Individual Assignment of Adult Diffuse Gliomas into the EM/PM Molecular Subtypes Using a TaqMan Low-Density Array

Jiuyi Li1, Yang Xue1, Anna Wenger2, Yingyu Sun1, Zheng Wang3,4,5, Chuanbao Zhang3,4,5, Yunqiu Zhang1, Boglarka Fekete Rydenhag6, Bertil Rydenhag6, Asgeir Store Jakola3,6, Tao Jiang4,5,7, Helena Carén2, and Xiaolong Fan1,7

Abstract

Purpose: We aimed to develop a diagnostic platform to capture the transcriptomic resemblance of individual adult diffuse gliomas of WHO grades II to IV to neural development and the genomic signature associated with glioma progression.

Experiment Design: Based on the EM/PM classification scheme, we designed a RT-PCR-based TaqMan low-density array (TLDA) containing 44 classifier and 4 reference genes. Samples of a training dataset (GSE48865), characterized by mRNA-seq, were utilized to optimize the TLDA design and to develop a support vector machine (SVM)-based prediction model. Complemented with Sanger sequencing for IDH1/2, and low coverage whole genome sequencing (WGS), the TLDA and SVM prediction model were tested in a validation (31 gliomas) and a test (121 gliomas) dataset.

Introduction

Gliomas are the most common primary tumors in the adult central nervous system (CNS), with a yearly incidence of ~5 cases per 100,000 people (1, 2). Despite intensive research over the past decades, the molecular mechanisms underlying glioma development and progression are inadequately understood and effective therapies are lacking (3). Gliomas manifest a high degree of clinical and biological heterogeneity. Based on their microscopic similarities with different putative cell types of origin, and their presumed levels of differentiation, current standard glioma diagnosis defines gliomas as astrocytomas or oligodendrogliomas; the severity of gliomas is further distinguished by malignant grades I to IV according to features of cellular atypia, cell proliferation, angiogenesis, and necrosis (4, 5). The most malignant form of gliomas is the grade IV astrocytoma (also termed as glioblastoma, GBM), which account for more than 50% of all adult gliomas. The integration of canonical genomic alterations including mutations in IDH1 or IDH2, and co-deletion of 1p/19q has improved the power of risk stratification and the objectiveness of glioma diagnosis (4), and patients with 1p/19q co-deletion may benefit from radiotherapy combined with a procarbazine, lomustine, and vincristine chemotherapy regime (6, 7).

IDH mutation and 1p19q co-deletion represent the initiating genomic abnormalities in gliomas with transcriptomic resemblance to oligodendrocyte progenitor cells (OPC; ref. 8). However, adult gliomas are constantly evolving and progressing, so that a given glioma may harbor tens or hundreds of copy number variations (CNV) and single-nucleotide alterations, which most likely also contribute to glioma pathogenesis and clinical behavior. In fact, a subset of IDH mutated gliomas showed poor prognosis, largely due to the activation of cell-cycle genes (9). Furthermore, pathways conferred by the cell(s)-of-origin can be
**Translational Relevance**

Pathways of neural development and differentiation and genomic abnormalities including IDH mutations and 1p19q codeletion are important determinants of glioma pathogenesis and progression. We have recently defined morphology-independent EM and PM glioma subtypes that show transcriptomic resemblance to neural stem cells and oligodendrocyte progenitor cells, respectively. They also exhibit distinctions in genomic alterations, prognosis, and ages at diagnosis. Here, we have designed a TaqMan low-density RT-PCR array and trained a support vector machine-based prediction model for individual diagnosis of adult diffuse gliomas of WHO grades II to IV into the EM or PM subtype. Complemented with low-coverage whole genome sequencing and detection of IDH mutations, adult gliomas are individually diagnosed as being of the EM or PM subtypes, resembling distinct neural lineages, and carrying a relatively quiescent or actively evolving genome. We believe that application of this diagnostic platform may improve risk stratification and treatment decision for glioma.

