

# Molecular Subtypes of Primary Glioblastoma Identified by Gene Expression Profiling

Gheeyoung Choe · Paul S. Mischel<sup>1</sup>

Department of Pathology, Seoul National University College of Medicine, Seoul, Korea; <sup>1</sup>Department of Pathology and Laboratory Medicine, University of California Los Angeles School of Medicine, Los Angeles, California, U.S.A.

Received : July 26, 2002  
Accepted : September 11, 2002

## Corresponding Author

Gheeyoung Choe, M.D.  
Department of Pathology, Seoul National University  
College of Medicine, 28 Yongon-dong, Chongno-gu,  
Seoul 110-744, Korea  
Tel: 02-740-8278  
Fax: 02-743-5530  
E-mail: gychoe@plaza.snu.ac.kr

\*This study was in part supported by a postdoctoral fellowship from Korea Science & Engineering Foundation (KOSEF) (to G.C.) and by U01 CA88127 from the National Cancer Institute (to P.S.M.).

**Background :** The over-expression of the epidermal growth factor receptor (EGFR) occurs in nearly 50% of primary glioblastoma multiforme (GBM). Disruption of multiple signaling pathways is a critical factor in regulating the biological and clinical behavior of GBMs. In the future, therapy that specifically targets these disrupted pathways may represent the best potential treatment for patients with GBM. Large scale gene expression profiling provides a powerful approach to identify these disrupted genetic pathways and to uncover previously unknown molecular subtypes. **Methods :** We used 13 cases of primary GBM biopsy samples obtained from untreated patients and Affymetrix high-density oligonucleotide arrays to identify novel subsets of primary GBMs. **Results :** We showed that the expression of 90 genes differentiate EGFR+ from EGFR non-expressing (EGFR-) *de novo* GBMs, including expression of a number of potentially targetable molecules that act as growth/survival factors for GBMs. We also demonstrated the presence of two additional molecular subtypes of primary GBMs, including one characterized by the coordinate upregulation of contiguous genes on chromosome 12q13-15, which has a distinct global gene expression profile and expresses both astrocytic and oligodendroglial genes. **Conclusion :** We have shown that there are EGFR+ primary GBMs, GBMs with coordinate upregulation of genes on chromosome 12q13-15, and primary GBMs lacking either alteration. Moreover, they have distinct transcriptional profiles. Our findings strongly suggest that the three GBMs are biologically different tumor types, despite their identical microscopic appearance, and provide an important first step in developing a molecular taxonomy of GBMs.

**Key Words :** Glioblastoma-Genes, erbB-1-Gene Expression Profiling-Oligonucleotide Array

## INTRODUCTION

Glioblastoma (GBM) is the most common malignant primary brain tumor in adults.<sup>1</sup> It is also among the most lethal cancers, with a median survival rate of approximately one year despite surgery and intensive radiation and chemotherapy. One possibility accounting for this poor outcome is that GBMs, despite their morphological commonalities, may actually be a diverse set of tumors that require different targeted therapies. Most researchers agree that GBMs are highly heterogeneous tumors, pathologically, clinically and molecularly, but the clinical implications of this heterogeneity are not clear.<sup>1</sup> One distinction among GBMs is based on clinical presentation. Primary GBMs arise *de novo* as grade IV tumors; secondary GBMs progress from a lower grade glioma.<sup>2</sup> In primary GBMs, epidermal growth factor receptor (EGFR) overexpression/amplification is detected in nearly two-

thirds of cases, suggesting a major role of EGFR in GBM pathogenesis.<sup>3</sup> Currently, the clinical and therapeutic implications of EGFR overexpression are far from clear. More importantly, the considerable heterogeneity in clinical and molecular features suggests that there may be multiple biological subtypes of primary GBMs.

A new approach to cancer therapy focuses on inhibiting signal transduction pathways that are constitutively activated in the tumor cells.<sup>4</sup> Pharmacological agents that specifically target these signaling pathways demonstrate considerable promise.<sup>4,5</sup> For example, the kinase inhibitor STI-571 (Gleevec) that targets the constitutively active Bcr-Abl kinase promotes clinical remission in patients with chronic myelogenous leukemia.<sup>6</sup> Because GBMs have a number of clearly defined signal transduction abnormalities, among which cooperative disruption appears to be a critical factor in regulating their biological and clinical behavior,<sup>7</sup> GBMs

may be appropriate for targeted molecular therapy. In fact, targeted molecular therapy has shown considerable efficacy and specificity in pre-clinical GBM models,<sup>5</sup> suggesting the feasibility of this approach. Therefore, the current challenge is to identify distinct molecular subgroups of primary GBMs with characteristic signal transduction pathway abnormalities that can be targeted.

