Translational Oncology
www.transonc.com

Volume 13 Number 2 February 2020 pp. 233–240 233

Expression of Alpha-type Platelet-derived Growth Factor Receptor–influenced Genes Predicts Clinical Outcome in Glioma

Eun-A Ko, Haeyeong Lee, Kenton M. Sanders, Sang Don Koh and Tong Zhou

Department of Physiology and Cell Biology, University of Nevada, Reno School of Medicine, Reno, NV 89557, USA

Abstract

BACKGROUND: Alpha-type platelet-derived growth factor receptor (PDGFRα) is a cell surface tyrosine kinase receptor for members of the platelet-derived growth factor family. PDGFRα plays an important role in the regulation of several biological processes and contributes to the pathophysiology of a broad range of human cancers, including glioma. Here, we hypothesize that the genes directly or indirectly influenced by PDGFRα might be useful for prognosis in glioma. *METHODS:* By comparing the genome-wide gene expression pattern between PDGFRα$^+$ and PDGFRα$^-$ cells from human oligodendrocyte progenitor, we defined the genes potentially influenced by PDGFRα. *RESULTS:* The PDGFRα-influenced genes are strongly associated with cancer-related pathways. We subsequently developed a prognostic gene signature derived from the PDGFRα-influenced genes. This gene signature is able to predict clinical outcome of glioma. This signature is also independent of traditional prognostic factors of glioma. Resampling tests indicate that the prognostic power of this gene signature outperforms random gene sets selected from human genome. More importantly, this signature is superior to the random gene signatures selected from glioma related genes. *CONCLUSIONS:* Despite the absence of clear elucidation of molecular mechanisms, this study suggests the vital role of PDGFRα in carcinogenesis. Furthermore, the PDGFRα-based gene signature provides a promising prognostic tool for glioma and validates PDGFRα as a novel and effective therapeutic target in human cancers.

Translational Oncology (2020) 13, 233–240

Introduction

Alpha-type platelet-derived growth factor receptor (PDGFRα) is a cell surface tyrosine kinase receptor for members of the platelet-derived growth factor family, which is encoded by the gene *PDGFRA*. Platelet-derived growth factor receptors (PDGFRs) are a family of catalytic receptors that play important roles in the regulation of several biological processes including embryonic development [1], angiogenesis [1,2], and cell proliferation and migration [1,3]. PDGFRs contribute to the pathophysiology of a broad range of human diseases, including cancers. For example, PDGFRα was found to be expressed more frequently by tumor-associated stromal cells in lung cancer [4]; overexpression of PDGFRα is associated with tumor progression in breast cancer [5]; significant correlation was identified between PDGFRα expression and lymph node metastasis in colon cancer [6]. Particularly, PDGFR signaling has been proposed to play a key role in malignant brain tumor of glial origin, such as glioma [1]. Glioma is the most common primary tumors in the central nervous system. Glioma can be categorized based on pathological evaluation of the tumor. According to the World Health Organization (WHO) grading system, gliomas are classified into four grades (I, II, III, and IV) with the best prognosis in grade I and the most malignancy in grade IV. WHO grade I and II tumors are termed low-grade glioma. The term high-grade glioma refers to tumors that are classified as

Address all correspondence to: Tong Zhou or Sang Don Koh, Department of Physiology and Cell Biology, University of Nevada, Reno School of Medicine, Reno, Nevada 89557, USA. E-mail: skoh@med.unr.edu or tongz@med.unr.edu.
Received 5 September 2019; Revised 4 October 2019; Accepted 7 October 2019

© 2019 The Authors. Published by Elsevier Inc. on behalf of Neoplasia Press, Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
1936-5233/19
https://doi.org/10.1016/j.tranon.2019.10.002

WHO grade III, such as anaplastic astrocytoma, anaplastic oligodendroglioma, anaplastic oligoastrocytoma, and anaplastic ependymoma, and WHO grade IV, such as glioblastoma. Amplification of PDGFRα has been observed in both low-grade [7] and high-grade [8,9] gliomas. In addition, overexpression of the gene *PDGFRA* in glioma at the time of the first diagnosis was found to be associated with poor overall survival [10]. Inhibition of PDGFRs has been shown to slow down glioma cell growth in experimental models [1]. Therefore, inhibition of PDGFR signaling has become one of the targeted therapeutic strategies for glioma [11].

Although the pathological function of PDGFRs in gliomas remains controversial, coexpression module based on the gene *PDGFRA* has been developed to enable the molecular classification of glioma for clinical diagnosis [12]. It is also reasonable to hypothesize that the genes directly or indirectly influenced by PDGFRα might be useful for prognostic purpose in glioma. Here, we utilized high-throughput gene expression data to identify the genes potentially influenced by PDGFRα in glioma. We compared the genome-wide gene expression pattern between PDGFRα$^+$ and PDGFRα$^-$ cells from human oligodendrocyte progenitor [13]. The genes deregulated in PDGFRα$^+$ cells were considered as PDGFRα influenced genes. Gene ontology analysis indicates that the PDGFRα-influenced genes are strongly associated with cancer-related pathways. We subsequently developed a prognostic gene signature derived from the PDGFRα influenced genes. This gene signature is able to predict clinical outcome in two independent glioma cohorts. This signature is also independent of traditional prognostic factors in glioma. Our study suggests that the PDGFRα-influenced genes potentially serve as biomarkers and therapeutic targets in clinical and pharmacological contexts, respectively.

