Classification of Human Astrocytic Gliomas on the Basis of Gene Expression: A Correlated Group of Genes with Angiogenic Activity Emerges As a Strong Predictor of Subtypes[1,2]

Sophie Godard, Gad Getz, Mauro Delorenzi, Pierre Farmer, Hiroyuki Kobayashi, Isabelle Desbaillets, Michimasa Nozaki, Annie-Claire Diserens, Marie-France Hamou, Pierre-Yves Dietrich, Luca Regli, Robert C. Janzer, Philipp Bucher, Roger Stupp, Nicolas de Tribolet, Eytan Domany, and Monika E. Hegi[3]

Laboratory of Tumor Biology and Genetics [S. G., H. K., I. D., M. N., A-C. D., M-F. H., N. d. T., M. E. H.] of the Department of Neurosurgery [L. R., N. d. T.], Multidisciplinary Oncology Center [R. S.], and Division of Neuropathology [R. C. J.], University Hospital (CHUV), 1011 Lausanne, Switzerland; Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot 76100, Israel [G. G., E. D.]; Swiss Institute of Bioinformatics [M. D., P. F., P. B.]; and NCCR Molecular Oncology [M. D., P. F., M. E. H.], Swiss Institute for Experimental Cancer Research, 1066 Epalinges, Switzerland; and Department of Oncology, Hôpital Universitaire Genève, Switzerland [P-Y. D.]

ABSTRACT

The development of targeted treatment strategies adapted to individual patients requires identification of the different tumor classes according to their biology and prognosis. We focus here on the molecular aspects underlying these differences, in terms of sets of genes that control pathogenesis of the different subtypes of astrocytic glioma. By performing cDNA-array analysis of 53 patient biopsies, comprising low-grade astrocytoma, secondary glioblastoma (respective recurrent high-grade tumors), and newly diagnosed primary glioblastoma, we demonstrate that human gliomas can be differentiated according to their gene expression. We found that low-grade astrocytoma have the most specific and similar expression profiles, whereas primary glioblastoma exhibit much larger variation between tumors. Secondary glioblastoma display features of both other groups. We identified several sets of genes with relatively highly correlated expression within groups that: (*a*) can be associated with specific biological functions; and (*b*) effectively differentiate tumor class. One prominent gene cluster discriminating primary *versus* nonprimary glioblastoma comprises mostly genes involved in angiogenesis, including *VEGF fms-related tyrosine kinase* 1 but also *IGFBP2,* that has not yet been directly linked to angiogenesis. *In situ* hybridization demonstrating coexpression of *IGFBP2* and *VEGF* in pseudopalisading cells surrounding tumor necrosis provided further evidence for a possible involvement of *IGFBP2* in angiogenesis. The separating groups of genes were found by the unsupervised coupled two-way clustering method, and their classification power was validated by a supervised construction of a nearly perfect glioma classifier.

INTRODUCTION

Because of their diffusely infiltrating behavior, LGA[4] (WHO grade II) cannot be resected completely and will usually recur. At relapse, progression to anaplastic astrocytoma (WHO grade III) or glioblastoma multiforme, the most malignant form of gliomas (WHO grade IV), is common. Most glioblastoma arise *de novo* without evidence of a less malignant precursor lesion and are termed PrGBM. Glioblastoma evolving from a previous lower grade astrocytoma are defined as ScGBM. Although PrGBM are indistinguishable from ScGBM by histology, the two types of tumors exhibit distinct genetic alterations and occur in different age groups. The mean age for PrGBM is ~55 years, whereas ScGBM typically occur in younger patients (<45 years). Thus, PrGBM and ScGBM can be considered as two different diseases (1), despite a similarly grim outcome with a median survival of <1 year after diagnosis and no effective therapy. PrGBM are characterized by amplification/rearrangement and overexpression of the EGFR gene (in 40 and 60% of the patients, respectively) often in association with deletion of the *INK4a/p14ARF* gene locus (2). The hallmarks for ScGBM are *TP53* mutations (60%) and overexpression of *PDGF* and *PDGF* receptor (3, 4). Development of mouse glioma models and developmental neurobiology have allowed for recent advances in the understanding of the molecular bases of these two distinct genetic pathways and their implication for tumor initiation and progression as well as the cell of origin (5). However, a substantial number of glial tumors cannot be characterized by either of the two pathways depicted above, suggesting additional not yet recognized pathogenetic pathways. The current knowledge of tumor genetics does not allow identifying clinically relevant factors predictive for outcome or response to therapy. More detailed knowledge of underlying mechanisms and their relevance for the cancer process will allow treating cancer specifically by targeting deregulated pathways, leading to rational design of future treatment modalities tailored according to the biology of the individual tumors (6).

An essential initial step toward this goal is the establishment of a taxonomy of tumors on the basis of their gene expression profiles. A search for alternative pathways must be based on identification of genes whose expression differs significantly between the various tumor classes. In particular, it is important to look for a group of (possibly) coregulated genes, some of which share some known biological function, and whose expression differentiates tumor classes. Identification of such groups leads to better understanding of the biological processes that underlie the distinction between the tumors and may provide clues for the roles of such genes in initiation and progression of cancer.

We aimed at identifying expression profiles that differentiate three groups of astrocytic glioma: (*a*) LGA; (*b*) their respective ScGBM; and (*c*) PrGBM. In a novel gene selection approach, we combined supervised statistical analysis with CTWC (7–9), an unsupervised method, to identify correlated groups of genes that distinguish between the various tumor subtypes. Here, we demonstrate that gliomas can be separated according to their gene expression profiles, with PrGBM exhibiting a much higher variation of expression profiles than LGA. A cluster of correlated genes was identified that separates PrGBM from the other tumors and contains genes that are known to

be involved in angiogenesis, such as *VEGF,* but also *IGFBP2,* whose implication in angiogenesis is novel. By using expression data of the most informative separating gene clusters, we were able to construct an almost perfect tumor classifier. Our main findings, based on cDNA arrays, were validated on an independent set of glioma and by other methods, respectively.

MATERIALS AND METHODS

Specimen and RNA Expression Analysis

Tumor Biopsies and Cell Lines. Tumor biopsies, obtained from patients who underwent surgery at University Hospitals in Lausanne or Geneva or the Cantonal Hospital in Fribourg (Switzerland), were shock frozen and stored at $-70°C$. The use of biopsies and respective clinical data have been approved by the local ethics committee and the respective federal agency. The tumors were diagnosed according to the WHO classification 2000 (1). Twenty-one biopsies originated from 20 patients enrolled in a prospective pilot trial for newly diagnosed glioblastoma (10). From the tumor bank, an additional 32 gliomas from 22 patients were included, comprising 24 LGA and 8 respective high-grade recurrent tumors (comprising two astrocytoma WHO grade III and 6 grade IV). These tumors have been analyzed previously for TP53 mutations (11). The tumor samples were organized into two data sets according to their date of analysis: (*a*) subsequently used as a training set, comprising 14 PrGBM, 5 ScGBM, and 12 LGA; and (*b*) used as validation set, 4 PrGBM, 4 ScGBM, and 12 LGA. A summary of the information on the individual tumors, including TP53 status, age, and gender of the patients, and their organization into data sets, is available in Table 1. Normal brain tissue for RNA isolation was obtained from a lobectomy after brain edema, and additional samples of human normal brain total RNA were obtained from Clontech (human total RNA panel IV).

TAs have been constructed from archived paraffin blocs at the University Hospital in Lausanne (1984–2000) as described (12). All cases have been reviewed according to the WHO classification 2000 (1) by the neuropathologist (R. C. J.). The GBM-TA comprises 190 GBM, and the non-GBM glioma-TA includes 158 gliomas WHO I–III of different glioma subtypes.

Cell Lines. Culture conditions for the glioblastoma cell line U87, LN229, LNZ308, and the TP53-inducible glioblastoma cell line 2024, which is derived from LNZ308 after introducing the Tet-On System, have been described before (13). Expression of wild-type TP53 was induced with 2 μg/ml doxocycline in cell line 2024 during 24 h before RNA isolation (2024+). For anoxia treatment, cells were cultured for 20 h in an anaerobic culture incubator (O_2 <1%; Scholzen, Microbiology Syst AG) filled with a mixture of N_2, H_2, and CO_2.

Isolation of Total RNA. Before RNA isolation, a section of the frozen tumor biopsy was reevaluated after H&E staining by the neuropathologist (R. C. J.) to estimate the proportion of solid tumor, contaminating normal tissue, or infiltration zone. Pieces comprising >30% normal tissue were excluded. The setting of thresholds on the tumor fraction is treated differently by various groups, an issue that has been reviewed recently by Ramaswamy and Golub (14), who suggest to use "tumor cell-enriched material." Fifty to 100 mg of frozen tumor tissue were homogenized in TRIzol solution (Life Technologies). RNA phase was purified in saturated phenol solution [60% phenol, 15% glycerol, and 0.1 M sodium acetate (pH 4)]. The quality of the RNA was evaluated on agarose gel. RNA from cell lines was isolated similarly.

In Situ **Hybridization and Northern Blot Analysis.** *In situ* hybridization was performed according to the protocol supplied by Roche (Roche Applied Science) using cRNA probes labeled with digoxigenin during the *in vitro* transcription reaction. The plasmid pcDNA3 containing the 1400-bp human *IGFBP2* cDNA (a kind gift from S. Babajko; Ref. 15) was linearized with *Eco*RI or *Hind*III and transcribed with T7 or Sp6 RNA polymerase, respectively, to obtain sense or antisense probes. The plasmid pBluescript-KS-M13+ containing the 650-bp human *VEGF165* cDNA (pBspt-KS-VEGF165, a kind gift from K. Plate; Ref. 16) was linearized with *Eco*RI or *Bam*HI and transcribed with T7 or T3 RNA polymerase, respectively, to obtain sense and antisense probes. Probes were further reduced to an average size of 100 bp by limited alkaline hydrolysis. Northern blot analysis was performed as described before using 10 μg of total RNA (13). The membrane was sequentially hybridized to the plasmid-derived probe for *VEGF165* (*Eco*RI/*Bam*HI-frag-

ment of pBspt-KS-VEGF165) and the PCR-derived probes for *IGFBP2, 3,* and *5* (respective primer sequences provided by Clontech). Probes were radioactively labeled using the random primed DNA labeling kit (Boehringer) using [α-^{32}P]dCTP (3000 Ci/mmol, Amersham). Expression was quantified by phosphorimager (Fuji, BAS 1000).

cDNA Synthesis and Hybridization on cDNA Array. Atlas Human Cancer 1.2 Array membranes (Clontech) were used for all experiments described. These nylon filters are spotted with 1185 genes, including reference genes, and 1176 genes related to cancer. Three to 5 μg of DNase-treated total RNA were used to prepare a labeled first strand cDNA using the Clontech kit, basically as recommended. Briefly, RNA in a volume of 2 μl was mixed with 1 μl of specific CDS primers and 1 μl of RNasin (Promega; 40 units/ml) and denatured at 70°C. Subsequently, 1 μl of Superscript II reverse transcriptase (Life Technologies, Inc.; 200 units/ml) and 3.5–5 μl of α-^{32}P-dATP (3000 Ci/mmol; Amersham) were added, and the reaction was performed at 48°C for 30 min. The probe was purified with Clontech Atlas Nucleospin extraction kit following the manufacturer's recommendations. Before use, the membranes were boiled in 0.5% SDS solution for stripping and also at first use. Membranes were exposed to an imaging plate (BAS MS 2040; Fuji) for 1–8 days. The Atlas Cancer Arrays were used three times.

