Molecular Classification of Human Diffuse Gliomas by Multidimensional Scaling Analysis of Gene Expression Profiles Parallels Morphology-Based Classification, Correlates with Survival, and Reveals Clinically-Relevant Novel Glioma Subsets

Gregory N. Fuller[1]; Kenneth R. Hess[2]; Chang Hun Rhee[*]; W. K. Alfred Yung[3]; Raymond A. Sawaya[4]; Janet M. Bruner[1]; Wei Zhang[1]

Departments of [1]Pathology, [2]Biostatistics, [3]Neuro-Oncology, and [4]Neurosurgery, The University of Texas M. D. Anderson Cancer Center, Houston

* Present address: Department of Neurosurgery, Korea Cancer Center Hospital, Seoul, Korea

A preliminary report of some of the data in this study was presented at the annual meeting of the American Association of Neuropathologists in Atlanta on June 10, 2000.

There are several currently employed classification systems for diffuse gliomas that sort tumors based on histological features. Contemporary molecular techniques, however, offer the promise of improved tumor classification and resultant patient stratification for treatment and prognosis. In particular, gene expression profiling has shown exceptional promise for providing an alternative and more objective molecular approach to glioma classification. In this study, we used cDNA array technology to profile the gene expression of 30 primary human glioma tissue samples comprising 4 different glioma subtypes as defined by current World Health Organization (WHO 2000) criteria: glioblastoma (GM, WHO grade IV), anaplastic astrocytoma (AA, WHO grade III), anaplastic oligodendroglioma (AO, WHO grade III), and oligodendroglioma (OL, WHO grade II). Gene expression data alone were used to group the tumors using multidimensional scaling, which is an unsupervised statistical method. Results show that impressive separation of the 4 glioma subtypes can be achieved solely on the basis of molecular data. In addition, a subcluster of 3 glioblastomas was identified as distinct from other GMs and from the oligodendroglial tumors. These 3 patients have shown extended survival compared to other GMs in the study. Survival analysis of the full data set revealed a good correlation with the molecular classification. Results of this proof-of-principle study demonstrate that molecular profiling alone can recapitulate conventional histologic classification and grading with high fidelity. In addition, results show that the molecular approach to tumor classification can generate clinically meaningful patient stratification, and, more importantly, is an efficient class-discovery tool for human gliomas, permitting the identification of previously unrecognized, clinically relevant tumor subsets.

Brain Pathol 2002;12:108-116

Introduction

The current paradigm for brain tumor diagnosis and classification, as exemplified by the recently revised World Health Organization Classification of Tumours of the Nervous System (17), is based primarily on morphologic pattern recognition: the identification of similarities between the phenotypic characteristics expressed by tumor cells compared to those of normal central nervous system constituents as assessed by light microscopic examination of H&E-stained tissue sections, immunohistochemistry, and transmission electron microscopy. Although the morphologic approach has unquestionably been of considerable utility, there are nevertheless a number of shortcomings. Using traditional phenotypic criteria, for example, the identification and classification of some tumor types, such as mixed oligoastrocytomas, is highly subjective and overly dependent upon the individual pathologist's relative weighting of various morphologic characteristics. Currently prevailing histology-based classification methods also do not permit accurate prediction of clinical behavior or response to specific therapeutic agents and regimens for individual patients within a given histologic

Corresponding author:
Gregory N. Fuller, MD, PhD, Department of Pathology-085, M. D. Anderson Cancer Center, 1515 Holcombe Blvd., Houston, TX 77030 (e-mail: gfuller@mdanderson.org)

rubric, as, for example, is the case for anaplastic astrocytoma, in which individual patient response to treatment and survival varies significantly. Another problem is the failure of morphology-based classifications to accurately predict individual patient sensitivity to the toxic effects of various therapies, such as necrosis of irradiated brain parenchyma.

Thus, there is a need on many levels for a more precise, effective, and objective approach to brain tumor diagnosis, classification, grading, and prognostication. Although a large corpus of information has been amassed on the molecular characteristics of different types of gliomas, such as the different alterations seen in primary versus secondary glioblastomas, most of this knowledge has not translated into tangible advancements in either treatment or prognosis. The first significant development in clinically-relevant molecular classification of brain tumors has been the recently recognized association between the combined loss of heterozygosity (LOH) for chromosomes 1p and 19q and therapeutic responsiveness of a subset of gliomas that display oligodendroglial differentiation features (5, 25). This significant discovery is undoubtedly only the beginning of meaningful stratification of the diffuse gliomas based on molecular characterization.

