBM samples than LGG samples

7   (Figure 2H to 2J; *p*-value < 0.01, Wilcoxon test). Collectively, these results demonstrate

8   that the reduced expression of *PRKCG* is not only associated with glioma, but also

9   associated with glioma progression in a quantitative manner, highlighting its potential for

10 glioma diagnosis and prognosis.

11 Since *PRKCG* harbors two CpG sites (namely, cg26626089 and cg04518808) that are

12 located in the promoter region and covered in both HumanMethylation27 (27K) and

13 HumanMethylation450 (450K) BeadChip datasets, we systematically investigated DNA

14 methylation profiles of these two sites (Figure 3). Using the four independent datasets

15 from D5-GSE36278-M (Sturm et al., 2012), V8-GSE50923-M (Lai et al., 2014), V9-

16 GSE61160-M (Mur et al., 2013) and V10-CGGA-M (Zhang et al., 2013), we found that

17 cg26626089 is significantly hyper-methylated in GBM patients compared with normal

18 samples (Figure 3A and 3C; *p*-value < 0.01, Wilcoxon test). The other CpG site

19 cg04518808, albeit not significantly, is also observed to be hyper-methylated in GBM

20 patients (Figure 3B and 3D; *p*-value > 0.05, Wilcoxon test). Furthermore, we examined

21 the variation of methylation level using whole-genome bisulfite sequencing data of six

22 GBM samples from TCGA and one normal sample from UCSC (2017 version;

23 http://genome.ucsc.edu, last accessed on 12 May 2019). Consistently, we observed that

1   most GBM patients show higher methylated promoter region than normal samples

2   (Figure S5). In addition, considering different grading gliomas, we observed that both

3   sites present much lower methylation levels in GBM samples than LGG samples (Figure

4   3E to 3H; p-value < 0.01, Wilcoxon test).

5   Collectively, *PRKCG* exhibits differential molecular profiles in normal and glioma

6   samples. Compared with normal samples, *PRKCG* presents lower expression and higher

7   methylation in glioma samples. With tumor malignancy, *PRKCG* expression and

8   methylation are both on the decrease (discussed later). These results suggest *PRKCG* as a

9   potential biomarker for glioma diagnosis and prognosis.

10   ***PRKCG is significantly associated with survival***

11   To further investigate the prognostic potential of *PRKCG*, we conducted the survival

12   analysis with samples' expression data and survival information. We discovered that

13   higher *PRKCG* expression is associated with longer overall survival (Figure 4A; p-value

14   < 0.01, log-rank test). Remarkably, this is consistently observed in three independent

15   validation datasets (Figure 4B to 4D; p-value < 0.01, log-rank test). The univariate and

16   multivariate Cox regression analyses also revealed that *PRKCG* expression is statistically

17   significantly associated with patients' survival (Table 2; p-value < 0.01).

18   DNA methylation is implicated in transcriptional regulation and may play a central role in

19   the generation of phenotypic instability (Baylin, 2005). Accordingly, we further

20   investigated whether *PRKCG* methylation is associated with clinical outcome. Based on

21   two independent datasets from TCGA ($n = 862$) and CGGA ($n = 151$), we observed that

22   higher methylation level of the site cg04518808 is significantly associated with better

14

1   survival (Figure S6A and S6B; p-value < 0.01, log-rank test). The other site cg26626089

2   exhibits similar effect, but is statistically insignificant in the CGGA dataset presumably

3   due to smaller sample size (Figure S6C, p-value < 0.01, log-rank test; Figure S6D, p-

4   value = 0.3, log-rank test). However, we found that the averaged methylation level of the

5   two sites is a more powerful and robust prognositc biomarker; higher *PRKCG*

6   methylation indicates better survival outcome, which is less affected by the sample size

7   (Figure 5A and 5B; p-value < 0.01, log-rank test). Moreover, the univariate and

8   multivariate Cox analyses excluded the age and gender as confounding variables

9   affecting patients' survival as well (Table 2; p value < 0.01).

10   Since *PRKCG* is located on 19q13.42, *PRKCG* CNV is most likely correlated with the

11   status of 19q and thus may be associated with clinical outcome. Therefore, we

12   investigated these relationships using 1,018 samples from TCGA (Figure 6). Consistent

13   with our expectation, the copy number status of *PRKCG* is accompanied by 19q gain/loss

14   (Figure 6A). As 1p/19q codel is an acknowledged prognostic biomarker for glioma, we

15   adopted both 1p/19q codel and *PRKCG* to further divide gliomas into four groups:

16   *PRKCG* normal, *PRKCG* gain, *PRKCG* loss & 1p/19q codel, and *PRKCG* loss & 1p/19q

17   non-codel. Interestingly, patients under these four groups present significantly different 5-

18   year overall survival (OS) rates, which, from better to worse, are 82% for 1p/19q codel,

19   42% for 1p/19q non-codel, 20% for *PRKCG* (19q) normal, and <8% for *PRKCG* (19q)

20   gain, respectively (Figure 6A; p value < 0.01, log-rank test). These results indicate that

21   one single biomarker might not be sufficient for accurate and reliable prognosis. We

22   illustrated the survival rates and the distributions of age, WHO grade and histology across

23   the four groups (Figure 6B). Intriguingly, *PRKCG* CNV status is closely associated with

1 different grading gliomas; nearly all LGG patients are *PRKCG* CNV loss and have

2 relatively young ages, whereas most GBM patients are older and possess *PRKCG* CNV

3 gain. Collectively, multi-omics analyses on expression, methylation and CNV datasets

4 demonstrated that *PRKCG* is a potential prognosis biomarker for glioma.

5 ***Combined methylation signatures of PRKCG and MGMT are effective in treatment***

6 ***prediction***

7 It is known that *MGMT* encodes a DNA-repair protein and hypermethylation of *MGMT* is

8 associated with diminished DNA-repair activity, accordingly allowing the alkylating drug

9 temozolomide (TMZ) to have greater effect in GBM treatment (Donson et al., 2007;

10 Fukushima et al., 2009; Hegi et al., 2005; Stupp et al., 2005). In our study, consistently,

11 we found that patients with methylated *MGMT* are more sensitive to TMZ treatment than

12 those with unmethylated *MGMT* (Figure 7A; p-value < 0.01, log-rank test).

13 Considering that a single biomarker might be lack of sufficient prediction power and thus

14 fail to determine the clinical therapeutic efficacy due to tumor heterogeneity (Li et al.,

15 2018), we sought to examine the predictive potential of *PRKCG* for TMZ using 228

16 glioma samples with matched DNA methylation and clinical data from TCGA. We

17 discovered that among the two CpG sites of *PRKCG* (cg26626089 and cg04518808), the

18 methylation level of cg26626089 is able to classify patients into two groups with distinct

19 survival advantages, as patients with methylated cg26626089 have significantly longer

20 survival than those with unmethylated cg26626089 (Figure 7B; p-value < 0.01, log-rank

21 test; Table 2, p-value < 0.01). In contrast, cg04518808 is unable to be used for glioma

22 classification (Figure S7A, p-value > 0.05, log-rank test). By combining *PRKCG*

1   (cg26626089) with *MGMT*, intriguingly, GBM patients receiving TMZ treatment can be

2   classified into four groups that exhibit significantly different survivals (Figures 7C and

3   S7B; *p*-value < 0.01, log-rank test). The four groups, namely, *MGMT*-unmethylated +

4   *PRKCG*-unmethylated, *MGMT*-unmethylated + *PRKCG*-methylated, *MGMT*-methylated

5   + *PRKCG*-unmethylated, and *MGMT*-methylated + *PRKCG*-methylated, present

6   gradually improved longer survivals, as their 20-month OS rates are 0.18, 0.29, 0.39 and

7   0.51 (Figure 7C), respectively. Therefore, the combined methylation signatures of

8   *PRKCG* and *MGMT* can be utilized to guide more accurate glioma stratification and

9   achieve better personalized therapeutic decisions.

10   **Discussion**

11   With the rapid advancement of sequencing technologies, there has been an increasing

12   number of high-throughput studies on glioma, resulting in massive multi-omics multi-

13   cohort datasets generated from different projects and laboratories throughout the world.

14   Therefore, it has become crucially significant on how to make full use of these valuable

15   data for comprehensive molecular identification of glioma biomarkers. In this study, we

16   for the first time, assembled the most comprehensive collection of public glioma datasets

17   with multi-omics data types and different populations/countries and established a

18   methodological strategy on integrative identification of biomarkers with higher

19   specificity and feasible detectability from periphery. Based on these, we performed

20   comprehensive molecular characterization of *PRKCG* as a biomarker for glioma

21   diagnosis, prognosis and treatment prediction, which have been consistently verified

22   across multiple independent discovery and validation datasets (Figures 2 to 5).

23   Specifically, 1) *PRKCG* presents lower expression and higher methylation in glioma

17

1 samples, which, with tumor malignancy, are both decreasing; 2) *PRKCG* is associated

2 with glioma progression, as its expression change from high to low is indicative of

3 glioma progression from low-grade to high-grade; 3) High RNA expression, high DNA

4 methylation, and low copy number, are all suggestive of good survivals (discussed

5 below); 4) Patients with DNA methylation in both *PRKCG* and *MGMT* are more sensitive

6 to TMZ treatment, suggesting the effectiveness of combined application of *PRKCG* with

7 the classical biomarker *MGMT* to achieve more precise survival after TMZ

8 chemotherapy. Although it has been documented that *PRKCG* is up-regulated in colon

9 cancer and loss of *PRKCG* inhibits cell migration and enhances the proliferation

10 (Catriona M. Dowling and Kiely, 2017), the up-regulation in colon cancer, as a matter of

11 fact, is extremely lower by comparison with glioma (LGG and GBM) (Figure S8). Thus,

12 unlike previous efforts that most discovered biomarkers at single omics level and with

13 limited samples, our findings highlight the importance of integrative multi-omics multi-

14 cohort data analysis and represents a data-driven blueprint toward accurate therapeutic

15 decisions and precision healthcare in cancer research.

16 Taking advantage of the comprehensive datasets, we explored how *PRKCG* is regulated

17 in glioma. Obviously, *PRKCG* is a glioma tumor suppressor gene significantly repressed

18 from normal brain tissue to glioma (Figure 2). At the same time, we found that most of

19 the GBM patients show higher methylated *PRKCG* than normal samples (Figure 3).

20 Noting that high methylation on the promoter region may lead to low expression (Herman

21 and Baylin, 2003), it is most likely that *PRKCG* expression is significantly negatively

22 affected by its methylation. However, it does not follow this rule in the comparison

23 between LGG and GBM; LGG exhibits both higher methylation level and higher

18

1   expression level than GBM (Figures 2 and 3). In fact, multi-omics profiles of *PRKCG* are

2   in harmony with classical biomarkers (Figure 8A); *PRKCG* methylation (cg26626089) is

3   associated with *IDH* status, consistent with a previous finding that *IDH*-mut is associated

4   with high methylation (Christensen et al., 2011). As LGG samples are always associated

5   with *IDH*-mut and GBM samples are associated with *IDH*-WT, it is not difficult to

6   understand why the methylation level of *PRKCG* is significantly lower in GBM than in

7   LGG. Such higher expression level and higher methylation level lead to the suspicion

8   whether *PRKCG* expression in glioma is positively regulated by its DNA methylation or

9   is attributable to its CNV.

10   Surprisingly, *PRKCG* CNV is positively correlated with its expression, as expected

11   within the CNV loss/gain group (Figure 8B and 8C; *p*-value < 0.01, Spearman correlation

12   = 0.26/0.32), while it exhibits a contradictory negative association in the group of all

13   CNV status (Figure 8D). According to the dosage effect theory (Henrichsen et al., 2009),

14   the CNV loss group should not express more *PRKCG* than the CNV gain group. This

15   implies that there is probably another factor rather than CNV to dominantly regulate

16   *PRKCG* expression in glioma. Although it contradicts the commonly accepted negative

17   association between gene expression and promoter CpG methylation, a large-scale pan-

18   cancer analysis has revealed positive correlation between promoter CpG methylation and

19   gene expression (John CG Spainhour, 2019). Consistently, we did observe significant

20   positive correlations between *PRKCG* expression and CpG methylation within the

21   promoter region (Figure 8E and 8F). This positive regulation of CpG methylation could

22   be quite strong, which significantly improves *PRKCG* expression in LGG samples; even

23   these samples exhibit obvious CNV loss (Figure 8A).

19

1 Therefore, *PRKCG* is most likely regulated in different directions by DNA methylation,

2 which negatively regulates *PRKCG* expression from normal to tumor, while positively

3 regulates the expression within tumor. This finding does challenge our understanding of

4 methylation regulation. To analyze whether this kind of regulation has any biological

5 significance, we characterized and identified 462 glioma suppressor genes that tend to be

6 regulated by DNA methylation (Table S4) (described in Methods). Among the 462 genes,

7 269 genes show negative correlations between methylation level and expression level,

8 whereas 193 genes show positive correlations (Figure S9A; Table S4) that are exactly

9 *PRKCG*-like. Intriguingly, the two groups' genes are enriched in different pathways. The

10 negatively correlated genes are significantly enriched in the pathways of "Neuroactive

11 ligand-receptor interaction", "cAMP signaling pathway", "cGMP-PKG signaling" and

12 "Morphine addiction", whereas the *PRKCG*-like genes are significantly enriched in the

13 pathways of "Ras signaling", "Dopaminergic synapse" and "Glutamatergic synapse"

14 (Figure S9B). Thus, *PRKCG*-like genes presumably present heterogeneous roles in

15 tumorigenesis with complex molecular mechanisms that need further extensive

16 explorations both bioinformatically and experimentally.

