

# Gene Expression Profiling of Gliomas Strongly Predicts Survival

William A. Freije,<sup>1,2</sup> F. Edmundo Castro-Vargas,<sup>1</sup> Zixing Fang,<sup>1</sup> Steve Horvath,<sup>1,3</sup> Timothy Cloughesy,<sup>4,5</sup> Linda M. Liao,<sup>5,6</sup> Paul S. Mischel,<sup>5,7</sup> and Stanley F. Nelson<sup>1,5</sup>

<sup>1</sup>Department of Human Genetics, <sup>2</sup>Department of Obstetrics and Gynecology, <sup>3</sup>Department of Biostatistics, <sup>4</sup>Department of Neurology, <sup>5</sup>Henry E. Singleton Brain Tumor Program, <sup>6</sup>Department of Surgery, Division of Neurosurgery, and <sup>7</sup>Department of Pathology, David Geffen School of Medicine, University of California at Los Angeles, Los Angeles, California

## ABSTRACT

In current clinical practice, histology-based grading of diffuse infiltrative gliomas is the best predictor of patient survival time. Yet histology provides little insight into the underlying biology of gliomas and is limited in its ability to identify and guide new molecularly targeted therapies. We have performed large-scale gene expression analysis using the Affymetrix HG U133 oligonucleotide arrays on 85 diffuse infiltrating gliomas of all histologic types to assess whether a gene expression-based, histology-independent classifier is predictive of survival and to determine whether gene expression signatures provide insight into the biology of gliomas. We found that gene expression-based grouping of tumors is a more powerful survival predictor than histologic grade or age. The poor prognosis samples could be grouped into three different poor prognosis groups, each with distinct molecular signatures. We further describe a list of 44 genes whose expression patterns reliably classify gliomas into previously unrecognized biological and prognostic groups: these genes are outstanding candidates for use in histology-independent classification of high-grade gliomas. The ability of the large scale and 44 gene set expression signatures to group tumors into strong survival groups was validated with an additional external and independent data set from another institution composed of 50 additional gliomas. This demonstrates that large-scale gene expression analysis and subset analysis of gliomas reveals unrecognized heterogeneity of tumors and is efficient at selecting prognosis-related gene expression differences which are able to be applied across institutions.

## INTRODUCTION

Diffuse infiltrative gliomas are the most common malignant primary brain tumor. Gliomas are histologically defined based on whether they exhibit primarily astrocytic or oligodendroglial morphology, and are graded by cellularity, cytological atypia, necrosis, mitotic figures, and vascular proliferation, features associated with biologically aggressive behavior (1). This system of diagnosis has been developed over decades of experience and forms the cornerstone of neuro-oncology. Patient prognosis and therapeutic decisions rely on accurate pathological grading. It is largely reproducible, although substantial disagreement between neuropathologists can occur with respect to both type and grade (2–4), and the exact method of grading changes over time (5). Because it is based on morphology, a biolog-

ical end state, it is limited in its ability to identify new potential treatment strategies.

Abundant evidence suggests the presence of unrecognized, clinically relevant subclasses of the diffuse gliomas, both with respect to their underlying molecular phenotype and their clinical response to therapy (6). It is also apparent from clinical response to a variety of treatment regimens that histologically identical tumors can behave in highly different manners (6). This can result in underpowered studies that fail to detect potentially promising new therapies (7). For instance, a small subset of patients can have a substantial response to irinotecan, whereas the majority of patients do not respond (8). These findings underscore the fact that histopathologic evaluation may not be sensitive to much of the variation in underlying biology. As oncologists move toward molecularly targeted therapies, identification of distinct molecularly defined subgroups becomes critical (9). To additionally clarify the underlying biology and to refine survival prognostication, adjunctive studies of tumor-associated genes such as the epidermal growth factor receptor (*EGFR*), tumor protein p53 (*TP53*), phosphatase and tensin homologue (*PTEN*), proliferation related Ki-67 antigen (*Ki-67*), murine double-minute 2 (*MDM2*), mut-s homologue 2 (*MSH2*), cyclin-dependent kinase inhibitor 2A (*CDKN2A*), sarcoma amplified sequence (*SAS*), and platelet-derived growth factor  $\alpha$  (*PDGFA*) have been used (10–14). In general, individual gene/protein assays alone or in combination with histologic features are neither predictive of survival of glioma patients nor able to guide therapeutic decisions. Thus, although these genes may indeed play a role in the biology of gliomas, their utility as diagnostic markers is not yet clear, perhaps due to heterogeneity within the tumor groupings.

Microarray analysis offers the option of unbiased, quantitative, and reproducible tumor evaluation by simultaneously evaluating thousands of individual gene expression measurements. This approach has been applied to many different cancers including gliomas (9, 15–21). Unlike pathological evaluation, microarray analysis provides the means to describe the underlying variation in genetic composition of the tumors, allowing for novel tumor class identification and patient prognostication. Efforts at dissecting gliomas into more homogeneous groups with clear relationships to underlying molecular defects using microarrays have suggested that: (A) different types of gliomas have different gene expression profiles (20), (B) there is evidence for distinct molecular subsets of morphologically identical gliomas (9), and (C) a gene expression-based predictor more robustly correlates with survival than does histologic type (17).

This study addresses two central issues that remain outstanding: (i) the use of gene expression profiling to identify predictors of survival in gliomas that can be robustly applied across different institutions and technological platforms, and (ii) the identification of relevant subclasses of gliomas that behave as distinct prognostic groups and that contain transcriptional signatures that can be used to guide and develop new treatments. This is the broadest study of malignant gliomas to date and has resulted in the delineation of clear prognostic groups based on gene expression.

Received 2/20/04; revised 7/7/04; accepted 7/26/04.

**Grant support:** U01CA88173 from the National Cancer Institute, the Henry Singleton Brain Tumor Program, the Elliott Family Foundation in memory of Neal Elliott, the Ziering Family Foundation in memory of Sigi Ziering, and Art of the Brain. In addition, microarray support was provided by the National Institute of Neurological Disorders and Stroke/National Institute of Mental Health Microarray Consortium within the University of California Los Angeles DNA Microarray Facility. P. S. Mischel was also supported by NS43147 and a grant from Accelerate Brain Cancer Cure. W. A. Freije is a scholar of the Women's Reproductive Health Research Center (Grant No. 5-K12-HD001281).

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.

**Note:** In an effort to stimulate further exploration of genomic data, the data, including raw cell files, have been submitted to GEO for distribution to interested researchers. Supplementary data for this article can be found at Cancer Research Online (<http://cancerres.aacrjournals.org>).

**Requests for reprints:** Stanley F. Nelson, Room 5506b, 695 Young Drive South, University of California Los Angeles Medical Center, Los Angeles, CA 90095. Phone: 310-794-7981; E-mail: [snelson@ucla.edu](mailto:snelson@ucla.edu).

©2004 American Association for Cancer Research.

## MATERIALS AND METHODS

**Patient Selection and Inclusion Criteria.** All patients undergoing surgical treatment at the University of California, Los Angeles for primary brain cancers between 1996 and 2003 were invited to participate in this Institutional Review Board-approved study. Of the patients participating in this broad protocol, 74 were analyzed as part of this study if their tumor was diagnosed as a grade III ( $n = 24$ ) or IV ( $n = 50$ ) glioma of any histologic type and fresh-frozen material was obtained. Only grade III and IV gliomas were included in this study, because the distinction between these grades is subtle and prone to misclassification. The time in days elapsed from resection to the day of death, or if the patient has remained alive, the time in days elapsed from resection to the current day was recorded for all of the samples studied. Patient ages at diagnosis varied from 18 to 82 years. There were 46 females and 28 males. A total of 11 repeat samples were available for 7 patients and included in the partial least squares/partition around medioids analysis, which used 85 total expression profiles. However, only survival data from the initial 74 resections were used in cox regression and survival prediction. Supplementary Table S1 lists demographic, treatment, and survival data of the patients and their samples.