To facilitate the molecular classification of tumors, a number of contemporary technologies are available that permit detailed, exhaustive characterization of the tumor genome, transcriptome and proteome. Of these, transcriptome profiling is currently the most widely employed for molecular classification studies (1, 2, 6, 8, 9, 12-14, 16, 20, 21, 23, 27). There are several expression array methodologies in use, which differ in physical medium and number of genes analyzed. Based on the relative density of arrayed sequences, and hence the number of genes that can be analyzed in a single hybridization experiment, expression arrays can be roughly separated into 2 categories: low-density and high-density. Low-density arrays, commonly nylon-membrane-based, typically contain hundreds to a few thousand genes, whereas high-density arrays, commonly glass slide-based, usually contain between a few thousand and tens of thousands of genes arrayed on a single microscopy slide. Although high-density arrays may provide much more raw data on gene expression levels compared to low-density arrays, this is not necessarily a requirement, or even a desirable attribute, for meaningful molecular classification and tumor stratification studies. Even a few genes that exhibit highly-differential expression patterns may serve as robust separators for class distinction and class discovery experiments. Glioma genomics research employing expression array profiling followed by tissue microarray immunohistochemical confirmation has already identified one progression- and survival-associated marker, insulin-like growth factor binding protein 2 (IGFBP2), which is uniformly and differentially overexpressed only in glioblastoma (13, 23).

In a proof-of-principle study, we carried out a gene expression profiling experiment using low-density arrays containing 588 genes with 30 primary glioma tissue samples comprising 4 histologic phenotypes: low-grade oligodendroglioma, O (7 tumors); anaplastic oligodendroglioma, AO (6 tumors); anaplastic astrocytoma, AA (4 tumors); and glioblastoma, GM (13 tumors). For the purposes of this study, only tumors with these 4 classical histologies were selected for inclusion in the study in order to establish a baseline for molecular classification and to address the question of whether the members of histologically-uniform tumor categories share a common molecular basis. Gliomas diagnosed as having mixed oligoastrocytic features were specifically excluded from the study. In this study, we present the results of a data mining algorithm analysis of the complete gene expression data set performed to determine if gene expression profiling can be used to generate a molecular classification of diffuse gliomas that meaningfully stratifies patients with respect to survival.

Methods

Primary glioma tissue samples. All primary glioma tissues were acquired from the Brain Tumor Program tissue bank of The University of Texas M. D. Anderson Cancer Center. Tissue bank specimens were quick-frozen shortly after surgical removal and stored at -80°C. Although it is not known whether or to what extent the time delay between tumor removal and tumor freezing affects gene expression, all of the tumor tissue samples used in this study experienced a similar length of delay. Thus, the tumor harvesting procedure would affect all samples in a similar and unbiased manner and would not be expected to contribute to the difference in gene expression patterns among samples. Hematoxylin-eosin (H&E)-stained frozen tissue sections are routinely prepared from all tissue bank specimens for screening purposes. All tissue specimens for cDNA array analysis were screened by a board-certified neuropathologist (GNF) and the diagnoses were independently confirmed by a second board-certified neuropathologist (JMB). The glioma tissue blocks were specifically selected for study based on dense tumor cellularity and relative puri-

ty. There was minimal normal contamination and minimal variation between samples in this regard. Glioma classification and grading were performed according to current World Health Organization criteria (WHO 2000) (17).

Isolation of total RNA and mRNA. Tissue samples were ground to powder under frozen conditions and 0.3 to 1.5 g of tissue powder was lysed in the lysis buffer TRI Reagent (Molecular Research Center, Cincinnati, OH). RNA isolation was then performed as described in detail previously (13).

Hybridization to cDNA expression array blots. The cDNA fragments representing 588 human genes with known functions and known tight transcriptional controls were immobilized in duplicate onto a nylon membrane (Clontech Laboratories, Inc.). Microarray hybridizations were performed as described previously (13).

Statistical analysis. The statistical algorithms used for our analyses are available as program libraries that can be called from S-Plus (version 3.4), which is an object-oriented, highly graphical statistical computing environment (MathSoft, Seattle, Wash). S-Plus (26) is running on a Digital AlphaServer 4100 Model 5/400 with Digital Unix Version 3.2. Prior to data analysis, the values generated from the hybridized images were preprocessed. The values were thresholded at one (ie, values less than 1 were set to 1), missing values were set to the median values, values from a given array were normalized to the overall median for that array (ie, values were divided by the overall median), and base 10 logarithms were computed for each value.

Multidimensional scaling. There are a number of statistical methods available for cluster analysis of large multidimensional data sets (3, 10, 11, 15, 19, 24). Multidimensional scaling (MDS), also called principal coordinate analysis, is similar to principal component analysis (3, 15). MDS is an iterative process that shares with principal component analysis the same goal of constructing a low-dimensional representation of high-dimensional data. Multidimensional scaling attempts to find configurations of the data in a lower dimensional representation such that the relative distance between objects is similar to that in the original, higher dimensional representation. Having a lower dimensional representation of the data facilitates visualization and analysis of patterns in the data. MDS has an advantage over hierarchical clustering in that multiple dimensions are used to relate objects to one another (in an unconstrained manner). "Clusters" of objects can be identified on an MDS plot by their relative proximity (ie, clusters consist of relatively tightly bunched points that are separated from other points and other point clusters on the MDS plot). Typically, 2 dimensions are adequate for visualization. Several permutations of the MDS approach are available. In the analysis presented here, we used metric MDS with the complement of the correlation coefficient as the distance metric.