17 *PRKCG* is located on the chromosome 19q13.42, unifying previous findings that 1p/19q

18 codel is closely associated with glioma. Consistently, *PRKCG* CNV is associated with

19 19q status (Figure 6A). To our knowledge, in addition to *PRKCG*, other genes (e.g.,

20 *TTYH1*, *UBE2S*) (Hu et al., 2017; Jung et al., 2017) that are located in 19q, are also

21 linked with glioma. However, it is still unclear why 19q is associated with patients' status

22 and what is the driving force responsible for the variation of these associated genes.

23 Interestingly, different from other genes as mentioned above, *PRKCG* (as well as other

1   four genes identified in this study, viz., *GFAP, BCAN, CAMK2A* and *OPCML*) is a highly

2   brain-specifically expressed gene and its protein is abundant in CSF, accordingly

3   achieving higher specificity and enabling practical detectability from periphery without

4   surgery operation. Nevertheless, it is noted that the CSF data used in this study involves

5   1,126 proteins across 133 normal samples (Sasayama et al., 2017). In fact, it would be

6   more desirable and helpful to compare the *PRKCG* protein pattern in CSF across normal

7   and glioma samples (ideally with different grades). Therefore, future inclusion of more

8   glioma CSF datasets is needed to systematically characterize the diagnostic variables and

9   their association to more precision clinical subtypes.

10   Particularly, although it is well-known that *MGMT* hypermethylation status is associated

11   with longer survival (Donson et al., 2007), here for the first time, we revealed that within

12   different status of *MGMT* (methylated or unmethylated), patients with methylated

13   *PRKCG* (cg26626089) always show better treatment outcome. *MGMT* encodes a DNA-

14   repair protein, and elevated *MGMT* expression is associated with TMZ resistance

15   (Donson et al., 2007). Not surprisingly, the methylated groups with better survivals show

16   significantly lower expression of *MGMT* than the unmethylated groups (Figure S7C).

17   Why would the combination of *MGMT* methylation with *PRKCG* methylation indicate an

18   even better survival? According to our results, *PRKCG* is potentially a glioma suppressor

19   gene. Therefore, the elevated expression of *PRKCG* positively regulated by DNA

20   methylation would promote the inhibition of tumor cells and thus exhibit a better effect

21   when combined with *MGMT*. In the future, it is necessary to collect a large number of

22   samples to elucidate the mechanisms behind. Also, it is crucial to identify more predictive

21

1   biomarkers for accurate stratification, precise clinical treatment, and improved healthcare

2   in oncology (La Thangue and Kerr, 2011).

3   The occurrence and development of glioma is a complicated heterogeneous process. In

4   our study, *PRKCG* performs well in differentiating normal, LGG, and GBM samples and

5   predicting patients' survival states. However, in pGBM and rGBM, we obtained

6   contradictory results when applied to different populations; *PRKCG* methylation shows

7   no significant difference between pGBM and rGBM in the Chinese population (V10-

8   CGGA-M) (Figure S10A and S10B; *p*-value > 0.05, Wilcoxon test) but significantly

9   difference in the Switzerland population (V14-GSE60274-M) (Kurscheid et al., 2015)

10   (Figure S10C and S10D; *p*-value < 0.05, Wilcoxon test). This is most likely caused by the

11   population genetic difference and/or the small sample size (both datasets have < 5 rGBM

12   samples). Undoubtedly, comprehensive integrated analysis across multi-omics datasets

13   and different populations is an inevitable trend in the era of big data, which would greatly

14   benefit systematic exploration of optimal biomarkers as well as characterization of their

15   molecular profiles in aid of accurate therapeutic decisions and precision healthcare.

5   **Author Contributions**

6   Conceptualization, L.L. and W.G.Y.; Methodology, L.L., W.G.Y., and Z.Z.; Validation,

7   L.L. and W.L.G.; Formal Analysis, L.L., W.G.Y., W.L.G., Y.C.L. and L.M.W.;

8   Investigation, L.L. and W.G.Y.; Writing – Original Draft, L.L.; Writing – Review &

9   Editing, M.L.N., W.L.G., S.S.H., H.L.L., and Z.Z.; Funding Acquisition, M.L.N. and
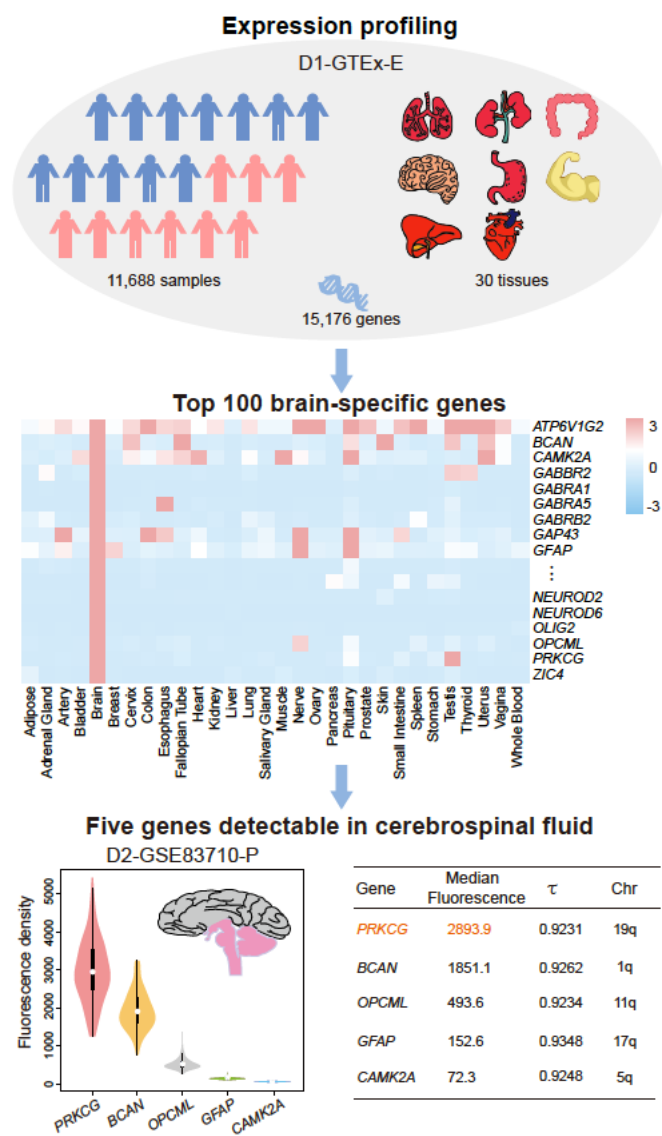
10   Z.Z.; Supervision, M.L.N. and Z.Z.

11   **Declaration of Interests**

12   The authors declare no competing interests.

13

14   **Figures**

15

1

Figure 1 Identification of brain-specific and CSF-detectable genes. Three steps were involved, namely, detection of brain-specific genes, identification of CSF-detectable genes, and ranking of candidate genes in light of protein fluorescence. See also Figures S1 and S2, and Tables S1 and S2.

1
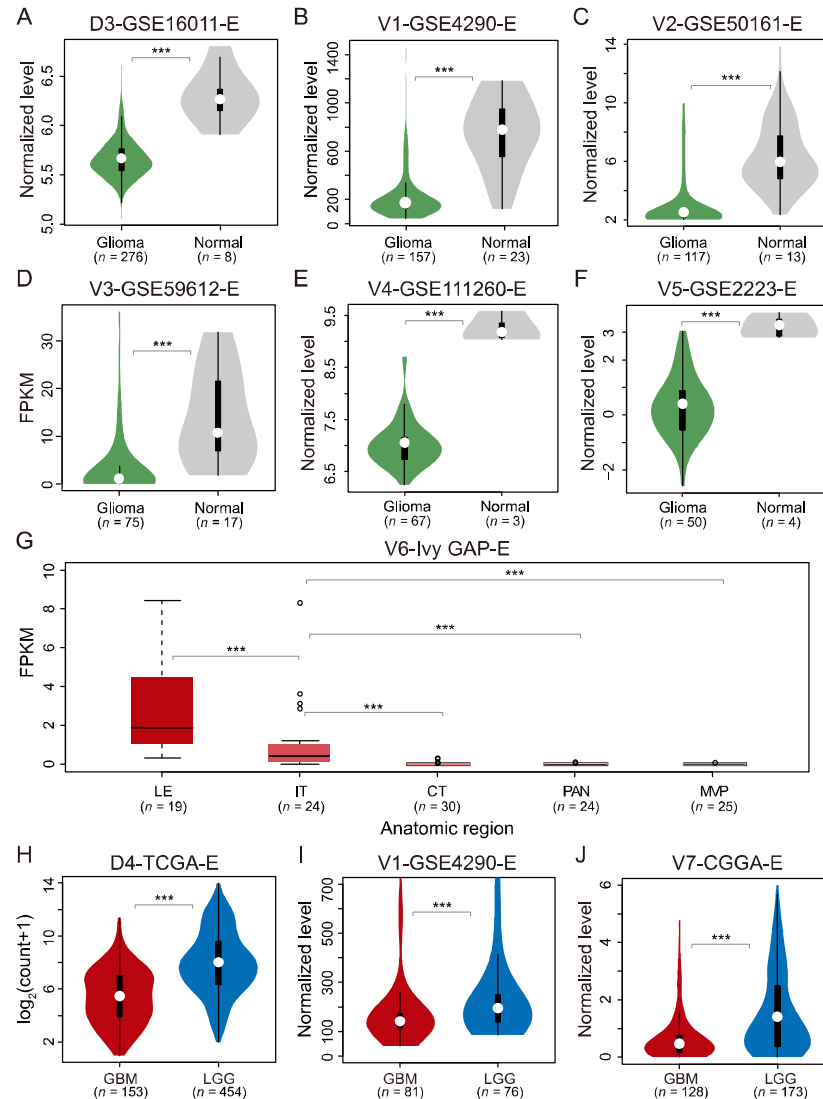
Figure 2 Expression profiles of *PRKCG* for glioma diagnosis and prognosis. *PRKCG*

expression profiles were compared between glioma and normal samples (D3-GSE16011-

E in panel A [RMA normalized], V1-GSE4290-E in panel B [MAS5 normalized], V2-

GSE50161-E in panel C [gcRMA normalized], V3-GSE59612-E in panel D, V4-

GSE111260-E in panel E [RMA normalized], V5-GSE2223-E in panel F [Lowess

normalized]), between different anatomic regions (V6-Ivy GAP-E in panel G), and

between GBM and LGG samples (D4-TCGA-E in panel H, V1-GSE4290-E in panel I

[MAS5 normalized] and V7-CGGA-E in panel J [Lowess normalized]). All the

1   normalization methods labeled above were derived from and detailed in their

2   corresponding publications, and all these datasets were made publicly accessible at

3   ftp://download.big.ac.cn/glioma_data/. The Wilcoxon tests were performed and the

4   statistical significance levels were coded by: *ns* $p > 0.05$, * $p < 0.05$, ** $p < 0.01$ and ***

5    $p < 0.001$.

1
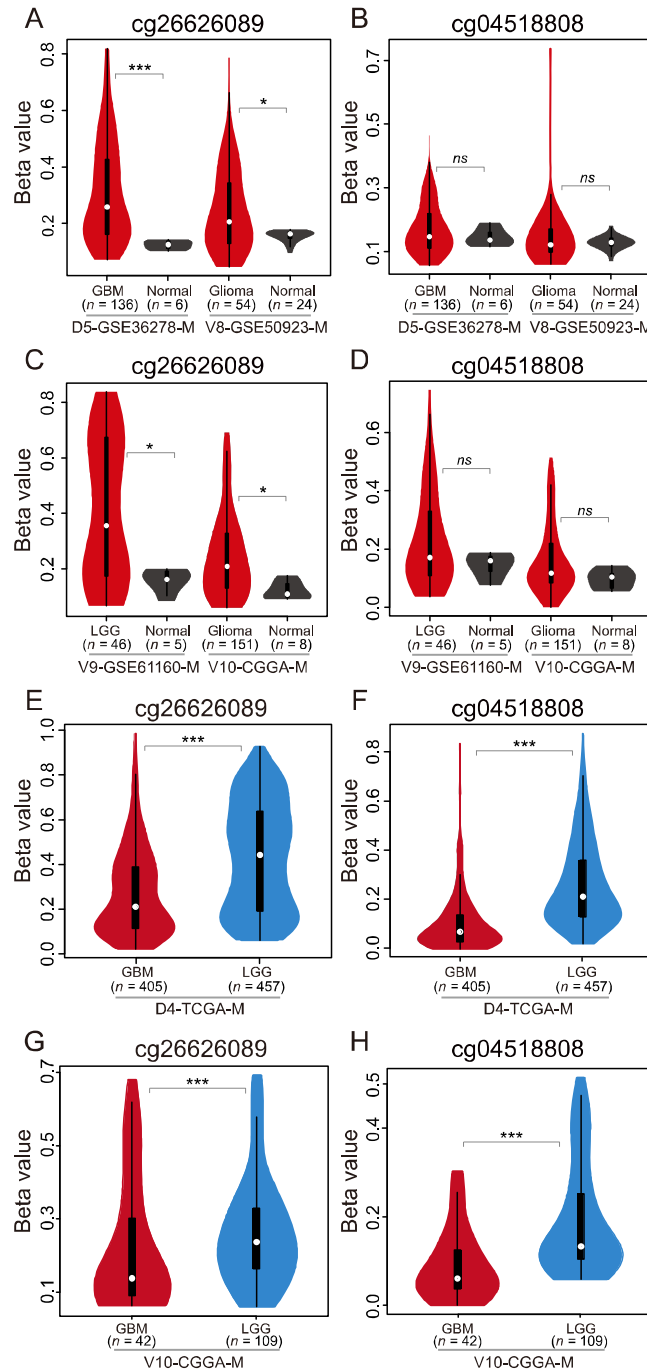
Figure 3 DNA methylation profiles of *PRKCG* for glioma diagnosis and prognosis.