There are multiple approaches to identifying biologically distinct tumor populations. One approach is to analyze patterns of gene expression to identify coordinately upregulated transcriptional pathways, which may influence the growth and survival of cancer cells as well as their response to therapy.<sup>8,9</sup> Large scale gene expression profiling provides a powerful approach to identifying these upregulated signaling pathways. It may also promote the discovery of previously unrecognized tumor subtypes, characterized by specific signal transduction abnormalities. Gene expression profiling, combined with appropriate analytic methods, can be used to detect distinct molecular signatures associated with specific mutations.<sup>10</sup> These approaches can also be used to discover previously unrecognized tumor subtypes with distinct molecular and/or clinical phenotypes or responses to therapy.<sup>8-11</sup> We hypothesized that the overexpression of EGFR has a major effect on the transcriptional profile of primary GBMs, including upregulation of genes whose products are involved in promoting tumor cell growth/survival, and which may ultimately provide targets for therapy. We further hypothesized that there are additional molecular subtypes of EGFR-negative GBMs that are characterized by their own transcriptional profiles, which may provide a different set of potentially therapeutic targets. In this paper, we demonstrate a global gene expression signature for EGFR-expressing GBMs. We show that the signature can be characterized by a relatively small number of genes, many of which are signal transduction molecules. We also find evidence for two types of EGFR negative GBMs, based on their gene expression signatures. The transcriptional profiles of these EGFR negative GBMs suggest the upregulation of different signaling pathways, which may require a distinct set of targeted molecular therapies.

## MATERIALS AND METHODS

Sample selection, immunohistochemistry and RNA isolation

To exclude any potential effects of treatment on gene expression, we studied GBMs in untreated patients who had developed *de novo* (primary GBMs). We collected GBM cases from UCLA neuropathology files dating between January 1995 and June 2001. We reviewed medical records of the glioblastoma patients and

selected cases that were clinically *de novo* and had never been treated. We collected 48 cases of untreated primary GBM. We made tissue microarray blocks from 48 cases. Three representative 0.6 mm cores (two tumor, one normal) were obtained from diagnostic areas of paraffin-embedded biopsy tissue from primary GBM patients and inserted into a grid pattern in a recipient paraffin block using a tissue arrayer.

Five (m sections were cut from the tissue microarray block and original (donor) paraffin blocks, and immunohistochemistry was performed using the mouse monoclonal antibody EGFR (Clone 31G7, Zymed, San Francisco, CA, U.S.A.). The sections were baked at 60°C and deparaffinized with xylenes and graded ethanol. Slides were treated with 3% hydrogen peroxide in methyl alcohol to quench endogenous peroxidase activity and then with pronase (0.03 g/mL of 0.05 M Tris buffer, pH 7.4) at 37°C for 8 minutes. Immunostaining involved sequential applications of primary antibody (EGFR at 1:150) for 16 hours at 4°C, followed by anti-mouse biotinylated immunoglobulins (Vector) at 1:100 dilution for one hour, and finally, avidin-biotin complex (Elite ABC, Vector) for one hour. Negative control slides received normal mouse serum (DAKO) as the primary antibody. Diaminobenzidine tetrahydrochloride was used as the enzyme substrate to visualize specific antibody localization and Harris hematoxylin was used as a nuclear counterstain. Tumors demonstrating strong EGFR immunopositivity in greater than 20% of tumor cells were considered to be positive.

We recruited snap frozen tumor samples of the selected cases, from which high quality RNA could be obtained. Thirteen cases consisting of 9 males and 4 females were available for analysis. Patients' ages ranged from 40 to 75 (median:60). All patients participating in this study gave informed consent prior to surgery. At the time of resection, the tumors were examined by a neuropathologist and dissected into two portions, one for tissue diagnosis and the other for RNA extraction. This procedure was done within 15 minutes after surgical resection. The portion for RNA extraction was snap frozen in liquid nitrogen and stored at -80°C. Total RNA was extracted from 100-150 mg of frozen tissue by using Trizol (Gibco BRL, Carlsbad, CA, U.S.A.) and one round of cleanup by Qiagen RNeasy total RNA isolation kit (Valencia, CA, U.S.A.). 10 µg of total RNA was used to generate double stranded cDNA. Afterwards, biotin labeled antisense cRNA was produced by in vitro transcription using the ENZO BioArray HighYield kit (New York, NY, U.S.A.). 20 µg of cRNA were fragmented and was hybridized to Affymetrix U95Av2 GeneChip arrays (Santa Clara, CA, U.S.A.) containing more than 12,000 characterized genes. The GeneChips were washed, stained with