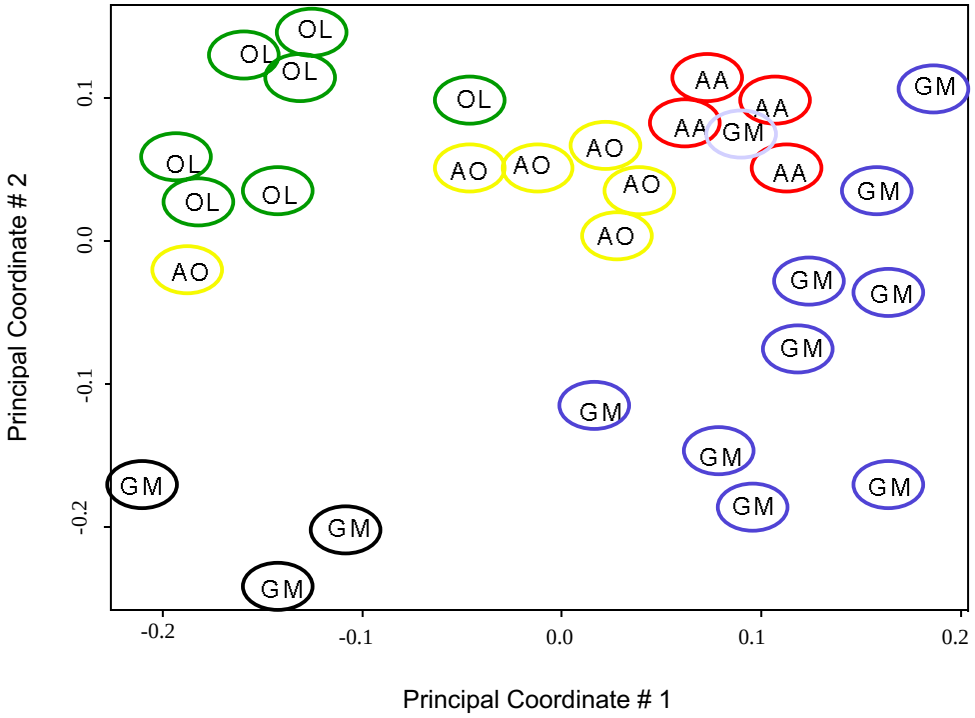
Survival analysis. To assess the relationship between the gene expression profiles and survival, we included the first 3 principal coordinates in a Cox proportional hazards regression model (with and without the standard clinical variables). We report the likelihood ratio p-values for the estimated parameters. We also report likelihood-based r-squared values that approximate the amount of variation in survival time that is explained by a particular model. We use graphs to visualize the association between survival and gene profile by plotting survival time against the linear predictor from the Cox model with the first 3 principal coordinates. We then add a smooth curve to this plot that shows how the median survival changes as a function of the linear predictor. This smooth curve is generated by using a moving average technique that moves a window over the range of the linear predictor. Within the window, the median survival time is computed with a weighted Kaplan-Meier calculation such that points closer to the middle of the window are given more weight.

Results

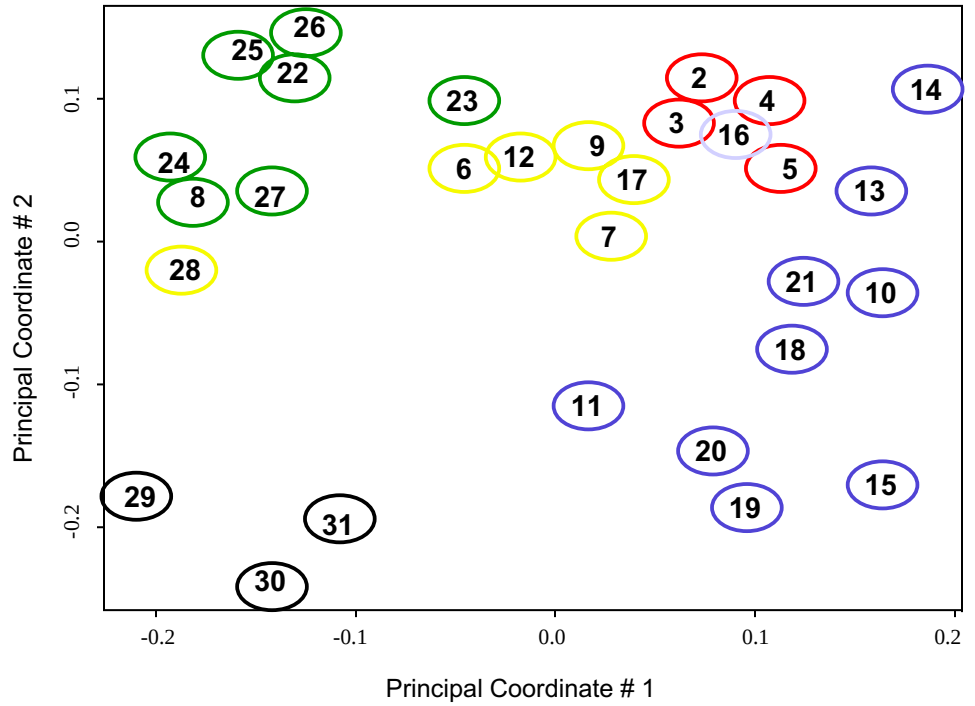
The 30 glioma tissue samples used for this study were classified according to current WHO criteria and subjected to gene expression profiling using cDNA microarrays. The microarrays were quantified and a spreadsheet was generated containing expression results of all 588 genes analyzed for the 30 samples. Our goal was to determine whether the gene expression data alone were sufficient to stratify the tumors in a clinically meaningful manner.