PRKCG methylation profiles were compared between GBM and normal samples (panels A to D), and between LGG and GBM samples (panels E to H). All these datasets can be publicly accessible at ftp://download.big.ac.cn/glioma_data/. The Wilcoxon tests were

1   used and their statistical significance levels were coded by: *ns p>0.05*, * *p<0.05*, **

2   *p<0.01* and *** *p<0.001*.
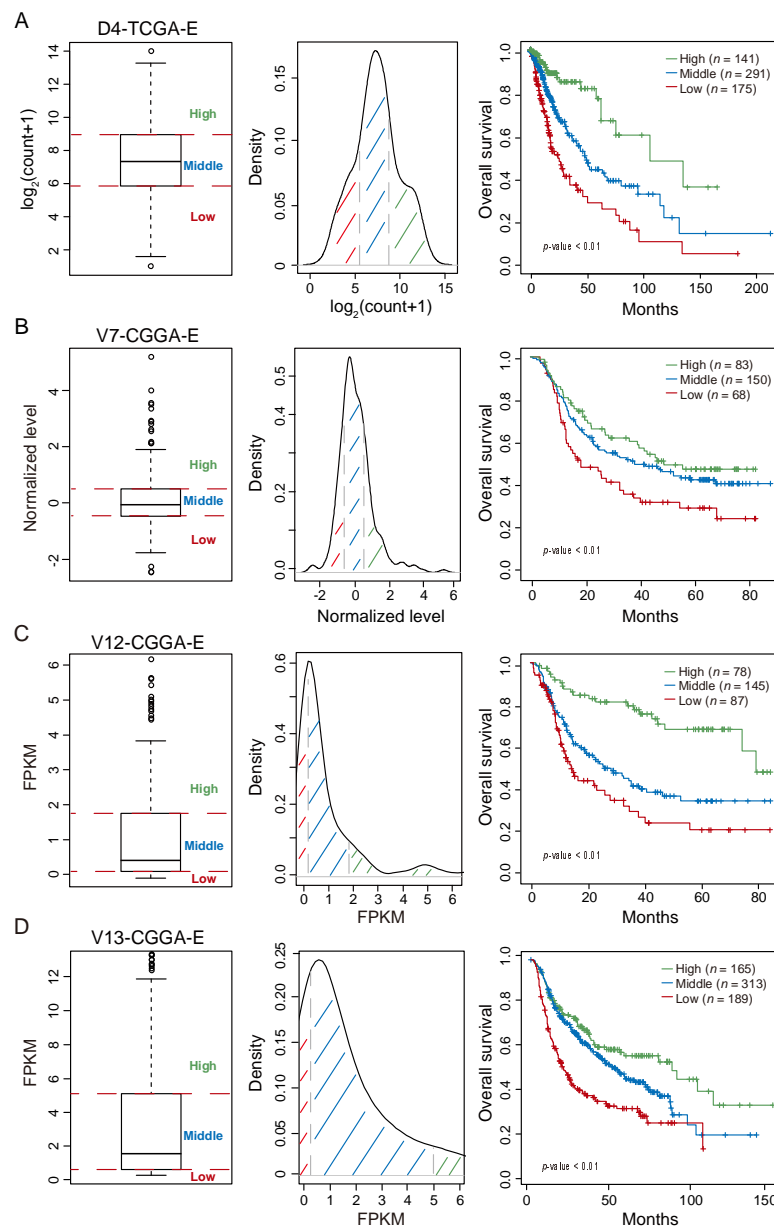
3

4   **Figure 4 *PRKCG* expression associated with survival.** Glioma patients were divided

5   into three groups based on the first and third quartile of *PRKCG* expression level in one

6   discovery dataset (D4-TCGA-E in panel A) and three validation datasets (V7-CGGA-E in

7   panel B, V12-CGGA-E in panel C, and V13-CGGA-E in panel D). All these datasets can

1   be publicly accessible at ftp://download.big.ac.cn/glioma_data/. The log-rank tests were

2   used to examine the statistical significance between different survival curves.
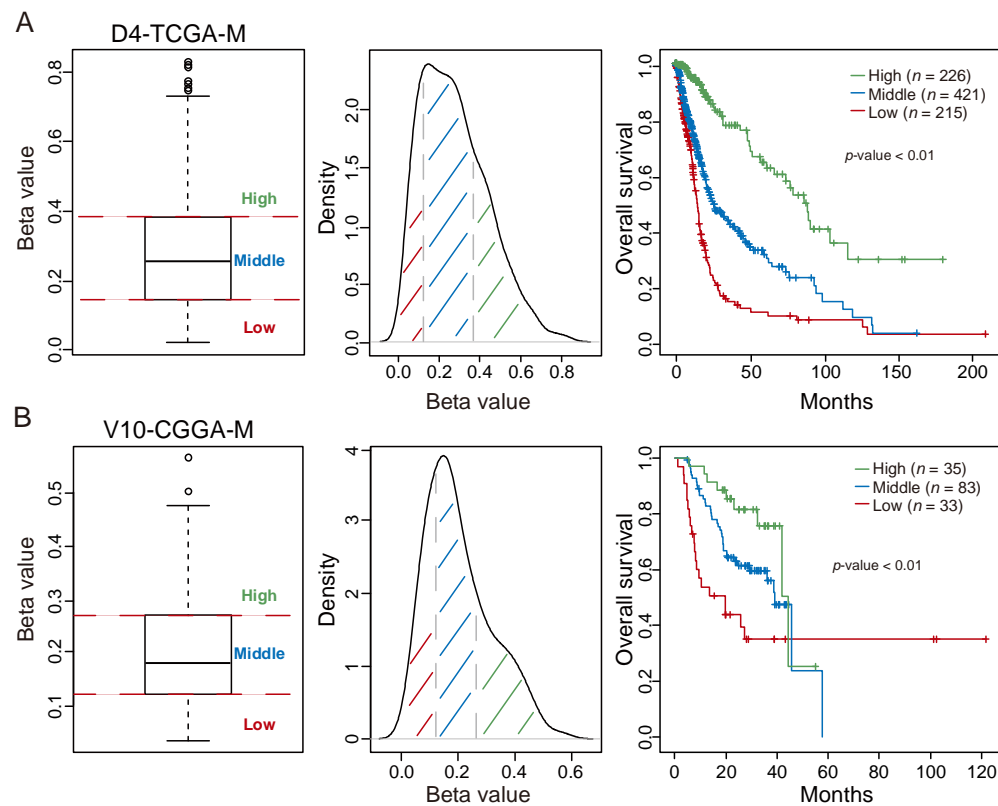
3

4   **Figure 5 *PRKCG* DNA methylation associated with survival.** Glioma patients were

5   divided into three groups based on the first and third quartile of *PRKCG* methylation

6   level in one discovery dataset (D4-TCGA-M in panel A) and one validation dataset (V10-

7   CGGA-M in panel B). All these datasets can be publicly accessible at

8   ftp://download.big.ac.cn/glioma_data/. The log-rank tests were used to examine the

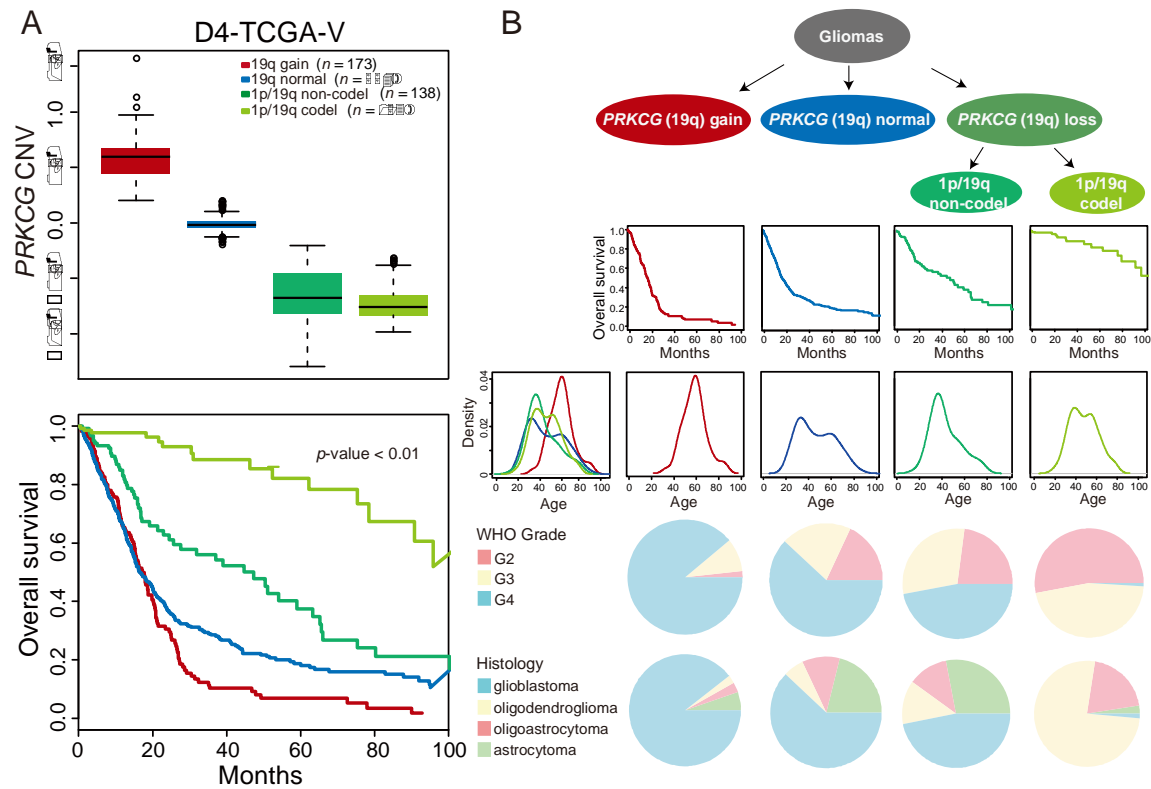9   statistical significance between different survival curves. See also Figure S6.

10

11

Figure 6 *PRKCG* CNV associated with survival. (A) Four groups of glioma patients were divided based on the 1p/19q status (19q gain, 19q normal, 1p/19q non-codel, and 1p/19q codel). **(B)** Kaplan-Meier survival probability, age, WHO grade and histology of the four groups.
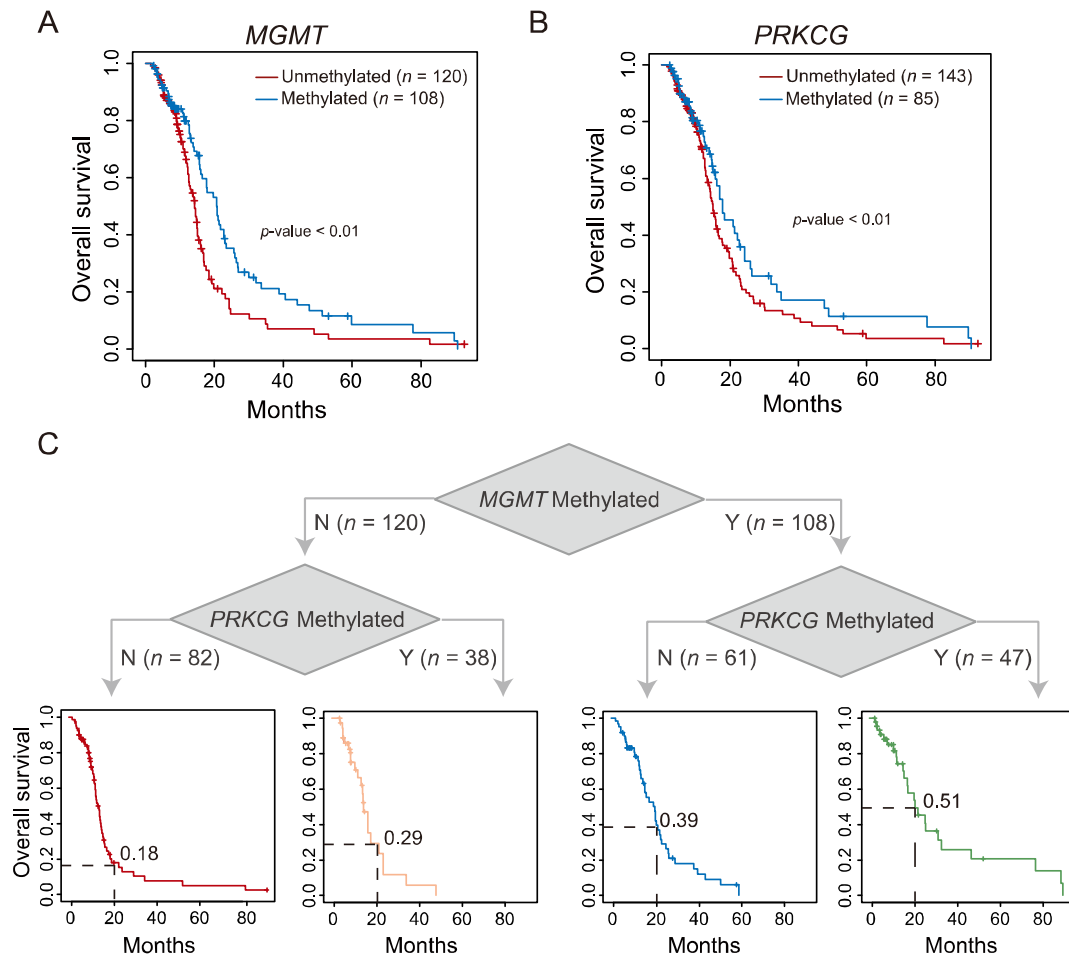
1