**Tissue Acquisition, RNA Extraction, and Gene Expression Array Analysis.** Surgical resection was attempted on all of the patients. Tumor tissue removed, which was not required for diagnosis, was snap frozen in liquid nitrogen. Representative thin sections of the removed tissue were diagnosed and graded independently by two board-certified neuropathologists using World Health Organization criteria. Furthermore, tumors were centrally re-analyzed by a board certified neuropathologists (P. S. M.). RNA extraction was performed as described previously (9) and analyzed on both a spectrophotometer and Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). Only samples with 28S/18S ratio of  $>1.5$  and no evidence of ribosomal peak degradation were included.

Probes were prepared using standard Affymetrix protocols as described previously (9) and hybridized to Affymetrix HG-U133A and HG-U133B arrays (Affymetrix, Santa Clara, CA). All of the microarrays were examined for surface defects, grid placement, background intensity, housekeeping gene expression, and 3 prime/5 prime ratio of probe sets from genes of varying length (signal 3'/5' ratio  $<3$ ). Four percent of arrays failed quality check on one of these points and were repeated successfully on the second attempt.

**Real-time Quantitative Reverse Transcription-PCR.** Two micrograms of total RNA were used in a reverse transcription reaction (volume 100  $\mu$ L) using MultiScribe Reverse Transcriptase (Applied Biosystems, Foster City, CA). Real-time PCR was carried out with MJ Opticon PCR Analyzer (MJ Research Inc., Waltham, MA) using SYBR Green PCR Core Reagents (Applied Biosystems Technology). The reactions were cycled 30 times [50°C, 2 minutes and 95°C, 10 minutes, (94°C, 15 seconds; 60°C, 1 minute; and 72°C, 1 minute)] with fluorescence measurements every 15 seconds at the end of each cycle to construct an amplification curve. A melting curve was performed at the end of amplification cycles to verify the specificity of the PCR products. All of the determinations were performed in duplicate. Sixteen genes identified in our screen were selected to validate the microarray results. Supplementary Fig. S1 lists the genes and primers used.

**Statistical Analysis.** Microarray Suite 5.0 was used to define absent/present calls and generate cel files using the default settings. Data files (cel) were uploaded into the dCHIP program (22) and normalized to the median intensity array. Quantification was performed using model-based expression and the perfect match minus mismatch method in dCHIP. Probe sets were initially filtered to select for a coefficient of variation  $>0.2$  with at least 10% of the samples having an expression intensity  $>500$ . This yielded 8,011 probe sets, which contain most of the variation in gene expression information across all of the samples. There were 5,826 probe sets from the HG-U133A array and 2,185 probe sets from the HG-U133B array.

The Kaplan-Meier method was used to estimate the survival distributions (23). Log rank tests were used to test the difference between stratified survival groups. To assess which covariates affect survival, we used multivariable Cox proportional hazard models (24). The proportional hazard assumption was tested using scaled Schoenfeld residuals. For each covariate, the relative hazard rate and the associated  $P$  values were examined. For all of the analyses,

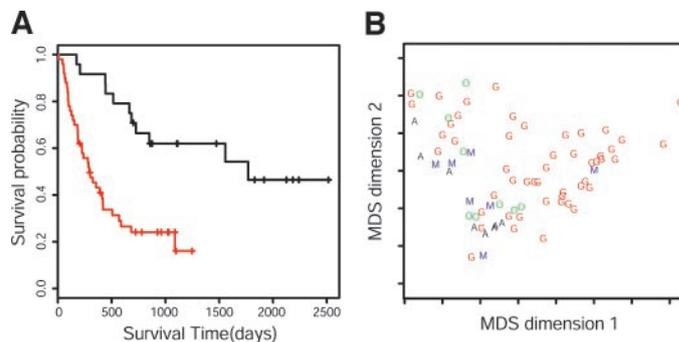


Fig. 1. Survival analysis of histology based classification of gliomas. A, Kaplan-Meier survival plot of the patients entered in this study partitioned by histologic grade. *Black line*, survival probability of the grade III (anaplastic astrocytoma, anaplastic mixed oligo-astrocytoma, and anaplastic oligodendroglioma) patients; *red line*, survival probability of grade IV (glioblastoma) patients. B, multidimensional scaling (MDS) plot of 74 diffuse infiltrating gliomas using 8,011 significantly variable genes. The glioblastoma multiforme samples are indicated by a *red G*, and the grade III tumors are a *black A*, a *green O*, and a *blue M* denoting anaplastic astrocytoma, anaplastic oligodendroglioma, and anaplastic mixed oligo-astrocytoma samples, respectively.

a  $P < 0.05$  was accepted as significant. Statistical analyses were carried out with the freely available software packages R.<sup>8</sup>

The straightforward implementation of partial least squares (SIMPLS) procedure with 8 components was used for supervised dimension reduction using survival outcomes in R (25). Because the SIMPLS approach cannot deal with censored survival outcomes directly, we replaced the survival times by the deviance residuals that result from fitting an intercept only Cox regression model. The deviance residual, which is a symmetrized version of a martingale residual, is a measure of “excess” death (26). When using partial least squares to predict the deviance residual in a test set, it can be interpreted as a measure of hazard. We used the partial least squares procedure in conjunction with partition around medioids clustering (27) to group all 85 of the tumors into two major clusters. We used the Euclidean distance in hierarchical and partition around medioids clustering. To visualize the data, we used classical multidimensional scaling, which takes as input a distance matrix between samples and returns a set of points in a lower dimensional space such that the distances between the points are approximately equivalent to the original distances (28).

## RESULTS

**High-Grade Gliomas Sampled in This Study Are Typical.** To assess the association between grade and survival, we plotted Kaplan-Meier survival curves of the patients entered in this study partitioned by histologic grade (III or IV; Fig. 1A). Patients with grade III tumors had a median survival of 4.8 years, whereas patients with grade IV tumors had a median survival of 293 days. These results are consistent with the literature (2, 29) and confirm that histologic grade is a good predictor of survival ( $P = 9.3 \times 10^{-5}$ ). Similarly, as expected, younger age was also associated with longer patient survival ( $P = 0.026$ ). A multidimensional scaling plot of all 74 of the initial tumors using the 8,011 most variable genes indicates that whereas there is some aggregation of the tumors by grade, the tumors types as defined by pathological grade or histology are not well distinguished by an unselected set of genes (Fig. 1B). The majority of patients received radiation therapy and chemotherapy regimens per standard protocols. There was no difference in the frequency of different types of chemotherapy regimens between the identified groups. Thus, therapeutic regimens do not confound the results presented here.

**Global Gene Expression Classifiers Outperform Histopathology in Survival Prediction.** To begin developing a gene expression-based predictor of high-grade gliomas, we first sought to determine which probe sets correlated with survival and were able to group the

<sup>8</sup> Internet address: <http://cran.r-project.org/>.

85 tumors into two distinct survival groups. We used partial least squares regression to arrive at eight components that explain 96% of the variation in survival time. We used these components in conjunction with partition around medoids clustering to group the samples into two prognosis-related gene expression-based groups: a good prognosis group entitled survival cluster group 1 and a poor prognosis group entitled survival cluster group 2. Kaplan-Meier plots of these two major groups show that survival cluster group 1 had a median survival of 4.8 years, whereas survival cluster group 2 had a median survival of 237 days (log rank  $P = 2.1 \times 10^{-8}$ ; Fig. 2). Furthermore, there was substantial regrouping of the histology defined groups. survival cluster group 1 contained 10 glioblastomas, 7 anaplastic astrocytomas, 6 anaplastic oligodendrogliomas, and 6 anaplastic mixed oligo-astrocytoma samples. The survival cluster group 2 group contained 40 glioblastomas, 1 anaplastic astrocytoma, 3 anaplastic oligodendrogliomas, and 1 anaplastic mixed oligo-astrocytoma samples. A multidimensional scaling plot of the first two partial least squares components shows separation of the two survival groups (Fig. 2B).