Figure 1. (Opposing page) *Molecular classification of diffuse gliomas by multidimensional scaling analysis of gene expression profiles.* A scatter plot is shown of the first two MDS principal coordinates with each point labeled with the tumor's histologic subtype (**A**) and sample identification (**B**). Unsupervised analysis of gene expression profiles produced clusters of samples largely corresponding to recognized histologic categories. See text for discussion of clusters and "outlier" cases. Abbreviations: AA, anaplastic astrocytoma (red ovals); AO, anaplastic oligodendroglioma (yellow ovals); GM, glioblastoma (dark blue, light blue and black ovals).

A


Principal Coordinate # 2
Principal Coordinate # 1
0.1
0.0
-0.1
-0.2
-0.2
-0.1
0.0
0.1
0.2
OL
AA
GM
AO
GM

B


Principal Coordinate # 2
Principal Coordinate # 1
0.1
0.0
-0.1
-0.2
-0.2
-0.1
0.0
0.1
0.2

Multidimensional scaling analysis. In the 2-dimensional MDS graph (Figure 1A, B), OL tissues form a fairly tight cluster in the upper left quadrant of the graph (green ovals). The anaplastic gliomas (AOs and AAs) form 2 distinct clusters in the middle of the graph (yellow and red ovals, respectively). The GMs, not surprisingly, are least tightly clustered, with a large group on the right-hand side of the graph centered roughly in the lower right quadrant (dark blue ovals), a separate cluster of 3 tumors in the left lower corner of the graph (black ovals), and a single case located in the center of the AA cluster (light blue oval). The impressive separation of the 4 diffuse glioma histologic subtypes, based solely on the unsupervised analysis of gene expression profiles, reflects a molecular basis for the morphologic and immunophenotypic differentiation features that are conventionally used to subclassify gliomas.

In addition to the separate cluster of 3 GMs (black ovals in lower left quadrant), 3 "outliers" are seen in the MDS plot (Figure 1A, B): case 28 (an AO clustering with the OLs), case 23 (an OL clustering with the AOs), and case 16 (a GM clustering with the AAs). In each of these cases, retrospective review of the histologic slides and clinical record was informative. For case 28 (Figure 1A, B), a review of all H&E-stained tissue sections showed a densely cellular oligodendroglioma with classical morphologic features (regular round nuclei with perinuclear halos). Although mild vascular prominence was seen focally, true microvascular proliferation was not present. A review of the MRI studies showed lack of contrast enhancement, consistent with the histologic absence of vascular proliferation. The MIB-1 (Ki-67 antigen) labeling index (LI) for this tumor was 5.1%. This compares with a mean LI of 5.2% for the other OLs evaluated in the study versus 11.1% for the AOs. Thus, a review of the histology, proliferation indices and neuroimaging studies strongly suggest that this oligodendroglioma is most appropriately classified as a low-grade oligodendroglioma (OL), rather than an AO as originally diagnosed at the time of surgery.

For case 23 (Figure 1A, B), review of the histologic slides revealed a glioma in which areas of classical oligodendroglial differentiation were present. It was upon these areas that the original diagnosis of OL was based. In addition, however, in other regions of the tumor, features suggestive of astrocytic differentiation were present, including pleomorphic nuclei and eosinophilic cytoplasm with multiple delicate cell processes. These latter cells did not possess the typical cytoplasmic configuration of either minigemistocytes or gliofibrillary oligodendrocytes (which are 2 GFAP-positive neoplastic cell types whose presence is accepted in "pure" oligodendrogliomas). Thus, although the original diagnosis of oligodendroglioma was reasonable, in retrospect the features taken as a whole suggest the possibility of a mixed oligoastrocytoma, which may account for the MDS sorting of this case to a position intermediate between the low-grade oligodendrogliomas (OL) and the astrocytic tumors (AAs and GMs).