TP53-Mutation Analysis. The tumors were screened for *TP53* mutations using the yeast functional assay as described (17, 18), followed by direct sequencing, if the test was positive (Microsynth, Balgach, Switzerland).

Immunohistochemistry. Immunohistochemical determination for EGFR (Novo Castra; NCL-EGFR; dilution 1:40) and tenascin R [Santa Cruz Biotechnology; Tenascin-R (N20) sc-9874; dilution 1:1000] on paraffin sections was performed according to standard procedures using a high temperature epitope retrieval technique in citrate buffer (pH 6.0; pressure cooker, 3 min). Semiquantitative evaluation was performed independently by two researchers.

Data Analysis

Preprocessing and Analysis of Expression Data. Expression was quantified by phosphorimager (Fuji BAS1000) and analyzed with Atlasimage 1.5 software (Clontech). After background subtraction signals were normalized using the "sum method" comparing the sum of the intensity of all genes in the experiment to the sum calculated from the reference sample, yielding the coefficient of normalization, "c." The reference used in all experiments represents the calculated average of five independent expression profiles derived from "normal" brain. The ratio (R) was calculated for every gene as follows: $r = [(s - s0) \times c + K]/[(r - r0) + K]$, where $K = b \times [s0 + (c \times r0)]$; s, gene intensity in sample; s0, background in sample; c, coefficient of normalization obtained using the sum method; r, gene intensity in reference experiment; r0, background in reference experiment; and b = 2. For s $\gg$ s0 and r $\gg$ r0, this formula generates normalized ratios R shrunk toward 1 for low expressed genes in both experiments, sample and reference, respectively. Furthermore, it circumvents the problem of losing valuable information if the experiment or the reference display no expression for a given gene (division by 0).

Distance and MDS. This method projects the data points from the high dimension in which it is embedded to a low (two or three) dimensional space, in a way that best preserves their relative distances. Euclidian distances in the space of logarithms of the ratios R were used. MDS was performed with the implementation for the classical metric scaling (also known as principal coordinate analysis; Ref. 19) available in the mva package for R available online at the Comprehensive R Archive Network.[5]

CTWC. A variation filter was applied; only those genes were kept, for which the ratio of the maximal and minimal R values (obtained for the 36 experiments) exceeded 2. For each gene (row), the log of the ratio R was mean-centered (subtracting the average) and normalized. Euclidean distances, measured between all pairs of genes and between all pairs of tumors, served as the input to our clustering procedure. CTWC has been described elsewhere; see Getz *et al.* (7) for full details and comparisons with other methods; (8) for its applications to leukemia, colon, and breast cancer data analysis; and (9)

---

[5] The URLs referred to are: The Comprehensive R Archive Network: http://www. R-project.org; CTWC-Server: http://ctwc.weizmann.ac.il; homepage of complete CTWC data analysis: http://www.hospvd.ch/itbg; GeneCards, encyclopedia for genes, proteins, and diseases: http://genecards.weizmann.ac.il.

Table 1 *Summary information on 56 experiments analyzed by gene expression profiling*

This set comprises 53 gliomas from 44 patients and three experiments with cell lines.

Sample[a]	Pathology[b]	Gender	Age[c]	TP53 status			Data analysis[g]		
					Codon	TP53 mutation	CTWC	TRN-Set	VAL-Set
1284	PrGBM	F	48	wt			1	1	0
1437	PrGBM	M	48	wt			1	1	0
1316	PrGBM	M	68	wt			1	1	0
1399	PrGBM	M	38	wt			1	1	0
G204	PrGBM	F	51	wt			1	1	0
1430	PrGBM	M	37	wt			1	1	0
G197	PrGBM	M	46	wt			1	1	0
1419	PrGBM	F	48	wt			1	1	0
1308	PrGBM	F	45	wt			1	1	0
1453	PrGBM	F	53	mut	244	GGC to GTC	1	1	0
1317	PrGBM	M	26	mut	175	CGC to CAC	1	1	0
1297	PrGBM	M	36	mut	163	TAC to TGC	1	1	0
1303	PrGBM	M	55	mut	273	CGC to CAT	1	1	0
1360	PrGBM	M	65	mut	173	GTG to ATG	1	1	0
G205	PrGBM	F	56	wt			0	0	1
G216	PrGBM	M	53	wt			0	0	1
1621	PrGBM	M	62	wt			0	0	1
G226	PrGBM	M	45	mut	273	CGT to CAT	0	0	1
1342	ScGBM	M	51	wt			1	1	0
749	ScGBM	F	47	mut	175	Ref. 11[f]	1	1	0
946	ScGBM	M	41	mut	258, 267, 283	Ref. 11[f]	1	1	0
809	ScGBM	M	53	mut	273	Ref. 11[f]	1	1	0
978	ScGBM	M	29	mut	241	Ref. 11[f]	1	1	0
413	ScGBM	F	53	wt		Ref. 11[f]	0	0	1
633	ScGBM	M	28	wt		Ref. 11[f]	0	0	1
735	ScGBM	M	34	mut	261	Ref. 11[f]	0	0	1
722	ScGBM	M	39	mut	248	Ref. 11[f]	0	0	1
421	LGA	M	2	wt		Ref. 11[f]	1	1	0
698	LGA	M	39	wt		Ref. 11[f]	1	1	0
1070	LGA	M	58	wt		Ref. 11[f]	1	1	0
80	LGA	M	28	wt		Ref. 11[f]	1	1	0
246	LGA	F	53	wt		Ref. 11[f]	1	1	0
328	LGA	M	28	wt		Ref. 11[f]	1	1	0
416	LGA	M	29	mut	241	Ref. 11[f]	1	1	0
92	LGA	M	34	mut	261	Ref. 11[f]	1	1	0
289	LGA	M	27	mut	248	Ref. 11[f]	1	1	0
460	LGA	F	47	mut	175	Ref. 11[f]	1	1	0
736	LGA	M	41	mut	258, 267, 283	Ref. 11[f]	1	1	0
635	LGA	M	53	mut	273	Ref. 11[f]	1	1	0
676	LGA	M	41	wt		Ref. 11[f]	0	0	1
355	LGA	M	35	wt		Ref. 11[f]	0	0	1
374	LGA	M	15	wt		Ref. 11[f]	0	0	1
875	LGA	M	50	wt		Ref. 11[f]	0	0	1
501	LGA	F	57	wt		Ref. 11[f]	0	0	1
898	LGA	F	35	mut	248	Ref. 11[f]	0	0	1
528	LGA	M	52	mut	248	Ref. 11[f]	0	0	1
551	LGA	F	31	mut	155	Ref. 11[f]	0	0	1
510	LGA	F	35	mut	248	Ref. 11[f]	0	0	1
210	LGA	M	39	mut	248	Ref. 11[f]	0	0	1
589	LGA	M	33	mut	220	Ref. 11[f]	0	0	1
552	LGA	M	26	mut	234	Ref. 11[f]	0	0	1
1497	RecGBM	M		wt			1	0	0
1357	OAIII[d]	M	36	mut	273	CGT to TGT	1	0	0
2024−	CL			null			1	0	0
2024+	CL			wt[e]			1	0	0
U87	CL			wt			1	0	0

[a] Tumors from same patients: PrGBM to recurrent GBM: 1430/1497; LGA to recurrent LGA: 80/328; LGA to ScGBM: 736/946, 635/809, 416/978, 460/749, 92/735, 210/722; LGA to anaplastic astrocytoma (WHO grade III): 416/978, 246/413.

[b] PrGBM (WHO grade IV); ScGBM (WHO grade IV); LGA (WHO grade II); OAIII (WHO grade III).

[c] Age at diagnosis.

[d] 1357 was originally diagnosed as PrGBM.

[e] Induction of wild-type *TP53* with Tet-On system (13).

[f] Published previously by Ishii *et al.* (11).

[g] Samples used for respective data analysis: CTWC; TRN-set, training set; VAL-set, validation set for tumor predictor.

describing the use of the publicly available CTWC–Server.[5] CTWC starts with clustering all genes on the basis of the data from all tumors and clustering all tumors, using data from all genes. In both resulting dendrograms, we identify stable (*i.e.,* statistically significant) gene and sample clusters; these are denoted, respectively, as GX or SY (where X,Y are running indices). Note that G1 represents the set of all genes and S1 that of all samples. In the second iterative step, each stable gene cluster GX is used to characterize and cluster the members of every sample cluster SY and *vice versa*.

To use CTWC, we must be able to identify stable clusters. One of the few algorithms that provide a stability index to each cluster is Super Paramagnetic

Clustering, a physics-based method that has been described in full detail in Blatt *et al.* (20). We register a cluster C as stable only if it exceeds a certain size and when *Stab(C),* its stability index defined in terms of the range of resolution parameters, T, through which cluster C "lives" (21), exceeds a certain threshold.

Classification. For binary class comparisons, two standard statistical tests were used: (*a*) the two-sample t test (with unknown but equal variances); and (*b*) the Wilcoxon rank-sum test. To address contamination with false positive genes associated with multiple comparisons, we use the method of Benjamini and Hochberg (22) that bounds the average FDR. The outcome of this method

Table 2 *Summary of results for two-way comparisons between tumor classes using the tumors of the training set and all genes (1185)*

Comparison tested (no. of tumors)	No. of genes selected (FDR q = 0.05)[a]				
	t test	Rank-sum	Shared	Union	Highest P
PrGBM (14) *vs.* ScGBM + LGA (5 + 12)	191	174	160	205	0.008
PrGBM + ScGBM (14 + 5) *vs.* LGA (12)	126	98	90	132	0.005
PrGBM (14) *vs.* LGA (12)	167	163	143	187	0.007

[a] FDR, false discovery rate.

is a list of differentiating genes; the expected fraction of false positive genes is at most *q*.

The class discrimination power of the selected sets of genes was validated by training a k-nearest neighbor (*k*-NN) classifier on one set of tumor tissue samples and testing the class prediction on an independent validation set. Computations were performed with the class package for R and with *k* = 3. Euclidian distance in the space of the ratios R was used, classification was decided by majority vote, with ties broken at random. When ties occurred, classification was repeated 100 times, and *P*s were averaged. *P*s for the significance of the deviation from independence between true and predicted labels were obtained with Fisher's exact test (ctest package for R) and with the alternative hypothesis set to "greater" in the 2 by 2 case.

RESULTS

Gene Expression Profiles Separate Tumor Classes

To characterize and classify gliomas by their gene expression profiles, RNA isolated from frozen gliomas and three glioblastoma cell lines was analyzed using cDNA arrays comprising 1185 genes.

MDS. The configuration of the 51 astrocytic gliomas in Euclidian space of overall gene expression, as visualized in Fig. 1 using MDS, clearly suggests that gene expression profiles contain information discriminating the three classes of astrocytic gliomas. The LGA and ScGBM show a higher degree of spatial intermingling, whereas the best separation is between LGA and PrGBM. This correctly reflects other biological characteristics of the ScGBM, in that they progress from LGA but share the malignancy grade (WHO IV) with PrGBM and are indistinguishable from them histopathologically. Interestingly, the pairwise distances in space of all genes (1185) are highest among PrGBM (mean, 10.06; SD, 2.1) and shortest between LGA (mean,

Fig. 1. MDS based on overall gene expression (1185 genes) of 51 astrocytic gliomas. The color code indicates the tumor subtype. ▼, samples from the training set; ●, gliomas from the validation set.