Model validation was performed to determine whether gene expression data contained survival predictive information that was distinct from the prediction embedded within histologic grade. Univariate and multivariate Cox regression analysis was used in fitted and validated models (Table 1). Univariate Cox regression analysis evaluating grade, age, gender, and microarray expression-based predictors (survival cluster), and multivariate Cox regression analysis evaluating survival cluster/grade and survival cluster/age were performed. The fitted univariate Cox regression model encompassing 74 samples revealed that survival cluster outperformed histologic grade prediction, with fitted  $P$ s of  $5.7 \times 10^{-7}$  and  $2.6 \times 10^{-4}$  for survival cluster and grade, respectively. To protect against overfitting, internal model validation was performed by randomly splitting two thirds of the data into training sets and one third into test data sets, thereby creating a validated model. The splitting was done in a way to preserve the ratio of grade III and grade IV patients in each data set. A partial least squares model of survival was fitted in the training data set. This training model was used to predict a microarray gene expression-based hazard score survival cluster in the test data. One-sided Wald test  $P$  values were recorded for each of the covariates. The random splitting into training and test data was repeated 50 times and the results averaged to define a validated  $P$  value (Table 1). This analysis showed that the microarray-based information was superior at pre-

Table 1 Significance of univariate and multivariate Cox regression analysis in fitted and validated models

Covariates in the model	Covariate	Hazard ratio	Fitted $P$ value	Validated $P$ value
Grade	Grade	4.00	$2.6e^{-4}$	0.05
Age	Age	1.03	$2.6e^{-3}$	0.08
Gender	Gender	0.66	0.174	0.44
SC	SC	7.03	$5.7e^{-7}$	$3.5e^{-3}$
SC, grade	SC	5.26	$1.1e^{-4}$	0.029
	Grade	1.85	0.144	0.38
SC, age	SC	6.35	$4.7e^{-6}$	$6.5e^{-3}$
	Age	1.01	0.216	0.39

NOTE. The fitted hazard ratios and fitted  $P$  values are obtained by univariate and multivariate Cox regression analysis using 74 tissue samples with the covariates of age, grade, gender, and SC. SC is a two group classifier based PLS/PAM of the microarray data. Internal model validation was performed by splitting the data into training and test sets, and calculating validated  $P$  values. The directions of the hazard ratios are as follows: grade, grade IV versus grade III; age, older versus younger; gender, female versus male; SC, SC2 versus SC1.

Abbreviations: SC, survival cluster; PLS, partial least squares; PAM, partition around medoids.

dicting survival to histologic grading. The validated  $P$  value of survival cluster was 0.0035, the grade validated  $P$  value was 0.05, and the multivariate validated  $P$  values of grade and survival cluster were 0.38 and 0.029, respectively. Thus, when including both predictors in a Cox model, grade becomes insignificant at the 0.05 significance level. Thus, the gene expression information is a stronger predictor of survival than patient age or histologic grade. Indeed, patient age and grade are not independent predictors of survival in the context of gene expression classification.

**Genetic Heterogeneity of High-Grade Gliomas Revealed by Hierarchical Clustering.** Whereas the survival cluster group 1 versus survival cluster group 2 grouping is strong and shows that a gene expression-based classification can perform well relative to current clinical predictors of survival, it would be difficult to use this information clinically, because it relies on data from 8,011 probe sets. Furthermore, this information is limited, because biological inferences from partial least squares analysis of the gene expression data are not obvious. To reduce the scale of genes identified and select the most strongly correlated genes, we selected genes with fold difference between survival cluster group 1 and survival cluster group 2 ( $>2$ -fold) and significance (two-sided  $t$  test  $P < 0.01$ ). Gene expression values from probe sets with absolute difference between the mean of the groups that were  $<100$  (10<sup>th</sup> percentile) were removed to reduce noise. This selection resulted in the identification of 595 probe sets. Of the probe sets, 397 were from the HG-U133A array, and 198 were from the HG-U133B array (Supplementary Table S2). This set of genes is termed the prediction set.

Hierarchical clustering of the samples with the prediction set reveals four dominant clusters of tumors defined by four groups of probe sets (Fig. 3). To characterize the four groups we performed gene ontology classification of the gene clusters that define each hierarchical cluster group through the EASE program (30). The longer survival group was named hierarchical cluster 1A and is defined by overexpression of genes involved in neurogenesis ( $P = 0.013$ ). Tumor members of hierarchical cluster 1A were almost exclusively from survival cluster group 1. Hierarchical cluster 1B was defined by overexpression of genes involved in synaptic transmission ( $P = 1.1 \times 10^{-4}$ ) and is clearly heterogeneous, because tumors in this group were equally derived from survival cluster group 1 and survival cluster group 2. Hierarchical cluster 2A and hierarchical cluster 2B contained most of the higher-grade tumors (previously grouped as survival cluster group 2). Hierarchical cluster 2A is defined by overexpression of genes involved in mitosis ( $P = 7.3 \times 10^{-16}$ ). Hierarchical cluster 2B is defined by overexpression of extracellular matrix components and regulators ( $P = 1.95 \times 10^{-15}$ ). The lower red/green

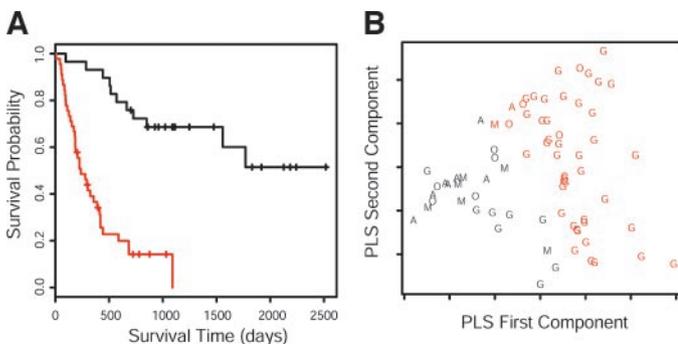


Fig. 2. Survival analysis of gliomas grouped by partitioning around the medoids. A, Kaplan-Meier survival plot of the two major groups of patients defined by the partial least squares (PLS)/partition around medoids. Survival cluster group 1 (black line) defines the good survival group, and survival cluster group 2 (red line) defines the poor survival group. B, multidimensional scaling plot of the first two principle components of the PLS analysis. Samples in black represent the survival cluster group 1 and samples in red represent the survival cluster group 2. The letters G, A, O, and M represent glioblastoma, anaplastic astrocytoma, anaplastic oligodendroglioma, and mixed oligo-astrocytoma samples, respectively.