The MDS plot shows one GM that sorts to the center of the AA cluster (case 16, light blue oval, in Figure 1A, B). Review of the histology showed a high-grade astrocytic neoplasm that met all of the WHO 2000 criteria for glioblastoma (WHO grade IV): pleomorphism, prominent mitotic activity, florid microvascular proliferation, and foci of necrosis with pseudopalisading. Review of the clinical record, however, did reveal one major difference between this tumor and the rest of the GMs included in the study: it was a secondary GM. The patient is a 32-year-old man who originally presented 4 years earlier with a non-contrast enhancing lesion biopsied as low-grade astrocytoma (WHO grade II). Based on the presence of newly developed contrast enhancement, anaplastic progression was suspected clinically and histologic evaluation of the subsequent surgical resection specimen showed glioblastoma, which was analyzed in this study. Current follow-up at the time of this report shows the patient to have stable residual disease 3 years after resection. In contrast, the other GMs included in the present study were primary GMs.

Not all of the subclusters within a given histologic rubric in the MDS plot could be explained (or even a tentative hypothesis derived) through re-examination of the histology and clinical record. For example, there appear to be 2 subclusters within the OL cluster, each comprising 3 tumors (Figure 1). Review of the tissue sections and clinical records, including MRI studies, for these 6 cases did not reveal any obvious clues as to the origin of this segregation. It is possible that the separation is an epiphenomenon. It is also possible that the subclustering is based on relevant molecular differences between the 2 groups that are not manifest at the routine light microscopic, immunophenotypic, or neuroimaging levels. Further molecular dissection of these associations, together with clinical follow-up, may prove fruitful.

It is obvious from even a cursory examination of the MDS plot that the tumors have sorted according to grade in a roughly diagonal fashion, running from the lowest grade gliomas located in the upper left quadrant of the graph to the highest grade tumors in the lower right quadrant. In this continuum, it proved informative

to examine tumors located at the borders or intersections between major clusters. For example, one AA was located at the border between the main AA group and the GM cluster (case 5 in Figure 1A, B). Review of the clinical history for this case suggested an AA in progression to GM; although no vascular proliferation was identified in the surgically resected tissue, the MRI scan showed the presence of a focus of contrast-enhancement. Radiologic-pathologic correlation studies have shown that the radiologic presence of contrast-enhancement in a glioma usually indicates the histologic presence of vascular proliferation. Thus, although the resected tissue was appropriately graded as AA based on the morphologic features seen in the material available for examination, it is likely that vascular proliferation was present in the unresected portion of the neoplasm, which if sampled would have warranted upgrading to GM.

Two AOs with foci of tumor necrosis, pseudopalisading and florid microvascular proliferation were included among the cases analyzed in this study (tumors 12 and 17 in Figure 1A, B). It is notable that they both sorted unequivocally with the other 3 AOs evaluated. Thus, the present molecular classification based on MDS analysis of expression profiles supports current WHO classification of such tumors as AOs, rather than as GMs as was often done in the past.

Perhaps the most striking "incongruency" or "discrepancy" seen in the MDS plot is the separate clustering of a group of 3 GMs (black ovals in the lower left quadrant of Figure 1A, B) away from the larger main GM cluster. Review of the histologies of these 3 tumors disclosed in each case a high-grade glioma meeting current WHO criteria for glioblastoma, including the presence of foci of tumor necrosis with pseudopalisading and florid microvascular proliferation. The 3 gliomas shared a vague morphological similarity, with large areas of dense vascularity. The histologic features were not those of classical anaplastic oligodendroglioma and these 3 tumors clustered remotely from the AOs as well as the GMs (Figure 1). These 3 patients were all still alive at last contact with follow-ups of 3, 21, and 26 months, compared to the median GM survival of 12 months, suggesting that they have better than average prognoses as compared to other GMs. Further molecular genetic analysis of these tumors is in progress.

Survival analysis and glioma gene expression data. Among the 30 glioma patients included in this study, we observed 15 deaths during the follow-up period, with a median follow-up of 34 months among the 15 patients alive at last contact. Using Kaplan-Meier estimates, the overall median survival was 27 months with 79% alive at 1 year and 54% alive at 2 years. Among the 13 GM patients, the median survival was 12 months (with 10 deaths), while among the 17 remaining patients the median survival was not reached (with 5 deaths). A Cox model with only the first principal coordinate had a likelihood ratio p-value of 0.0047 with a likelihood-based r-squared value of 0.30. (The r- squared value approximates the amount of variation in survival time that is explained by the model.) A Cox model with the first 3 principal coordinates had a likelihood ratio p-value of 0.0002 with an r-squared value of 0.48. In contrast, a Cox model with the 4 standard clinical covariates (histology [1 = GM, 0 = other], age, Karnofsky performance status, and extent of resection [1 = gross total, 0 = other]), had a likelihood ratio p-value of 0.0013 and an r-squared value of 0.45. After adjusting for these clinical covariates, the likelihood ratio p-value for the first 3 MDS principal coordinates was 0.0032 and the r-squared value for the combined model was 0.65. These results must of course be viewed with caution because Cox regression is a large sample method and fitting a 7-parameter model to data from 30 patients with 15 deaths may be spreading the data thin. To visualize the association between survival and gene profile we plotted survival time against the linear predictor from the Cox model with the 3 principal coordinates. We added a smooth curve to this plot that shows how the median survival changes as a function of the linear predictor (Figure 2). As seen in Figure 2, the median survival declines from more than 250 weeks for patients with the lowest values of the linear predictor to about 125 weeks for patients with linear predictor values near the median. The median survival continues to decline with increasing values of the linear predictor to about 50 weeks for patients with the highest values. These results, therefore, show a close correlation between MDS clustering and survival. As mentioned previously, the 3 GM patients that clustered separately from the rest were all still alive at last contact with follow-ups of 3, 21, and 26 months, compared to the median GM survival of 12 months.