**Materials and Methods**

Samples and clinical data

All research performed was approved by the institutional review board at Tiantan Hospital in Beijing and by the regional ethics committee in Gothenburg, and was in accordance with the principles expressed at the declaration at Helsinki. Written consent was obtained from each patient prior to specimen collection. The composition of the morphologically diagnosed subgroups in the training (52 gliomas and 7 nontumor brain tissues) were reported previously (26), the details on the validation (31 gliomas and 4 nontumor brain tissues) and test (121 gliomas) sets are presented in Supplementary Table S1. Samples from Beijing were diagnosed by 2 neuropathologists at the Department of Pathology, Beijing Tiantan Hospital; samples from Gothenburg were diagnosed by one neuropathologist at Department of Pathology, Sahlgrenska University Hospital. Samples from Beijing or Gothenburg were processed locally for nucleic acid extraction, reverse transcription, and TLD gene expression profiling. The data were analyzed in Beijing. For each patient, the following clinical data were collected: morphological diagnosis, gender, age, and OS. The median follow-up was 17.7 months (range: 1.1–44.1 months).

**RT-PCR TLDA design and training of SVM prediction model**

We set out to develop a RT-PCR–based glioma EM/PM subtyping assay. For each of the EM/PM subtypes, diagnostic classifier genes should be expressed at higher level, show high expression variability across tumor subtypes, and be relatively invariantly expressed within a given subtype. We calculated the coefficient of variation (CV) for the expression levels of all EM and PM genes across the EM/PM subtypes using transcriptome data from 521 adult diffuse gliomas of WHO grades II to IV and as controls 21 epileptic brain tissues of the REMBRANDT data set (19, 27). Through this step, we selected 27 EM and 32 PM genes with high expression CV as candidate diagnostic classifiers. Similarly, we also included 9 members of the FCGR2I coexpression module that are enriched in samples with the EM**^hi^/PM**^low^ feature. Fifteen highly connected hub genes identified in whole transcriptome weighted gene coexpression network analysis were also...
considered as candidate diagnostic markers. Together with 12 candidate housekeeping genes with relatively low expression, CV and 18S rRNA as the quality control, we constructed a 96-gene TLDA in 384-well format. The TaqMan probe sets and primers of the 96 transcripts were chosen from an online database (Thermo Fisher Scientific) and preloaded on the 384-well cards.

The expression profiles of candidate classifiers and reference genes were measured in 3 epileptic brain tissues and 37 glioma samples derived from a cohort of 272 gliomas previously characterized by mRNA sequencing (GSE48865; ref. 26). Their EM/PM subtypes were determined by non-negative matrix factorization (NMF; ref. 28). Only samples with positive silhouette values, namely the core samples, were included in subsequent analyses. Candidate classifier genes with low expression or false amplification were excluded. On the basis of the extent of differential expression across the EM/PM subtypes, the classifier genes were reduced to 44. Three reference genes, ACTB, GAPDH, and UBC showed the most invariable expression pattern across the glioma samples tested (Supplementary Table S2). Together with 18S rRNA, we designed a new 48-gene array in 384-well format, which enabled a simultaneous analysis of 8 samples.

The 40 aforementioned samples together with another 16 samples from GSE48865 were tested using the 48-gene TLDA. Based on the resulting expression profiles of the 44 classifiers, we trained an EM/PM subtype prediction model using a linear SVM model. The EM/PM subtyping array and the SVM prediction model were subsequently applied in a validation cohort consisting of 181 newly diagnosed gliomas and 4 non-tumor brain tissues. In parallel, these samples were profiled using Affymetrix Human Clariom D whole transcriptome arrays and their EM/PM subtypes were independently assessed using NMF. Finally, the EM/PM subtyping array and the SVM prediction model were applied to a test cohort consisting of 51 gliomas from Beijing and 70 gliomas from Gothenburg.