Materials and Methods

Gene Expression Data Sets

We obtained the gene expression data for both PDGFRα$^+$ and PDGFRα$^-$ oligodendrocyte progenitor cells from the Gene Expression Omnibus (GEO) database [14] (GEO accession: GSE29368), which was based on the Affymetrix Human Genome U133 Plus 2.0 Array [13]. In the original study, the PDGFRα$^+$ cells were defined based on expression of the PDGFRα epitope CD140a, which were sorted from the fetal human forebrain using FACS [13]. The gene expression data from University Hospital of Coimbra (UHC), Portugal (GEO accession: GSE43289) [15] and Henry Ford Hospital (HFH), USA (GEO accession: GSE4290) [16] were used to identify the glioma related genes, in which the gene expression level was significantly correlated with WHO glioma grade. Both the UHC and HFH data sets were based on the Affymetrix Human Genome U133 Plus 2.0 Array. The gene expression data from Shanghai Changzheng Hospital (SCH), China (GEO accession: GSE19728) [17], was used to validate the relationship between glioma grade and our PDGFRα-influenced gene signature (PIGS), which was also based on the Affymetrix Human Genome U133 Plus 2.0 Array. To validate the prognostic power of the gene signature, we collected two independent cohorts with available clinical outcome information: the EORTC (European Organisation for Research and Treatment of Cancer) cohort (GEO accession: GSE43107) [18] and the UCLA (University of California at Los Angeles) cohort (GEO accession: GSE4412) [19], which were based on Affymetrix Human Exon 1.0 ST Array and Affymetrix Human Genome U133A/B Array, respectively. Figure 1 indicates the working scheme of how all these transcriptomic data sets were implicated in this study.

Microarray Data Processing

We applied the robust multiarray average (RMA) function in the "affy" package of bioconductor [20] to summarize the expression level of each probeset for the data sets generated by Affymetrix Human Genome U133 Plus 2.0 Array and Affymetrix Human Genome U133A/B Array. For the data set based on Affymetrix Human Exon 1.0 ST Array, the gene expression values were summarized using the Affymetrix Power Tools Version 1.15.0. The function "mas5calls" in the "affy" package [21] was used to compute probeset present/absent call for the progenitor cell, UHC, and HFH data sets. For the progenitor cell data set, only the probesets that were present in all replicates of at least one group were used for further analysis. For the

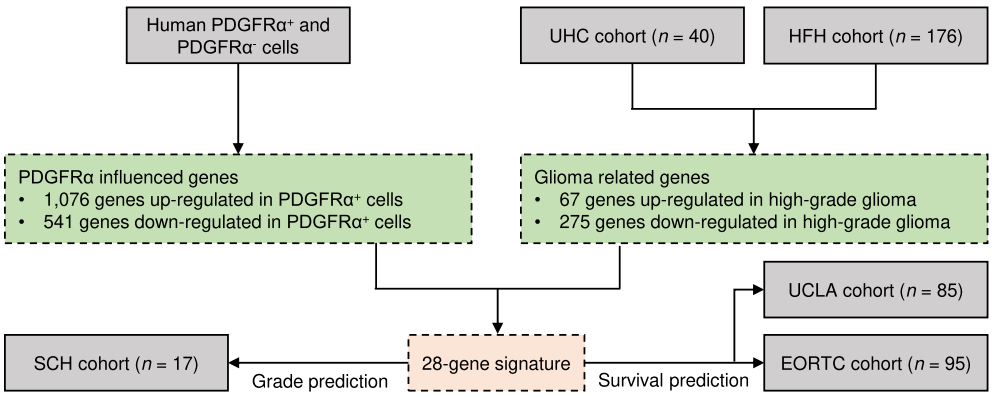

Figure 1. The working scheme of the study. We first compared the transcriptomic pattern between PDGFRα$^+$ and PDGFRα$^-$ cells from human oligodendrocyte progenitor, which yielded a list of PDGFRα-influenced genes. Next, the UHC and HFH cohorts were analyzed to infer the glioma-related genes, which were either positively or negatively correlated with glioma grade. We developed a gene signature based on the intersection between the upregulated genes in high-grade glioma and the overexpressed genes in PDGFRα$^+$ cells and the intersection between the downregulated genes in high-grade glioma and the underexpressed genes in PDGFRα$^+$ cells. We validated the power of this gene signature in grade prediction in the SCH cohort. We further validated the predictive power of the gene signature in differentiating glioma patients with distinct survivals in the EORTC and UCLA cohorts.

Table 1. The 28 genes of PIGS.

Gene symbol	Gene title	Weight
ANXA1	Annexin A1	1
ANXA2	Annexin A2	1
CD63	CD63 molecule	1
CHI3L1	Chitinase 3-like 1 (cartilage glycoprotein-39)	1
DTX3L	Deltex 3-like (*Drosophila*)	1
FNDC3B	Fibronectin type III domain containing 3B	1
HRH1	Histamine receptor H1	1
ITGA7	Integrin, alpha 7	1
MRC2	Mannose receptor, C type 2	1
PPIC	Peptidylprolyl isomerase C (cyclophilin C)	1
ATP6V1G2	ATPase, H+ transporting, lysosomal 13 kDa, V1 subunit G2	−1
ATRNL1	Attractin like 1	−1
C20orf194	Chromosome 20 open reading frame 194	−1
CAMTA1	Calmodulin-binding transcription activator 1	−1
CAND1	Cullin-associated and neddylation-dissociated 1	−1
CCDC90A	Coiled-coil domain containing 90A	−1
CLASP2	Cytoplasmic linker associated protein 2	−1
CPEB3	Cytoplasmic polyadenylation element binding protein 3	−1
ERBB4	v-erb-a erythroblastic leukemia viral oncogene homolog 4 (avian)	−1
GTF2H5	General transcription factor IIH, polypeptide 5	−1
NTN4	Netrin 4	−1
PLEKHM3	Pleckstrin homology domain containing, family M, member 3	−1
POU6F1	POU class 6 homeobox 1	−1
PTPN4	Protein tyrosine phosphatase, nonreceptor type 4 (megakaryocyte)	−1
RALGPS1	Ral GEF with PH domain and SH3 binding motif 1	−1
SGIP1	SH3-domain GRB2-like (endophilin) interacting protein 1	−1
TMEM57	Transmembrane protein 57	−1
ZNF33A	Zinc finger protein 33A	−1