6.09; SD, 1.18), whereas ScGBM are intermediate (mean, 7.27; SD, 1.18). Using 358 genes that pass a variation filter (see CTWC, below) yields a nearly similar MDS picture (data not shown). Thus, MDS of gene expression suggests more biological heterogeneity of PrGBM, without firm evidence for obvious subclasses, reflecting well the heterogeneous morphological and clinical characteristics of glioblastoma multiforme.

Supervised Analysis. We applied *t* test and Wilcoxon's rank-sum test to look for genes that differentiate two known classes, using all 1185 genes and the training set of 31 astrocytic gliomas. The threshold for the FDR was set at q = 0.05, *i.e.,* the expected number of false positives was kept at 5%. The results for three two-way comparisons are summarized in Table 2. The lists of differentiating genes determined by the two tests are very similar. Because of the low number of the ScGBM in the training set, comparisons involving ScGBM, ScGBM *versus* PrGBM, and ScGBM *versus* LGA yielded only very few separating genes (1 and 3, respectively). Full results of these supervised analyses, including a list of the respective genes selected by these binary comparisons, can be found in Table S1 of the supplementary information.

Our efforts to find genes correlated with *TP53* status failed, whether we used all tumors or only subsets thereof. Neither did we find any gene cluster separating groups of samples on the basis of their *TP53* status (see *TP53* status of biopsies in Table 1). Our results are in line with a previous report on a series of LGA, which also did not find evidence for genes correlated with *TP53* status (23). This result may not be surprising in consideration of the fact that the TP53 pathway has been found to be inactivated, in >70% of all astrocytic gliomas, regardless of tumor grade, by mutational alteration of *TP53* or alternative mechanisms, such as overexpression of mdm2, a negative regulator of TP53, or inactivation of p14ARF, a negative regulator of mdm2, by deletion or promoter methylation of the gene (24).

A Cluster of Angiogenesis-related Genes Separates PrGBM from the Other Tumors

First Level CTWC: Unsupervised Analysis. The data derived from 36 experiments (Table 1; of 33 gliomas and three experiments with glioblastoma cell lines) were subjected to the variation filter described in "Materials and Methods," reducing the number of genes (rows) to 358. This filtering reduces the "noise" generated by genes that do not vary significantly over the samples. Changing the filtering parameters from 2 to 1.5 (or 3) changed the number of genes that passed to 601 (or 190), but the changes induced in the observed gene clusters were comparable with the fluctuations observed in repeated runs (with 358 genes). The first operation, G1(S1), clustered the set of all 358 genes (G1), using the data from the set of all 36 experiments (S1); the second, S1(G1), clustered the tumors S1 using their expression profiles measured for all of the genes G1. The two resulting dendrograms are shown in Fig. 2*A* together with the correspondingly reordered expression matrix. G1(S1) identified 15 stable gene clusters (Stability > 8, corresponding to *P* < 0.01) that are indicated in the gene dendrogram of Fig. 2*A* and denoted G2 to G16. The operation S1(G1) yields a dendrogram with two stable clusters, S2 and S3. Details
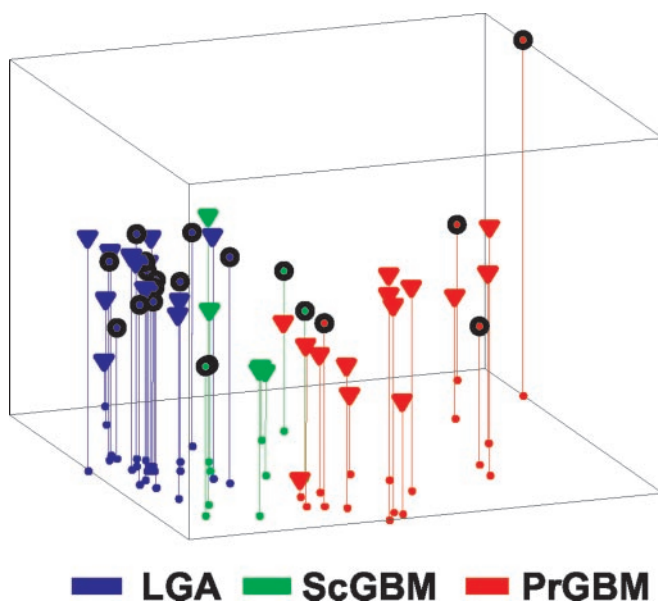
of the clusters obtained in the CTWC can be viewed and searched online at the respective Web page.[5]

Second Level CTWC: Clustering Tumors Using Selected Gene Clusters. We used all of the 15 stable gene clusters found above, one by one, to recluster S1. Whenever a stable partition of the samples was found, we checked whether it divided the tumors according to some biologically relevant attribute. In particular, we looked for gene clusters that partition the tumors into GBM *versus* LGA.

Separation into PrGBM *Versus* LGA and ScGBM on the basis of Angiogenic Activity in S1(G5). In the first level of CTWC, we identified G5 as a stable cluster of nine genes (Fig. 2, *A* and *B*). This cluster comprises hallmarks of angiogenesis, such as *VEGF, FLT1* (fms-related tyrosine kinase 1; also called *VEGFR1*), and pleiotrophin (25). Some of the other genes have also been related to some aspects of angiogenesis before (Table 3), namely, *IGFBP2* (insulin-like growth factor binding protein 2), which has been suggested to be activated by hypoxia-inducible factor-1, although in an indirect manner because of the absence of a defined hypoxia-response element (26). When the expression levels of only these genes were used to characterize the tumors, a large and stable cluster, S11, of 21 tumors emerged (Fig. 2*B*). This cluster contained all of the 12 LGA and 4 of 5 ScGBM. Of the remaining 5 samples of S11, three were cell lines, whose expression profiles were consistently different. Hence, the expression levels of G5 gave rise to a nearly perfect separation of clinically defined tumor entities, namely PrGBM from LGA and ScGBM. These genes were significantly up-regulated in PrGBM (Fig. 2*B*) as compared with LGA and ScGBM.

Other Separating Gene Clusters

Separation by Immune Response-related Genes in S1(G12). The gene cluster G12 contains 13 genes that are almost exclusively related to the immune system and inflammation (Table 3). Four of which are known to be regulated by γ-IFN (*FCGR1A, CYBB, IFITM1,* and *MX1*), and some of the Fc γ receptors and *DAP12* represent markers for natural killer cells, although their expression is not exclusive. The dendrogram obtained by clustering the tumors using these genes is presented in Fig. 2*C*. Of its two large stable clusters, S39 contains nine tumors, eight of which are PrGBM; the expression level of the G12 genes in these tumors is high. The other large cluster, S38, contains all but one of the LGA and ScGBM samples, together with four PrGBM. Three of these four tumors are from patients whose survival was relatively long (>18 months; Ref. 10). Hence, the immune system-related genes of G12 appear to have lower expression levels in LGA and ScGBM and in those PrGBM that have longer survival.

Gene Selection by Combining Supervised Analysis with Clustering. In a second step, we used the results of the supervised analysis described above to identify those clusters in the dendrogram yielded by G1(S1) that are rich in discriminating genes. Of the 205 genes that were found to differentiate PrGBM from ScGBM and LGA (Table 2; Table S1 of supplementary information), 91 passed the variation filter and were included in G1. As depicted in Fig. 3*A,* most separating genes belonged to one of four marked clusters. The same four clusters
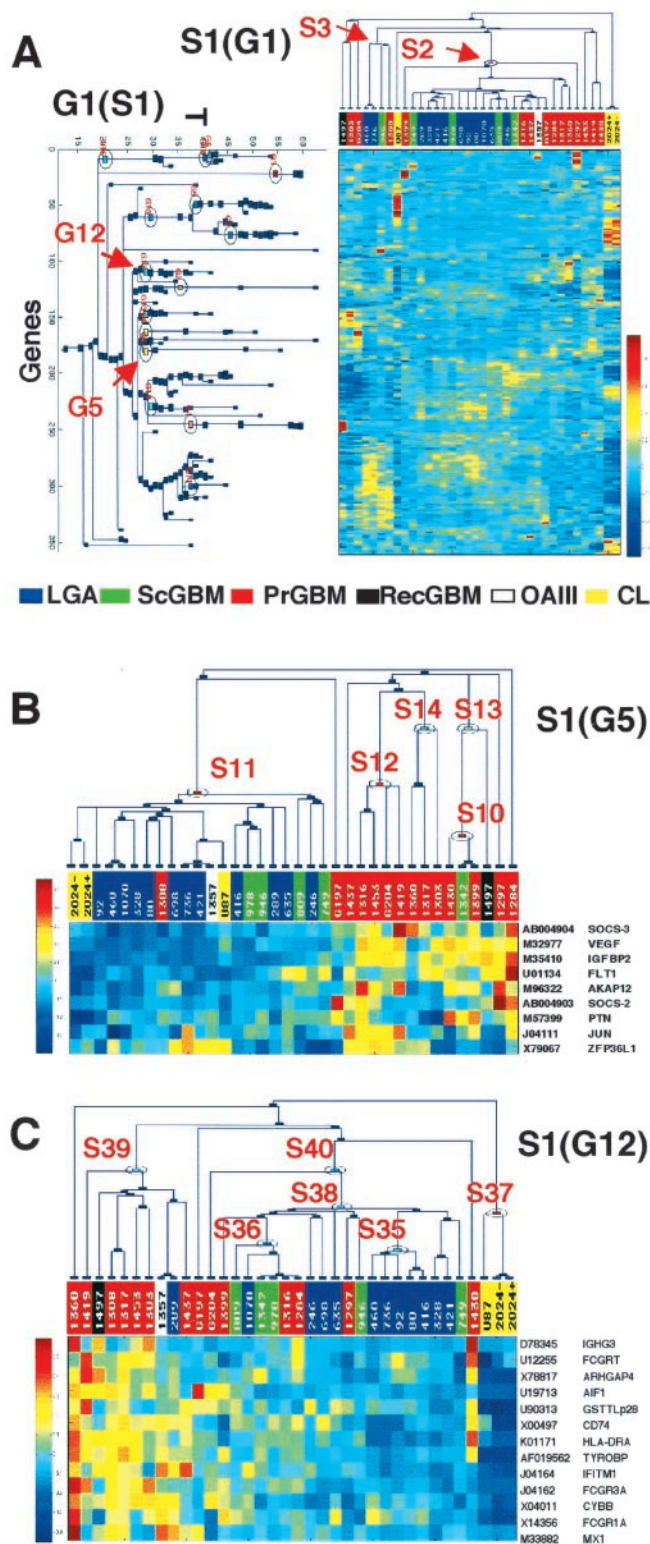
Fig. 2. Unsupervised analysis by CTWC. *A,* first level CTWC clusters G1 (the set of all 358 genes that have passed the filtering criteria), using all 36 experiments (S1), and clusters all samples (S1) according to the gene expression profiles of all genes G1. These clustering operations are called G1(S1) and S1(G1). G1(S1) yields a dendrogram of gene clusters, and S1(G1) is a dendrogram of sample clusters. The respectively reordered expression matrix is visualized using a color scale, representing centered and normalized values of the log₂R. Fifteen stable gene clusters emerged and are marked with a ring (G2–G16). Clustering the samples according to all genes, S1(G1), yielded two stable sample clusters, S2 and a smaller cluster S3. The tumor samples are denoted at the *top,* using *colors* to represent tumor subtypes. *RecGBM,* recurrent glioblastoma; *OAIII,* oligoastrocytoma WHO grade III; *CL,* glioblastoma cell line. In *B,* in second level CTWC, we cluster all samples S1 according to selected gene clusters, S1(G5). Clustering S1 according to gene cluster G5 [which was obtained from G1(S1); *A*] yields a dendrogram

for the samples. Note the separation of LGA and ScGBM from PrGBM. PrGBMs exhibit overexpression of the genes of G5 that are related to angiogenesis. In *C,* clustering S1 according to G12 yields a dendrogram S1(G12) separating a group of PrGBM in S39. G12 contains mostly genes related to the immune system. In the dendrograms, a *box* represents a cluster. The T value at which the box is placed corresponds to the temperature at which the cluster disintegrates. The sizes of the clusters are not reflected by their boxes. The parameters for a stable gene cluster have been set at a stability threshold of 8, a maximal dropout of 3 at a single step of T, and the minimal cluster size at 5. The criteria for stable sample clusters are: Stab(C) > 8, maximal dropout at a single step of T is 1, and minimal cluster size is 3.