Genomic profiling

To assess the genomic alterations of the gliomas in validation and test cohorts, we performed low coverage (2 ×) WGS to profile the landscape of somatic CNV. Further, mutations of IDH1 R132 and IDH2 R172 were evaluated using Sanger sequencing. Clinical data, in validation and test cohorts were combined to investigate whether the individually predicted EM or PM gliomas were associated with distinctions in prognosis, age at diagnosis and genomic alterations similar to the findings of the database analysis (19).

Nucleic acid extraction

Frozen tissues were transported in liquid nitrogen and were stored at −80°C until use. Fresh tissues were cut into cubes smaller than 5 mm × 5 mm × 5 mm, immersed in 2.5 mL RNAlater, transported on ice and stored at −20°C until use. Total RNA and genomic DNA were purified simultaneously from ~30 mg tissues using Allprep RNA/DNA Mini Kit (Qiagen) for RT-PCR assay and genome sequencing following the manufacturer’s instruction. 900 ng total RNA from each sample was reverse transcribed to cDNA in a 20 μL reaction system with high-capacity RNA to DNA Kit (Life Technology). The integrity of the cDNA was assessed by measuring the expression levels of the control genes ACTB and GAPDH. If both controls were properly amplified, the samples were assessed with the TLDA assay.

TaqMan gene expression profiling

Two microliters of cDNA template was diluted to 50 μL using RNase-free water and subsequently mixed with 50 μL 2× TaqMan Universal PCR Master mix with UNG (Life Technology). This reaction mix was then loaded to each slot of the TLDA card. The card was centrifuged twice at 1,500 rpm for 1 minute to transfer 1 μL of the mix to each reaction cell, sealed, and subsequently measured using ABI 7900HT fast real-time PCR system or Viia 7 Flex Real-Time PCR system following the universal cycling conditions: 2 minutes at 50°C for UNG incubation and 10 minutes at 95°C for polymerase activation. Forty cycles (15 seconds at 95°C for denaturation and 1 minute at 60°C for annealing and synthesis) were run for amplification.

Affymetrix human Clarion D array analysis

cDNA was synthesized using one-cycle target labeling and control reagents (Affymetrix), and cRNA was generated using a GeneChip WT Labeling Kit (Affymetrix). cRNA was fragmented and hybridized to Affymetrix human Clarion Darrays (Affymetrix). GeneChips were washed and stained in the Affymetrix Fluidics Station 450. All arrays were scanned using Affymetrix GeneChip Command Console (AGCC) installed in GeneChip Scanner 3000 7G. The data were analyzed with Robust Multichip Analysis (RMA) algorithm using Affymetrix default analysis settings and global scaling as normalization method. Values presented are log2 RMA signal intensity.

Low-coverage WGS

Genomic DNA was randomly fragmented to segments of 350 to 500 bp by Covaris sonication, end- repaired, A-tailed, adaptor ligated and PCR amplified using TrueSeq DNA LT Sample Prep Kit. The DNA libraries were then enriched, purified, and normalized using Qubit 4 Fluorometer. 150-bp paired-end sequencing was performed on the illumina HiSeq 4000 system. The average resulting genome coverage was about 2.1×.

IDH1/2 mutation detection

Mutation status of IDH1(R132) and IDH2(R172) were determined by Sanger sequencing using the following primers:

IDH1:

Forward primer: AACATGCAAATACACATTATIGCCAAC;
Reverse primer: GAATGCAAATACGACATAGC

IDH2:

Forward primer: AAAAAATGTTGGCTTGTACCTGC;
Reverse primer: CAAAAACATCCACCAGCTAGTC

Data analysis

Selection of reference genes and data normalization. The TLDAs were run on the 7900HT fast real-time PCR system or Viia 7 Real-Time PCR System, the resulting sds files or eds files were analyzed using Expression Suite Software v1.0.4 (Thermo Fisher Scientific). The cycle threshold (the CT value) for each gene was automatically determined. Relative expression levels were calculated as $2^{-\Delta \Delta CT}$ normalized against the geometric mean of the reference genes. The CT values for each sample were rescaled using quantile normalization before further analysis.