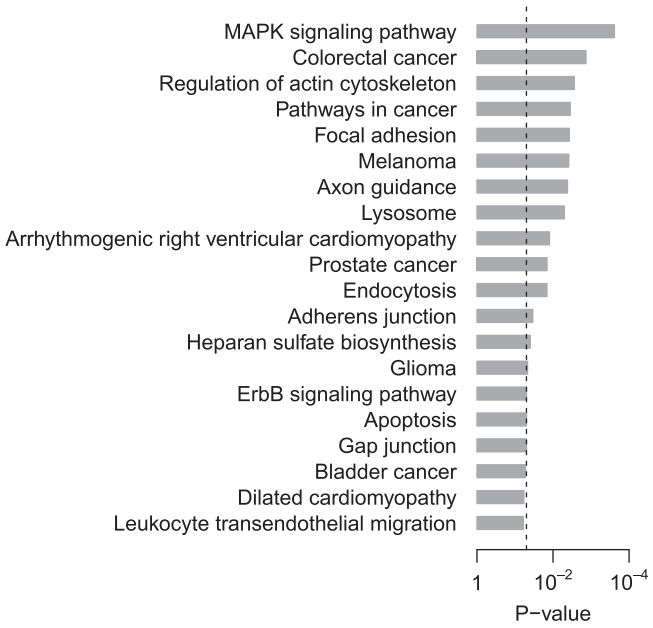

Figure 2. The top 20 KEGG pathways associated with the PDGFRα-influenced genes. The *P*-values were calculated by Fisher's exact test. The vertical dash line denotes the significance level of $\alpha = 0.05$.

UHC and HFH data sets, only the probesets that were present in at least two-thirds of the samples were retained. We limited our analysis to the probesets with unique annotations. The genes on chromosomes X and Y were removed. For the SCH, EORTC, and UCLA data sets, we used the geometric mean of expression values of all probesets mapping to a gene if the gene was encoded by multiple probesets. Significance analysis of microarrays [22], implemented in the samr library of the R Statistical Package, was used to compare $\log_2$-transformed gene expression levels between PDGFRα$^+$ and PDGFRα$^-$ progenitor cells. False discovery rate (FDR) was controlled using the q-value method [23,24]. Transcripts with FDR < 5% and FC > 2 were deemed differentially expressed.

Risk Score

We used a published scoring system to compute the risk score for each patient [25−27]. The risk score is a linear combination of gene expression values. The formula is shown below:

$$S = \sum_{i=1}^{n} W_i(e_i - \mu_i) / \tau_i$$

Here, S is the risk score of the patient; n is the number of genes in PIGS; W_i denotes the weight of gene i (as shown in Table 1); e_i denotes the expression level of gene i; and μ_i and τ_i are the mean and standard deviation of the expression values for gene i across all subjects, respectively. A higher risk score implies a poorer clinical outcome.

Results

Genes Influenced by PDGFRα

We compared the gene expression pattern between PDGFRα$^+$ and PDGFRα$^-$ cells from human oligodendrocyte progenitor. One PDGFRα$^+$ cells microarray data set containing gene expression information for both PDGFRα$^+$ and PDGFRα$^-$- cells was collected from the GEO database [14] (GEO accession: GSE29368) [13], which was based on the Affymetrix Human Genome U133 Plus 2.0 Array. At the specified significance level of FDR < 5% and fold change (FC) > 2 (see Materials and Methods for details), 1445 probesets encoding 1076 genes were found to be overexpressed in PDGFRα$^+$ cells (Supplementary Table S1 and Figure S1), while 741 probesets encoding 541 genes were underexpressed in PDGFRα$^+$ cells (Supplementary Table S2 and Figure S1). We considered these deregulated genes as PDGFRα-influenced genes. We next searched the enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) [28] pathways among the PDGFRα-influenced genes. We found that the PDGFRα-influenced genes are significantly associated with several cancer-related KEGG terms, such as "colorectal cancer," "pathways in cancer," "melanoma," "prostate cancer," and "glioma" (Figure 2). These findings suggest that the PDGFRα-influenced genes are involved in human cancer pathogenesis.

To determine how deep the PDGFRα-influenced genes are involved in glioma, we explored the genes that are associated with the severity of glioma. We obtained two gene expression data sets of glioma patients from the GEO database: the UHC cohort (GEO accession: GSE43289) [15] and the HFH cohort (GEO accession: GSE4290) [16]. Both data sets were based on the Affymetrix Human Genome U133 Plus 2.0 Array. There were 40 subjects in the UHC cohort, which included 3 grade I, 3 grade II, 6 grade III, and 28 grade IV patients. For the HFH cohort, there were in total 176 samples including 23 nontumor, 45 grade II, 31 grade III, and 77 grade IV subjects. Spearman's rank correlation test was used to identify the glioma related genes, in which the gene expression level was significantly correlated with glioma grade. Only the genes differentially expressed with glioma grade (Spearman's rank correlation test: adjusted $P < 0.005$ after Benjamini−Hochberg correction) in the both cohorts were retained for further analysis. In total, we found that