Table 3 *Genes[a] comprised in stable clusters G5(S1) and G12(S1) related to specific biological processes*

Gene cluster	GenBank acc. no.	Gene symbol	Location[b]	S[c]	Gene description	Putative function
G5						
1	AB004904	SOCS-3	17q25.3	1	suppressor of cytokine signaling 3; STAT-induced STAT inhibitor 3 (STATI3)	SOCS3 is involved in negative regulation of cytokines that signal through the JAK/STAT pathway. Inhibits cytokine signal transduction by binding to tyrosine kinase receptors including gp130, LIF, erythropoietin, insulin, and leptin receptors. Interacts with multiple activated proteins of the tyrosine kinase signaling pathway, including IGF1 receptor, insulin receptor, and JAK2 (40).
2	M32977	VEGF	6p21.1	1	VEGF-A	Growth factor active in angiogenesis, vasculogenesis, and endothelial cell growth. It induces endothelial cell proliferation, promotes cell migration, inhibits apoptosis, and induces permeabilization of blood vessels.
3	M35410	IGFBP2	2q35	1	IGFBP2	IGF-binding proteins prolong the half-life of the IGFs and have been shown to either inhibit or stimulate the growth-promoting effects of the IGFs on cell culture. They alter the interaction of IGFs with their cell surface receptors.
4	U01134	FLT1	3q12.2	1	fms-related tyrosine kinase 1; vascular endothelial growth factor receptor 1 (VEGFR1)	Receptor for VEGF, VEGFB, and PGF. Has a tyrosine-protein kinase activity. The VEGF-kinase ligand/receptor signaling system plays a key role in vascular development and regulation of vascular permeability.
5	M96322	AKAP12	6q25.1	1	A kinase anchor protein 12, gravin	Anchoring protein that mediates the subcellular compartmentation of protein kinase (PKA) and protein kinase C (PKC). May play a role in wound repair and vascular development (48).
6	AB004903	SOCS-2	12q22	1	Suppressor of cytokine signaling 2, STAT-induced STAT inhibitor 2 (STATI2)	Negative feedback system that regulates cytokine signal transduction. SOCS2 appears to be a negative regulator in the growth hormone/IGF1 signaling pathway (40).
7	M57399	PTN	7q33	0	Pleiotrophin; heparin-binding growth factor 8, neurite growth-promoting factor 1	Heparin-binding mitogenic protein. Has neurite extension activity. Angiogenic properties have been demonstrated (25).
8	J04111	JUN	1p32.1	0	c-jun proto-oncogene; transcription factor AP-1	Transcription factor, interacts with c-fos to form a dimer. Interacts with SMAD3/SMAD4 heterodimers. Interacts with TCF20. Cooperates with HIF-1 in hypoxia-induced gene transcription (49).
9	X79067	ZFP36L1	14q24.1	0	Zinc finger protein 36, C3H type-like 1, TIS11B protein; EGF response factor 1	Probable regulatory protein involved in regulating the response to growth factors.
G12						
1	D78345	IGHG3	14q32.33	1	Human DNA for Ig γ heavy-chain, membrane-bound-type and secrete-type	σ-region located between C mu and C Δ genes of human immunoglobulin heavy chain.
2	U12255	FCGRT	19q13.33	0	IgG receptor Fc large subunit P51; FcRN	Binds to the Fc region of monomeric immunoglobulins γ. Mediates the uptake of IgG from milk. Possible role in transfer of immunoglobulin G from mother to fetus.
3	X78817	ARHGAP4	Xq28	0	ρ-GAP hematopoietic protein C1 (RGC1); KIAA0131, (ρ GTPase activating protein 4)	Inhibitory effect on stress fiber organization. May down-regulate ρ-like GTPase in hematopoietic cells. Predominantly in hematopoietic cells.
4	U19713	AIF1	6p21.33	1	allograft inflammatory factor 1 (AIF1); glutathione-*S*-transferase (GST) homologue, glutathione transferase ω.	May play a role in macrophage activation and function.
5	U90313	GSTTLp28	10q25.1		Acts as a small stress response protein likely involved in cellular redox homeostasis.	
6	X00497	CD74	5q33.1	0	invariant polypeptide of major histocompatibility complex, class II antigen-associated	Plays a critical role in MHC class II antigen processing.
7	K01171	HLA-DRA	6p21.32	1	Major histocompatibility complex, class II, DR α, HLA class II histocompatibility antigen α chain	MHC class II receptor activity
8	AF019562	TYROBP	19q13.12	1	TYRO protein tyrosine kinase binding protein, DNAX activation protein 12 (DAP12)	Noncovalently associates with membrane glycoproteins of the killer-cell inhibitory receptor (KIR) family without an ITIM in their cytoplasmic domain.
9	J04164	IFITM1	CHR 11	0	IFN-induced transmembrane protein 1 (9–27), Leu13, CD225	Control of cell growth. Component of a multimeric complex involved in the transduction of antiproliferative and homotypic adhesion signals.
10	J04162	FCGR3A	1q23.3	1	Fc fragment of γG, low affinity IIIa, leukocyte IgG receptor (Fc-γ-R), CD16	Receptor for the Fc region of IgG. Mediates antibody-dependent cellular cytotoxicity (adcc) and other antibody-dependent responses, such as phagocytosis.
11	X04011	CYBB	Xp11.4	0	Cytochrome b-245, β polypeptide (chronic granulomatous disease), GP91-PHOX	Critical component of the membrane-bound oxidase of phagocytes that generates superoxide.
12	X14356	FCGR1A	1q21.2	0	Fc fragment of IgG, high affinity Ia, CD64	Binds to the Fc region of immunoglobulins γ.
13	M33882	MX1	21q22.3	0	IFN-regulated resistance GTP-binding protein MXA (IFI-78K); IFN-induced protein P78	Shows activity against influenza virus and vsv, a rhabdovirus.

[a] Information on genes listed in this table is taken from GeneCards.[5]
[b] Ensembl cytogenetic band.
[c] Separating gene according to statistical criteria, see "Materials and Methods" (Table S1 of supplementary information for complete gene list).
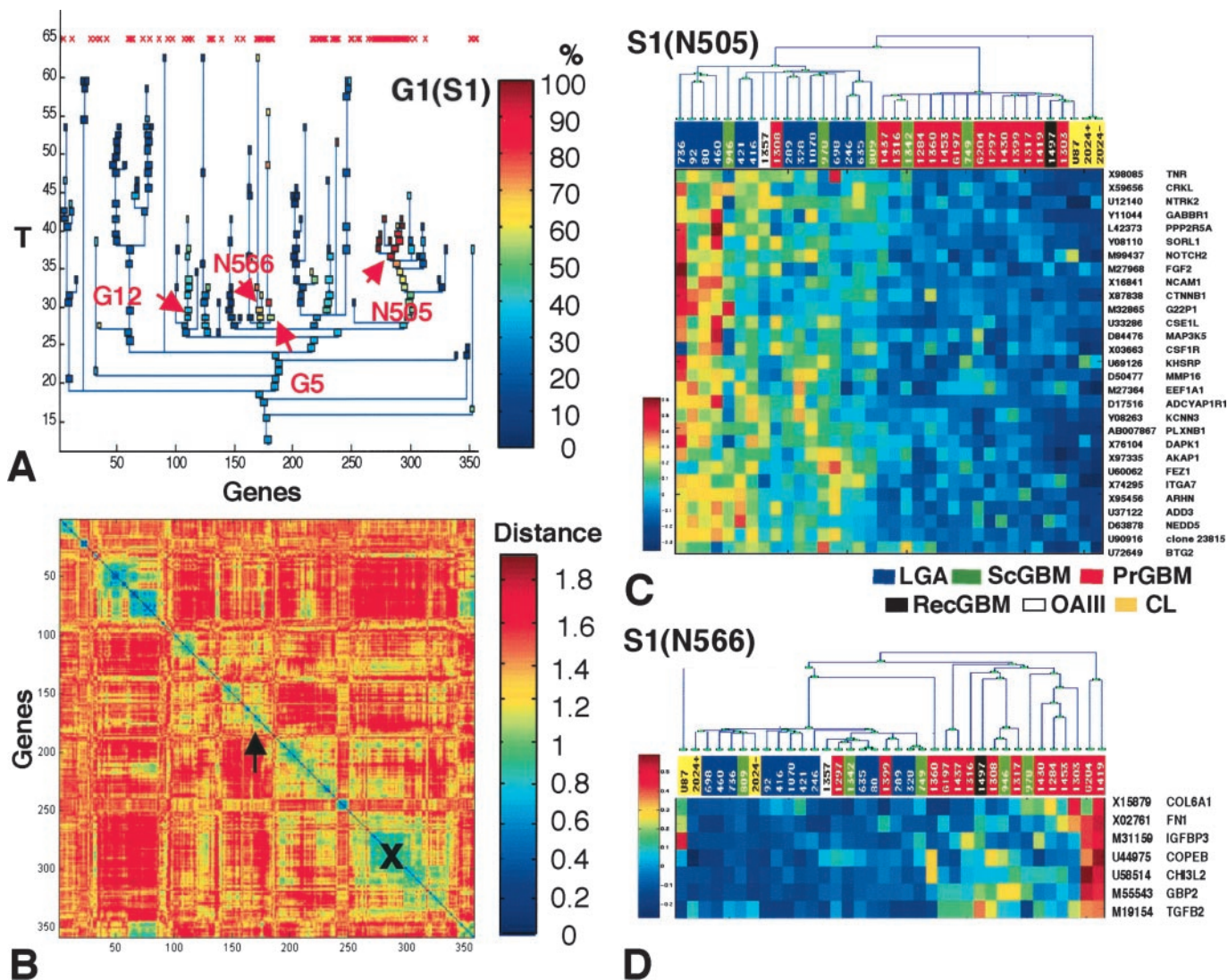
Fig. 3. Combination of supervised and unsupervised analysis. In *A,* the genes found by pairwise comparison of PrGBM *versus* LGA and ScGBM were identified in the gene dendrogram G1(S1) (see also Fig. 2*A*) and marked on *top* with a *red x.* The percentage of these genes in a given gene cluster is presented in a color code. Most of the identified genes are grouped in four to five clusters. Two of them had been found previously (G5 and G12) *arrow.* The two new clusters identified are particularly rich in these genes, denoted as N505 and N566. They were not stable according to the criteria set; however, the distance matrix (*B*) recognizes these nodes as comprising closely related genes, indicated by an *arrow* for N566 and an *X* for N505. The distance D is visualized as a color code, with *blue* indicating short distance. Note the order of the genes, and the scale is the same for *A* and *B.* Reclustering of all experiments S1 according to N505 and N566, respectively, is shown in *C* and *D.* Both nodes yield an almost perfect separation of the tumors according to their classification. This is not surprising because in both N505 and N566, all but one of the genes were identified as separating in the binary comparison.

were identified also as rich in genes that partition samples according to the other supervised binary comparisons mentioned above. Two of the identified clusters were indeed G5 (six of nine genes are among the 91 separating ones) and G12 (5 of 13). However, two additional clusters were identified as rich in separating genes; these are marked in Fig. 3*A* as nodes N505 (28 of 29) and N566 (6 of 7). These sets of genes (Table 4) were not selected as stable clusters by CTWC because they were either too small or they "leaked" too many genes as the control parameter *T* was increased. However, as clearly shown in the distance matrix (Fig. 3*B*), the genes of both of these nodes had relatively highly correlated expression profiles.