2 **Figure 7 Combined DNA methylation signatures of *MGMT* and *PRKCG* for**

3 **treatment prediction. (A)** Kaplan-Meier survival curves for GBM patients with TMZ

4 treatment based on *MGMT* methylation. **(B)** Kaplan-Meier survival curves for GBM

5 patients with TMZ treatment based on *PRKCG* (cg26626089) methylation. **(C)** Kaplan-

6 Meier survival curves for GBM patients with TMZ treatment based on *MGMT* and

7 *PRKCG* combined methylation signatures. See also Figure S7.
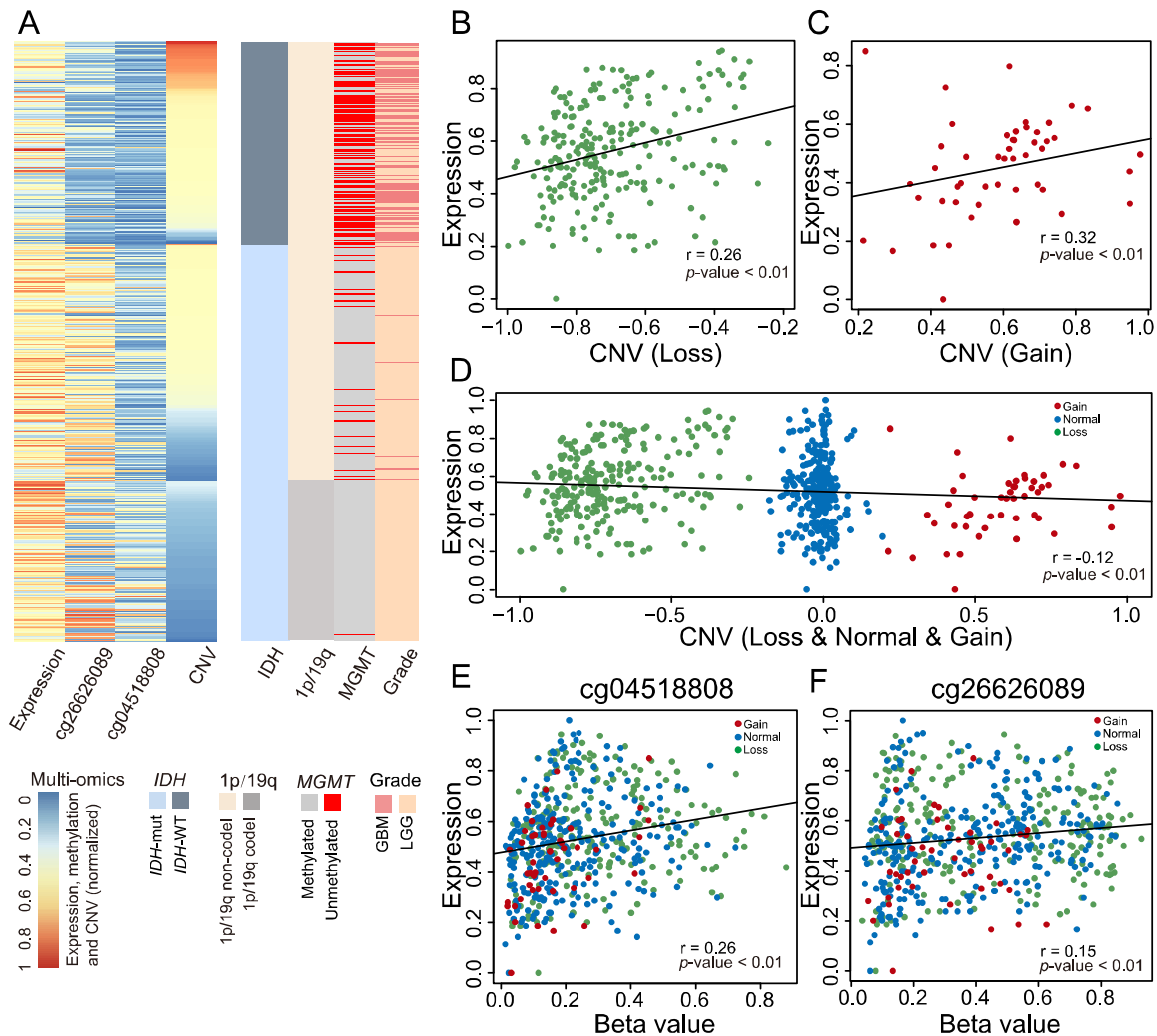
8

9

10

Figure 8 Multi-omics molecular profiles of *PRKCG*. (A) Association of *PRKCG*'s multi-omics signatures with *IDH*, 1p/19q status and WHO classification. **(B)** Correlation between *PRKCG* expression and CNV Loss. (**C**) Correlation between *PRKCG* expression and CNV Gain. (**D**) Correlation between *PRKCG* expression and all CNV status. **(E)** Correlation between *PRKCG* expression and DNA methylation of the CpG site cg04518808. **(F)** Correlation between *PRKCG* expression and DNA methylation of the CpG site cg26626089.

1   **Tables**

Table 1. Summary of multi-omics multi-cohort glioma datasets

Category	Accession number	Source	Omics type	# Sample	# Population country/race	Reference
Discovery	D1-GTEx-E	GTEx	Expression (RNA-Seq)	11,688	mostly white	(John Lonsdale, 2013)
	D2-GSE83710-P	GSE83710	Protein	133	Japan	(Sasayama et al., 2017)
	D3-GSE16011-E	GSE16011	Expression (Microarray)	284	Netherlands	(Gravendeel et al., 2009)
	D4-TCGA-V D4-TCGA-E D4-TCGA-M D4-TCGA-M (TMZ treatment)	TCGA	CNV Expression (RNA-Seq) Methylation (27K+450K) Methylation (27K+450K)	1,018 607 862 228	mostly white	(Ceccarelli et al., 2016)
	D5-GSE36278-M	GSE36278	Methylation (450K)	142	Germany	(Sturm et al., 2012)
Validation	V1-GSE4290-E	GSE4290	Expression (Microarray)	180	USA	(Sun et al., 2006)
	V2-GSE50161-E	GSE50161	Expression (Microarray)	130	USA	(Griesinger et al., 2013)
	V3-GSE59612-E	GSE59612	Expression (RNA-Seq)	92	USA	(Gill et al., 2014)
	V4-GSE111260-E	GSE111260	Expression (Microarray)	70	Norway	-
	V5-GSE2223-E	GSE2223	Expression (Microarray)	54	USA	(Bredel et al., 2006; Bredel et al., 2005)
	V6-Ivy GAP-E	Ivy GAP	Expression (RNA-Seq)	122	unknown	(Puchalski. et al., 2018)
	V7-CGGA-E	CGGA	Expression (Microarray)	301	China	(Sun et al., 2014; Yan et al., 2012)
	V8-GSE50923-M	GEO	Methylation (27K)	78	USA	(Lai et al., 2014)
	V9-GSE61160-M	GEO	Methylation (450K)	51	Spain	(Mur et al., 2013)
	V10-CGGA-M	CGGA	Methylation (27K)	159	China	(Zhang et al., 2013)
	V11-TCGA-M	TCGA	Methylation (WGBS)	6	white	-

33

V12-CGGA-E	CGGA	Expression (RNA-Seq)	310	China	(Bao et al., 2014)
V13-CGGA-E	CGGA	Expression (RNA-Seq)	667	China	-
V14-GSE60274-M	GEO	Methylation (450K)	68	Switzerland	(Kurscheid et al., 2015)

1 Note:
2 CGGA: Chinese Glioma Genome Atlas, http://www.cgga.org.cn
3 GEO: Gene Expression Omnibus, https://www.ncbi.nlm.nih.gov/geo/
4 GTEx: Genotype-Tissue Expression, https://www.gtexportal.org/
5 TCGA: The Cancer Genome Atlas, https://portal.gdc.cancer.gov
6 Ivy GAP: Ivy Glioblastoma Atlas Project, http://glioblastoma.alleninstitute.org/

1   **Table 2. Univariate and multivariate Cox regression analyses for *PRKCG***

Subgroup	Univariate			Multivariate		
	HR	**95% CI**	***P*-value**	**HR**	**95% CI**	***P*-value**
D4-TCGA-E (n = 607)						
PRKCG-expression	0.48	0.38-0.61	$p < 0.01$	0.58	0.45-0.73	$p < 0.01$
Age	1.07	1.06-1.09	$p < 0.01$	1.07	1.06-1.08	$p < 0.01$
Gender	1.01	0.75-1.36	$p = 0.9$	0.87	0.64-1.17	$p = 0.37$
V7-CGGA-E (*n* = 301)						
PRKCG-expression	0.73	0.60-0.88	$p < 0.01$	0.81	0.66-0.98	$p = 0.03$
Age	1.04	1.03-1.06	$p < 0.01$	1.04	1.02-1.05	$p < 0.01$
Gender	0.83	0.61-1.15	$p = 0.27$	0.87	0.63-1.20	$p = 0.39$
V12-CGGA-E (*n* = 310)						
PRKCG-expression	0.59	0.41-0.86	$p < 0.01$	0.59	0.40-0.86	$p < 0.01$
Age	1.04	1.02-1.05	$p < 0.01$	1.05	1.02-1.07	$p < 0.01$
Gender	0.85	0.60-1.19	$p = 0.35$	0.71	0.39-1.28	$p = 0.26$
V13-CGGA-E (*n* = 667)						
PRKCG-expression	0.64	0.5-0.74	$p < 0.01$	0.63	0.54-0.74	$p < 0.01$
Age	1.03	1.02-1.04	$p < 0.01$	1.03	1.02-1.04	$p < 0.01$
Gender	1.03	0.83-1.27	$p = 0.82$	1.07	0.86-1.32	$p = 0.57$
D4-TCGA-M (*n* = 862)						
PRKCG-methylation	0.45	0.39-0.52	$p < 0.01$	0.57	0.48-0.66	$p < 0.01$
Age	1.07	1.06-1.08	$p < 0.01$	1.06	1.05-1.07	$p < 0.01$
Gender	1.15	0.93-1.45	$p = 0.21$	1.23	0.99-1.52	$p = 0.06$
D4-TCGA-M (TMZ treatment; *n* = 228)						
PRKCG-methylation	0.69	0.49-0.98	$p = 0.04$	0.67	0.47-0.95	$p = 0.03$
MGMT	0.61	0.44-0.86	$p < 0.01$	0.66	0.47-0.94	$p = 0.02$
Age	1.03	1.02-1.05	$p < 0.01$	1.03	1.02-1.05	$p < 0.01$
Gender	0.76	0.54-1.07	$p = 0.11$	0.73	0.51-1.03	$p = 0.07$
V10-CCGA-M (*n* = 159)						
PRKCG-methylation	0.53	0.36-0.77	$p < 0.01$	0.50	0.34-0.73	$p < 0.01$
Age	1.04	1.02-1.05	$p < 0.01$	1.04	1.02-1.06	$p < 0.01$
Gender	1.23	0.75-2.02	$p = 0.42$	1.52	0.91-2.53	$p = 0.11$