www.aacrjournals.org Clin Cancer Res; 2019
To search for reference genes with relatively invariable expression across glioma samples and brain tissues, we first constructed an array including 13 candidate housekeeping genes and tested the performance of candidate reference genes using 37 glioma samples and 3 epileptic brain tissues. The geNORM algorithm (29) implemented in Data Assistant v3.1 software (Thermo Fisher Scientific) was used to analyze their expression and ACTB, GAPDH, and UBC were selected as reference genes due to high expression levels and invariant expression patterns across the samples. The expression levels of the classifier genes were normalized against the geometric average CT values of the 3 reference genes.

Training of classification model and subtype prediction

First designed a 2-step process to identify 44 classifier genes for the EM/PM glioma molecular subtypes. Based on our previous characterization of the REMBRANDT dataset (27), we selected 27 EM genes, 32 PM genes, 15 hub genes in the WGCNA network, 9 genes coexpressed with FGFR11, 12 candidate reference genes, and 18S ribosomal RNA to construct a RT-PCR array containing 96 assays. This array was tested with a training cohort of 40 core samples (11 EM and 22 PM gliomas, 4 gliomas with the EM<sup>low</sup>PM<sup>low</sup> phenotype and 3 epileptic brain tissues) from a RNA-seq based dataset, GSE48865 (26). Their EM/PM subtype label was obtained using NMF. Candidate classifier genes that failed to be amplified (6 genes) or did not exhibit EM/PM subtype-specific expression pattern (33 genes) were excluded. Candidate classifier genes with higher extent of variable expression within the EM/PM subtype were also excluded. Complemented with their biological relevance, 12 classifier genes for the EM gliomas, 28 classifiers for the PM gliomas, and 4 classifier genes for the gliomas with the EM<sup>low</sup>PM<sup>low</sup> phenotype were selected (Fig. 1; Supplementary Table S2). These classifiers, together with 3 invariably expressed reference genes (ACTB, GAPDH, and UBC) and 18S ribosomal RNA, constituted the TLDA for the EM/PM subtyping of glioma.

Notably, genes specific for the mature myelinating oligodendrocyte (ENPP2, MAL, and MOG) or neuron (GABRA1; ref. 33) were selected as the classifier genes for the nontumor brain tissues and the glioma samples with an EM<sup>low</sup>PM<sup>low</sup> phenotype. Pearson correlation coefficient analyses showed that the selected classifier genes for the EM and PM gliomas were tightly coexpressed on the RT-PCR platform (Fig. 2; Supplementary Fig. S1).

Training of an SVM model for assigning individual gliomas into the EM/PM subtype

The 59 samples of the training set containing 52 core samples from GSE48865 (13 EM and 34 PM gliomas, 5 samples with the EM<sup>low</sup>PM<sup>low</sup> profile; ref. 26) and 7 nontumor brain samples were profiled using the 48-gene EM/PM TLDA. Based on the normalized CT values of the classifier genes, we trained an EM/PM subtype prediction model using an SVM algorithm. PCA plot showed distinct expression profiles for samples with the EM, PM, and EM<sup>low</sup>PM<sup>low</sup> phenotype, respectively (Fig. 3A). Hundred rounds of cross-validation analysis showed that the median accuracy, precision, and recall rate of the prediction model were 83%, 85%, and 82%, respectively (Fig. 3B).