These additional clusters contain some interesting genes; members of the Wnt- and Notch-pathway, genes involved in adhesion/migration, and apoptosis-related genes in node N505 (Fig. 3*C*; Table 4) and in node N566 genes like *TGFβ2* and *IGFBP3* that have been reported before to be overexpressed in high-grade gliomas (Fig. 3*D*; Table 4).

Next, we used these two additional clusters to partition all of the samples S1. As shown in Fig. 3*C,* node N505 gives rise to a nearly

perfect separation of the samples into three clusters: (*a*) "Non-PrGBM"; of 17 samples, comprising 12 of 12 LGA, 3 of 5 ScGBM, and 1 PrGBM and the OAIII, both of which (1308 and 1357) joined the non-PrGBM cluster in S1(G5); (*b*) mainly PrGBM, 13 of 14 PrGBM, the RecGBM (1497), 2 of 5 ScGBM, one of which (1342) had joined the PrGBM in S1(G5), 1 cell line (U87); and (*c*) a cluster with the remaining cell line under two conditions (wild-type *TP53* on and off; 2024+, 2024−). Hence, node N505 gives a clear separation into LGA *versus* PrGBM, with the ScGBM split evenly. As we also see in Fig. 3*D,* node N566 identifies a cluster of 20 samples comprising 12 of 12 LGA; thus, this node also separates the samples largely into LGA *versus* PrGBM, with the ScGBM distributed evenly.

Constructing a Tumor Classifier

A gene expression based tumor classifier was built to validate our findings. A successful test of such a classifier based on the selected genes and applied to an independent validation set of 20 samples

Table 4 *Genes[a] comprised in nodes N505(S1) and N566(S1) rich in separating genes*

Gene cluster	GenBank acc. no.	Gene symbol	Location[b]	S[c]	Gene description	Putative function
N505						
1	X98085	TNR	1q25.1	1	tenascin-R; restrictin; janusin	Extracellular matix protein, secreted by oligodendrocytes during myelination, some astrocytes and neurons. Can modify adhesive substrate properties of fibronectin. Involved in cell adhesion, migration, and differentiation.
2	X59656	CRKL	22q11.21	1	v-crk sarcoma virus CT10 oncogene homologue (avian)-like; CRK-like protein	May mediate the transduction of intracellular signals.
3	U12140	NTRK2	9q21.33	1	neurotrophic tyrosine kinase, receptor, type 2; brain-derived neurotrophic factor (BDNF)/NT-3 growth factors receptor; TRKB tyrosine kinase receptor; GP145-TRKB	Receptor for brain-derived neurotrophic factor (BDNF), neurotrophin-3 and neurotrophin-4/5 but not nerve growth factor (NGF)
4	Y11044	GABBR1	6p22.1	1	γ-aminobutyric acid (GABA) B receptor, 1	Receptor for GABA.
5	L42373	PPP2R5A	1q32.3	1	protein phosphatase 2, regulatory subunit B (B56), alpha isoform); protein phosphatase 2A B56-α, PP2A	The b regulatory subunit might modulate substrate selectivity and catalytic activity, and also might direct the localization of the catalytic enzyme to a particular subcellular compartment.
6	Y08110	SORL1	11q23.3	1	sortilin-related receptor L(DLR class) A repeats-containing); low-density lipoprotein receptor-related protein LR11	Likely to be a multifunctional endocytic receptor, that may be implicated in the uptake of lipoproteins and of proteases. Could play a role in cell-cell interaction. Expressed mainly in brain.
7	M99437	NOTCH2	1p11.2	1	Notch homologue 2, neurogenic locus notch protein	Functions as a receptor for membrane-bound ligands Jagged1, Jagged2, and Δ1 to regulate cell-fate determination. Affects the implementation of differentiation, proliferation, and apoptotic programs.
8	M27968	FGF2	4q27	0	fibroblast growth factor, basic; FGFB; heparin-binding growth factor 2 precursor (HBGF2)	The heparin-binding growth factors are angiogenic agents *in vivo* and are potent mitogens for a variety of cell types *in vitro*. There are differences in the tissue distribution and concentration of these two growth factors.
9	X16841	NCAM1	11q23.1	1	neural cell adhesion molecule 1; NCAM120; CD56	Cell adhesion molecule involved in neuron-neuron adhesion, neurite fasciculation, outgrowth of neurites, etc.
10	X87838	CTNNB1	3p22.1	1	β-beta catenin	Involved in the regulation of cell adhesion and in signal transduction through the wnt pathway.
11	M32865	G22P1	22q13.2	1	thyroid autoantigen 70 kDa; Ku 70-kDa subunit;	Single stranded DNA-dependent ATP-dependent helicase. Has a role in chromosome translocation.
12	U33286	CSE1L	20q13.13	1	chromosome segregation 1-like (yeast); cellular apoptosis susceptibility protein (CAS);	Export receptor for importin α. Mediates importin-α reexport from the nucleus to the cytoplasm after import substrates have been released into the nucleoplasm. Highly expressed in proliferating cells.
13	D84476	MAP3K5	6q23.3	1	mitogen-activated protein kinase kinase kinase 5; MAP/ERK kinase kinase 5; MAPKKK5; MEKK5; ASK1	Phosphorylates and activates two different subgroups of MAP kinase kinases, MKK4/SEK1 and MKK3/MAPKK6 (OR MKK6). Overexpression induces apoptotic cell death.
14	X03663	CSF1R	5q33.1	1	colony-stimulating factor I receptor precursor; fms proto-oncogene (c-fms); CD115	This protein is the receptor for CSF-1, it is a protein tyrosine-kinase transmembrane receptor.
15	U69126	KHSRP	19p13.3	1	KH-type splicing regulatory protein; fuse-binding protein 2 (FBP2)	A new regulatory protein, KSRP, mediates exon inclusion through an intronic splicing enhancer.
16	D50477	MMP16	8q21.3	1	matrix metalloproteinase 16 (membrane-inserted); membrane-type matrix metalloproteinase 3 (MT-MMP3)	Endopeptidase, degrades components of the extracellular matrix, such as collagen type III and fibronectin. Activates progelatinase A. Involved in the matrix remodeling of blood vessels. No effect on type I, II, IV, and V collagen.
17	M27364	EEF1A1	6q14.1	1	eukaryotic translation elongation factor 1 α 1; EF1 α	This protein promotes the GTP-dependent binding of aminoacyl-trna to the a-site of ribosomes during protein biosynthesis.
18	D17516	ADCYAP1R1	7p14.3	1	adenylate cyclase activating polypeptide 1 (pituitary) receptor type I; PACAP I receptor; PACAPR1	Receptor for pacap-27 and pacap-38. The activity of this receptor is mediated by G proteins, which activate adenylyl cyclase.
19	Y08263	KCNN3	1q22	1	potassium intermediate/small conductance calcium-activated channel, subfamily N, member 3; AAD14	Forms a voltage-independent potassium channel activated by intracellular calcium.
20	AB007867	PLXNB1	3p21.31	1	plexin B1; plexin 5; KIAA0407	A family of transmembrane proteins with homology to the MET-hepatocyte growth factor receptor.
21	X76104	DAPK1	9q21.33	1	death-associated protein kinase 1; DAP kinase 1	Involved in mediating IFN-γ-induced cell death.
22	X97335	AKAP1	17q23.2	1	A-kinase anchoring protein 1; PRKA	Binds to type I and II regulatory subunits of protein kinase A and anchors them to the cytoplasmic face of the mitochondrial outer membrane.
23	U60062	FEZ1	11q24.2	1	fasciculation and elongation protein zeta 1; zygin I	May be involved in axonal outgrowth as component of the network of molecules that regulate cellular morphology and axon guidance machinery.
24	X74295	ITGA7	12q13.2	1	integrin α 7B precursor; IGA7B	Integrin α-7/β-1 is the primary laminin receptor on skeletal myoblasts and adult myofibers.
25	X95456	ARHN	17q23.3	1	ras homologue gene family, member N; Rho7 protein	May be specifically involved in neuronal and hepatic functions.
26	U37122	ADD3	10q25.1	1	adducin 3 γ	Membrane cytoskeleton-associated protein that promotes the assembly of the spectrin-actin network. Binds to calmodulin.
27	D63878	NEDD5	2q37.3	1	neural precursor cell expressed, developmentally down-regulated 5	Mammalian septin is a novel cytoskeletal component interacting with actin-based structures
28	U90916	clone 23815		1	Human clone 23815	Soares library 1NIB from IMAGE Consortium
29	U72649	BTG2	1q32.1	1	btg protein; NGF-inducible antiproliferative protein PC3	Antiproliferative protein.

Table 4 *Continued*

Gene cluster	GenBank acc. no.	Gene symbol	Location[b]	S[c]	Gene description	Putative function
N566						
1	X15879	COL6A1	21q22.3	0	collagen, type VI, α-1	Collagen VI acts as a cell-binding protein.
2	X02761	FN1	2q35	1	fibronectin 1	Fibronectins bind cell surfaces and various compounds including collagen, fibrin, heparin, DNA, and actin. Fibronectins are involved in cell adhesion, cell motility, opsonization, wound healing, and maintenance of cell shape.
3	M31159	IGFBP3	7p13	1	IGFBP 3	IGF-binding proteins prolong the half-life of the IGFs and have been shown to either inhibit or stimulate the growth-promoting effects of the IGFs on cell culture. They alter the interaction of IGFs with their cell surface receptors.
4	U44975	COPEB	10p15.2	1	core promoter element binding protein; DNA-binding protein CPBP	Transcriptional activator (by similarity). Binds a GC box motif. Could play a role in B-cell growth and development.
5	U58514	CHI3L2	1p13.3	1	chitinase 3-like 2	Not detected in brain, spleen, pancreas, and liver. Belongs to family 18 of glycosyl hydrolases.
6	M55543	GBP2	1p22.2	1	guanylate binding protein 2, IFN-inducible	Binds GTP, GDP, and GMP. Induction by IFN-γ during macrophage activation.
7	M19154	TGFB2	1q41	1	TGF-β2; TGFB2	TGF-β2 has suppressive effects on interleukin-2 dependent t-cell growth.

[a] Information on genes listed in this table is taken from GeneCards.[5]

[b] Ensembl cytogenetic band.

[c] Separating gene according to statistical criteria, see "Materials and Methods" (Table S1 of supplementary information for complete gene list).

indicates that the separation found previously is robust and was not caused by random fluctuations. For each separating cluster of genes (G5, G12, node N505, and node N566) and their combination, we generated a *k*-nearest neighbor classifier (k = 3) and tested its ability to perform class partitions. In all cases, the separation power was found to be significant (Table 5). In addition, we tested whether a three-way separation of the tumor types can be performed based on the combination of the four gene clusters. This gave rise to only one to two errors out of the 20 gliomas, thus validating our claim that the tumor types have different gene expression profiles that can be used to distinguish between them (*P* = 2.49E-06). When using the gene sets separately, the distinction of PrGBM *versus* ScGBM was the least successful. Most classification errors disappeared when all four selected clusters were used. This suggests that distinct sets discriminate between different aspects of the classes. In fact, genes of the same cluster were generally highly correlated across the samples of the same class as a result of the selection procedure, whereas genes of different clusters were not. Although the best performance was obtained by combining all genes, the G5 set alone provided already a good separation power in all cases, including three-way classification where only one to two (8%) errors were made (Table 5).