Note: HR=Hazard

Ratio

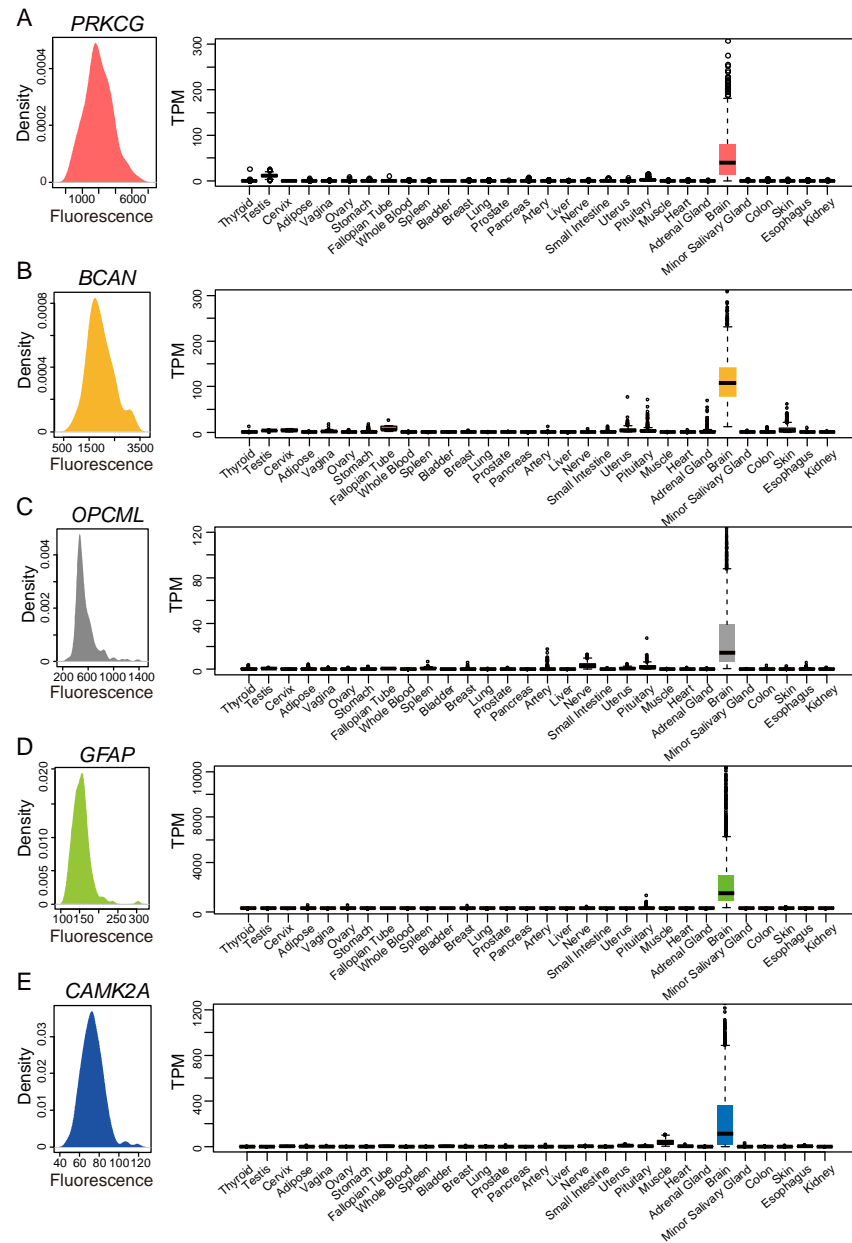1   **Supplemental Information**

2
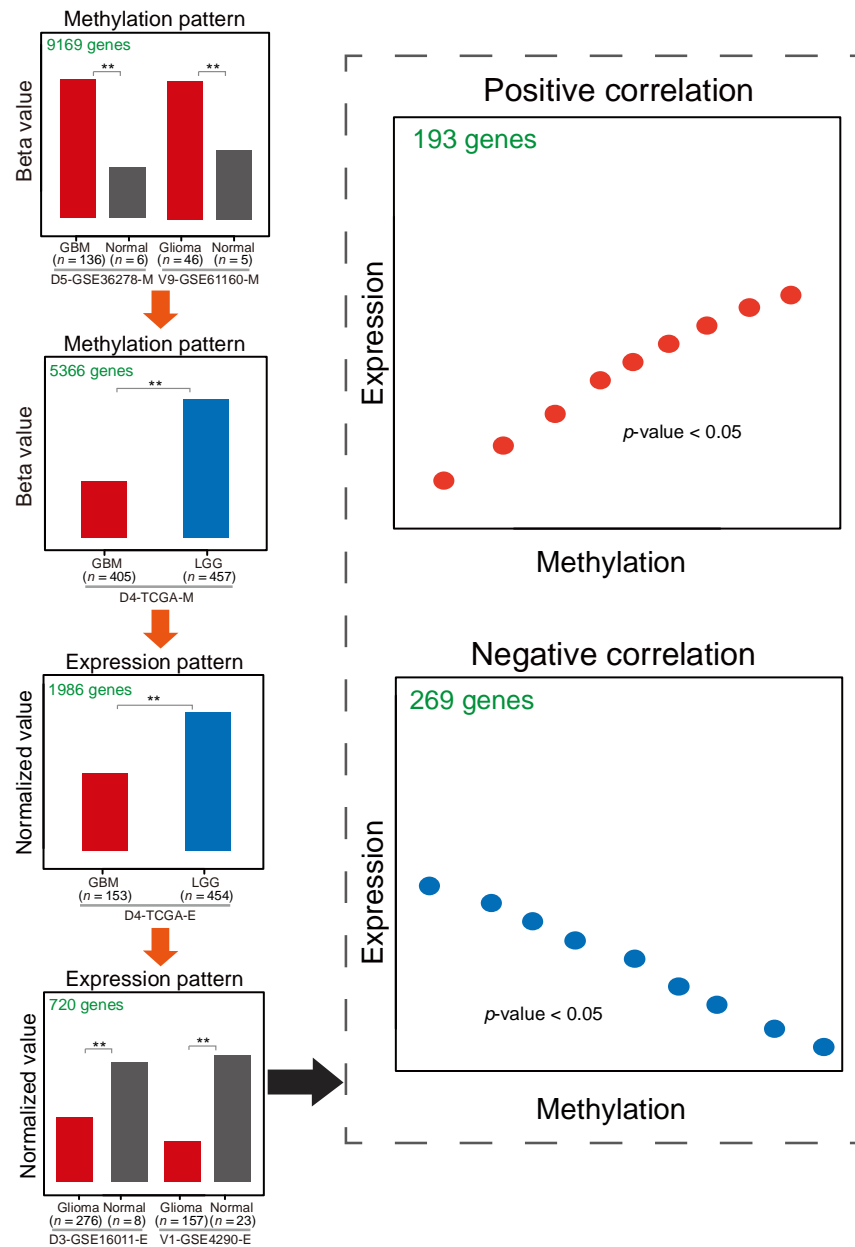
3

4   **Figure S1 Workflow for identification of *PRKCG*-like genes.** Four steps were adopted

5   to detect *PRKCG*-like genes in glioma. As a result, there are 193 and 269 genes that

1   present positive and negative correlation between expression and methylation,
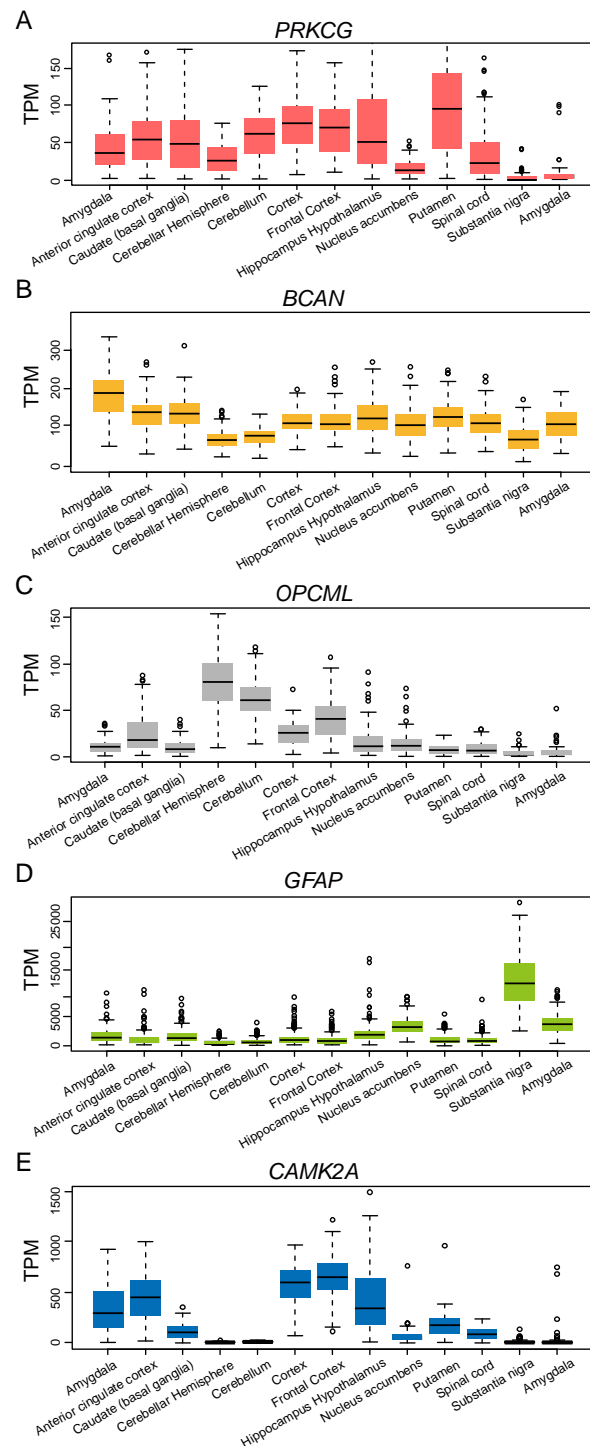
2   respectively.

3

4

5

1   **Figure S2 Protein level distributions in CSF and RNA expression profiles of *PRKCG***

2   **(A), *BCAN* (B), *OPCML* (C), *GFAP* (D), *CAMK2A* (E) across 30 normal human**

3   **tissues.** Related to Figure 1.

4

5

1

2 **Figure S3 Expression profiles of *PRKCG* (A), *BCAN* (B), *OPCML* (C), *GFAP* (D),**

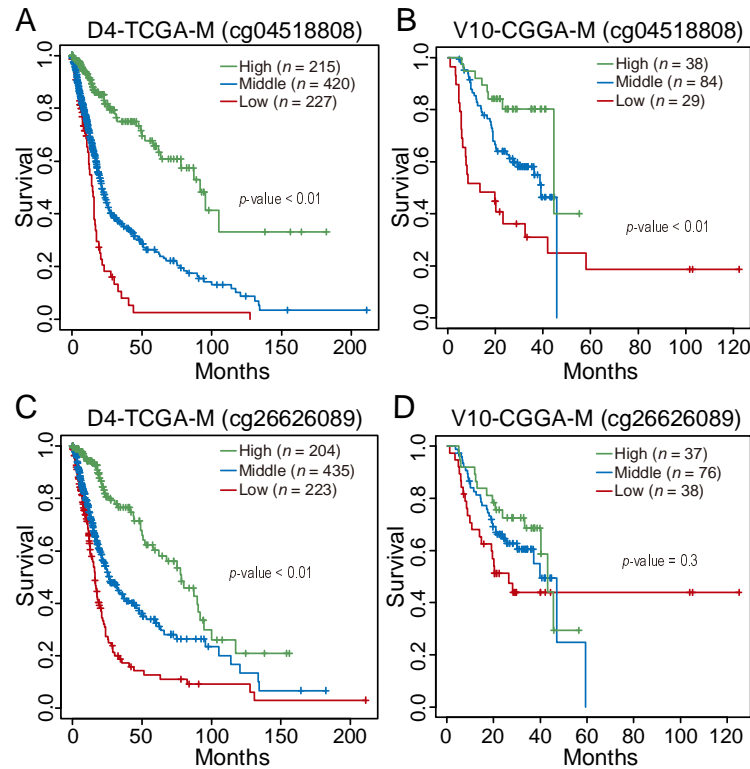3 **and *CAMK2A* (E) across 13 human brain regions.** Related to Figure 1.

1
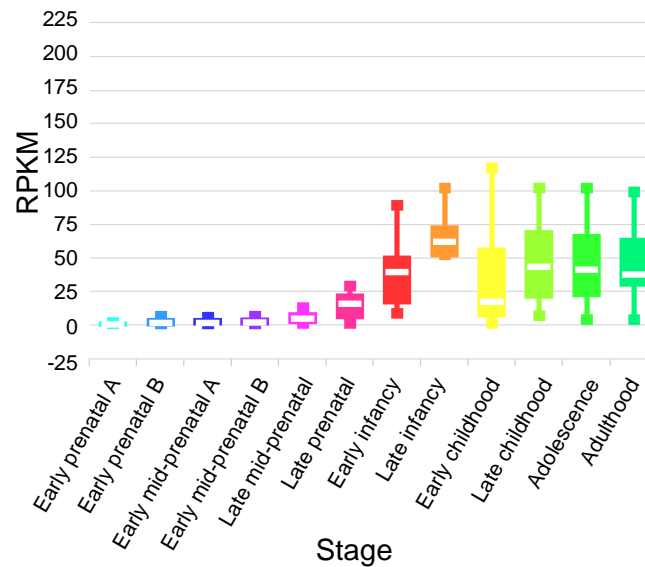
2

3   **Figure S4 Expression profiles of *PRKCG* during brain development.**
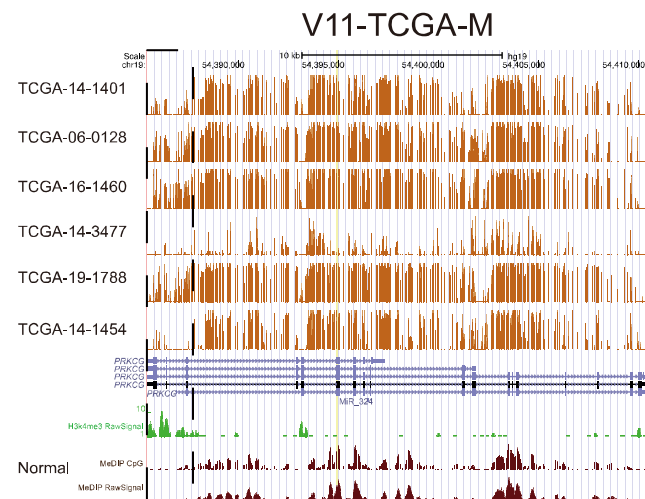
4

5   **Figure S5 Bisulfite DNA methylation profiles of *PRKCG* across six GBM samples**

6   **and one normal sample.** Related to Figure 3.