Verification of the TLDA and SVM prediction model in independent validation and test set

To evaluate whether the SVM prediction model would encounter overfitting or underfitting in new datasets, we built an independent validation cohort containing 35 samples (31 gliomas and 4 nontumor brain tissues). We profiled their transcription using Affymetrix Clariom D array and identified 9 EM, 19 PM, and 7 EM<sup>low</sup>PM<sup>low</sup> samples using NMF. In parallel, samples were also profiled using the TLDA EM/PM subtyping array. Using the SVM prediction model, EM/PM subtype was predicted for individual samples. As shown in the PCA plot (Fig. 4A), individually predicted EM, PM, or EM<sup>low</sup>PM<sup>low</sup> samples exhibited distinct expression pattern of the classifier genes. We compared the TLDA array and SVM prediction model-based subtyping results to that from Clariom D assay and NMF prediction model. Only 1 sample of the PM subtype was wrongly predicted as the EM subtype, and 1 EM<sup>low</sup>PM<sup>low</sup> sample wrongly as the PM subtype. EM/PM subtypes of the other 33 samples were correctly predicted. The AUIC in ROC analysis for the EM, PM, and EM<sup>low</sup>PM<sup>low</sup> subtypes were at 0.86,
0.85, and 0.91, respectively, confirming a high accuracy of the EM/PM subtype prediction (Fig. 4B). Next, we assessed the performance of the EM/PM subtyping array and SVM prediction model in a test cohort consisting of 51 samples from Beijing and 70 samples from Gothenburg; 39 EM, 64 PM, and 18 EMlowPMlow samples were predicted. The subtype-specific expression pattern of the classifier genes was evident in the PCA (Fig. 4C), again suggesting the robustness of the TLDA assay and SVM prediction model.

Distinct patterns of CNVs and IDH mutations between the individually predicted EM and PM gliomas

Our previous clustering analyses using data from TCGA and the REMBRANDT dataset showed that the EM and PM gliomas harbor distinct sets of genomic alterations (19). To assess whether the individually predicted EM and PM gliomas also harbored the same set of genomic alterations as in the database analysis, we performed low-coverage (~2×) WGS for 143 glioma samples and analyzed their CNV profile. Although chromosome 7 gain (43/46, 93.5%), chromosome 10 loss (43/46, 93.5%), amplification of EGFR, deletion of PTEN and CDKN2A/B were highly enriched in the EM gliomas, 1p19q codeletion occurred exclusively in the PM gliomas (32/82, 39%), 4q loss in 29/82 of PM gliomas, and 3/46 of EM gliomas. EGFR was amplified in half of the EM gliomas (Fig. 5A and B). We also assessed whether the expression data derived from the TLDA EM/PM subtyping assay could reflect EGFR amplification. Among the EM gliomas, further elevated EGFR expression was not observed in those samples with a trisomy of chromosome 7. However, elevated EGFR expression was found in those EM samples with focal amplifications at chromosome 7p11.2 (Fig. 5C); a similar pattern was also found for SEC61G (Fig. 5C), which is located in the same amplification as EGFR. In parallel, Sanger sequencing showed that the IDH1 R132 and IDH2 R172 mutations occurred only in the PM gliomas.
In agreement with the findings in our previous analyses of large data sets (19), these findings show that individually diagnosed EM and PM gliomas harbor distinct genomic alterations. Under the same conditions, gains and losses of large CNVs were not detected in 12 of the 15 samples with the EM\textsubscript{low}PM\textsubscript{low} phenotype because of low tumor purity (Fig. 5B; Supplementary Fig. S2), tumor purity was below 12% in 11 of the 15 samples analyzed (Supplementary Table S1). Similar to the nontumor brain tissues, these samples showed high expression of mature oligodendrocyte and neuron markers, indicating that samples with the EM\textsubscript{low}PM\textsubscript{low} phenotype contained a high extent of normal brain tissues (Fig. 1). Samples with the EM\textsubscript{low}PM\textsubscript{low} phenotype were thus excluded from further analysis.