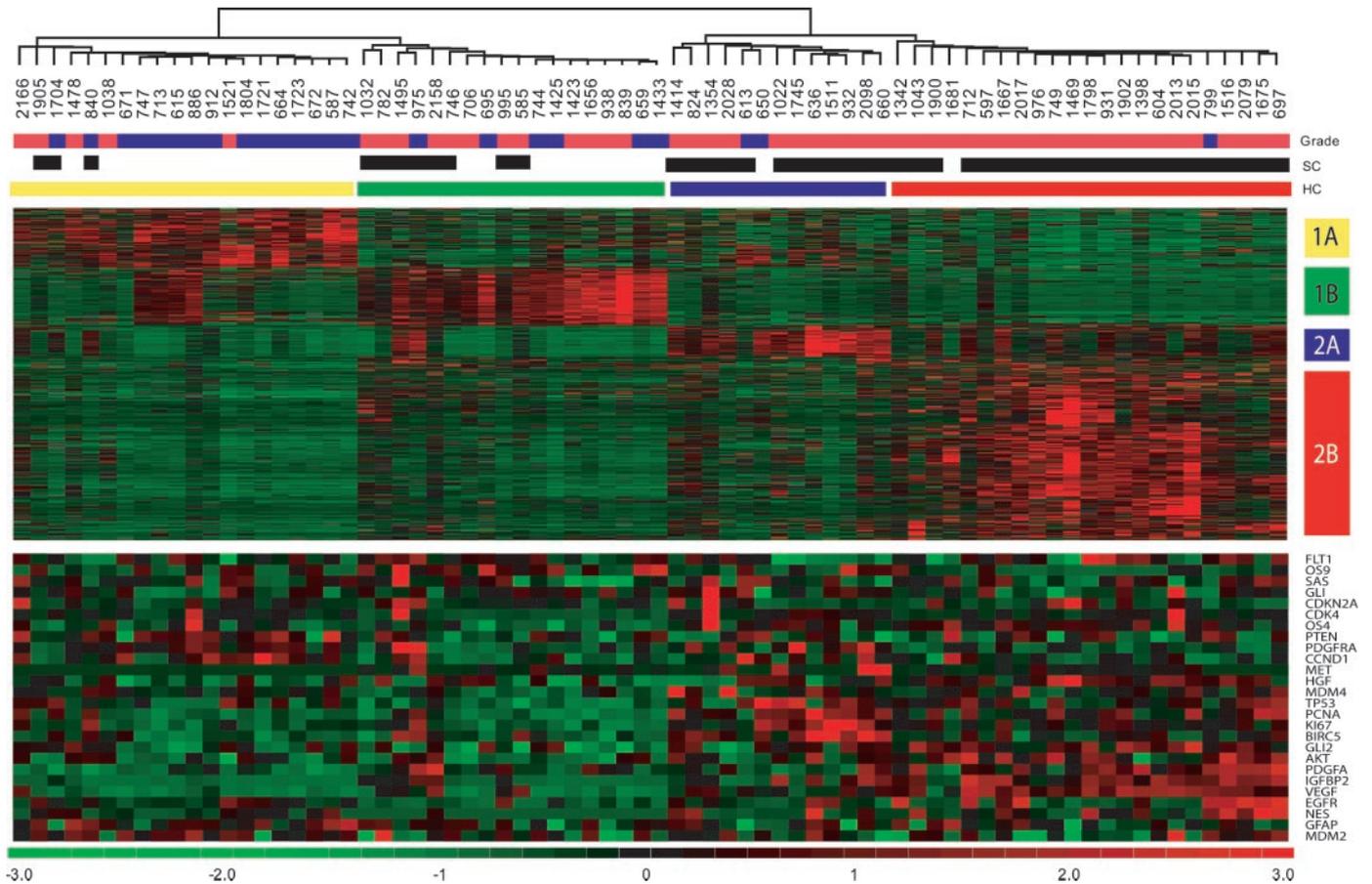


Fig. 3. Hierarchical clustering dendrogram of 74 gliomas with 595 genes significantly different between survival cluster group 1 and survival cluster group 2; 595 genes identified as significantly different between survival cluster group 1 and survival cluster group 2 were used for unsupervised hierarchical clustering of tumor specimens and genes. Gliomas are listed across the top by tumor identification number and are joined by a dendrogram. Tumors are also identified by group membership. Grade indicates grade III histology (blue) or grade IV histology (red). SC indicates membership in initial survival cluster analysis survival cluster group 1 (white,  $n = 29$ ) or survival cluster group 2 (black,  $n = 45$ ). Group membership by hierarchical clustering (HC) is indicated along the right edge of the gene clusters. The upper red/green expression diagram shows expression data from the 595 survival associated probe sets, indicating strong clustering into four classes corresponding to hierarchical cluster 1A, hierarchical cluster 1B, hierarchical cluster 2A, and hierarchical cluster 2B. Log 2-fold intensities and color relationships are at the bottom. Hierarchical cluster classification of genes is indicated along the right edge for the gene clusters. The lower red/green expression diagram represents genes examined previously and associated with gliomagenesis. From top to bottom the genes are vascular endothelial growth factor receptor (*FLT1*; ref. 39), osteosarcoma amplified 9 (*OS9*; ref. 47), *SAS* (48), glioma associated oncogene (*GLI*; ref. 49), *CDKN2A* (50), cyclin dependent kinase 4 (*CDK4*; ref. 48), osteosarcoma amplified 4 (*OS4*; ref. 47), *PTEN* (14), platelet derived growth factor receptor  $\alpha$  (*PDGFRA*; ref. 51), cyclin D1 (*CCND1*; ref. 52), hepatocyte growth factor receptor (*MET*; ref. 33), hepatocyte growth factor (*HGF*; ref. 53), murine double-minute 4 (*MDM4*; ref. 54), *TP53* (55), proliferating cell nuclear antigen (*PCNA*; ref. 56), *Ki67* (12), survivin (*BIRC5*; ref. 57), glioma associated-kruppel family member 2 (*GLI2*; ref. 58), v-AKT murine thymoma viral oncogene homologue 1 (*AKT*; ref. 59), *PDGFA* (60), insulin-like growth factor binding protein 2 (*IGFBP2*; ref. 61), vascular endothelial growth factor (*VEGF*; ref. 39), epidermal growth factor receptor (*EGFR*; ref. 62), nestin (*NES*; ref. 63), glial fibrillary acidic protein (*GFAP*; ref. 64), and *MDM2* (51).

expression diagram of Fig. 3 represents select genes examined previously in gliomagenesis and their relationship to the described hierarchical cluster groups.

**Description of a 44 Gene List That Robustly Classifies Gliomas into Prognostic Groups.** To define a robust list of genes for additional assay development that can determine individual tumor membership to the four hierarchical clustering groups, individual-named genes were rank ordered by pair-wise comparison between each hierarchical cluster group and all of the other samples. A balanced list of 44 of the most strongly and consistently differentially expressed genes that best define the hierarchical groups was identified (Table 2). The list of genes was determined by rank ordering all of the genes identified that minimally had a fold change  $>2$ , mean expression difference  $>100$ , and by  $P$  value (maximum  $P = 0.01$ , hierarchical clustering group of overexpression versus other three groups). Eleven named genes highly overexpressed in each group were selected to balance the gene lists. Using this gene list a simple gene voting protocol was developed that compared the signal intensity of the voting gene to the mean expression value over the entire study. If the value of the signal intensity of the voting gene was greater than the

study mean, a single positive vote was made, and if the signal intensity of the voting gene was lower than the study mean, a zero vote was made. The sum of the individual gene votes was tallied for each hierarchical cluster, and this total was divided by the number of voting genes for that particular hierarchical cluster group to arrive at a percentage of positive group-specific voting genes. The classification of the sample was determined by the hierarchical cluster with the largest percentage of genes with positive votes. Using this gene voting protocol, tumor samples were assigned to one of the four major hierarchical groups. Voting profiles from representative tumor sample from each of the four hierarchical groups are presented in Fig. 4. Kaplan-Meier plots of the hierarchical cluster groups and the gene voting groups parallel each other (Fig. 5). Fig. 5A depicts the Kaplan-Meier survival analysis of the hierarchical groups as determined by 595 gene hierarchical clustering (log rank  $P = 0.00011$ ). Kaplan-Meier plots for each of these gene expression-defined groups based on hierarchical cluster membership indicate that membership in hierarchical cluster 1A strongly predicts a favorable prognosis relative to the other groups (log rank  $P = 7 \times 10^{-6}$ ). Hierarchical cluster 2A and hierarchical cluster 2B strongly predict poor patient survival. In ag-