1

Figure S6 Prognostic potential of *PRKCG* DNA methylation. Glioma patients were divided into three groups based on the first and third quartile of cg04518808/cg26626089 methylation in discovery dataset (D4-TCGA-M in panel A/C) and validation dataset (V10-CGGA-M in panel B/D). All these datasets can be publicly accessible at ftp://download.big.ac.cn/glioma_data/. The log-rank tests were used to examine the statistical significance between different survival curves.

1
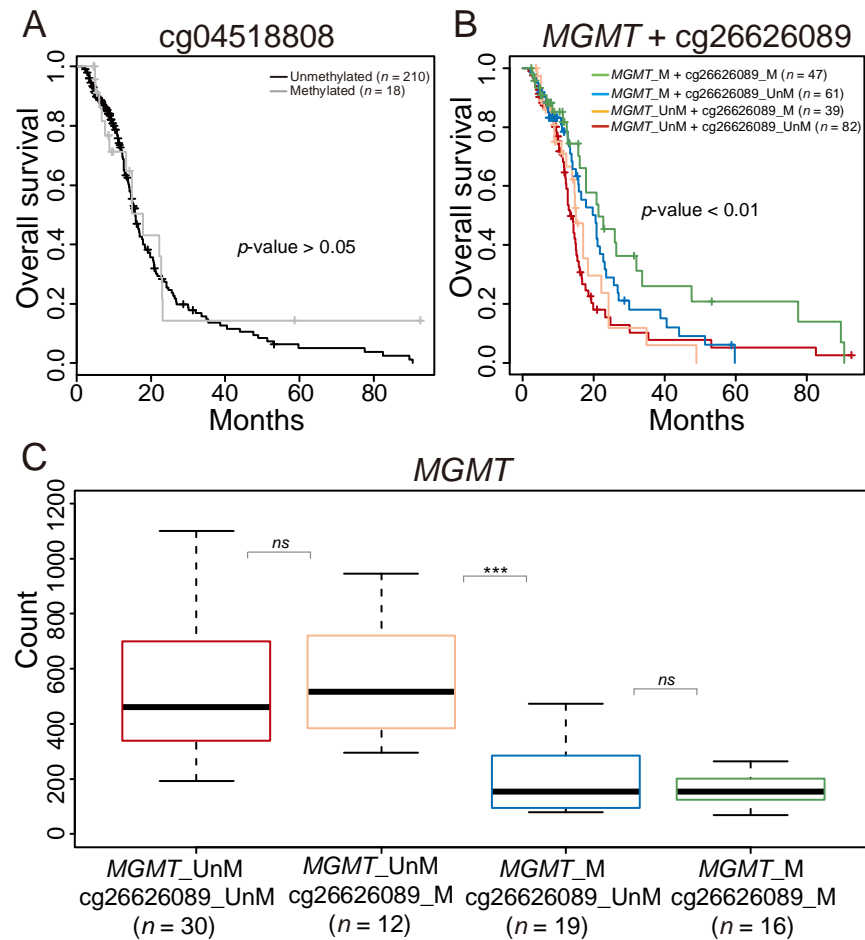
Figure S7 Predictive potential of *PRKCG* DNA methylation. (**A**) Kaplan-Meier survival curves for GBM patients with TMZ treatment based on *PRKCG* (cg04518808) methylation. (**B**) Methylation site cg26626089 in combination with *MGMT*, which were used to classify GBM patients into four groups. (**C**) Expression profiles of *MGMT* in the four groups.
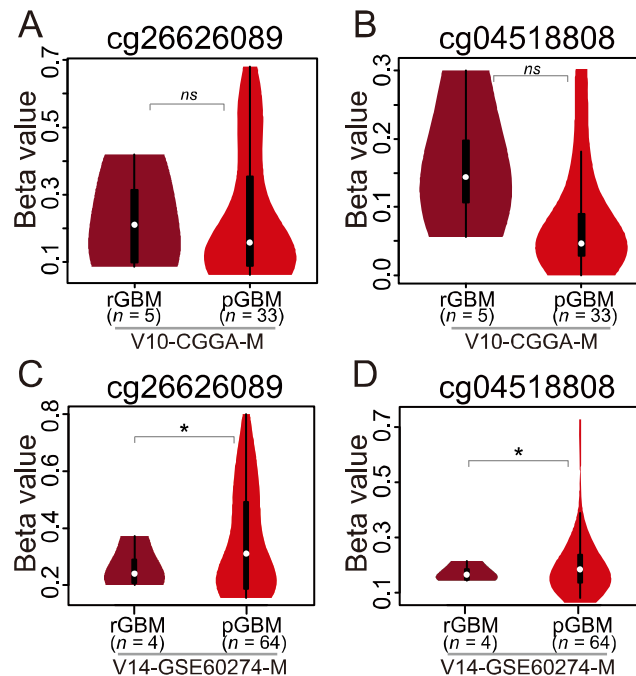
1
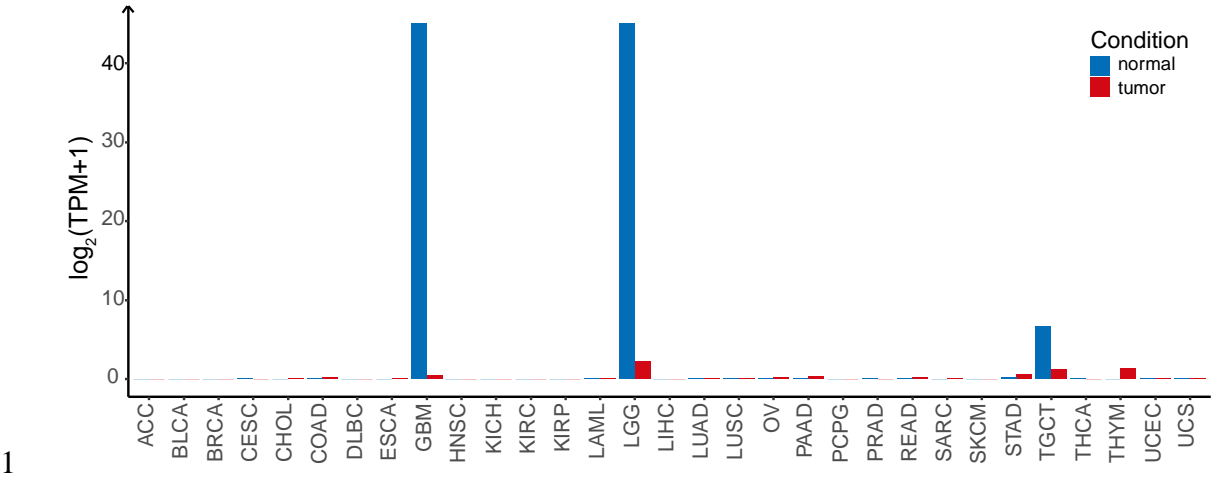
2 **Figure S8 Expression profiles of *PRKCG* across 31 human tumor and normal tissues.**

1
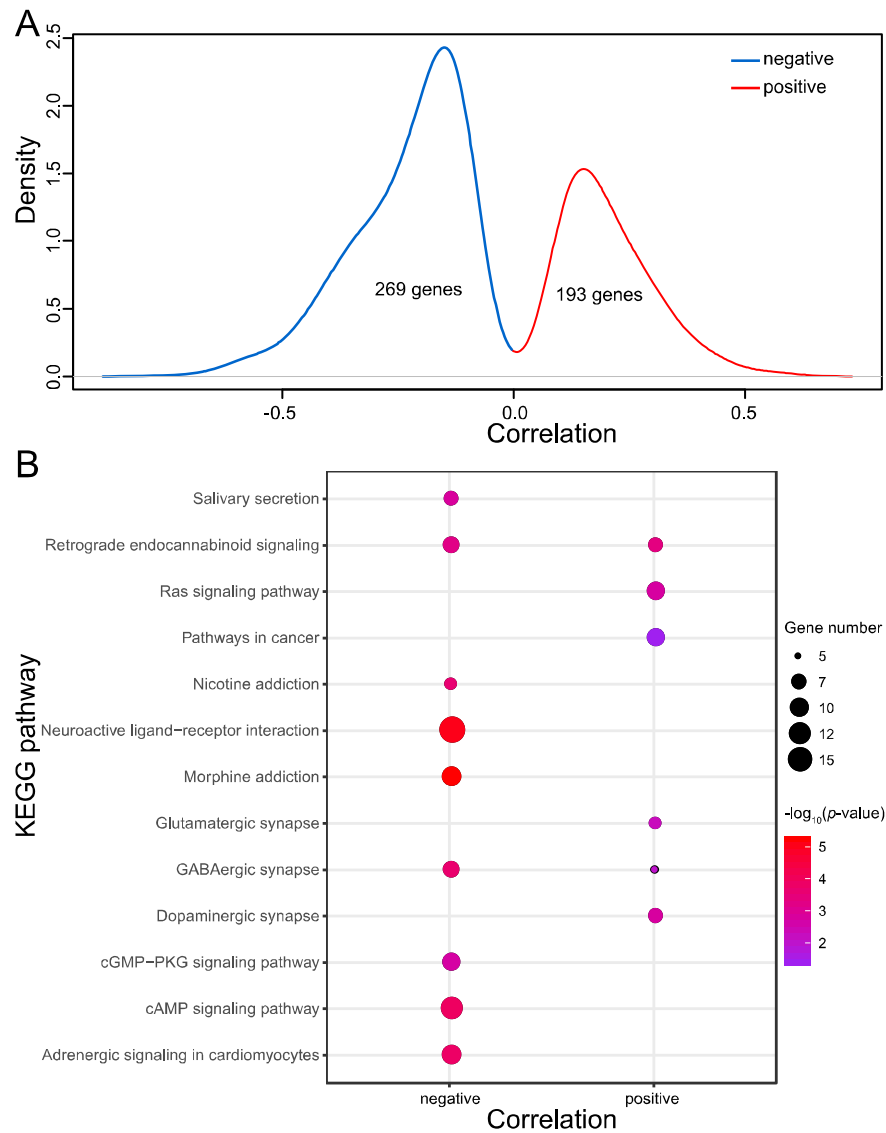
Figure S9 Comparative analysis of two groups' genes that present positive and

negative correlations between gene expression and methylation, respectively. (A) The

density plot of the Spearman correlation. (B) The KEGG pathway enrichment.

1

Figure S10 DNA methylation profiles of *PRKCG* in recurrent GBM (rGBM) and primary GBM (pGBM) samples. *PRKCG* methylation profiles were compared between rGBM and pGBM samples (V10-CGGA-M in panels A and B and V14-GSE60274-M in panels C and D). All these datasets can be publicly accessible at ftp://download.big.ac.cn/glioma_data/. The Wilcoxon tests were used and the statistical significance levels were coded by: *ns p>*0.05, * *p<*0.05, ** *p<*0.01 and *** *p<*0.001.

8

Table S1 15,176 genes' RNA expression levels across 30 normal human tissues.

10

Table S2 τ values and maximum expression levels of the 100 brain-specific protein-coding genes.

1

Table S3 Gene Ontology categories of genes co-expressed with *PRKCG*.

3

Table S4 Spearman correlation between gene expression and CpG site methylation (*PRKCG*-like genes).

6

7 **References**

8 A.Maher., E., B.Furnari., F., M.Bachoo., R., H.rowitch., D., N.Louis., D., Cavenee., W.
9 K., and A.DePinho, R. (2019). Malignant glioma:genetics and biology of a grave matter.
10 GENES & DEVELOPMENT *15*, 1311-1333.

11 Bao, Z. S., Chen, H. M., Yang, M. Y., Zhang, C. B., Yu, K., Ye, W. L., Hu, B. Q., Yan, W.,
12 Zhang, W., Akers, J.*, et al.* (2014). RNA-seq of 272 gliomas revealed a novel, recurrent
13 PTPRZ1-MET fusion transcript in secondary glioblastomas. Genome Res *24*, 1765-1773.

14 Baylin, S. B. (2005). DNA methylation and gene silencing in cancer. Nat Clin Pract
15 Oncol *2 Suppl 1*, S4-11.

16 Binabaj, M. M., Bahrami, A., ShahidSales, S., Joodi, M., Joudi Mashhad, M., Hassanian,
17 S. M., Anvari, K., and Avan, A. (2018). The prognostic value of MGMT promoter
18 methylation in glioblastoma: A meta-analysis of clinical trials. J Cell Physiol *233*, 378-
19 386.

20 Bredel, M., Bredel, C., Juric, D., Duran, G. E., Yu, R. X., Harsh, G. R., Vogel, H., Recht,
21 L. D., Scheck, A. C., and Sikic, B. I. (2006). Tumor necrosis factor-alpha-induced protein
22 3 as a putative regulator of nuclear factor-kappaB-mediated resistance to O6-alkylating
23 agents in human glioblastomas. J Clin Oncol *24*, 274-287.