Discussion

Historically, tumor classification has often been more of an art than a science. Varying morphologic criteria have been used. Without objective parameters, tumor grading and classification becomes significantly subjective and dependent on the individual diagnostician's personal biases. Classification systems for the diffuse gliomas, in particular, have undergone a number of revisions and permutations over the last 70 years. The 3 sys-

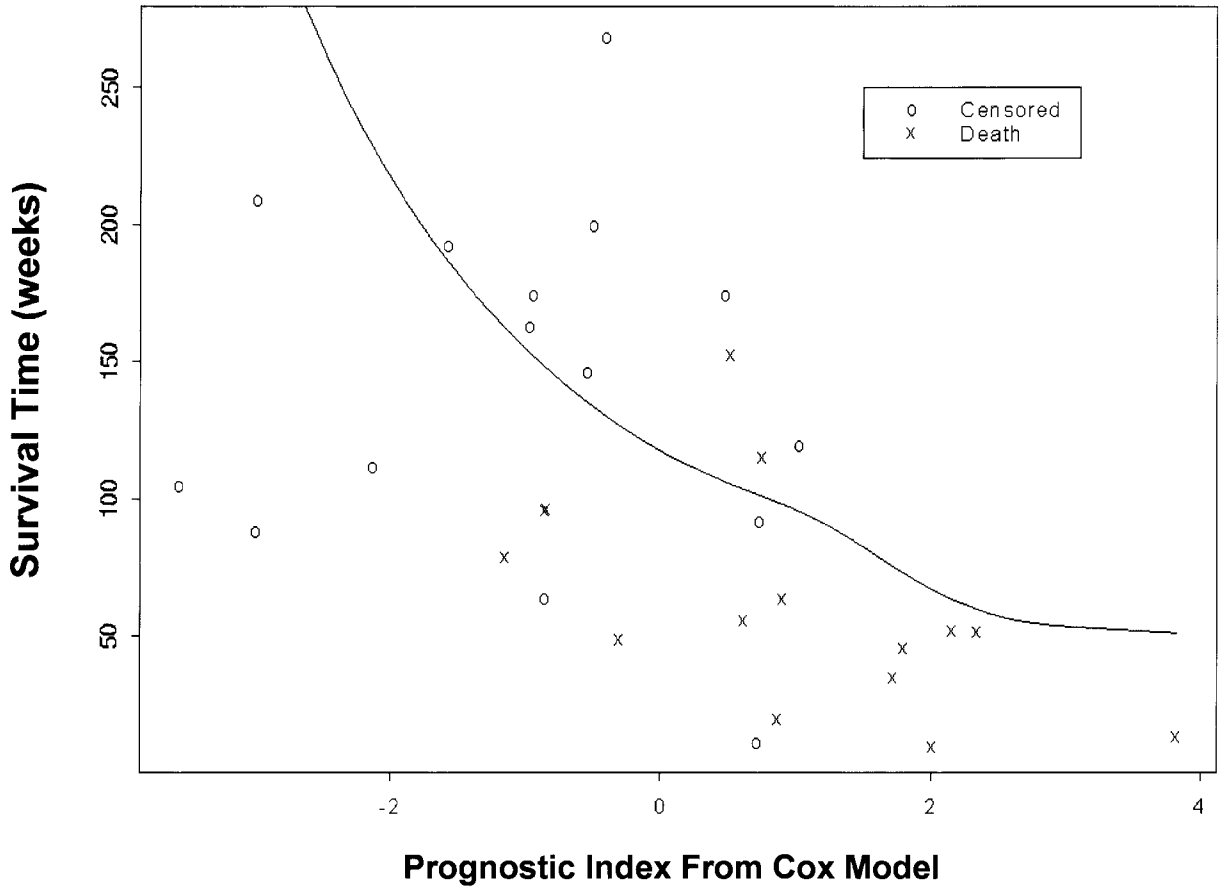


Figure 2. *Survival analysis.* A smoothed scatter plot is shown of follow-up time versus a prognostic index generated from the first three MDS principal coordinates, with points labeled according to the patients' vital status at last contact. The prognostic index significantly correlates with survival ($p = 0.0002$) even after adjustment for clinical prognostic factors ($p = 0.0032$).

tems currently in widespread use (WHO, St. Anne-Mayo, modified Ringertz) are all based on an assessment of morphologic features and pattern recognition (4, 7, 17, 18, 22).