Validation of Gene Expression Data by Other Methods

Genes found most significantly up-regulated in these glioma as compared with normal brain comprise genes involved in the cell cycle, angiogenesis, migration, and immune response, such as vimen-

tin, fibronectin, *EGFR, VEGF, CDK4, PDGFRA, SPARC, HLA-G,* and *HLA-DRA*. Similar lists have been published before (23, 27–29). An overview of the most differentially expressed genes is available in Table S2 of supplementary information.

To further validate our findings obtained by gene expression profiling, we confirmed a gene known to be overexpressed, *EGFR,* on respective paraffin sections by immunohistochemistry. Furthermore, we validated differential expression of *TNR* on a larger panel of tumors using immunohistochemistry on tissue arrays.

A good correlation between *EGFR* overexpression as measured on cDNA arrays and by immunohistochemical detection of the EGFR on respective paraffin sections (no/low expression *versus* high expression) was observed using a subset of 26 gliomas. The mean expression levels were 265 (SD 101) for samples with no/low expression as determined by immunohistochemistry and 1052 (SD 1369), respectively, for biopsies with high immunopositivity (unpaired *t* test, *P* = 0.035). Please note that expression of EGFR as determined by immunohistochemistry can be focal.

Relative overexpression of *TNR* in LGA as compared with PrGBM and ScGBM was indicated by supervised analysis and its presence in the separating gene cluster N505. We chose to validate TNR, a multifunctional extracellular matrix glycoprotein that is of interest for further evaluations because of its role in adhesion, migration, and differentiation during central nervous system morphogenesis, in particular, differentiation of oligodendrocytes (30). This member of the tenascin family is exclusively expressed in the central nervous system

Table 5 *Verification of k-NN-class prediction using selected gene clusters on validation set of astrocytic gliomas*

Comparisons Gene clusters used by the classifier[a]	Classification Two-Way, k = 3												Classification[c] Three-Way, k = 3	
	ScGBM *vs.* LGA		PrGBM *vs.* ScGBM		PrGBM *vs.* LGA		PrGBM *vs.* (ScGBM + LGA)		LGA *vs.* (ScGBM + PrGBM)				LGA *vs.* PrGBM *vs.* ScGBM	
	4 *vs.* 12		4 *vs.* 4		4 *vs.* 12		4 *vs.* (4 + 12)		12 *vs.* (4 + 4)				12 *vs.* 4 *vs.* 4	
	No. of errors (%)	*P*[b]	No. of errors (%)	*P*	No. of errors (%)	*P*	No. of errors (%)	*P*	No. of errors (%)	*P*			Mean no. of errors (%)	Mean *P*
G5	2 (12%)	0.05	1 (13%)	0.1	1 (6%)	0.007	1 (5%)	0.004	1 (5%)	0.0001			1–2 (8%)	6E-05
N505	2 (12%)	0.03	3 (37%)	1	2 (12%)	0.03	4 (20%)	0.06	3 (15%)	0.003			6 (30%)	0.002
N566	2 (12%)	0.05	2 (25%)	0.5	1 (6%)	0.007	2 (10%)	0.01	3 (15%)	0.004			4 (20%)	0.003
G12	3 (19%)	0.1	2 (25%)	0.5	2 (12%)	0.03	3 (15%)	0.03	7 (25%)	0.02			5 (27%)	0.05
G5 + G505 + G566 + G12	1 (6%)	0.007	0 (0%)	0.03	0 (0%)	0.0005	0 (0%)	0.0002	0 (0%)	8E-06			1–2 (7%)	2E-06

[a] The classifier was built on the training set of 31 astrocytic gliomas, employing the respective gene set and the k-nearest neighbor (k-NN, k = 3) method.

[b] Fisher's exact test.

[c] Classification was repeated 100 times, and *P*s were averaged, see "Materials and Methods" for explanation.
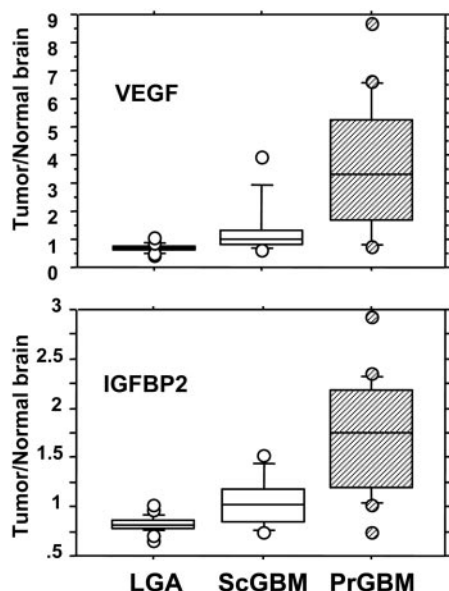
Fig. 4. Angiogenic activity correlates with tumor subtype. The relative expression of *VEGF* and *IGFBP2* as compared with normal brain is shown for the three tumor subtypes, using all samples of the training and validation set. Binary comparisons using the unpaired *t* test between any two tumor subtypes showed highly significant differences for *VEGF* and *IGFBP2,* respectively. LGA *versus* PrGBM: *VEGF,* $P < 0.0001$; *IGFBP2,* $P < 0.0001$; LGA *versus* ScGBM: *VEGF,* $P = 0.0061$; *IGFBP2,* $P = 0.0003$; PrGBM *versus* ScGBM: *VEGF,* $P = 0.0099$; *IGFBP2,* $P = 0.0027$.

in particular by oligodendrocytes. TNR expression was validated on a large panel of gliomas by immunohistochemistry on our GBM-TA and non-GBM glioma-TA, respectively. The criteria for semiquantitative analysis were no/low expression as opposed to high expression. Twenty-nine "pure" LGA were available for analysis on the array applying the same pathology criteria as for the LGA included for gene expression profiling. LGA exhibited significantly higher expression of TNR (83%, 24 of 29) than GBM (63%, 84 of 133 available for analysis; Fisher's exact test, $P = 0.03$).

IGFBP2 Is Coexpressed with VEGF Exerting Most Prominent Expression in Pseudopalisading Cells Surrounding Tumor Necrosis. Most of the genes in G5 are involved in angiogenesis in a broader sense as detailed above. Within this cluster, *IGFBP2* expression displayed the highest correlation ($r = 0.85$) with *VEGF* expression as shown in Fig. 2B. Both genes exhibited significant differential expression between any two tumor subclasses as visualized in a box-plot in Fig. 4, using all 51 astrocytic gliomas of the test and the validation set. To further understand the biology of *IGFBP2,* we addressed the question whether the two genes are involved in the same process. Therefore, *in situ* hybridization was performed on serial frozen sections of glioblastomas for *IGFBP2* and *VEGF,* respectively. Expression of *IFGBP2* and *VEGF* was most prominent in pseudopalisading cells surrounding tumor necrosis of GBM (Fig. 5; *VEGF in situ* hybridization; data not shown). Thus, the *IGFBP2* expression pattern in the glioblastoma tissue is superposable to the reported pattern for *VEGF* (16).

Anoxia Induces Expression of *IGFBPs* in Glioblastoma Cell Lines. Subsequently, we investigated the regulation of *IGFBP2* and two other family members, *IGFBP3* and *5,* under anoxic conditions. In PrGBM, increased expression of *IGFBP2* was often paralleled by overexpression of either *IGFBP3* (comprised in N566; Fig. 3D) or *IGFBP5* (comprised in the list of separating genes, see Table S1 of supplementary information). Glioblastoma cell lines LNZ308 and LN229 displayed *IGFBP2* up-regulation concomitant with *VEGF* after 20 h of anoxia (Fig. 6). Similarly, *IGFBP5* was induced in these

two cell lines, whereas U87 that is wild-type for TP53 exerted most prominent induction of *IGFBP3,* a TP53 target gene. Thus, *IGFBP3* and 5 might also be induced in the same process in gliomas. Induction of different members of the family in these cell lines is in line with a certain redundancy of these genes observed in *IGFBP2* knockout mice. These show increased levels of other *IGFBPs* and display a much less dramatic phenotype than the one initially predicted based on the fetal *IGFBP2* expression pattern (31).

DISCUSSION

Our characterization of astrocytic gliomas by their gene expression profiles revealed consistent inherent differences between LGA, ScGBM, and PrGBM. Thus, the previous histological and clinical recognition of these three glioma entities can be strongly supported by cDNA-array data as reflected by MDS of overall gene expression (Fig. 1). The LGA were found to be the most closely related group and well separated from PrGBM that displayed the most heterogeneous expression profiles. The ScGBM share some of the features of both subgroups, reflecting the common pathogenetic pathway with LGA and the malignant behavior with PrGBM.

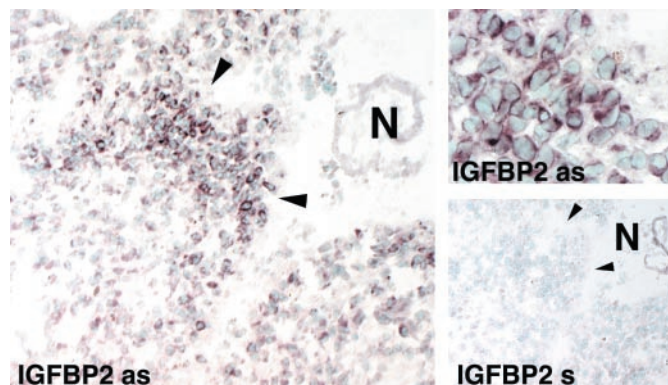Analysis of the expression profiles for correlated genes separating

Fig. 5. *IGFBP-2* mRNA is localized to pseudopalisading cells in human glioblastoma. Hybridization with an *IGFBP-2* antisense cRNA probe yields a strong signal in the pseudopalisading cells, as indicated with *arrows,* surrounding tumor necrosis (N) of a GBM, same region in higher magnification is displayed in the *top right panel. Bottom right panel, IGFBP-2* sense negative control.
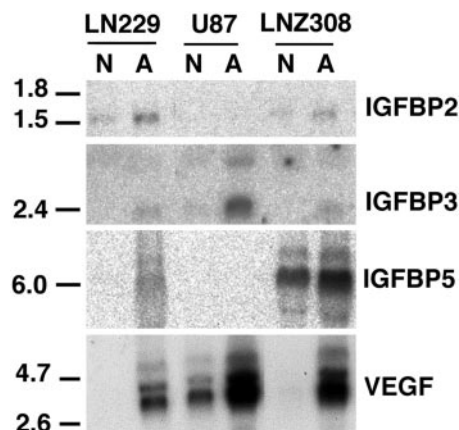

Fig. 6. Induction of *IGFBP2, 3,* or *5* by anoxia in glioblastoma cell lines. Glioblastoma cell lines LN229, U87, and LNZ308 were cultured under normoxic (*N*) or anoxic (*A*) conditions for 20 h. Expression *of IGFBP2, 3,* and *5* was evaluated by Northern blot analysis and compared with induction of *VEGF.* At least one of the analyzed *IGFBP* family members was induced on hypoxic treatment. Note, U87 the only cell line in this series harboring wild-type TP53 exhibits substantial induction of the TP53 target gene *IGFBP3.* Fragment sizes in kilobases are marked on the *left.*

tumor subtypes allowed identification of sets of genes implicating particular biological features that are associated with tumor-specific subtypes. Most strikingly, a cluster (G5) of correlated genes suggests that inherent angiogenic activity, manifested in overexpression of known angiogenic factors, including the most potent, *VEGF,* and *FLT1* and *PTN*, distinguishes PrGBM from the other two groups. In fact, this feature alone was sufficient to classify most astrocytic gliomas (92%) correctly according to their subtype. Glioblastoma are among the most vascularized tumors with characteristic, structurally and functionally abrogated blood vessels of multilayered glomeruloid pattern. Angiogenesis, an essential requisite for tumor growth, is a tightly controlled, complex process involving various pro and antiangiogenic factors that need to act in a concerted fashion to yield functional blood vessels (32). The angiogenic factors VEGF, angiopoietins, and Ephrins act specifically on endothelial cells, because their respective receptors FLT1 and KDR, Tie receptors, and Ephrin receptors are exclusively expressed on these cells. In contrast, other angiogenic factors highly abundant in glioma, such as FGF2, PDGF, and TGF-β, act also on other cell types (33).