Table 2 The top 44 genes that distinguish the major clusters of gliomas amenable to evaluation by immunohistochemistry

Probeset	Title	Gene symbol (ref.)	HC group	Gene ontology category determined by EASE
205289_at	Bone morphogenetic protein 2	BMP2 (31)	HC1A	Neurogenesis
222898_s_at	Delta-like 3	DLL3	HC1A	Neurogenesis
204225_at	Histone deacetylase 4	HDAC4	HC1A	Neurogenesis
204273_at	Endothelin receptor type B	EDNRB	HC1A	Neurogenesis
219743_at	Hairy/enhancer-of-split related	HEY2	HC1A	Neurogenesis
214680_at	Neurotrophic tyrosine kinase receptor type 2	NTRK2	HC1A	Neurogenesis
209981_at	Ortholog of rat pippin	PIPPIN	HC1A	Unclassified
210815_s_at	Calcitonin receptor-like	CALCRL	HC1A	Signal transducer
221204_s_at	Cartilage acidic protein 1	CRTAC1	HC1A	Cell adhesion
207145_at	Growth differentiation factor 8	GDF8	HC1A	Organogenesis
205874_at	Inositol 1,4,5-trisphosphate 3-kinase A	IP3K	HC1B	Signal transduction
204338_s_at	Regulator of G-protein signalling 4	RGS4(32)	HC1B	Calmodulin binding
203998_s_at	Synaptotagmin I	SYT1	HC1B	Synaptic vesicle
203798_s_at	Visinin-like 1	VSNL1	HC1B	Calcium ion binding
207853_s_at	Synuclein, $\beta$	SNCB	HC1B	Enzyme inhibitor
227053_at	Protein kinase C casein kinase substrate in neurons 1	PACSN1	HC1B	Phospholipase inhibitor
204467_s_at	Synuclein, $\alpha$	SNCA	HC1B	Development
206162_x_at	Synaptotagmin V	SYT5	HC1B	Synaptic vesicle
224650_at	Mal, T-cell differentiation protein 2	MAL2	HC1B	Integral membrane
203510_at	Met proto-oncogene	MET (33)	HC2A	Cell proliferation
201292_at	Topoisomerase II $\alpha$	TOP2A (34)	HC2A	Mitosis
202410_x_at	Insulin-like growth factor 2	IGF2 (35)	HC2A	Cell cycle/cell proliferation
203213_at	Cell division cycle 2	CDC2 (36)	HC2A	Mitosis
209464_at	Serine/threonine kinase 12	STK12 (65)	HC2A	Cell cycle/cell proliferation
209714_s_at	Cyclin-dependent kinase inhibitor 3	CDKN3	HC2A	Mitosis
202107_s_at	Minichromosome maintenance deficient 2	MCM2(66)	HC2A	Mitosis
202095_s_at	Survivin	BIRC5 (57)	HC2A	Mitosis
203554_x_at	Pituitary tumor-transforming 1	PTTG1	HC2A	Mitosis
209172_s_at	Centromere protein F	CENPF (37)	HC2A	Mitosis
201438_at	Collagen, type VI, $\alpha 3$	COL6A3	HC2B	Extracellular matrix
201508_at	Insulin-like growth factor binding protein 4	IGFBP4(37)	HC2B	Extracellular
203851_at	Insulin-like growth factor binding protein 6	IGFBP6	HC2B	Extracellular
215446_s_at	Lysyl oxidase	LOX	HC2B	Extracellular matrix
205266_at	Cholinergic differentiation factor	LIF (67)	HC2B	Extracellular
201110_s_at	Thrombospondin 1	THBS1(38)	HC2B	Extracellular matrix
201505_at	Laminin $\beta 1$	LAMB1	HC2B	Extracellular matrix
203186_s_at	S100 calcium binding protein A4	S100A4 (40)	HC2B	Morphogenesis/development
209335_at	Decorin	DCN (41)	HC2B	Extracellular matrix
208949_s_at	Galectin 3	LGALS3	HC2B	Extracellular/cell adhesion
210512_s_at	Vascular endothelial growth factor	VEGF (39)	HC2B	Extracellular/cell adhesion
212464_s_at	Fibronectin 1	FN1 (41)	HC2B	Extracellular matrix
203936_s_at	Matrix metalloproteinase 9	MMP9 (68)	HC2B	Extracellular matrix
202291_s_at	Matrix Gla protein	MGP(37)	HC2B	Extracellular matrix
202718_at	Insulin-like growth factor binding protein 2	IGFBP2 (67)	HC2B	Extracellular

Abbreviation: HC, hierarchical clustering.

gregate, hierarchical cluster 1B membership is more heterogeneous than the other groups but is generally a poor prognosis group. Fig. 5B represents the Kaplan-Meier survival analysis of the four voted groups as determined by the 44 gene voting algorithm (log rank  $P = 0.00022$ ).

**Confirmation of Microarray Results with Reverse Transcription-PCR.** Twenty six tumor samples representing characteristic members of each hierarchical cluster group (hierarchical cluster 1A,  $n = 7$ ; hierarchical cluster 1B,  $n = 6$ ; hierarchical cluster 2A,  $n = 7$ ; hierarchical cluster 2B,  $n = 6$ ) were evaluated by reverse transcription-PCR for 16 genes. The selected 16 genes were composed of four from each hierarchical cluster group and include BMP2 (31), DLL3, HDAC4, EDNRB, IP3K, RGS4 (32), SYT1, VSNL1, MET (33), TOP2A (34), IGF2 (35), CDC2 (36), COL6A3, IGFBP4 (37), LOX, and THBS1 (38). The mean Pearson correlation coefficient between microarray quantification and reverse transcription-PCR quantification was 0.89. Graphs of expression measurements determined by microarray and real-time reverse transcription-PCR are shown in Supplemental Fig. 1 for DLL3, SYT1, TOP2A, and COL6A3.

**Gene Expression Based Survival Predictors Are Supported by Analysis of an Independent and External Data Set.** A validation dataset was obtained from an independent analysis of gene expression in a series of 50 glioblastoma and anaplastic oligodendroglioma tumors (17). These tumors were analyzed on a distinct array platform (U95A), which surveys 12,626 probesets. We were able to identify

matches to 344 probesets of the 595 probesets used in our U133 based predictor. The image files (cel) were reanalyzed to quantify and normalize gene expression values in the same fashion as our U133 data. Using the 344 probesets we used partition around medoids clustering to divide the patients into two distinct survival associated groups ( $P = 6.1 \times 10^{-5}$ ). The gene expression predictor proposed here significantly outperformed the initial histopathology grading classification ( $P = 0.007$ ) used in the initial study strongly validating the prognostic power of the prediction set defined here (17). Furthermore, we find that each tumor strongly groups within a gene expression-based group. This dataset strongly supports the reproducibility of the prediction set of 595 genes to group gliomas based on microarray gene expression data into four meaningful and reproducible survival groups and to show a gene expression-based predictor as a rigorous means to classify gliomas.