streptavidin phycoerythrin and scanned to generate an image file. The quality, yield and size distribution of the total RNA, labeled transcripts, and fragmented cRNA were estimated by spectrophotometric analysis and the RNA 6000 Nano-LabChips (Agilent Technologies, Palo Alto, CA, U.S.A.).

### Preprocessing and Statistical Analysis

The expression of 12,533 probe sets encoding 10,000 genes (Affymetrix U95av2 oligonucleotide arrays) in each patient sample was detected and quantified using model based indices. The CEL files for all the microarray hybridizations generated by Affymetrix Microarray Suite Software were imported into the software dChip to compute the model based expression index for each gene.<sup>12</sup> All arrays were normalized against the array with median overall intensity. The genes showing a coefficient of variation of less than 0.5 across all samples were excluded from further analyses. For selecting differentially expressed genes (gene filtering), a thresholding approach was used. Genes with a fold change exceeding 1.5 and an absolute difference in their model based expression index bigger than 50 between two groups of samples were selected. To validate this gene-filtering criterion, we computed the leave-one-out cross-validation error rate of 2 prediction methods that used the same gene filtering criterion in their construction. To be specific, we trained (constructed) a k-nearest neighbor and a gene voting predictor on all the samples by using the same gene filtering criterion.<sup>8</sup> The resulting predictors were applied to the left-out observations, which comprised the test set, and the misclassification error rate was recorded. This was repeated for every sample and the leave-one-out cross-validation error rate was calculated as the average misclassification rate. We used dChip to perform a hierarchical clustering of the samples or genes using Euclidean distance.<sup>13</sup> Classical multidimensional scaling (MDS) is another unsupervised learning method, which takes the dissimilarities between tumor samples and returns a set of points in a lower dimensional space such that the distances between the points are approximately equivalent to the dissimilarities. Since Euclidean distances between samples were used, the MDS plot produced here is equivalent to plotting the samples with their first 3 principal components. We used the statistical software S-Plus ([www.insightful.com](http://www.insightful.com)).

## RESULTS

First, we determined the EGFR status of each tumor on a pro-

tein level (by immunohistochemistry) and on a transcript level (by microarray assay). Immunohistochemical overexpression of EGFR was noted in 21 out of 38 cases (55%) on the tissue microarray block. The cases used in gene expression profiling showed immunohistochemical overexpression of EGFR in 4 out of the 13 cases. The results of EGFR immunohistochemistry using the tissue microarray block corresponded with those using original (donor) paraffin blocks. By Affymetrix U95av2 oligonucleotide arrays, 8 of the thirteen samples did not have detectable EGFR transcripts and one sample contained a barely detectable EGFR signal; no EGFR protein expression was detected in these tumors by immunohistochemistry (Fig. 1). In contrast, four samples had very high levels of EGFR transcripts and were strongly immunoreactive for EGFR. In these EGFR immunopositive cases, the EGFR transcript level was greater than 27 fold increased ( $p < 0.0001$ ) (Table 1). Therefore, nine of the thirteen test tumor samples were EGFR negative and four were strongly EGFR positive.

According to the thresholding approach, ninety genes (101 probe sets) were differentially expressed between the EGFR+ and EGFR-GBMs (Fig. 2). By a cross validation analysis to assess the validity of this thresholding approach, the misclassification rate turned out to be zero. All 13 samples were correctly classified regarding EGFR status in the leave-one-out cross validation analysis.