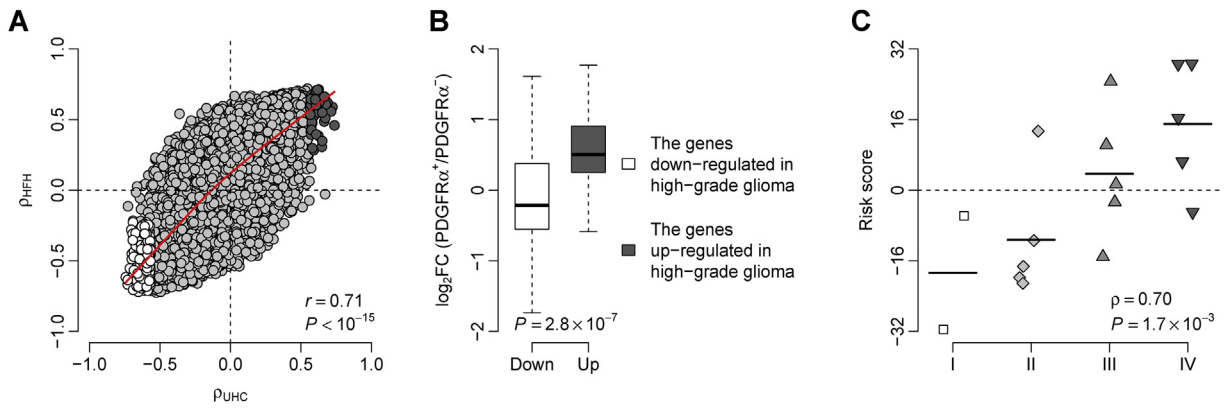

Figure 3. Relationship between glioma grade and the PDGFRα-influenced genes. (A) Glioma-related genes. Each dot stands for a gene. Spearman's rank correlation test was used to calculate the correlation between glioma grade and gene expression for the both UHC and HFH cohorts. The correlation coefficient of the UHC cohort (ρ_{UHC}) is significantly correlated with the correlation coefficient of the HFH cohort (ρ_{HFH}) (Pearson correlation test: $r = 0.71$ and $P < 10^{-10}$). Only the genes differentially expressed with glioma grade in the both cohorts were considered as glioma-related genes. The dark gray filled dots denote the genes upregulated in high-grade glioma, while the white filled dots represent the genes downregulated in high-grade glioma. The red line shows lowess-smoothed data. (B) Comparison between the genes upregulated and downregulated in high-grade glioma. Y-axis denotes the $\log_2$-transformed fold change (FC) in gene expression between PDGFRα$^+$ and PDGFRα$^-$ cells. Log_2FC is significantly higher for the genes upregulated in high-grade glioma (*t*-test: $P = 2.8 \times 10^{-7}$), which suggests that the genes upregulated in high-grade glioma are more likely to be overexpressed in PDGFRα$^+$ cells. (C) Positive correlation between glioma grade and PIGS-based risk score in the SCH cohort. Each dot stands for a patient from the SCH cohort. The horizontal line indicates the mean of each category. Spearman's rank correlation test: $\rho = 0.70$ and $P = 1.7 \times 10^{-3}$.

81 probesets encoding 67 genes were upregulated in high-grade glioma (Figure 3A and Supplementary Table S3) while 340 probesets encoding 275 genes were downregulated in high-grade glioma (Figure 3A and Supplementary Table S4). Interestingly, the genes that were upregulated in high-grade glioma were more likely to be overexpressed in PDGFRα$^+$ cells, compared with the downregulated genes (Figure 3B). Among the upregulated genes in high-grade glioma, 10 genes were found to overlap with the overexpressed genes in PDGFRα$^+$ cells, which is statistically significant (cumulative hypergeometric test: $P = 5.0 \times 10^{-3}$). For the downregulated genes in high-grade glioma, 18 genes significantly overlapped with the underexpressed genes in PDGFRα$^+$ cells (cumulative hypergeometric test: $P = 2.7 \times 10^{-2}$). All these results suggest that the PDGFRα-influenced genes are significantly enriched among the glioma associated gene set.

PDGFRα-influenced Gene Signature

Above, we identified 10 genes within the intersection between the upregulated genes in high-grade glioma and the overexpressed genes in PDGFRα$^+$ cells. In addition, 18 downregulated genes in high-grade glioma were found to overlap with the underexpressed genes in PDGFRα$^+$ cells. We designated all these 28 genes as PIGS (Table 1 and Supplementary Figures S2 and S3). Based on the PIGS, we constructed a scoring system to assign each subject a risk score, representing a linear combination of the PIGS expression values weighted by the direction of differential expression: 1 for the upregulated and −1 for the downregulated genes in PDGFRα$^+$ cells (see Materials and Methods for details). A higher risk score suggests a poorer clinical outcome.

We first tested whether the PIGS based risk score was able to predict glioma grade. For this purpose, we obtained an independent gene expression data sets, the SCH cohort, from the GEO database including 17 glioma patients (GEO accession: GSE19728) [17]. As we expected, there was a significant positive correlation (Spearman's rank correlation test: $P = 1.7 \times 10^{-3}$) between glioma grade and PIGS-based risk score (Figure 3C).

PIGS Predicts Survival in Glioma

We next tested whether the PIGS-based risk score can be used to predict survival in glioma. From the GEO database, we downloaded two independent gene expression data sets: the EORTC cohort including 95 high-grade glioma patients (GEO accession: GSE43107) [18] and the UCLA cohort composed of 85 high-grade glioma patients (GEO accession: GSE4412) [19]. These data sets were chosen based on two criteria: (i) the large number of samples (sample size ≥ 80) and (ii) the availability of clinical outcome data. We defined PIGS positive (PIGS$^+$) patients as those having a risk score > 0, while the other patients were assigned as PIGS negative (PIGS$^-$), as the median of risk score was approximately equal to zero in each validation cohort (Supplementary Figure S4). Univariate Cox proportional hazard regression of survival indicates that the PIGS$^+$ patients have a 2.91- and 2.46-increased risk of death in the EORTC and UCLA cohorts, respectively (Table 2). Kaplan−Meier survival curves demonstrate a significant difference in survival between the PIGS$^+$ and PIGS$^-$ glioma patients in the two validation cohorts (log-rank test: $P = 2.2 \times 10^{-6}$ for the EORTC cohort and $P = 8.1 \times 10^{-4}$ for the UCLA cohort) (Figure 4A).

Table 2. Univariate Cox proportional hazards regression of survival by PIGS status.

Cohort	*N*	HR	95% CI of HR	*P*-value
EORTC	95	2.91	(1.84, 4.61)	5.5×10^{-6}
UCLA	85	2.46	(1.43, 4.23)	1.2×10^{-3}

N: patient number; HR: hazard ratio; CI: confidence interval.