## DISCUSSION

We have performed large-scale gene expression analysis of all histologic types of high-grade gliomas to develop a robust gene expression-based, histology-independent classification scheme. The process of selecting genes capable of separating tumors into good and poor prognosis groups proved to be robust as judged by an independent dataset. This gene expression-based predictor provides insights into glioma biology and heterogeneity and may develop into a useful

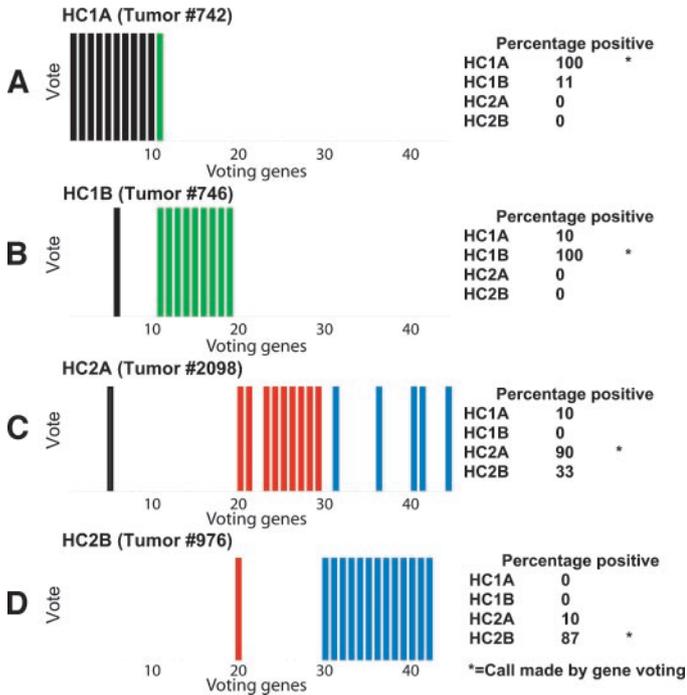


Fig. 4. Gene voting profiles and call of four tumors; 44 voting genes encompassing 10 for hierarchical cluster 1A (HC1A; black), 9 for hierarchical cluster 1B (green), 10 for hierarchical cluster 2A (HC2A; red), and 15 for hierarchical cluster 2B (HC2B; blue) were used to classify exemplary tumors from hierarchical cluster 1A (742, A), hierarchical cluster 1B (HC1B, 746, B), hierarchical cluster 2A (2098, C), and hierarchical cluster 2B (976, D). The percentage of positive voting genes particular to a hierarchical cluster group and the final call ascribed by the gene voting algorithm are displayed next to the voting profile. Although particular tumors may contain votes from several hierarchical cluster groups, each tumor is assigned membership to one of the four hierarchical cluster groups based on dominance of one particular hierarchical cluster group.

clinical tool. There are at least four classes of diffuse infiltrative gliomas, which are distinct based on the set of genes overexpressed in the tumors. The set of genes that are overexpressed provides key insights and leads for targeted molecular therapies.

High-grade gliomas are characterized by extensive diffuse local and distant infiltration into normal brain tissue, which contributes to the difficulty of treating high-grade gliomas. Among the poor survival clusters, group hierarchical cluster 2B is defined by overexpression of a set of extracellular matrix-related genes, and membership in this group is strongly correlated with poor patient survival. Overexpression of genes involved in matrix structural components, assembly, and modifications may facilitate local invasion and migration of hierarchical cluster 2B tumors through the central nervous system. Extracellular-associated growth factors or modulators are similarly increased in this group including vascular endothelial growth factor (39), IGFBP4 (37), and IGFBP6. In general the inferences from this large-scale expression study support and extend earlier, more focused analyses in gliomas. Overexpressed tumor-associated proteins include S100A4 (40), TGFBI, and MBP1. Extracellular components have been shown to be differentially expressed in gliomas and culture gliomas (41). MGP has been identified previously in large-scale screens (37). Degree of malignancy has been correlated with expression of THBS1 (38). Furthermore, several of these genes have been shown to have specific roles in cell migration/invasion. For instance, S100A4 can control astrocytic cell migration (40). The hierarchical cluster 2B tumors are also characterized by overexpression of well-studied genes *EGFR*, *AKT1*, and *IGFBP2* (Fig. 3). In contrast, the hierarchical cluster 2A poor prognosis glioma group is characterized by overexpression of mitotic and cellular proliferation genes including

antigen of *Ki-67* and *PCNA*. *TOP2A* serves as a similar proliferation antigen, and has been shown to correlate with poor survival (34).

Our finding that there are subclasses within the poor survival gliomas patients, including one with a molecular signature indicative of invasion (hierarchical cluster 2B) and one with a molecular signature associated with proliferation (hierarchical cluster 2A) is in line with previous suggestions that there may be different subclasses of biologically aggressive gliomas (42) and suggests that these subclasses may be susceptible to different molecular inhibitors. These data also highlight the lack of therapies focused on the underlying genetic driving forces of hierarchical cluster 2B. In the future, it will be important to determine whether these molecular subclasses show different growth or invasion patterns and whether they respond differently to anti-invasive *versus* antiproliferative therapies.

The better surviving classes of high-grade gliomas suggest a potentially more differentiated phenotype. Hierarchical cluster 1A is defined by overexpression of genes classified as being involved in neurogenesis, and membership portends the longest survival, with median survival times >6 years. Neurogenesis-related genes highly expressed within this group include *BMP2* (31), *DLL3*, *HDAC4*, *EDNRB*, *HEY2*, and *NTRK2*. These genes play a role in neurogenesis and may preferentially lead to terminal differentiation of the malignant cells. The gliomas characterized by overexpression of genes related to synaptic transmission (hierarchical cluster 1B) are the most heterogeneous grouping of tumors based on survival, and it is unclear what the interpretation of the genes identified should be. This normal neuronal signature may be driven by: (1) a pattern of neuronal gene expression within the glioma cells, (2) an increased content of intact neuronal structures within the less aggressive tumors, or (3) the possibility that genes implicated previously in neuronal development may also play a role in glial tumor development. These results, along with previous studies demonstrating the potential for stem cells within tumors and that tumors of one morphologic phenotype can differentiate into a different morphologic phenotype under growth factor exposure (43–45) suggest important areas for future study. Frequency of treatment, both chemotherapeutic and radiation, was examined for all of the hierarchical cluster groups, and there were no substantial differences.

Chromosome 1p/19q loss has been detected in oligodendrogliomas and associated with better survival (46). To address the impact of this favorable prognostic group, 1p and 19q loss was assessed by loss of heterozygosity. 1p/19q loss was not frequently detected in our sample set: 2 of 11 anaplastic oligodendroglioma specimens (from 9 patients),

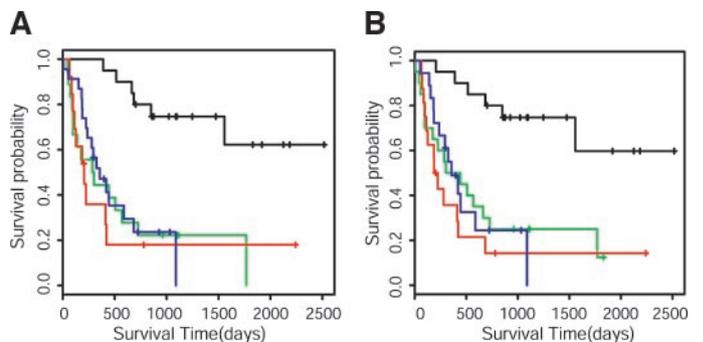


Fig. 5. Survival analysis of gliomas grouped by gene expression. A, Kaplan-Meier survival plots of patients whose tumors were grouped into hierarchical cluster 1A, hierarchical cluster 1B, hierarchical cluster 2A, and hierarchical cluster 2B by hierarchical clustering. Black line, hierarchical cluster 1A group (n = 20). Green line, hierarchical cluster 1B group (n = 18). Red, hierarchical cluster 2A (n = 13). Blue, hierarchical cluster 2B group (n = 23). Log rank P = 0.00011. B, Kaplan-Meier survival plot of the patients as classified by the 44 gene voting protocol. Black line, hierarchical cluster 1A group (n = 20). Green line, hierarchical cluster 1B group (n = 20). Red, hierarchical cluster 2A (n = 16). Blue, hierarchical cluster 2B group (n = 18; log rank P = 0.00022).