According to our analysis, PrGBM appear to have higher angiogenic activity mediated by the genes comprised in G5 than LGA but also higher activity than ScGBM. The difference between PrGBM and ScGBM is unexpected, because they have the same malignancy grade (WHO IV) and cannot be discriminated by classical histology. This may reflect the fact that PrGBM are rapidly growing tumors that cannot keep up with their increasing need for blood supply and thus may suffer from more severe hypoxic conditions triggering angiogenic activity. In contrast, ScGBM may progress over years from LGA and thus may use different pathways for sustained angiogenesis. As compared with PrGBM, LGA display relative overexpression of the angiogenic factor *FGF2* that emerged in the separating gene node N505 (Fig. 3*C*; all LGA *versus* all PrGBM, $P < 0.01$, unpaired *t* test).

An important role for *FGF2* in LGA is also supported by findings from expression profiling efforts reported from small sets of pediatric gliomas (6 LGA *versus* 7 high-grade astrocytoma; Ref. 34) revealing significantly higher *FGF2* expression in LGA. Differential expression of *VEGF,* however, did not reach statistical significance, possibly because of the small numbers of samples. Interestingly, in this study (34) that appeared during revision of this manuscript, the authors used the opposite strategy to analyze their gene expression data obtained from these pediatric gliomas. A knowledge-based list of 133 genes related in some way to angiogenesis was constructed and successfully used to separate LGA from GBM according to hierarchical clustering and PCA. This supervised gene reduction yielded similar separation power as obtained using their full set of 9198 genes present on the chip, thus again emphasizing the important role of angiogenesis in tumor progression.

The biological differences between mechanisms of angiogenesis may have important implications for response to antiangiogenic therapy, which is about to enter the clinics (35).

An important implication of the CTWC method concerns investigation of the genes that belong to identified clusters, here in particular G5, for their biological function, assuming that coexpressed genes might be coregulated and therefore might be involved in the same biological process. The fact that *IGFBP2* is strongly correlated with the angiogenesis genes came as a surprise that will be worth detailed additional studies. The presented association supported by our finding of coexpression of *IGFBP2* and *VEGF* in pseudopalisading cells surrounding necrosis in GBM (Fig. 5) provides strong evidence that *IGFBP2* overexpression in astrocytic gliomas may reflect its implication in angiogenesis. These correlations add to the fact the *IGFBP2* can be induced under anoxic conditions in glioblastoma cell lines (Fig. 6) and mouse embryonic stem cells (26).

Overexpression of *IGFBP2* in GBM has first been discovered by Fuller *et al.* (36) using gene expression profiling. This finding was confirmed (23, 27, 34, 37) and extended to other tumor types. Subsequent immunohistochemical investigations for IGFBP2, a secreted protein, suggested a significant association with the histopathological malignancy grade in glioma (37) with more pronounced immunopositivity in macrophage/microglial and glioma cells around focal necrosis (38). The IGFBPs were first identified and characterized based on their role to bind and modulate the *in vivo* bioactivity of the mitogenic and anabolic peptides IGF-I and IGF-II. More recently, it has become clear that IGFBPs are multifunctional proteins with IGF-dependent and -independent functions in controlling growth and metabolism. The intrinsic bioactivity of IGFBPs, in a positive or negative fashion, depends not only on the cell type but also on their differentiation state and physiological/pathophysiologic condition (39). Furthermore, additional genes of cluster G5 have been linked to the IGF/IGFBP system (*SOCS2* and *3, ZFP36L1*; Refs. 40 and 41). For IGF-I itself, a role in angiogenesis has been established (26, 42). Thus, our observations support the idea to develop a therapeutic approach targeting the IGF-I/IGFBP system in cancer.

Another separating gene cluster emerging from CTWC analysis associated with known biological function is G12 (Fig. 2; Table 3). This gene cluster partitioned a subgroup of PrGBM with increased expression from LGA and comprises only genes related to the immune system. Half of the genes are known to be IFN inducible and are likely expressed by lymphoid cells and/or macrophages and microglia; others represent markers of natural killer cells. These results confirm the observation by Fathallah-Shaykh *et al.* (29) who reported elevated expression of MHC class I and II genes. However, some of the genes are potentially expressed by tumor cells according to our preliminary results. It is worth to note that studies trying to relate immune response in gliomas with survival have not yielded conclusive results. The notion that gliomas produce an immuno-compromised environment by secreting cytokines, such as TGFβ-2 (43, 44), is also supported by this study; TGFβ-2 emerged in node N566 that separates PrGBM from LGA (Fig. 3*D*). Along the same lines, we found HLA-G to be overexpressed (over normal brain, mean 2.9) in PrGBM. This nonclassical class I antigen is mainly expressed at the materno-fetal interface during pregnancy where it plays an important role in maternal-fetal tolerance. HLA-G expression has been associated with tumor progression in several tumor types and has been suggested to contribute to escape from immune surveillance (45, 46). Furthermore, functional assays with human glioma cell lines indicated that few HLA-G-positive cells within a population of HLA-G-negative tumor cells *in vitro* are sufficient to exert a significant immune inhibitory effect (46).

The tumor classifier using expression data of all four identified discriminatory clusters yielded correct tumor prediction in 93% of an independent set of astrocytic gliomas. This remarkable success has to be appreciated in light of the inherent inaccuracy of the initial tumor diagnosis for gliomas, because of sampling errors and subjective histological criteria that lead to striking interobserver differences among neuropathologists (47). In this context, three tumors originally designated as PrGBM attracted our attention (1357, 1297, and 1308), because they appear evenly in clusters characteristic for PrGBM and LGA, respectively, depending on the analysis (Figs. 2 and 3). Central pathology review revised the first (1357) as OAIII, whereas the other two cases were confirmed. Tumor 1297, clinically defined as a PrGBM, presented with genetic characteristics of a ScGBM, namely TP53 mutation and most prominent overexpression of PDGF receptor. Finally, glioblastoma 1308 exhibited overexpression of EGFR, characteristic for PrGBM. However, the biopsy used for gene expression profiling comprised a part of the tumor displaying only features of

WHO grade II (the most malignant part of tumor defines diagnosis). This demonstrates that regions of a glioblastoma with the appearance of a lower grade astrocytoma exert already features only found in PrGBM, exemplified here as overexpression of the EGFR gene and the genes comprised in G12 and N566, characteristic for PrGBM. In contrast, after reclustering S1 (all samples) according to G5, comprising genes indicating high angiogenic activity, this tumor (PrGBM 1308) partitions with LGA. This is biologically convincing, because lower grade astrocytoma have no necrosis and thus are expected to suffer from less hypoxic stress. This observation might be of clinical relevance, knowing the difficulties in tumor sampling at surgery in particular stereotactic biopsies. Thus, the true tumor grade may be underestimated. It follows, the identification of markers specific for glioblastoma that are detectable in apparent lower grade areas of such tumors would allow to improve the diagnostic reliability and ultimately the choice of therapy. Interestingly, in MDS, both 1297 and 1308 were located at the interface of ScGBM and PrGBM.

Combined unsupervised and supervised analysis is a novel approach of gene selection that allows identification of clusters rich in genes informative (by supervised analysis) for tumor classification. CTWC was particularly suitable for this task, because it is designed to go beyond clustering of all genes on the basis of the data from all tumors, and clustering of all tumors, using data from all genes. This is of importance because most of the genes for which expression levels have been measured are irrelevant for the partition sought. CTWC proposes identification of correlated groups (clusters) of genes and using only data from one such group at a time to recluster the tumors, or *vice versa*.

Our gene selection approach has some interesting features, and allowed us: (*a*) to reduce the many discriminatory genes obtained by binary class comparisons; (*b*) enabled an almost correct classification of tumor entities, namely discrimination of LGA from ScGBM, and ScGBM from PrGBM, although pairwise comparisons yielded only few separating genes (partially because of small numbers of ScGBM in the training set, $n = 5$; threshold of FDR of q < 0.05); (*c*) by using the signal of a group of correlated genes, the noise of the individual measurements averages out and is reduced; and most importantly (*d*) the identified gene clusters yielded information on the biological context of coexpressed and possibly coregulated genes. Such insight may give some indication on biological processes determining tumor entities. This fact is highlighted in this study in particular by identification of G5, with high discriminatory power on its own, featuring angiogenesis-related genes.

Analyzing gene expression profiles has yielded biologically relevant information using a limited set of genes (1185). This approach is a first step toward molecular diagnostics for astrocytic gliomas that may improve tumor diagnoses in the future by adding objective criteria. Gene expression profiles may ultimately provide a tool for the identification of patients who are most likely to benefit from targeted therapy. It follows that such approaches for outcome prediction need to be established in clinical trials that in turn will allow rational design of future therapies.

Here, we discussed mostly findings demonstrating similarities within and differences between tumor subtypes for classification purposes. Numerous interesting clusters emerging from this analysis are awaiting attention to gain additional insights into the biology of astrocytic gliomas, *e.g.,* in regard to differentiation and migration (*e.g., TNR*, node 505), and the involvement of the wnt/notch pathways as indicated in N505, and other clusters that have not been discussed in this study [found in G1(S2), see full CTWC analysis on the web[5]].