In this study, we used an unsupervised statistical method, multidimensional scaling (MDS), to group 30 gliomas solely on the basis of their gene expression profiles. MDS-based analysis resulted in good separation of the 4 classes of diffuse gliomas evaluated. The impressive molecular subtyping included the clustering of 2 cases of anaplastic oligodendroglioma with florid microvascular proliferation, necrosis and pseudopalisading (features shared with glioblastoma) with other AOs rather than with GMs. Gene expression profiling also separated low-grade oligodendrogliomas (OLs) from anaplastic oligodendrogliomas (AOs), as has also been reported by Watson et al (27). Gliomas with oligodendroglial differentiation features were well separated from those with astrocytic features. Most importantly, molecular classification of the diffuse gliomas correlated well with survival.

Review of the morphologic features and clinical records of the few "outliers" that clustered with tumors of other histologic diagnoses was informative, with a feature or association identified in every case that might explain the discrepant MDS locus, and also provided a stimulus and direction for further investigation. One of the most significant observations of this molecular classification experiment was the identification of a clustered subset of 3 GMs that had not been recognized on standard histologic evaluation. The potential clinical relevance of this molecular subgroup is indicated by longer survivals at last follow-up compared to other GMs in the study. This illustrates the "class discovery" function of gene expression profiling, ie, expression-based MDS analysis may serve to identify unique and potentially

clinically important features of otherwise histologically unremarkable members of a given tumor type.

MDS analysis reduces multidimensional data sets (588 dimensions in the present case) to 2 dimensions in which the clustered objects are graphically represented. The exact meaning of these 2 dimensions in terms of genes is not clear. The identification of genes that yield the maximum separation of the 4 groups evaluated in this study would require a supervised statistical approach such as linear discriminant analysis. However, use of supervised approaches is based on the assumption that the groupings (histologic tumor diagnoses in this case) are correct to begin with, and this may not be a valid assumption. For example, the 3 GMs in the present study that clustered away from the other GMs and also away from the high-grade oligodendrogliomas likely warrant separate recognition based on the patients' increased survival. The morphology-based lumping of these tumors with other GMs or with AOs would only coarsen our nosologic resolution.

It is also evident from a perusal of the MDS plot that the glioma patients form a continuum and there are cases positioned in the boundary zones between adjacent groups. For example, one OL located approximately halfway between the central OL cluster and the AA cluster upon histologic re-examination showed morphologic features suggestive of possible astrocytic differentiation (mixed oligoastrocytoma). Another example is provided by an AA located at the boundary between the AA cluster and the GM cluster, in which review of the clinical record revealed the presence of focal contrast enhancement on the MRI scan, suggesting the presence of vascular proliferation which, had it been included in the resected tissue, would have resulted in upgrading to GM. This case was morphologically at the interface of AA and GM, likely representing an AA in anaplastic progression to GM, and the MDS plot might be reflecting this fact.

This proof of-principle study clearly demonstrated that molecular classification of the gliomas based solely on gene expression data is feasible and potentially informative. The challenge is how to exploit this capability to derive more effective treatment strategies. It is promising that MDS-based molecular subtyping of the diffuse gliomas correlates with patient survival. Molecular classification also offers an independent prognostic assessment separate from that provided by traditional histologic classification.

Summary

Gene expression profiling and other molecular studies performed in our laboratory and others provide general support for the current morphology-based WHO classification of the diffuse gliomas. Molecular classification by MDS analysis of expression data, however, provides information that is much more highly nuanced and objective compared to that obtained by histologic assessment. In addition, molecular classification studies are likely to identify novel, clinically-relevant glioma subsets that warrant additional detailed investigation.