1 of 8 anaplastic astrocytoma, and 1 of 7 anaplastic mixed oligo-astrocytoma were identified with 1p/19q loss. There was a trend toward longer survival in those patients with anaplastic gliomas that harbor a 1p/19q loss; however, the power to detect any survival differences based on this specific deletion is low due to the small number of samples. The predictor set presented here is not impacted by presence of the 1p/19q loss, and it is possible that 1p/19q loss is an independent predictor of prognosis.

The strong correlation of microarray gene expression measurements to reverse transcription-PCR measurements highlights the accuracy of current high-density oligonucleotide arrays to quantify gene expression. Furthermore, data from 50 tumors analyzed previously by others on U95A arrays confirm the reliability of the groupings and support that group membership is strongly able to predict patient survival. In both sample sets, tumors were sorted into groups with stronger prognostic information than histologic grading using either hierarchical clustering or a selected gene voting list. This is remarkable and implies that gene expression-based predictors are robust, can be consistently applied at multiple institutions and array platforms, and can still retain information about survival prediction. This is true whether evaluating the principal partial least squares components summarizing the expression of 8,011 probesets, the expression of the 595 survival predictor probesets, or the expression of the 44 voting genes. Thus, the approaches demonstrated here are efficient at generating reliable gene lists for cancer classification even in a highly restricted histologic group of cancers.

Antisera or monoclonal antibodies have already been raised to many of the protein products of the 44 voting gene set. Thus, the proteins encoded by these voting genes may make excellent targets for immunohistochemical development to supplement current histologic practice. We have begun to develop such a panel of reagents for screening high-grade gliomas based on the results presented here. As an example, the well-used antigen detected by Ki-67 antibody is substantially more highly detected in the hierarchical cluster 2A tumors relative to the other groups ( $P = 7 \times 10^{-6}$ ; data not shown). A substantially smaller panel of immunohistochemical reagents may prove as effective for tumor classification and prognosis. Alternatively, the identified 44 genes are an enriched set of genes on which to develop a small panel microarray for focused classification of high-grade gliomas, which may be used in a directed, cost-effective manner.

## ACKNOWLEDGMENTS

The authors wish to express their gratitude to patients participating in this study.

## REFERENCES

- Kleihues P, Sobin LH. World Health Organization classification of tumors. *Cancer* 2000;88:2887.
- Louis DN, Holland EC, Cairncross JG. Glioma classification: a molecular reappraisal. *Am J Pathol* 2001;159:779–86.
- Prayson RA, Agamanolis DP, Cohen ML, et al. Interobserver reproducibility among neuropathologists and surgical pathologists in fibrillary astrocytoma grading. *J Neurol Sci* 2000;175:33–9.
- Coons SW, Johnson PC, Scheithauer BW, Yates AJ, Pearl DK. Improving diagnostic accuracy and interobserver concordance in the classification and grading of primary gliomas. *Cancer* 1997;79:1381–93.
- Burger PC. What is an oligodendroglioma? *Brain Pathol* 2002;12:257–9.
- Mischel PS, Nelson SF, Cloughesy TF. Molecular analysis of glioblastoma: pathway profiling and its implications for patient therapy. *Cancer Biol Ther* 2003;2:242–7.
- Betensky RA, Louis DN, Cairncross JG. Influence of unrecognized molecular heterogeneity on randomized clinical trials. *J Clin Oncol* 2002;20:2495–9.
- Cloughesy TF, Filka E, Kuhn J, et al. Two studies evaluating irinotecan treatment for recurrent malignant glioma using an every-3-week regimen. *Cancer* 2003;97:2381–6.
- Mischel PS, Shai R, Shi T, et al. Identification of molecular subtypes of glioblastoma by gene expression profiling. *Oncogene* 2003;22:2361–73.
- Tortosa A, Vinolas N, Villa S, et al. Prognostic implication of clinical, radiologic, and pathologic features in patients with anaplastic gliomas. *Cancer* 2003;97:1063–71.
- Reavey-Cantwell JF, Haroun RI, Zahurak M, et al. The prognostic value of tumor markers in patients with glioblastoma multiforme: analysis of 32 patients and review of the literature. *J Neurooncol* 2001;55:195–204.
- Bouvier-Labit C, Chinot O, Ochi C, et al. Prognostic significance of Ki67, p53 and epidermal growth factor receptor immunostaining in human glioblastomas. *Neuropathol Appl Neurobiol* 1998;24:381–8.
- Stark AM, Hugo HH, Witzel P, Mihajlovic Z, Mehdorn HM. Age-related expression of p53, Mdm2, EGFR and Msh2 in glioblastoma multiforme. *Zentralbl Neurochir* 2003;64:30–6.
- Li J, Yen C, Liaw D, et al. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. *Science* 1997;275:1943–7.
- Kim S, Dougherty ER, Shmulevich L, et al. Identification of combination gene sets for glioma classification. *Mol Cancer Ther* 2002;1:1229–36.
- Ljubimova JY, Lakhter AJ, Loksh A, et al. Overexpression of alpha4 chain-containing laminins in human glial tumors identified by gene microarray analysis. *Cancer Res* 2001;61:5601–10.
- Nutt CL, Mani DR, Betensky RA, et al. Gene expression-based classification of malignant gliomas correlates better with survival than histological classification. *Cancer Res* 2003;63:1602–7.
- Rickman DS, Bobek MP, Misek DE, et al. Distinctive molecular profiles of high-grade and low-grade gliomas based on oligonucleotide microarray analysis. *Cancer Res* 2001;61:6885–91.
- Sallinen SL, Sallinen PK, Haapasalo HK, et al. Identification of differentially expressed genes in human gliomas by DNA microarray and tissue chip techniques. *Cancer Res* 2000;60:6617–22.
- Shai R, Shi T, Kremen TJ, et al. Gene expression profiling identifies molecular subtypes of gliomas. *Oncogene* 2003;22:4918–23.
- Tanwar MK, Gilbert MR, Holland EC. Gene expression microarray analysis reveals YKL-40 to be a potential serum marker for malignant character in human glioma. *Cancer Res* 2002;62:4364–8.
- Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc Natl Acad Sci USA* 2001;98:31–6.
- Kaplan E, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958;53:457–81.
- Cox DR, Oakes D. *Analysis of survival data*. London: Chapman and Hall; 1990.
- de Jong S. SIMPLS: an alternative approach to partial least squares regression. *Chemometrics and Intelligent Laboratory Systems* 1993;18:251–63.
- Therneau TM, Grambsch PM, Fleming TR. Martingale-based residuals and survival models. *Biometrika* 1990;77:147–60.
- Kaufman L, Rousseeuw PJ. *Finding groups in data, an introduction to cluster analysis*. Brussels: John Wiley & Sons; 1990.
- Cox T F, Cox, MA. *Multidimensional scaling*. London: Chapman and Hall; 1994.
- Burger PC, Vogel FS, Green SB, Strike TA. Glioblastoma multiforme and anaplastic astrocytoma. Pathologic criteria and prognostic implications. *Cancer* 1985;56:1106–11.
- Hosack DA, Dennis G Jr, Sherman BT, Lane HC, Lempicki RA. Identifying biological themes within lists of genes with EASE. *Genome Biol* 2003;4:R70.
- Zhang X, Wu J, Li X, et al. Effects of recombinant human bone morphogenic protein-2 and hyaluronic acid on invasion of brain glioma in vivo. *Zhonghua Yi Xue Za Zhi* 2002;82:90–3.
- Tatenhorst L, Senner V, Puttmann S, Paulus W. Regulators of G-protein signaling 3 and 4 (RGS3, RGS4) are associated with glioma cell motility. *J Neuropathol Exp Neurol* 2004;63:210–22.
- Fischer U, Muller HW, Sattler HP, et al. Amplification of the MET gene in glioma. *Genes Chromosomes Cancer* 1995;12:63–5.
- Holden JA, Townsend JJ. DNA topoisomerase II-alpha as a proliferation marker in astrocytic neoplasms of the central nervous system: correlation with MIB1 expression and patient survival. *Mod Pathol* 1999;12:1094–100.
- Sandberg AC, Engberg C, Lake M, von Holst H, Sara VR. The expression of insulin-like growth factor I and insulin-like growth factor II genes in the human fetal and adult brain and in glioma. *Neurosci Lett* 1988;93:114–9.
- Fiano V, Ghimentti C, Schiffer D. Expression of cyclins, cyclin-dependent kinases and cyclin-dependent kinase inhibitors in oligodendrogliomas in humans. *Neurosci Lett* 2003;347:111–5.
- van den Boom J, Wolter M, Kuick R, et al. Characterization of gene expression profiles associated with glioma progression using oligonucleotide-based microarray analysis and real-time reverse transcription-polymerase chain reaction. *Am J Pathol* 2003;163:1033–43.
- Kawataki T, Naganuma H, Sasaki A, et al. Correlation of thrombospondin-1 and transforming growth factor-beta expression with malignancy of glioma. *Neuropathology* 2000;20:161–9.
- Plate KH, Breier G, Weich HA, Risau W. Vascular endothelial growth factor is a potential tumour angiogenesis factor in human gliomas in vivo. *Nature* 1992;359:845–8.
- Belot N, Pochet R, Heizmann CW, Kiss R, Decaestecker C. Extracellular S100A4 stimulates the migration rate of astrocytic tumor cells by modifying the organization of their actin cytoskeleton. *Biochim Biophys Acta* 2002;1600:74–83.
- Bouterfa H, Darlapp AR, Klein E, et al. Expression of different extracellular matrix components in human brain tumor and melanoma cells in respect to variant culture conditions. *J Neurooncol* 1999;44:23–33.
- von Deimling A, von Ammon K, Schoenfeld D, et al. Subsets of glioblastoma multiforme defined by molecular genetic analysis. *Brain Pathol* 1993;3:19–26.