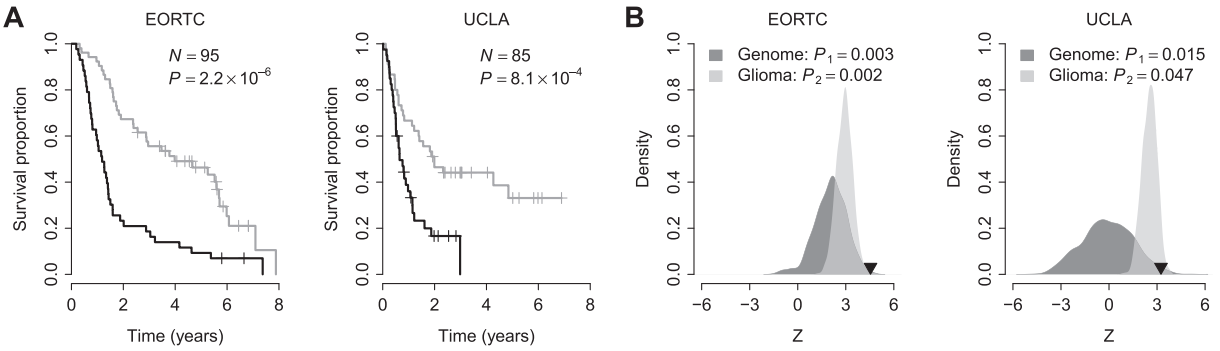

Figure 4. PIGS predicts survival in glioma in the validation cohorts. (A) Kaplan–Meier curves for the glioma patients from the EORTC and UCLA cohorts. The black curves are for the PIGS$^+$ patients, while the gray curves are for the PIGS$^-$ patients. *P*-values were calculated by log-rank test. (B) Superior prognostic power of PIGS compared with random gene signature. The dark gray area shows the distribution of Z for the 1000 resampled gene signatures picked up from human genome with identical size as PIGS. The dark gray area shows the distribution of Z for the 1000 resampled gene signatures selected from the glioma-related genes. The black triangle stands for the Z-value of PIGS. Right-tailed *P*-values of the sampling distributions were calculated.

Nonrandom Prognostic Power of PIGS

Venet et al. suggested that most published prognostic gene signatures were not significantly better than random gene sets of identical size that were randomly picked up from human genome [29]. Here, resampling tests were used to address this issue. We generated 1000 random gene signatures by randomly selecting 28 genes from human genome (the same size as PIGS). For each random gene signature, we calculated the risk score for each glioma patient. Univariate Cox proportional hazard regression of survival was used to evaluate the association between the random gene signature and clinical outcome. For each round of randomization (i.e., each randomly generated 28-gene list), we calculated the Wald statistic (Z), the ratio of Cox regression coefficient to its standard error, which stands for the prognostic power of the 28 random genes. Our alternative hypothesis was that the Z of PIGS should be more positive than expected by chance if the prognostic power of PIGS was significantly better than the random gene signatures. Figure 4B demonstrates that the Z of PIGS is significantly larger than that of the random gene sets in the two validation cohorts (right-tailed: $P = 0.003$ for the EORTC cohort and $P = 0.015$ for the UCLA cohort).

PIGS is Better than the Random Gene Signatures Selected from Glioma-related Genes

Here, we address why we focused on the PDGFRα-influenced genes to develop prognostic signature. To answer this question, we compared the performance of PIGS against the gene set associated with glioma by a second resampling test. We limited the resampling pool to the genes that were differentially expressed with glioma grade (Supplementary Tables S3 and S4) and defined these genes as glioma related. We then randomly selected 28 genes from the glioma-related gene pool and tested the predictive power of the random gene set. The performance of the random gene signature was also quantified by the Z-value computed by univariate Cox proportional hazard regression of survival. We found that the prognostic power of PIGS is significantly better than that of 1000 random glioma-related gene signatures in the both validation cohorts (right-tailed: $P = 0.002$ for the EORTC cohort and $P = 0.047$ for the UCLA cohort) (Figure 4B). These results suggest that the prognostic signature derived from the PDGFRα-influenced genes is superior to the gene sets filtered out by unbiased approaches.

PIGS is Independent of Standard Prognostic Covariates

To confirm the role of PIGS as an independent prognostic factor, we investigated the performance of PIGS in comparison with the traditional prognostic variables in glioma. Because of the limited clinical and pathological information, we did not consider the UCLA cohort. Only the EORTC cohort was investigated by multivariate model.

Firstly, we only considered the prognostic variables including age, gender, type of surgery (biopsy, partial, or total resection), Eastern Cooperative Oncology Group (ECOG) performance score [30], loss of heterozygosity (LOH) on chromosome 1p and 19q [31], and histological status (anaplastic oligoastrocytoma or anaplastic oligodendroglioma). There were 89 patients without missing data. Multivariate Cox proportional hazards regression of survival indicates that PIGS is the most significant covariate compared with the other clinical and pathological factors (Table 3).

Secondly, we took more molecular factors into account, including epidermal growth factor receptor (*EGFR*) amplification [32], isocitrate dehydrogenase 1 (*IDH1*) mutation [32], and O-6-methylguanine-DNA methyltransferase (*MGMT*) promoter methylation [33]. Because of missing observations, only 53 patients were included in this round. Multivariate Cox proportional hazards regression reveals that PIGS is still the most significant covariate in the new multivariate model (Table 4).

Table 3. Multivariate Cox proportional hazards regression conducted on 89 patients from the EORTC cohort.

Covariate	HR	95% CI of HR	*P*-value
PIGS, + vs. −	3.46	(2.04, 5.84)	3.7×10^{-6}
Age (per year)	1.03	(1.00, 1.06)	6.7×10^{-2}
Gender male vs. female	0.94	(0.55, 1.61)	8.4×10^{-1}
Type of surgery (biopsy, partial, or total resection)	0.65	(0.43, 0.97)	3.7×10^{-2}
ECOG performance score (0, 1, or 2)	1.51	(1.08, 2.10)	1.7×10^{-2}
1p/19q LOH, + vs. −	0.87	(0.43, 1.75)	6.9×10^{-1}
Histology AOA vs. AOD	1.70	(0.95, 3.04)	7.4×10^{-2}

HR: hazard ratio; CI: confidence interval; AOA: anaplastic oligoastrocytoma; AOD: anaplastic oligodendroglioma.

Table 4. Multivariate Cox proportional hazards regression conducted on 53 patients from the EORTC cohort.