Clinical relevance

Individually predicted EM/PM subtypes were independent of the morphological diagnosis. Among the 78 GBM samples, 40 showed the EM and 29 the PM expression signature. Among the 16 lower-grade (grade II or III) astrocytomas, 1 showed the EM and 10 showed the PM expression, whereas the 7 grade III oligodendrogliomas and 17 grade II oligodendrogliomas were all identified as the PM glioma. Among the 18 mixed oligoastrocytomas, 3 were identified as the EM glioma and 13 as the PM glioma (Fig. 6A). Patients of the PM subtype were of younger age at diagnosis (46.1 ± 13.5 years for PM gliomas vs. 59.1 ± 12.4 years for EM gliomas, P < 0.001, t test; Fig. 6B). Patients with EM gliomas had a median OS of 15.6 months, whereas the median OS for patients with PM gliomas was not yet reached as of the publication date.
over 50% of the patients were alive at the end of this study (P < 0.001, Fig. 6B). Compared with the PM subtype, a hazard ratio of 3.55 with 95% confidence interval (CI) of 1.97 to 6.45 (P < 0.001) was found for the EM subtype.

In agreement with previous reports (13, 34), our results show that grade II to III gliomas (the lower grade gliomas, LGG) with an EM phenotype (without IDH mutation) progressed more rapidly compared with the PM gliomas of the same morphologic subtype and grades (with IDH mutation; Supplementary Fig. S3). However, the patients of GBM with a PM phenotype had shorter survival times compared with the patients of LGG with a PM phenotype, although high expression of PM genes (indicating commitment in the oligodendrocyte lineage), IDH mutation, and 1p19q codeletion were detected in both subsets of the PM gliomas. The percentage of CNV in the whole genome of the GBMs with a PM phenotype was significantly elevated compared with the LGG with a PM phenotype (Fig. 5D and E). As shown with the 3 GBMs harboring 1p19q codeletion in our dataset (Supplementary Fig. S4A) and the analysis of 100 K SNP array data for the PM gliomas in the REMBRANDT dataset (Supplementary Fig. S4B and S4C), PM gliomas with the GBM morphology frequently harbored homozygous deletion of CDKN2A/B at 9p21.3, loss of 10q, and complex karyotypes containing high numbers of subclonal CNVs and chromothripsis; they were also associated with poorer survival times (Supplementary Fig. S4D). These results together suggest that although IDH mutations and 1p19q codeletion represent the initiating genomic abnormalities in the PM glioma committed in the oligodendrocyte lineage, elevated CNV burden, as measured with low-coverage WGS, could serve as a surrogate marker of glioma progression.

Discussion

On the basis of microscopic similarities with different putative cells of origin and their presumed levels of differentiation, gliomas have been traditionally diagnosed as astrocytomas or oligodendrogliomas and further divided into grades I to IV (4, 5). The integration of IDH mutation and 1p19q codeletion has markedly improved the prognostification of glioma diagnosis (4, 5). Gliomas that harbor a mutation in IDH1 or IDH2 show distinct transcriptomic, genomic, and clinical features compared with gliomas with wild-type IDH genes (8). However, the resemblance of morphologically defined glioma subtypes, astrocytomas in particular, to neural cell lineages and differentiation stages is as yet unclear, and genomic abnormalities other than IDH mutations and 1p19q codeletion are not yet incorporated into glioma diagnosis. Our study aimed to incorporate the signatures correlating glioma subtypes to neural development lineage and the extent of glioma progression into glioma molecular diagnosis.

Compared with the genomic abnormalities-based molecular classification (9), transcriptome profiling (11, 16, 19) may capture signatures derived from both the cell(s) of origin and the genomic and epigenetic abnormalities of the tumor. The EM/PM classification scheme defines glioma molecular subtypes both by their relationship to neural development signatures as well as with respect to genomic abnormalities (19). However, translation of transcriptome profiling-based classification scheme into a diagnostic tool requires the development of an assay for robust and rapid measurement of expression profiles in a cost-effective manner and an algorithm for assigning individual samples into the molecular subtypes.