43. Dai C, Celestino JC, Okada Y, et al. PDGF autocrine stimulation dedifferentiates cultured astrocytes and induces oligodendrogliomas and oligoastrocytomas from neural progenitors and astrocytes in vivo. *Genes Dev* 2001;15:1913–25.
44. Holland EC, Li Y, Celestino J, et al. Astrocytes give rise to oligodendrogliomas and astrocytomas after gene transfer of polyoma virus middle T antigen in vivo. *Am J Pathol* 2000;157:1031–7.
45. Holland EC. Progenitor cells and glioma formation. *Curr Opin Neurol* 2001;14:683–8.
46. Cairncross JG, Ueki K, Zlatescu MC, et al. Specific genetic predictors of chemotherapeutic response and survival in patients with anaplastic oligodendrogliomas. *J Natl Cancer Inst* 1998;90:1473–9.
47. Gracia E, Fischer U, elKahloun A, et al. Isolation of genes amplified in human cancers by microdissection mediated cDNA capture. *Hum Mol Genet* 1996;5:595–600.
48. Reifenberger G, Reifenberger J, Ichimura K, Meltzer PS, Collins VP. Amplification of multiple genes from chromosomal region 12q13–14 in human malignant gliomas: preliminary mapping of the amplicons shows preferential involvement of CDK4, SAS, and MDM2. *Cancer Res* 1994;54:4299–303.
49. Arheden K, Ronne M, Mandahl N, et al. In situ hybridization localizes the human putative oncogene GLI to chromosome subbands 12q13.3–14.1. *Hum Genet* 1989;82:1–2.
50. Ichimura K, Schmidt EE, Goike HM, Collins VP. Human glioblastomas with no alterations of the CDKN2A (p16INK4A, MTS1) and CDK4 genes have frequent mutations of the retinoblastoma gene. *Oncogene* 1996;13:1065–72.
51. Collins VP. Amplified genes in human gliomas. *Semin Cancer Biol* 1993;4:27–32.
52. Buschges R, Weber RG, Actor B, et al. Amplification and expression of cyclin D genes (CCND1, CCND2 and CCND3) in human malignant gliomas. *Brain Pathol* 1999;9:435–42; discussion 2–3.
53. Moriyama T, Kataoka H, Seguchi K, Tsubouchi H, Koono M. Effects of hepatocyte growth factor (HGF) on human glioma cells in vitro: HGF acts as a motility factor in glioma cells. *Int J Cancer* 1996;66:678–85.
54. Riemenschneider MJ, Buschges R, Wolter M, et al. Amplification and overexpression of the MDM4 (MDMX) gene from 1q32 in a subset of malignant gliomas without TP53 mutation or MDM2 amplification. *Cancer Res* 1999;59:6091–6.
55. Steck PA, Saya H. Pathways of oncogenesis in primary brain tumors. *Curr Opin Oncol* 1991;3:476–84.
56. Fitzgibbons PL, Turner RR, Appley AJ, et al. Flow cytometric DNA and nuclear antigen content in astrocytic neoplasms. *Am J Clin Pathol* 1988;89:640–4.
57. Chakravarti A, Delaney MA, Noll E, et al. Prognostic and pathologic significance of quantitative protein expression profiling in human gliomas. *Clin Cancer Res* 2001;7:2387–95.
58. Katayam M, Yoshida K, Ishimori H, et al. Patched and smoothed mRNA expression in human astrocytic tumors inversely correlates with histological malignancy. *J Neurooncol* 2002;59:107–15.
59. Haas-Kogan D, Shalev N, Wong M, et al. Protein kinase B (PKB/Akt) activity is elevated in glioblastoma cells due to mutation of the tumor suppressor PTEN/MMAC. *Curr Biol* 1998;8:1195–8.
60. Mapstone TB. Expression of platelet-derived growth factor and transforming growth factor and their correlation with cellular morphology in glial tumors. *J Neurosurg* 1991;75:447–51.
61. Unterman TG, Glick RP, Waites GT, Bell SC. Production of insulin-like growth factor-binding proteins by human central nervous system tumors. *Cancer Res* 1991;51:3030–6.
62. Westermark B, Magnusson A, Heldin CH. Effect of epidermal growth factor on membrane motility and cell locomotion in cultures of human clonal glioma cells. *J Neurosci Res* 1982;8:491–507.
63. Dahlstrand J, Collins VP, Lendahl U. Expression of the class VI intermediate filament nestin in human central nervous system tumors. *Cancer Res* 1992;52:5334–41.
64. Dittmann L, Axelsen NH, Norgaard-Pedersen B, Bock E. Antigens in human glioblastomas and meningiomas: Search for tumour and onco-foetal antigens. Estimation of S-100 and GFA protein. *Br J Cancer* 1977;35:135–41.
65. Araki K, Nozaki K, Ueba T, Tatsuka M, Hashimoto N. High expression of Aurora-B/Aurora and Ipl1-like midbody-associated protein (AIM-1) in astrocytomas. *J Neurooncol* 2004;67:53–64.
66. Wharton SB, Chan KK, Anderson JR, Stoeber K, Williams GH. Replicative Mcm2 protein as a novel proliferation marker in oligodendrogliomas and its relationship to Ki67 labelling index, histological grade and prognosis. *Neuropathol Appl Neurobiol* 2001;27:305–13.
67. Halfter H, Kremerskothen J, Weber J, et al. Growth inhibition of newly established human glioma cell lines by leukemia inhibitory factor. *J Neurooncol* 1998;39:1–18.
68. Rao JS, Steck PA, Mohanam S, et al. Elevated levels of M(r) 92,000 type IV collagenase in human brain tumors. *Cancer Res* 1993;53:2208–11.