Covariate	HR	95% CI of HR	*P*-value
PIGS, + vs. −	6.10	(2.48, 15.00)	8.2×10^{-5}
Age (per year)	1.05	(1.00, 1.10)	4.5×10^{-2}
Gender male vs. female	1.75	(0.76, 4.05)	1.9×10^{-1}
Type of surgery (biopsy, partial, or total resection)	0.31	(0.16, 0.58)	3.2×10^{-4}
ECOG performance status (0, 1, or 2)	1.01	(0.59, 1.72)	9.7×10^{-1}
1p/19q LOH, + vs. −	0.57	(0.15, 2.10)	3.9×10^{-1}
EGFR amplification, + vs. −	1.37	(0.58, 3.25)	4.8×10^{-1}
IDH1 mutation, + vs. −	0.40	(0.16, 1.01)	5.2×10^{-2}
MGMT methylation, + vs. −	3.53	(1.19, 10.53)	2.4×10^{-2}
Histology AOA vs. AOD	1.18	(0.49, 2.86)	7.1×10^{-1}

HR: hazard ratio; CI: confidence interval; AOA: anaplastic oligoastrocytoma; AOD: anaplastic oligodendroglioma.

From Tables 3 and 4, we can notice that patient age, type of surgery, ECOG performance score, and *MGMT* methylation status are also significant variables in multivariate model. Therefore, we further stratified the patients in the EORTC cohort according to respective significant factors and redid Cox proportional hazards regression. For patients with age ≤ 45, $45 <$ age ≤ 55, and age ≥ 55, PIGS$^+$ patients had 2.78-fold ($P = 1.2 \times 10^{-2}$), 3.04-fold ($P = 7.7 \times 10^{-3}$), and 3.09-fold ($P = 1.4 \times 10^{-2}$) increased risk for death, respectively. For patients with biopsy, partial resection, and total resection, PIGS$^+$ patients had 9.15-fold ($P = 4.2 \times 10^{-2}$), 3.18-fold ($P = 3.9 \times 10^{-4}$), and 2.73-fold ($P = 1.7 \times 10^{-2}$) increased risk for death, respectively. For patients with ECOG performance score 0, 1, and 2, PIGS$^+$ patients had 2.24-fold ($P = 4.2 \times 10^{-2}$), 3.99-fold ($P = 2.8 \times 10^{-4}$), and 4.69-fold ($P = 1.7 \times 10^{-2}$) increased risk for death, respectively. For patients with methylated *MGMT* promoter, PIGS$^+$ patients had 3.52-fold ($P = 3.2 \times 10^{-4}$) increased risk for death. Kaplan–Meier survival curves also demonstrate significantly reduced survival for PIGS$^+$ patients in each subset grouped by age, type of surgery, ECOG performance score, and *MGMT* methylation status (Figure 5). Taken together, these results suggest that PIGS is an independent prognostic variable and enhances the identification of glioma patients at greater risk for death.

Discussion

PDGFRα plays a role in organ development, wound healing, and tumor progression. Mutations in the gene PDGFRA have been