In this study, we developed a RT-PCR–based TLDA platform and an SVM prediction model for robust assignment of individual gliomas into the EM and PM subtype. In agreement with our previous database analyses (19), individually predicted EM and PM gliomas were distinct in age at diagnosis and prognosis, and exhibited unique patterns of genomic alterations. Compared with the microarray- or RNA-seq based assays, our RT-PCR TLDA assay and SVM prediction model enables EM/PM subtyping within 24 hours. The accuracy of EM and PM subtype prediction could be...
improved by enlarging the size of the training set and further refined selection of the classifier genes, particularly the classifier genes specific for the EM subtype.

Our findings show that incorporation of low-coverage WGS enabled the detection of a broad spectrum of CNVs across the entire genome and can thereby improve the power of molecular classification. Characteristic chromosome alterations, including chromosome 7 gain and chromosome 10 loss, homozygous loss of 9p21.3, and high copy number amplification of EGFR were reaffirmed in EM gliomas. Although PM gliomas with WHO grades II and III showed relatively quiescent genomes, PM gliomas with GBM features harbored a large number of additional chromosomal abnormalities, coinciding with poor prognosis in patients with such PM gliomas. Although 1p19q codeletion was clonal, many of the additional chromosomal abnormalities were subclonal, suggesting that these gliomas may have evolved from gliomas harboring 1p19q as an initiating event. Our recent analysis using the TCGA database indicates that the extent of chromosome abnormalities and mutation burden is concordant in the vast majority of gliomas (35), the 1p19q codeleted PM gliomas with GBM features may carry an unstable genome and harbor a high mutation burden, indicating an advanced stage of tumor progression. These findings are consistent with recent reports that the presence of an IDH mutation does not always indicate a favorable prognosis (9, 11), activation of cell-cycle genes in the PM gliomas with GBM morphology may play a significant role in their prognosis (11). PM gliomas should therefore be further stratified according to the extent and patterns of the CNV, PM gliomas with extensive CNVs and unstable karyotype are likely at more advanced stage(s) of tumor progression, and show aggressive clinical behavior.

Diagnosing gliomas into the EM/PM subtype may improve glioma risk stratification and treatment design. Because the EM gliomas resemble neural stem cells/developing astrocytes, and PM gliomas resemble OPC (19), dys-regulated proliferation, and differentiation mechanisms of neural stem cells/developing astrocytes and OPC could be respectively involved in the pathogenesis of the EM and PM gliomas. Furthermore, as the classifier genes are

---

**Figure 5.**
Distinct genomic alterations in individually identified EM or PM gliomas. A, Patterns of CNVs for representative EM and PM samples. Concomitant gain of chromosome 7 and loss of chromosome 10 were enriched in the EM gliomas, whereas codeletion of 1p/19q occurred exclusively in the PM gliomas. Representative EM samples with focal amplification of EGFR and deletion of CDKN2A/B, PM samples with or without 1p19q co-deletion are shown. B, Summary of the occurrence of the canonical CNVs and IDH mutations in the EM/PM glioma samples tested. C, Elevated EGFR expression in gliomas with high copy number gain. EGFR expression was found significantly elevated in gliomas with high focal copy number gain but not in gliomas with trisomy of chromosome 7. A similar pattern was found for SEC61G located in the same amplification. D, CNV landscape of the PM gliomas with the lower grade (grade II and III) or GBM morphology. PM gliomas with or without 1p19q co-deletion were presented in the top and bottom panels, respectively. E, Elevated CNV burden in the PM glioma with the GBM morphology compared with LGG gliomas. PM gliomas with or without 1p19q co-deletion are shown separately.
implicated in both glial development and glioma pathogenesis, diagnosing gliomas into the EM/PM subtype may facilitate the elucidation of glioma etiology and the assessment of new biomarkers and treatment regimens.