24 Bredel, M., Bredel, C., Juric, D., Harsh, G. R., Vogel, H., Recht, L. D., and Sikic, B. I.
25 (2005). Functional network analysis reveals extended gliomagenesis pathway maps and
26 three novel MYC-interacting genes in human gliomas. Cancer Res *65*, 8679-8689.

27 Cairncross, G., Wang, M., Shaw, E., Jenkins, R., Brachman, D., Buckner, J., Fink, K.,
28 Souhami, L., Laperriere, N., Curran, W., and Mehta, M. (2013). Phase III trial of
29 chemoradiotherapy for anaplastic oligodendroglioma: long-term results of RTOG 9402. J
30 Clin Oncol *31*, 337-343.

Cancer Genome Atlas Research, N., Brat, D. J., Verhaak, R. G., Aldape, K. D., Yung, W. K., Salama, S. R., Cooper, L. A., Rheinbay, E., Miller, C. R., Vitucci, M.*, et al.* (2015). Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. N Engl J Med *372*, 2481-2498.

Carminati., P. O., Mello., S. S., Fachin., A. L., Junta., C. M., Sandrin-Garcia., P., Carlotti., C. G., Donadi., E. A., Passos., G. A. S., and Sakamoto-Hojo, E. T. (2010). Alterations in gene expression profiles correlated with cisplatin cytotoxicity in the glioma U343 cell line. Genetics and molecular biology *33*, 159-168.

Catriona M. Dowling, S. L. H., James J. Phelan, Mary Clare Cathcart, Stephen P. Finn, Brian Mehigan, Paul McCormick, John C. Coffey, Jacintha, and Kiely, O. S. a. P. A. (2017). Expression of protein kinase C gamma promotes cell migration in colon cancer. Oncotarget *8*, 72096-72107.

Ceccarelli, M., Barthel, F. P., Malta, T. M., Sabedot, T. S., Salama, S. R., Murray, B. A., Morozova, O., Newton, Y., Radenbaugh, A., Pagnotta, S. M.*, et al.* (2016). Molecular Profiling Reveals Biologically Discrete Subsets and Pathways of Progression in Diffuse Glioma. Cell *164*, 550-563.

Christensen, B. C., Smith, A. A., Zheng, S., Koestler, D. C., Houseman, E. A., Marsit, C. J., Wiemels, J. L., Nelson, H. H., Karagas, M. R., Wrensch, M. R.*, et al.* (2011). DNA methylation, isocitrate dehydrogenase mutation, and survival in glioma. J Natl Cancer Inst *103*, 143-153.

Cohen, A. L., Holmen, S. L., and Colman, H. (2013). IDH1 and IDH2 mutations in gliomas. Curr Neurol Neurosci Rep *13*, 345.

Cook, P. J., Thomas, R., Kannan, R., de Leon, E. S., Drilon, A., Rosenblum, M. K., Scaltriti, M., Benezra, R., and Ventura, A. (2018). Author Correction: Somatic chromosomal engineering identifies BCAN-NTRK1 as a potent glioma driver and therapeutic target. Nat Commun *9*, 16187.

Dang, L., Jin, S., and Su, S. M. (2010). IDH mutations in glioma and acute myeloid leukemia. Trends Mol Med *16*, 387-397.

Dang, L., Yen, K., and Attar, E. C. (2016). IDH mutations in cancer and progress toward development of targeted therapeutics. Ann Oncol *27*, 599-608.

do Carmo, A., Balca-Silva, J., Matias, D., and Lopes, M. C. (2013). PKC signaling in glioblastoma. Cancer Biol Ther *14*, 287-294.

Donson, A. M., Addo-Yobo, S. O., Handler, M. H., Gore, L., and Foreman, N. K. (2007). MGMT promoter methylation correlates with survival benefit and sensitivity to temozolomide in pediatric glioblastoma. Pediatr Blood Cancer *48*, 403-407.

1   Eckel-Passow, J. E., Lachance, D. H., Molinaro, A. M., Walsh, K. M., Decker, P. A.,
2   Sicotte, H., Pekmezci, M., Rice, T., Kosel, M. L., Smirnov, I. V.*, et al.* (2015). Glioma
3   Groups Based on 1p/19q, IDH, and TERT Promoter Mutations in Tumors. N Engl J Med
4   *372*, 2499-2508.

5   Esteller, M. (2007). Cancer epigenomics: DNA methylomes and histone-modification
6   maps. Nat Rev Genet *8*, 286-298.

7   Fan, Q. W., Cheng, C., Knight, Z. A., Haas-Kogan, D., Stokoe, D., James, C. D.,
8   McCormick, F., Shokat, K. M., and Weiss, W. A. (2009). EGFR signals to mTOR through
9   PKC and independently of Akt in glioma. Sci Signal *2*, ra4.

10 Flynn, J. R., Wang, L., Gillespie, D. L., Stoddard, G. J., Reid, J. K., Owens, J., Ellsworth,
11 G. B., Salzman, K. L., Kinney, A. Y., and Jensen, R. L. (2008). Hypoxia-regulated protein
12 expression, patient characteristics, and preoperative imaging as predictors of survival in
13 adults with glioblastoma multiforme. Cancer *113*, 1032-1042.

14 Fukushima, T., Takeshima, H., and Kataoka, H. (2009). Anti-glioma therapy with
15 temozolomide and status of the DNA-repair gene MGMT. Anticancer Res *29*, 4845-4854.

16 Gill, B. J., Pisapia, D. J., Malone, H. R., Goldstein, H., Lei, L., Sonabend, A., Yun, J.,
17 Samanamud, J., Sims, J. S., Banu, M.*, et al.* (2014). MRI-localized biopsies reveal
18 subtype-specific differences in molecular and cellular composition at the margins of
19 glioblastoma. Proc Natl Acad Sci U S A *111*, 12550-12555.

20 Gravendeel, L. A., Kouwenhoven, M. C., Gevaert, O., de Rooi, J. J., Stubbs, A. P., Duijm,
21 J. E., Daemen, A., Bleeker, F. E., Bralten, L. B., Kloosterhof, N. K.*, et al.* (2009).
22 Intrinsic gene expression profiles of gliomas are a better predictor of survival than
23 histology. Cancer Res *69*, 9065-9072.

24 Griesinger, A. M., Birks, D. K., Donson, A. M., Amani, V., Hoffman, L. M., Waziri, A.,
25 Wang, M., Handler, M. H., and Foreman, N. K. (2013). Characterization of distinct
26 immunophenotypes across pediatric brain tumor types. J Immunol *191*, 4880-4888.

27 Guo, C., Pirozzi, C. J., Lopez, G. Y., and Yan, H. (2011). Isocitrate dehydrogenase
28 mutations in gliomas: mechanisms, biomarkers and therapeutic target. Curr Opin Neurol
29 *24*, 648-652.

30 Hegi, M. E., Diserens, A. C., Gorlia, T., Hamou, M. F., de Tribolet, N., Weller, M., Kros,
31 J. M., Hainfellner, J. A., Mason, W., Mariani, L.*, et al.* (2005). MGMT gene silencing and
32 benefit from temozolomide in glioblastoma. N Engl J Med *352*, 997-1003.

33 Henrichsen, C. N., Chaignat, E., and Reymond, A. (2009). Copy number variants,
34 diseases and gene expression. Hum Mol Genet *18*, R1-8.

35 Herman, J. G., and Baylin, S. B. (2003). Gene Silencing in Cancer in Association with
36 Promoter Hypermethylation. New England Journal of Medicine *349*, 2042-2054.

48

Horst, M., Brouwer, E., Verwijnen, S., Rodijk, M., de Jong, M., Hoeben, R., de Leeuw, B., and Smitt, P. S. (2007). Targeting malignant gliomas with a glial fibrillary acidic protein (GFAP)-selective oncolytic adenovirus. J Gene Med *9*, 1071-1079.

Hu, H., Mu, Q., Bao, Z., Chen, Y., Liu, Y., Chen, J., Wang, K., Wang, Z., Nam, Y., Jiang, B.*, et al.* (2018). Mutational Landscape of Secondary Glioblastoma Guides MET-Targeted Trial in Brain Tumor. Cell *175*, 1665-1678 e1618.

Hu, L., Li, X., Liu, Q., Xu, J., Ge, H., Wang, Z., Wang, H., Wang, Z., Shi, C., Xu, X.*, et al.* (2017). UBE2S, a novel substrate of Akt1, associates with Ku70 and regulates DNA repair and glioblastoma multiforme resistance to chemotherapy. Oncogene *36*, 1145-1156.

Ichimura, K., Pearson, D. M., Kocialkowski, S., Backlund, L. M., Chan, R., Jones, D. T., and Collins, V. P. (2009). IDH1 mutations are present in the majority of common adult gliomas but rare in primary glioblastomas. Neuro Oncol *11*, 341-347.

Jenkins, R. B., Blair, H., Ballman, K. V., Giannini, C., Arusell, R. M., Law, M., Flynn, H., Passe, S., Felten, S., Brown, P. D.*, et al.* (2006). A t(1;19)(q10;p10) mediates the combined deletions of 1p and 19q and predicts a better prognosis of patients with oligodendroglioma. Cancer Res *66*, 9852-9861.

John CG Spainhour, H. S. L., Soojin V Yi and Peng Qiu (2019). Correlation Patterns Between DNA Methylation and Gene Expression in The Cancer Genome Atlas. Cancer Informatics *18*, 1176935119828776.

John Lonsdale, J. T., Mike Salvatore, Rebecca Phillips (2013). The Genotype-Tissue Expression (GTEx) project. Nat Genet *45*, 580-585.

John, S., Sivakumar, K. C., and Mishra, R. (2017). Bacoside A Induces Tumor Cell Death in Human Glioblastoma Cell Lines through Catastrophic Macropinocytosis. Front Mol Neurosci *10*, 171.

Jung, E., Osswald, M., Blaes, J., Wiestler, B., Sahm, F., Schmenger, T., Solecki, G., Deumelandt, K., Kurz, F. T., Xie, R.*, et al.* (2017). Tweety-Homolog 1 Drives Brain Colonization of Gliomas. J Neurosci *37*, 6837-6850.

Kerr, D. A., Lopez, H. U., Deshpande, V., Hornicek, F. J., Duan, Z., Zhang, Y., Rosenberg, A. E., Borger, D. R., and Nielsen, G. P. (2013). Molecular distinction of chondrosarcoma from chondroblastic osteosarcoma through IDH1/2 mutations. Am J Surg Pathol *37*, 787-795.

Kim, Y. W., Koul, D., Kim, S. H., Lucio-Eterovic, A. K., Freire, P. R., Yao, J., Wang, J., Almeida, J. S., Aldape, K., and Yung, W. K. (2013). Identification of prognostic gene signatures of glioblastoma: a study based on TCGA data analysis. Neuro-oncology *15*, 829-839.

1 Klebe, S., Durr, A., Rentschler, A., Hahn-Barma, V., Abele, M., Bouslam, N., Schols, L.,
2 Jedynak, P., Forlani, S., Denis, E.*, et al.* (2005). New mutations in protein kinase
3 Cgamma associated with spinocerebellar ataxia type 14. Ann Neurol *58*, 720-729.

4 Kloosterhof, N. K., Bralten, L. B. C., Dubbink, H. J., French, P. J., and van den Bent, M.
5 J. (2011). Isocitrate dehydrogenase-1 mutations: a fundamentally new understanding of
6 diffuse glioma? The Lancet Oncology *12*, 83-91.

7 Kurscheid, S., Bady, P., Sciuscio, D., Samarzija, I., Shay, T., Vassallo, I., Criekinge, W.
8 V., Daniel, R. T., van den Bent, M. J., Marosi, C.*, et al.* (2015). Chromosome 7 gain and
9 DNA hypermethylation at the HOXA10 locus are associated with expression of a stem
10 cell related HOX-signature in glioblastoma. Genome Biol *16*, 16.

11 La Thangue, N. B., and Kerr, D. J. (2011). Predictive biomarkers: a paradigm shift
12 towards personalized cancer medicine. Nat Rev Clin Oncol *8*, 587-596.

13 Lai, R. K., Chen, Y., Guan, X., Nousome, D., Sharma, C., Canoll, P., Bruce, J., Sloan, A.
14 E., Cortes, E., Vonsattel, J. P.*, et al.* (2014). Genome-wide methylation analyses in
15 glioblastoma multiforme. PLoS One *9*, e89376.

16 Li, Y. H., Yu, C. Y., Li, X. X., Zhang, P., Tang, J., Yang, Q., Fu, T., Zhang, X., Cui, X., Tu,
17 G.*, et al.* (2018). Therapeutic target database update 2018: enriched resource for
18 facilitating bench-to-clinic research of targeted therapeutics. Nucleic Acids Res *46*,
19 D1121-D1127.

20 Long, H., Liang, C., Zhang, X., Fang, L., Wang, G., Qi, S., Huo, H., and Song, Y. (2017).
21 Prediction and Analysis of Key Genes in Glioblastoma Based on Bioinformatics. Biomed
22 Res Int *2017*, 7653101.

23 Louis, D. N., Ohgaki, H., Wiestler, O. D., Cavenee, W. K., Burger, P. C., Jouvet, A.,
24 Scheithauer, B. W., and Kleihues, P. (2007). The 2007 WHO classification of tumours of
25 the central nervous system. Acta Neuropathol *114*, 97-109.