References

1. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J Jr, Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Staudt LM (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403:503-11.
2. Bittner M, Meltzer P, Chen Y, Jiang Y, Seftor E, Hendrix M, Radmacher M, Simon R, Yakhini Z, Ben-Dor A, Sampas N, Dougherty E, Wang E, Marincola F, Gooden C, Lueders J, Glatfelter A, Pollock P, Carpten J, Gillanders E, Leja D, Dietrich K, Beaudry C, Berens M, Alberts D, Sondak V (2000) Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature* 406:536-540.
3. Borg I, Groenen P (1977) *Modern multidimensional scaling: theory and applications.* New York: Springer-Verlag.
4. Burger PC, Vogel FS, Green SB, Strike TA (1985) Glioblastoma multiforme and anaplastic astrocytoma. Pathologic criteria and prognostic implications. *Cancer* 56:1106-1111.
5. Cairncross JG, Ueki K, Zlatescu MC, Lisle DK, Finkelstein DM, Hammond RR, Silver JS, Stark PC, MacDonald DR, Ino Y, Ramsay DA, Louis DN (1998) Specific genetic predictors of chemotherapeutic response and survival in pateints with anaplastic oligodendrogliomas. *Journal of the National Cancer Institute* 90:1473-1479.
6. Caskey LS, Fuller GN, Bruner JM, Yung WKA, Yung WKA, Sawaya RA, Holland EC, Zhang W (2000) Toward a molecular classification of the gliomas: histopathologies, molecular genetics, and gene expression profiling. *Histology and Histopathology* 15:971-981.
7. Daumas-Duport C, Scheithauer BW, O'Fallon J, Kelly P (1988) Grading of astrocytomas. A simple and reproducible method. *Cancer* 62:2152-2165.
8. DeRisi J, Penland L, Brown PO, Bittner ML, Meltzer PS, Ray M, Chen Y, Su YA, Trent JM (1996) Use of a cDNA microarray to analyze gene expression patterns in human cancer. *Nat Genet* 14:457-460.
9. Duggan DJ, Bittner M, Chen Y, Meltzer P, Trent JM (1999) Expression profiling using cDNA microarrays. *Nat Genet* (suppl) 21:10-14.

10. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 95:14863-14868.
11. Everitt BS (1993) *Cluster Analysis*, 3rd edition. London: Arnold.
12. Fuller GN, Hess KR, Rhee, Bruner JM, Sawaya R, Yung WKA, Zhang W (2000) Molecular classification of human gliomas by gene expression profiling. *J Neuropathol Exp Neurol* 59:422 (Abstract)
13. Fuller GN, Rhee CH, Hess K, Caskey L, Wang RP, Bruner JM, Yung A, Zhang W (1999) Reactivation of insulin-like growth factor binding protein II expression during glioblastoma transformation revealed by parallel gene expression profiling. *Cancer Res* 59:4228-4232.
14. Golub TR, Slonim D, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 285:531-537.
15. Johnson RA, Wichern DW (1982) *Applied Multivariate Statistical Analysis.* Englewood Cliffs: Prentice-Hall.
16. Khan J, Simon R, Bittner M, Chen Y, Leighton SB, Pohida T, Smith PD, Jiang Y, Gooden GC, Trent JM, Meltzer PS (1999) Gene expression profiling of alveolar rhabdomysarcoma with cDNA microarrays. *Cancer Res* 58:5009-5013.
17. Kleihues P, Cavenee WK (2000) *Pathology and genetics of tumours of the nervous system. World Health Organization Classification of Tumours of the Nervous System.* Lyon: IARC Press.
18. Nelson JS, Tsukada Y, Schoenfeld D, Fulling K, Lamarche J, Peress N. Necrosis as a prognostic criterion in malignant supratentorial astrocytic gliomas. *Cancer* (1983) 52:550-554.
19. Pison G, Struyf A, Rousseeuw PJ (1999) Displaying a clustering with CLUSPLOT. *Computational Statistics and Data Analysis* 30:381-391.
20. Rhee CH, Hess K, Jabbur J, Ruiz M, Yang Y, Chen S, Chenchik A, Fuller GN, Zhang W (1999) cDNA expression array reveals heterogeneous gene expression profiles in three glioblastoma cell lines. *Oncogene* 18:2711-2717.
21. Rhee CH, Ruan S, Chen S, Chenchik A, Levin VA, Yung AW, Fuller GN, Zhang W (1999) Characterization of cellular pathways involved in glioblastoma response to the chemotherapeutic agent 1,3-bis(2-chloroethyl)-1-nitrosourea (BCNU) by gene expression profiling. *Oncology Reports* 6:393-401.
22. Ringertz N (1950) Grading of gliomas. *Acta Pathol Microbiol Scand* 27:51-64.
23. Sallinen S-L, Sallinen PK, Haapasalo HK, Helin HJ, Helen PT, Schraml P, Kallioniemi O-P, Kononen J (2000) Identification of differentially expressed genes in human gliomas by DNA microarray and tissue chip techniques. *Cancer Res* 60:6617-6622.
24. Scheibler D, Schneider W (1985) Monte Carlo Tests of the accuracy of cluster analysis algorithms: a comparison of hierarchical and nonhierarchial methods. *Multivar Behav Res* 20:282-304.
25. Smith JS, Perry A, Borell TJ, Lee HK, O'Fallon J, Hosek SM, Kimmel D, Yates A, Burger PC, Scheithauer BW, Jenkins RB (2000) Alterations of chromosome arms 1p and 19q as predictors of survival in oligodendrogliomas, astrocytomas, and mixed oligoastrocytomas. *J Clin Oncol* 18:636-645.
26. Venables WN, Ripley BD (1999) *Modern Applied Statistics with S-Plus*. London: Springer.
27. Watson MA, Perry A, Budhjara V, Hicks C, Shannon WD, Rich KM. (2001) Gene expression profiling with oligonucleotide microarrays distinguishes World Health Organization grade of oligodendrogliomas. *Cancer Research* 61: 1825-1829.