In summary, EM/PM subtyping, integrated with Sanger sequencing of IDH1 mutations and low-coverage WGS, constitutes a valuable platform to capture the glioma transcriptomic resemblance to neural development, IDH1/2 mutations, and canonical CNVs, and the extent of tumor progression in a resource-, time-, and cost-effective manner. Risk stratification and development of subtype-specific therapy for individual glioma patients could be benefited from an implementation of this platform.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors’ Contributions
Conception and design: J. Li, T. Jiang, X. Fan
Development of methodology: J. Li, Y. Sun
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): J. Li, Y. Xue, A. Wenger, Z. Wang, C. Zhang, Y. Zang, B.F. Rydenhag, B. Rydenhag, A.S. Jakola, H. Carén
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): J. Li, Y. Xue, Y. Sun, Z. Wang, X. Fan
Writing, review, and/or revision of the manuscript: J. Li, A.S. Jakola, T. Jiang, H. Carén, X. Fan
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): J. Li, A. Wenger, Z. Wang, C. Zhang
Study supervision: T. Jiang, X. Fan

Acknowledgments
This study was supported by the National Natural Science Foundation of China (Grant Nos. 81773015, 81472348, and 31711530156).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received January 25, 2019; revised June 24, 2019; accepted August 23, 2019; published first xx xx, xxxx.

References
glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell 2010;17:98–110.


AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

Q1: Page: 1: Author: Per journal style, genes, alleles, loci, and oncogenes are italicized; proteins are roman. Please check throughout to see that the words are styled correctly. AACR journals have developed explicit instructions about reporting results from experiments involving the use of animal models as well as the use of approved gene and protein nomenclature at their first mention in the manuscript. Please review the instructions at http://aacrjournals.org/content/authors/editorial-policies#genomen to ensure that your article is in compliance. If your article is not in compliance, please make the appropriate changes in your proof.

Q2: Page: 1: Author: Please verify the drug names and their dosages used in the article.

Q3: Page: 1: Author: Please verify the affiliations and their corresponding author links.

Q4: Page: 5: Author: Please confirm quality/labeling of all images included within this article. Thank you.

Q5: Page: 9: AU;/PE: The conflict-of-interest disclosure statement that appears in the proof incorporates the information from forms completed and signed off on by each individual author. No factual changes can be made to disclosure information at the proof stage. However, typographical errors or misspelling of author names should be noted on the proof and will be corrected before publication. Please note if any such errors need to be corrected. Is the disclosure statement correct?

Q6: Page: 10: Author: Please provide volume number and age range for ref. 32.

AU: Below is a summary of the name segmentation for the authors according to our records. The First Name and the Surname data will be provided to PubMed when the article is indexed for searching. Please check each name carefully and verify that the First Name and Surname are correct. If a name is not segmented correctly, please write the correct First Name and Surname on this page and return it with your proofs. If no changes are made to this list, we will assume that the names are segmented correctly, and the names will be indexed as is by PubMed and other indexing services.
<table>
<thead>
<tr>
<th>First Name</th>
<th>Surname</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jiuyi</td>
<td>Li</td>
</tr>
<tr>
<td>Yang</td>
<td>Xue</td>
</tr>
<tr>
<td>Anna</td>
<td>Wenger</td>
</tr>
<tr>
<td>Yingyu</td>
<td>Sun</td>
</tr>
<tr>
<td>Zheng</td>
<td>Wang</td>
</tr>
<tr>
<td>Chuanbao</td>
<td>Zhang</td>
</tr>
<tr>
<td>Yunqiu</td>
<td>Zhang</td>
</tr>
<tr>
<td>Boglarka</td>
<td>Fekete Rydenhag</td>
</tr>
<tr>
<td>Bertil</td>
<td>Rydenhag</td>
</tr>
<tr>
<td>Asgeir Store</td>
<td>Jakola</td>
</tr>
<tr>
<td>Tao</td>
<td>Jiang</td>
</tr>
<tr>
<td>Helena</td>
<td>Carén</td>
</tr>
<tr>
<td>Xiaolong</td>
<td>Fan</td>
</tr>
</tbody>
</table>