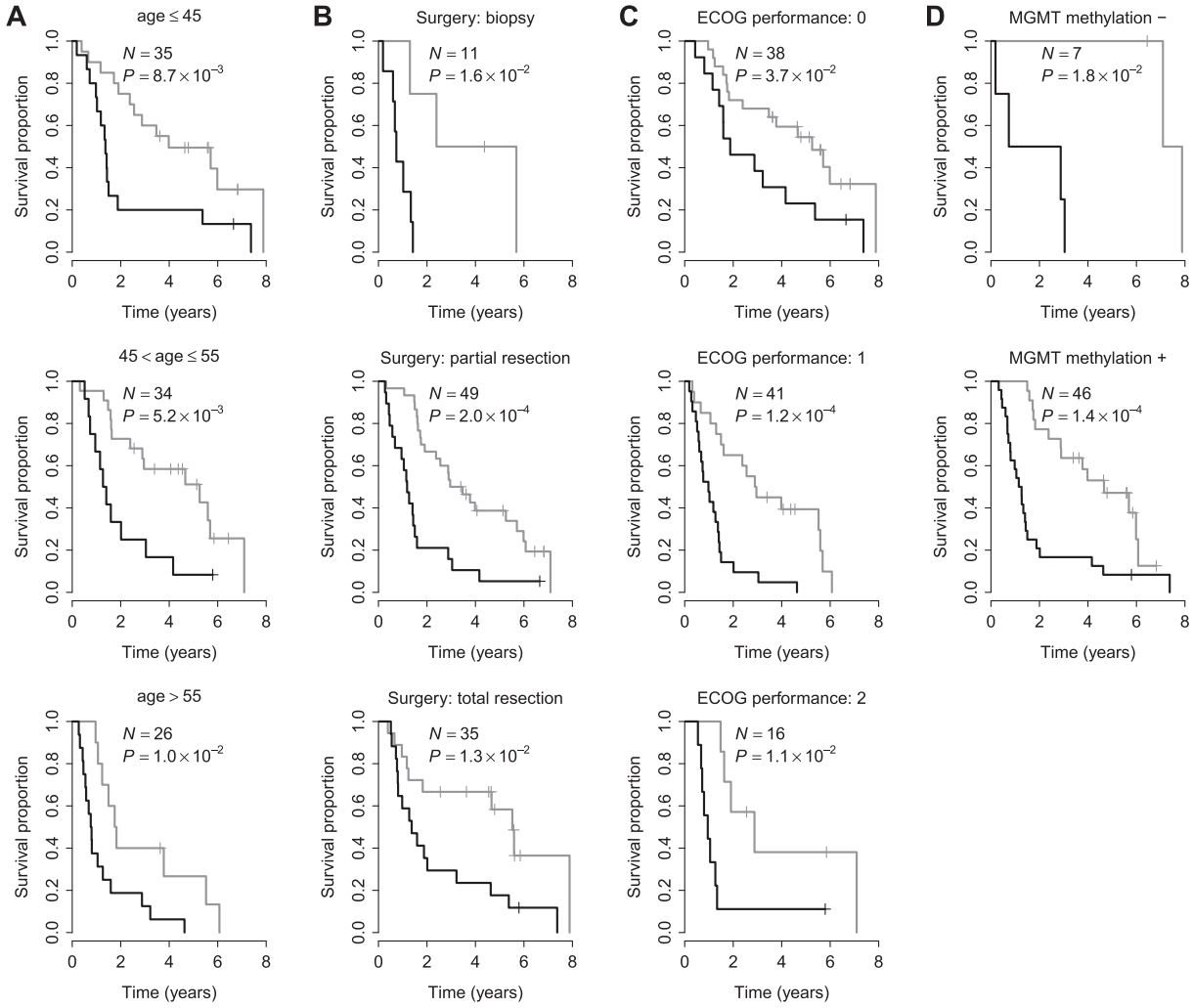

Figure 5. Kaplan–Meier curves for glioma patients grouped by clinical and pathological factors. (A) Patients were stratified by age. (B) Patients were stratified by surgery type. (C) Patients were stratified by ECOG performance score. (D) Patients were stratified by MGMT promoter methylation status. The black curves are for the PIGS$^+$ patients, while the gray curves are for the PIGS$^-$ patients. *P*-values were calculated by log-rank test.

associated with a variety of human cancers [34–36]. In addition, elevated PDGFRα expression was found in several human tumors [4–6,35], particularly in glioma [1,7,10]. In this study, we investigated the prognostic power of PDGFRα-influenced genes in glioma. By comparing the genome-wide gene expression pattern between PDGFRα$^+$ and PDGFRα$^-$ cells, we defined the genes potentially influenced by PDGFRα. These genes are strongly associated with cancer-related KEGG pathways. We subsequently developed a prognostic gene signature, PIGS, which was composed of 28 PDGFRα-influenced protein-coding genes. We indicate that PIGS-based risk score can be used to predict glioma grade. In addition, PIGS is able to predict clinical outcome in two independent glioma cohorts.

Multivariate Cox regression indicates that PIGS outperforms the traditional prognostic factors of glioma. Besides PIGS status, we considered nine clinical and pathological variables in the multivariate model, including age, gender, surgery type, ECOG performance score, 1p/19q LOH status, EGFR amplification status, IDH1 mutation status, MGMT promoter methylation status, and histological status. PIGS-based risk score is the most significant covariate compared with all the other factors. Even we stratified the data sets according to the other significant covariates, PIGS was still able to differentiate the patients with poor outcome from the long survival ones in each subgroup. These results confirm that PIGS is not dependent on specific values of the respective covariates. PIGS working cooperatively with traditional clinical and pathological factors may increase prognostic accuracy when identifying patients at higher risk of death in glioma.

A controversial computational study by Venet et al. suggested that the majority of published prognostic gene signatures of breast cancer were not significantly better than random gene sets of identical size that were randomly selected from human genome [29]. To address this issue in our study, we conducted a resampling test by randomly selecting 28 genes (the same size as PIGS) from human transcriptome and calculated the prognostic power for the random gene signature. The resampling test indicates that PIGS is superior to the random gene sets selected from human genome. However, it should also be noted that, in the EORTC cohort, the *Z*-value is larger than two (2.3% percentile in normal distribution) for almost half of the random gene signatures (Figure 4B). Therefore, the performance of a prognostic signature for glioma should not only be measured by the nominal *P*-values generated by Cox regression or log-rank test as many randomly generated gene signatures could also classify subjects with a fairly significant nominal *P*-value. Therefore, we suggest that resampling test should be a standard procedure when generating prognostic biomarkers for specific human disease. Nominal *P*-values only address the statistical question as to whether the given gene set is related to disease, but not the question whether the gene set is more related to disease than random gene sets [29]. Resampling test also demonstrates that the prognostic power of PIGS is even better than the random signatures selected from the gene pool that are differentially expressed in glioma. This result addressed the question why we only developed gene signature around PDGFRα instead of using unbiased screening.

Conclusions

Despite the absence of clear elucidation of molecular mechanisms, this study suggests the vital role of PDGFRα in carcinogenesis. Furthermore, the PDGFRα-based gene signature provides a promising prognostic tool for glioma and validates PDGFRα as a novel and effective therapeutic target in human cancers.

Declarations of Interest

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tranon.2019.10.002.