REFERENCES

1. Kleihues, P., and Cavenee, W. K. Pathology & Genetics. Tumours of the Nervous System. Lyon, France: IARC, 2000.
2. Hegi, M. E., zur Hausen, A., Rüedi, D., Malin, G., and Kleihues, P. Hemizygous or homozygous deletion of the chromosomal region containing the p16INK4a gene is associated with amplification of the EGF receptor gene in glioblastomas. Int. J. Cancer, *73:* 57–63, 1997.
3. Watanabe, K., Tachibana, O., Sato, K., Yonekawa, Y., Kleihues, P., and Ohgaki, H. Overexpression of the EGF receptor and *p53* mutations are mutually exclusive in the evolution of primary and secondary glioblastomas. Brain Pathol., *6:* 217–224, 1996.
4. Hermanson, M., Funa, K., Koopmann, J., Maintz, D., Waha, A., Westermark, B., Heldin, C-H., Wiestler, O. D., Louis, D. N., von Deimling, A., and Nistér, M. Association of loss of heterozygosity on chromosome 17p with high platelet-derived growth factor α receptor expression in human malignant gliomas. Cancer Res., *56:* 164–171, 1996.
5. Zhu, Y., and Parada, L. F. The molecular and genetic basis of neurological tumours. Nat. Rev. Cancer, *2:* 616–626, 2002.
6. Shawver, L. K., Slamon, D., and Ullrich, A. Smart drugs: tyrosine kinase inhibitors in cancer therapy. Cancer Cell, *1:* 117–123, 2002.
7. Getz, G., Levine, E., and Domany, E. Coupled two-way clustering analysis of gene microarray data. Proc. Natl. Acad. Sci. USA, *97:* 12079–12084, 2000.
8. Getz, G., Gal, H., Kela, I., Notterman, D. A., and Domany, E. Coupled two-way clustering analysis of breast cancer and colon cancer gene expression data. Bioinformatics, *19:* 1079–1089, 2003.
9. Getz, G., and Domany, E. Coupled two-way clustering server. Bioinformatics, *19:* 1153–1154, 2003.
10. Stupp, R., Dietrich, P.-Y., Ostermann Kraljevic, S., Pica, A., Maillard, I., Maeder, P., Meuli, R., Janzer, R., Pizzolato, G., Miralbell, R., Porchet, F., Regli, L., de Tribolet, N., Mirimanoff, R. O., and Leyvraz, S. Promising survival for patients with newly diagnosed glioblastoma multiforme treated with concomitant radiation plus temozolomide followed by adjuvant temozolomide. J. Clin. Oncol., *20:* 1375–1382, 2002.
11. Ishii, N., Tada, M., Hamou, M. F., Janzer, R. C., Meagher-Villemure, K., Wiestler, O. D., Tribolet, N., and Van Meir, E. G. Cells with TP53 mutations in low grade astrocytic tumors evolve clonally to malignancy and are an unfavorable prognostic factor. Oncogene, *18:* 5870–5878, 1999.
12. Kononen, J., Bubendorf, L., Kallioniemi, A., Bärlund, M., Schraml, P., Leighton, S., Torhoorst, J., Mihatsch, M. J., Sauter, G., and Kallioniemi, O-P. Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat. Med., *4:* 844–847, 1998.
13. Albertoni, M., Shaw, P. H., Nozaki, M., Godard, S., Tenan, M., Hamou, M-F., Fairlie, W. D., Breit, S. N., Paralkar, V. M., de Tribolet, N., Van Meir, E. G., and Hegi, M. E. Anoxia induces the macrophage inhibitory cytokine-1 (MIC-1) in glioblastoma cells independently of p53 and HIF-1. Oncogene, *21:* 4212–4219, 2002.
14. Ramaswamy, S., and Golub, T. R. DNA microarrays in clinical oncology. J. Clin. Oncol., *20:* 1932–1941, 2002.
15. Menoury, M., Binoux, M., and Babajko, S. IGFBP-2 expression in a human cell line is associated with increased IGFBP-3 proteolysis, decreased IGFBP-1 expression and increased tumorigenicity. Int. J. Cancer, *77:* 874–879, 1998.
16. Plate, K. H., Breier, G., Weich, H. A., and Risau, W. Vascular endothelial growth factor is a potential tumour angiogenesis factor in human gliomas in vivo. Nature (Lond.), *359:* 845–848, 1992.
17. Flaman, J-M., Frebourg, T., Moreau, V., Charbonnier, F., Martin, C., Chappuis, P., Sappino, A-P., Limacher, J-M., Bron, L., Benhattar, J., Tada, M., Van Meir, E. G., Estreicher, A., and Iggo, R. D. A simple p53 functional assay for screening cell lines, blood, and tumors. Proc. Natl. Acad. Sci. USA, *92:* 3963–3967, 1995.
18. Nozaki, M., Tada, M., Kashiwazaki, H., Hamou, M-F., Diserens, A-C., Shinoe, Y., Sawamura, Y., Iwaki, Y., de Tribolet, N., and Hegi, M. E. p73 is not mutated in meningiomas as determined with a functional yeast assay but p73 expression increases with tumor grade. Brain Pathol., *11:* 296–305, 2001.
19. Venables, W. N., and Ripley, B. D. (eds.). Modern Applied Statistics with S-PLUS, Ed. 3, p. 333. New York: Springer Verlag, 1999.
20. Blatt, M., Wiseman, S., and Domany, E. Superparamagnetic clustering of data. Phys. Rev. Lett., *76:* 3251–3254, 1997.
21. Levine, E., and Domany, E. Resampling method for unsupervised estimation of cluster validity. Neural. Comput., *13:* 2573–2593, 2001.
22. Benjamini, Y., and Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B, *57:* 289–300, 1995.
23. Huang, H., Colella, S., Kurrer, M., Yonekawa, Y., Kleihues, P., and Ohgaki, H. Gene expression profiling of low-grade diffuse astrocytomas by cDNA arrays. Cancer Res., *60:* 6868–6874, 2000.
24. Ichimura, K., Bolin, M. B., Goike, H. M., Schmidt, E. E., Moshref, A., and Collins, V. P. Deregulation of the p14ARF/MDM2/p53 pathway is a prerequisite for human astrocytic gliomas with G1-S transition control gene abnormalities. Cancer Res., *60:* 417–424, 2000.

25. Choudhuri, R., Zhang, H. T., Donnini, S., Ziche, M., and Bicknell, R. An angiogenic role for the neurokines midkine and pleiotrophin in tumorigenesis. Cancer Res., *57:* 1814–1819, 1997.

26. Feldser, D., Agani, F., Iyer, N. V., Pak, B., Ferreira, G., and Semenza, G. L. Reciprocal positive regulation of hypoxia-inducible factor 1α and insulin-like growth factor 2. Cancer Res., *59:* 3915–3918, 1999.

27. Rickman, D. S., Bobek, M. P., Misek, D. E., Kuick, R., Blaivas, M., Kurnit, D. M., Taylor, J., and Hanash, S. M. Distinctive molecular profiles of high-grade and low-grade gliomas based on oligonucleotide microarray analysis. Cancer Res., *61:* 6885–6891, 2001.

28. Ljubimova, J. Y., Lakhter, A. J., Loksh, A., Yong, W. H., Riedinger, M. S., Miner, J. H., Sorokin, L. M., Ljubimov, A. V., and Black, K. L. Overexpression of α4 chain-containing laminins in human glial tumors identified by gene microarray analysis. Cancer Res., *61:* 5601–5610, 2001.

29. Fathallah-Shaykh, H. M., Rigen, M., Zhao, L. J., Bansal, K., He, B., Engelhard, H. H., Cerullo, L., Roenn, K. V., Byrne, R., Munoz, L., Rosseau, G. L., Glick, R., Lichtor, T., and DiSavino, E. Mathematical modeling of noise and discovery of genetic expression classes in gliomas. Oncogene, *21:* 7164–7174, 2002.

30. Pesheva, P., and Probstmeier, R. The yin and yang of tenascin-R in CNS development and pathology. Prog. Neurobiol., *61:* 465–493, 2000.

31. Pintar, J. E., Schuller, A., Cerro, J. A., Czick, M., Grewal, A., and Green, B. Genetic ablation of IGFBP-2 suggests functional redundancy in the IGFBP family. Prog. Growth Factor Res., *6:* 437–445, 1995.

32. Yancopoulos, G. D., Davis, S., Gale, N. W., Rudge, J. S., Wiegand, S. J., and Holash, J. Vascular-specific growth factors and blood vessel formation. Nature (Lond.), *407:* 242–248, 2000.

33. Zadeh, G., and Guha, A. Neoangiogenesis in human astrocytomas: expression and functional role of angiopoietins and their cognate receptors. Front. Biosci., *8:* E128–E137, 2003.

34. Khatua, S., Peterson, K. M., Brown, K. M., Lawlor, C., Santi, M. R., LaFleur, B., Dressman, D., Stephan, D. A., and MacDonald, T. J. Overexpression of the EGFR/FKBP12/HIF-2α pathway identified in childhood astrocytomas by angiogenesis gene profiling. Cancer Res., *63:* 1865–1870, 2003.

35. Kerbel, R., and Folkman, J. Clinical translation of angiogenesis inhibitors. Nat. Rev. Cancer, *2:* 727–739, 2002.

36. Fuller, G. N., Rhee, C. H., Hess, K. R., Caskey, L. S., Wang, R., Bruner, J. M., Yung, W. K., and Zhang, W. Reactivation of insulin-like growth factor binding protein 2 expression in glioblastoma multiforme: a revelation by parallel gene expression profiling. Cancer Res., *59:* 4228–4232, 1999.

37. Sallinen, S. L., Sallinen, P. K., Haapasalo, H. K., Helin, H. J., Helen, P. T., Schraml, P., Kallioniemi, O. P., and Kononen, J. Identification of differentially expressed genes in human gliomas by DNA microarray and tissue chip techniques. Cancer Res., *60:* 6617–6622, 2000.

38. Elmlinger, M. W., Deininger, M. H., Schuett, B. S., Meyermann, R., Duffner, F., Grote, E. H., and Ranke, M. B. In vivo expression of insulin-like growth factor-binding protein-2 in human gliomas increases with the tumor grade. Endocrinology, *142:* 1652–1658, 2001.

39. Hoeflich, A., Reisinger, R., Lahm, H., Kiess, W., Blum, W. F., Kolb, H. J., Weber, M. M., and Wolf, E. Insulin-like growth factor-binding protein 2 in tumorigenesis: protector or promoter? Cancer Res., *61:* 8601–8610, 2001.

40. O'Shea, J. J., Gadina, M., and Schreiber, R. D. Cytokine signaling in 2002: new surprises in the Jak/Stat pathway. Cell, *109:* S121–S131, 2002.

41. Corps, A. N., and Brown, K. D. Insulin and insulin-like growth factor I stimulate expression of the primary response gene cMG1/TIS11b by a wortmannin-sensitive pathway in RIE-1 cells. FEBS Lett., *368:* 160–164, 1995.

42. Warren, R. S., Yuan, H., Matli, M. R., Ferrara, N., and Donner, D. B. Induction of vascular endothelial growth factor by insulin-like growth factor 1 in colorectal carcinoma. J. Biol. Chem., *271:* 29483–29488, 1996.

43. Bodmer, S., Strommer, K., Frei, K., Siepl, C., de Tribolet, N., Heid, I., and Fontana, A. Immunosuppression and transforming growth factor-beta in glioblastoma. Preferential production of transforming growth factor-beta 2. J. Immunol., *143:* 3222–3229, 1989.

44. Walker, P. R., and Dietrich, P-Y. Immune escape of gliomas. Prog. Brain Res., *132:* 685–698, 2001.

45. Carosella, E. D., Paul, P., Moreau, P., and Rouas-Freiss, N. HLA-G and HLA-E: fundamental and pathophysiological aspects. Immunol. Today, *21:* 532–534, 2000.

46. Wiendl, H., Mitsdoerffer, M., Hofmeister, V., Wischhusen, J., Bornemann, A., Meyermann, R., Weiss, E. H., Melms, A., and Weller, M. A functional role of HLA-G expression in human gliomas: an alternative strategy of immune escape. J. Immunol., *168:* 4772–4780, 2002.

47. Coons, S. W., Johnson, P. C., Scheithauer, B. W., Yates, A. J., and Pearl, D. K. Improving diagnostic accuracy and interobserver concordance in the classification and grading of primary glioma. Cancer (Phila.), *79:* 1381–1393, 1997.

48. Grove, B. D., and Bruchey, A. K. Intracellular distribution of gravin, a PKA and PKC binding protein, in vascular endothelial cells. J. Vasc. Res., *38:* 163–175, 2001.

49. Alfranca, A., Gutierrez, M. D., Vara, A., Aragones, J., Vidal, F., and Landazuri, M. O. c-Jun and hypoxia-inducible factor 1 functionally cooperate in hypoxia-induced gene transcription. Mol. Cell. Biol., *22:* 12–22, 2002.