26 Louis, D. N., Perry, A., Reifenberger, G., von Deimling, A., Figarella-Branger, D.,
27 Cavenee, W. K., Ohgaki, H., Wiestler, O. D., Kleihues, P., and Ellison, D. W. (2016). The
28 2016 World Health Organization Classification of Tumors of the Central Nervous
29 System: a summary. Acta Neuropathol *131*, 803-820.

30 McNamara, M. G., Sahebjam, S., and Mason, W. P. (2013). Emerging biomarkers in
31 glioblastoma. Cancers (Basel) *5*, 1103-1119.

32 Miller, A. M., Shah, R. H., Pentsova, E. I., Pourmaleki, M., Briggs, S., Distefano, N.,
33 Zheng, Y., Skakodub, A., Mehta, S. A., Campos, C.*, et al.* (2019). Tracking tumour
34 evolution in glioma through liquid biopsies of cerebrospinal fluid. Nature *565*, 654-658.

35 Molinaro, A. M., Taylor, J. W., Wiencke, J. K., and Wrensch, M. R. (2019). Genetic and
36 molecular epidemiology of adult diffuse glioma. Nat Rev Neurol *15*, 405-417.

1   Mouliere, F., Mair, R., Chandrananda, D., Marass, F., Smith, C. G., Su, J., Morris, J.,
2   Watts, C., Brindle, K. M., and Rosenfeld, N. (2018). Detection of cell-free DNA
3   fragmentation and copy number alterations in cerebrospinal fluid from glioma patients.
4   EMBO Mol Med *10*.

5   Mur, P., Mollejo, M., Hernandez-Iglesias, T., de Lope, A. R., Castresana, J. S., Garcia, J.
6   F., Fiano, C., Ribalta, T., Rey, J. A., and Melendez, B. (2015). Molecular classification
7   defines 4 prognostically distinct glioma groups irrespective of diagnosis and grade. J
8   Neuropathol Exp Neurol *74*, 241-249.

9   Mur, P., Mollejo, M., Ruano, Y., de Lope, A. R., Fiano, C., Garcia, J. F., Castresana, J. S.,
10 Hernandez-Lain, A., Rey, J. A., and Melendez, B. (2013). Codeletion of 1p and 19q
11 determines distinct gene methylation and expression profiles in IDH-mutated
12 oligodendroglial tumors. Acta Neuropathol *126*, 277-289.

13 Park, J. Y., Lee, J. E., Park, J. B., Yoo, H., Lee, S. H., and Kim, J. H. (2014). Roles of
14 Long Non-Coding RNAs on Tumorigenesis and Glioma Development. Brain Tumor Res
15 Treat *2*, 1-6.

16 Phillips, H. S., Kharbanda, S., Chen, R., Forrest, W. F., Soriano, R. H., Wu, T. D., Misra,
17 A., Nigro, J. M., Colman, H., Soroceanu, L*., et al.* (2006). Molecular subclasses of high-
18 grade glioma predict prognosis, delineate a pattern of disease progression, and resemble
19 stages in neurogenesis. Cancer Cell *9*, 157-173.

20 Puchalski., R. B., Shah., N., Miller., J., Dalley., R., Nomura., S. R., Yoon., J.-G., Smith.,
21 K. A., Lankerovich., M., Bertagnolli., D., Bickley., K*., et al.* (2018). An anatomic
22 transcriptional atlas of human glioblastoma. Science *360*, 660-663.

23 Reed, J. E., Dunn, J. R., du Plessis, D. G., Shaw, E. J., Reeves, P., Gee, A. L., Warnke, P.
24 C., Sellar, G. C., Moss, D. J., and Walker, C. (2007). Expression of cellular adhesion
25 molecule 'OPCML' is down-regulated in gliomas and other brain tumours. Neuropathol
26 Appl Neurobiol *33*, 77-85.

27 Reina-Campos, M., Diaz-Meco, M. T., and Moscat, J. (2019). The Dual Roles of the
28 Atypical Protein Kinase Cs in Cancer. Cancer Cell *36*, 218-235.

29 Rivera, A. L., Pelloski, C. E., Gilbert, M. R., Colman, H., De La Cruz, C., Sulman, E. P.,
30 Bekele, B. N., and Aldape, K. D. (2010). MGMT promoter methylation is predictive of
31 response to radiotherapy and prognostic in the absence of adjuvant alkylating
32 chemotherapy for glioblastoma. Neuro Oncol *12*, 116-121.

33 Saito, N., and Shirai, Y. (2002). Protein kinase C gamma (PKC gamma): function of
34 neuron specific isotype. J Biochem *132*, 683-687.

35 Sasayama, D., Hattori, K., Ogawa, S., Yokota, Y., Matsumura, R., Teraishi, T., Hori, H.,
36 Ota, M., Yoshida, S., and Kunugi, H. (2017). Genome-wide quantitative trait loci
37 mapping of the human cerebrospinal fluid proteome. Hum Mol Genet *26*, 44-51.

1   Schwartzbaum, J. A., Fisher, J. L., Aldape, K. D., and Wrensch, M. (2006). Epidemiology
2   and molecular pathology of glioma. Nat Clin Pract Neurol *2*, 494-503; quiz 491 p
3   following 516.

4   Shao, Y., Chen, C., Shen, H., He, B. Z., Yu, D., Jiang, S., Zhao, S., Gao, Z., Zhu, Z.,
5   Chen, X.*, et al.* (2019). GenTree, an integrated resource for analyzing the evolution and
6   function of primate-specific coding genes. Genome Res *29*, 682-696.

7   Stewart, L. A. (2002). Chemotherapy in adult high-grade glioma: a systematic review and
8   meta-analysis of individual patient data from 12 randomised trials. Lancet *359*, 1011-
9   1018.

10   Stupp, R., Mason, W. P., van den Bent, M. J., Weller, M., Fisher, B., Taphoorn, M. J. B.,
11   Belanger, K., Brandes, A. A., Marosi, C., Bogdahn, U.*, et al.* (2005). Radiotherapy plus
12   Concomitant and Adjuvant Temozolomide for Glioblastoma. New England Journal of
13   Medicine *352*, 987-996.

14   Sturm, D., Witt, H., Hovestadt, V., Khuong-Quang, D. A., Jones, D. T., Konermann, C.,
15   Pfaff, E., Tonjes, M., Sill, M., Bender, S.*, et al.* (2012). Hotspot mutations in H3F3A and
16   IDH1 define distinct epigenetic and biological subgroups of glioblastoma. Cancer Cell
17   *22*, 425-437.

18   Sun, L., Hui, A. M., Su, Q., Vortmeyer, A., Kotliarov, Y., Pastorino, S., Passaniti, A.,
19   Menon, J., Walling, J., Bailey, R.*, et al.* (2006). Neuronal and glioma-derived stem cell
20   factor induces angiogenesis within the brain. Cancer Cell *9*, 287-300.

21   Sun, Y., Zhang, W., Chen, D., Lv, Y., Zheng, J., Lilljebjorn, H., Ran, L., Bao, Z., Soneson,
22   C., Sjogren, H. O.*, et al.* (2014). A glioma classification scheme based on coexpression
23   modules of EGFR and PDGFRA. Proc Natl Acad Sci U S A *111*, 3538-3543.

24   Sydney C. Gary, C. A. Z., Veronica L. Chiang, Janette U. Gaw, Grace Gray, Susan
25   Hockfield (2000). cDNA cloning, chromosomal localization, and expression analysis of
26   human BEHAB/brevican, a brain specific proteoglycan regulated during cortical
27   development and in glioma. GENE *256*, 139-147.

28   Turcan, S., Rohle, D., Goenka, A., Walsh, L. A., Fang, F., Yilmaz, E., Campos, C., Fabius,
29   A. W., Lu, C., Ward, P. S.*, et al.* (2012). IDH1 mutation is sufficient to establish the
30   glioma hypermethylator phenotype. Nature *483*, 479-483.

31   van Bodegraven, E. J., van Asperen, J. V., Robe, P. A. J., and Hol, E. M. (2019).
32   Importance of GFAP isoform-specific analyses in astrocytoma. Glia *67*, 1417-1433.

33   van den Bent, M. J., Brandes, A. A., Taphoorn, M. J., Kros, J. M., Kouwenhoven, M. C.,
34   Delattre, J. Y., Bernsen, H. J., Frenay, M., Tijssen, C. C., Grisold, W.*, et al.* (2013).
35   Adjuvant procarbazine, lomustine, and vincristine chemotherapy in newly diagnosed
36   anaplastic oligodendroglioma: long-term follow-up of EORTC brain tumor group study
37   26951. J Clin Oncol *31*, 344-350.

1 van den Bent, M. J., Gao, Y., Kerkhof, M., Kros, J. M., Gorlia, T., van Zwieten, K.,
2 Prince, J., van Duinen, S., Sillevis Smitt, P. A., Taphoorn, M., and French, P. J. (2015).
3 Changes in the EGFR amplification and EGFRvIII expression between paired primary
4 and recurrent glioblastomas. Neuro Oncol *17*, 935-941.

5 Verhaak, R. G., Hoadley, K. A., Purdom, E., Wang, V., Qi, Y., Wilkerson, M. D., Miller,
6 C. R., Ding, L., Golub, T., Mesirov, J. P.*, et al.* (2010). Integrated genomic analysis
7 identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in
8 PDGFRA, IDH1, EGFR, and NF1. Cancer Cell *17*, 98-110.

9 Waitkus, M. S., Diplas, B. H., and Yan, H. (2018). Biological Role and Therapeutic
10 Potential of IDH Mutations in Cancer. Cancer Cell *34*, 186-195.

11 Wang, Q., Zhang, J., Liu, Y., Zhang, W., Zhou, J., Duan, R., Pu, P., Kang, C., and Han, L.
12 (2016). A novel cell cycle-associated lncRNA, HOXA11-AS, is transcribed from the 5-
13 prime end of the HOXA transcript and is a biomarker of progression in glioma. Cancer
14 Lett *373*, 251-259.

15 Wesseling, P., Kros, J. M., and Jeuken, J. W. M. (2011). The pathological diagnosis of
16 diffuse gliomas: towards a smart synthesis of microscopic and molecular information in a
17 multidisciplinary context. Diagnostic Histopathology *17*, 486-494.

18 Wesseling, P., van den Bent, M., and Perry, A. (2015). Oligodendroglioma: pathology,
19 molecular mechanisms and markers. Acta Neuropathol *129*, 809-827.

20 Wick, W., Weller, M., van den Bent, M., Sanson, M., Weiler, M., von Deimling, A., Plass,
21 C., Hegi, M., Platten, M., and Reifenberger, G. (2014). MGMT testing--the challenges for
22 biomarker-based glioma treatment. Nat Rev Neurol *10*, 372-385.

23 Wiestler, B., Capper, D., Holland-Letz, T., Korshunov, A., von Deimling, A., Pfister, S.
24 M., Platten, M., Weller, M., and Wick, W. (2013). ATRX loss refines the classification of
25 anaplastic gliomas and identifies a subgroup of IDH mutant astrocytic tumors with better
26 prognosis. Acta Neuropathol *126*, 443-451.

27 Yabe, I., Sasaki, H., Chen, D. H., Raskind, W. H., Bird, T. D., Yamashita, I., Tsuji, S.,
28 Kikuchi, S., and Tashiro, K. (2003). Spinocerebellar ataxia type 14 caused by a mutation
29 in protein kinase C gamma. Arch Neurol *60*, 1749-1751.

30 Yan, W., Zhang, W., You, G., Zhang, J., Han, L., Bao, Z., Wang, Y., Liu, Y., Jiang, C.,
31 Kang, C.*, et al.* (2012). Molecular classification of gliomas based on whole genome gene
32 expression: a systematic report of 225 samples from the Chinese Glioma Cooperative
33 Group. Neuro Oncol *14*, 1432-1440.

34 Yanai, I., Benjamin, H., Shmoish, M., Chalifa-Caspi, V., Shklar, M., Ophir, R., Bar-Even,
35 A., Horn-Saban, S., Safran, M., Domany, E.*, et al.* (2005). Genome-wide midrange
36 transcription profiles reveal expression level relationships in human tissue specification.
37 Bioinformatics *21*, 650-659.

1 Zhang, S., Zhao, B. S., Zhou, A., Lin, K., Zheng, S., Lu, Z., Chen, Y., Sulman, E. P., Xie,
2 K., Bogler, O., *et al.* (2017). m(6)A Demethylase ALKBH5 Maintains Tumorigenicity of
3 Glioblastoma Stem-like Cells by Sustaining FOXM1 Expression and Cell Proliferation
4 Program. Cancer Cell *31*, 591-606 e596.

5 Zhang, W., Yan, W., You, G., Bao, Z., Wang, Y., Liu, Y., You, Y., and Jiang, T. (2013).
6 Genome-wide DNA methylation profiling identifies ALDH1A3 promoter methylation as
7 a prognostic predictor in G-CIMP- primary glioblastoma. Cancer Lett *328*, 120-125.

8 Zhao, Z., Zhang, K., Wang, Z., Wang, K., Liu, X., Wu, F., and Chen, J. (2019). A
9 comprehensive review of available omics data resources and molecular profiling for
10 precision glioma studies. Biomed Rep *10*, 3-9.

11