References

[1] Andrae J, Gallini R and Betsholtz C (2008). Role of platelet-derived growth factors in physiology and medicine. *Genes Dev* **22**, 1276–1312.
[2] Zhang J, Cao R, Zhang Y, Jia T, Cao Y and Wahlberg E (2009). Differential roles of PDGFR-alpha and PDGFR-beta in angiogenesis and vessel stability. *FASEB J* **23**, 153–163.
[3] Kim SJ, Kim SY, Kwon CH and Kim YK (2007). Differential effect of FGF and PDGF on cell proliferation and migration in osteoblastic cells. *Growth Factors* **25**, 77–86.
[4] Gerber DE, Gupta P, Dellinger MT, Toombs JE, Peyton M and Duignan I, et al (2012). Stromal platelet-derived growth factor receptor alpha (PDGFRalpha) provides a therapeutic target independent of tumor cell PDGFRalpha expression in lung cancer xenografts. *Mol Cancer Ther* **11**, 2473–2482.
[5] Carvalho I, Milanezi F, Martins A, Reis RM and Schmitt F (2005). Overexpression of platelet-derived growth factor receptor alpha in breast cancer is associated with tumour progression. *Breast Cancer Res* **7**, R788–R795.
[6] Wehler TC, Frerichs K, Graf C, Drescher D, Schimanski K and Biesterfeld S, et al (2008). PDGFRalpha/beta expression correlates with the metastatic behavior of human colorectal cancer: a possible rationale for a molecular targeting strategy. *Oncol Rep* **19**, 697–704.
[7] Motomura K, Mittelbronn M, Paulus W, Brokinkel B, Keyvani K and Sure U, et al (2013). PDGFRA gain in low-grade diffuse gliomas. *J Neuropathol Exp Neurol* **72**, 61–66.
[8] Paugh BS, Broniscer A, Qu C, Miller CP, Zhang J and Tatevossian RG, et al (2011). Genome-wide analyses identify recurrent amplifications of receptor tyrosine kinases and cell-cycle regulatory genes in diffuse intrinsic pontine glioma. *J Clin Oncol* **29**, 3999–4006.
[9] Puget S, Philippe C, Bax DA, Job B, Varlet P and Junier MP, et al (2012). Mesenchymal transition and PDGFRA amplification/mutation are key distinct oncogenic events in pediatric diffuse intrinsic pontine gliomas. *PLoS One* **7**:e30313.
[10] Puputti M, Tynninen O, Sihto H, Blom T, Maenpaa H and Isola J, et al (2006). Amplification of KIT, PDGFRA, VEGFR2, and EGFR in gliomas. *Mol Cancer Res* **4**, 927–934.
[11] Newton HB (2003). Molecular neuro-oncology and development of targeted therapeutic strategies for brain tumors. Part 1: growth factor and Ras signaling pathways. *Expert Rev Anticancer Ther* **3**, 595–614.
[12] Sun Y, Zhang W, Chen D, Lv Y, Zheng J and Lilljebjorn H, et al (2014). A glioma classification scheme based on coexpression modules of EGFR and PDGFRA. *Proc Natl Acad Sci U S A* **111**, 3538–3543.
[13] Sim FJ, McClain CR, Schanz SJ, Protack TL, Windrem MS and Goldman SA (2011). CD140a identifies a population of highly myelinogenic, migration-competent and efficiently engrafting human oligodendrocyte progenitor cells. *Nat Biotechnol* **29**, 934–941.
[14] Edgar R, Domrachev M and Lash AE (2002). Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* **30**, 207–210.
[15] Vital AL, Tabernero MD, Castrillo A, Rebelo O, Tao H and Gomes F, et al (2010). Gene expression profiles of human glioblastomas are associated with both tumor cytogenetics and histopathology. *Neuro Oncol* **12**, 991–1003.
[16] Sun L, Hui AM, Su Q, Vortmeyer A, Kotliarov Y and Pastorino S, et al (2006). Neuronal and glioma-derived stem cell factor induces angiogenesis within the brain. *Cancer Cell* **9**, 287–300.
[17] Liu Z, Yao Z, Li C, Lu Y and Gao C (2011). Gene expression profiling in human high-grade astrocytomas. *Comp Funct Genomics* **2011**, 245137.
[18] Erdem-Eraslan L, Gravendeel LA, de Rooi J, Eilers PH, Idbaih A and Spliet WG, et al (2013). Intrinsic molecular subtypes of glioma are prognostic

and predict benefit from adjuvant procarbazine, lomustine, and vincristine chemotherapy in combination with other prognostic factors in anaplastic oligodendroglial brain tumors: a report from EORTC study 26951. *J Clin Oncol* **31**, 328–336.

[19] Freije WA, Castro-Vargas FE, Fang Z, Horvath S, Cloughesy T and Liau LM, et al (2004). Gene expression profiling of gliomas strongly predicts survival. *Cancer Res* **64**, 6503–6510.

[20] Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B and Speed TP (2003). Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* **31**, e15.

[21] Liu WM, Mei R, Di X, Ryder TB, Hubbell E and Dee S, et al (2002). Analysis of high density expression microarrays with signed-rank call algorithms. *Bioinformatics* **18**, 1593–1599.

[22] Tusher VG, Tibshirani R and Chu G (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* **98**, 5116–5121.

[23] Storey JD (2002). A direct approach to false discovery rates. *J R Stat Soc Ser B Stat Methodol* **64**, 479–498.

[24] Taylor J, Tibshirani R and Efron B (2005). The 'miss rate' for the analysis of gene expression data. *Biostatistics* **6**, 111–117.

[25] Pitroda SP, Zhou T, Sweis RF, Filippo M, Labay E and Beckett MA, et al (2012). Tumor endothelial inflammation predicts clinical outcome in diverse human cancers. *PLoS One* **7**:e46104.

[26] Zhou T, Wang T and Garcia JG (2014). Genes influenced by the non-muscle isoform of Myosin light chain kinase impact human cancer prognosis. *PLoS One* **9**:e94325.

[27] Zhou T, Wang T and Garcia JG (2014). Expression of nicotinamide phosphoribosyltransferase-influenced genes predicts recurrence-free survival in lung and breast cancers. *Sci Rep* **4**, 6107.

[28] Kanehisa M, Goto S, Kawashima S, Okuno Y and Hattori M (2004). The KEGG resource for deciphering the genome. *Nucleic Acids Res* **32**, D277–D280.

[29] Venet D, Dumont JE and Detours V (2011). Most random gene expression signatures are significantly associated with breast cancer outcome. *PLoS Comput Biol* **7**:e1002240.

[30] Dahlrot RH (2014). The prognostic value of clinical factors and cancer stem cell-related markers in gliomas. *Dan Med J* **61**, B4944.

[31] Smith JS, Alderete B, Minn Y, Borell TJ, Perry A and Mohapatra G, et al (1999). Localization of common deletion regions on 1p and 19q in human gliomas and their association with histological subtype. *Oncogene* **18**, 4144–4152.

[32] Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y and Wilkerson MD, et al (2010). Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell* **17**, 98–110.

[33] Weller M, Stupp R, Reifenberger G, Brandes AA, van den Bent MJ and Wick W, et al (2010). MGMT promoter methylation in malignant gliomas: ready for personalized medicine? *Nat Rev Neurol* **6**, 39–51.

[34] Velghe AI, Van Cauwenberghe S, Polyansky AA, Chand D, Montano-Almendras CP and Charni S, et al (2014). PDGFRA alterations in cancer: characterization of a gain-of-function V536E transmembrane mutant as well as loss-of-function and passenger mutations. *Oncogene* **33**, 2568–2576.

[35] Ozawa T, Brennan CW, Wang L, Squatrito M, Sasayama T and Nakada M, et al (2010). PDGFRA gene rearrangements are frequent genetic events in PDGFRA-amplified glioblastomas. *Genes Dev* **24**, 2205–2218.

[36] Corless CL, Schroeder A, Griffith D, Town A, McGreevey L and Harrell P, et al (2005). PDGFRA mutations in gastrointestinal stromal tumors: frequency, spectrum and in vitro sensitivity to imatinib. *J Clin Oncol* **23**, 5357–5364.