The EGFR+ GBMs were clearly separable from the EGFR-

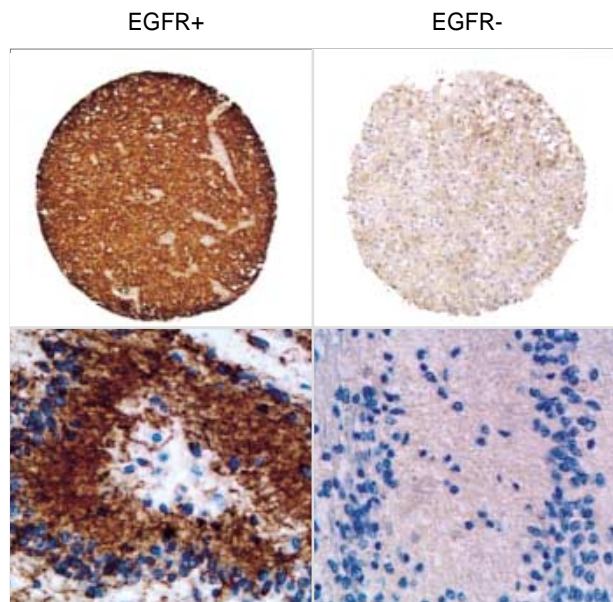


Fig. 1. Immunohistochemical staining for epidermal growth factor receptor (EGFR). Left column shows EGFR+ cases and right column shows EGFR- cases. Upper row shows immunostains using tissue microarray block and lower row shows those using original (donor) paraffin blocks.

tumors by hierarchical clustering, a standard unsupervised learning method.<sup>13</sup> One of the clusters was significantly enriched for EGFR+ GBMs (Fisher's exact test;  $p < 0.007$ ). Further, the dendrogram suggested the presence of two subtypes of EGFR negative GBMs (Fig. 2). One subset (right branch of EGFR-GBMs in dendrogram) was characterized by increased expression of a set of contiguous genes on chromosome 12q13-15, including *cyclin dependent kinase 4 (CDK4)* (12q14), *sarcoma amplified sequence (SAS)* (12q13-14), *amplified in osteosarcoma (OS-9)* (12q13.2) and *conserved*

*gene amplified in osteosarcoma (OS-4)* (12q13-15) (Fisher's exact test;  $p < 0.0014$ ) (Fig. 2). These genes have previously been shown to be overexpressed in approximately 10-20% of high grade gliomas, usually in association with a DNA amplification event.<sup>14,15</sup> Therefore, we assessed the expression of additional probe sets from this chromosomal locus that are represented on the Affymetrix U95av2 oligonucleotide arrays. Expression of all of the contiguous genes in this chromosomal locus that were present on the high density oligonucleotide array (*OS-9*,

**Table 1.** Genes upregulated in EGFR overexpressing glioblastomas

Gene	Ratio of EGFR+: EGFR-	p value	Function
EGFR	27	0.0001	Growth factor receptor for GBM cells-signal transduction
Pleiotrophin (PTN)	1.9	0.002	Growth factor for GBM cells-signal transduction
PTRPZ1	1.8	0.03	PTN receptor-signal transduction
VEGF	3.2	0.05	Angiogenesis in GBMs-signal transduction
Endothelin B receptor	2.6	0.05	Survival factor for GBM cells-signal transduction
GS3955	2.3	0.003	Putative serine/threonine kinase-signal transduction
TEGT (Bax inhibitor 1)	2.2	0.006	Anti-apoptotic factor-signal transduction
SRI (sorcin)	2.3	0.0002	Chemoresistance-putative multidrug resistance gene
MYO10 (Myosin X)	3.0	0.0004	Motility factor
SLC1A3	2.4	0.0004	High affinity glutamate transporter
Na+/K+ ATPase, Alpha 2 subunit	4.0	0.01	Na/K ATPase subunit-transporter
MLC1	3.3	0.006	Novel brain-expressed cell surface protein-suggested transporter function
AEBP1	2.4	0.005	Transcriptional repressor
Cyclin D2	2.0	0.03	Progression through G1 phase of the mammalian cell cycle
CD99/MIC2	1.5	0.06	Cell surface maker

EGFR: epidermal growth factor receptor, PTRPZ1: protein-tyrosine phosphatase zeta 1, VEGF: vascular endothelial growth factor, TEGT: testis-enhanced gene transcript, SLC1A3: swine leukocyte C locus-1 A3, Na+/K+ ATPase: Na+/K+ adenosine triphosphatase, MLC1: megalencephalic leukoencephalopathy with subcortical cysts 1, AEBP1: adipocyte-enhancer binding protein 1.

**Table 2.** Genes upregulated in 12q13-15 overexpressing glioblastomas

Gene	Fold increase	p value	Function
OS-9	4.4	<0.0001	12q: amplified in osteosarcoma and in GBM (unknown function)
OS-4	3.2	0.001	12q: amplified in osteosarcoma and GBM (unknown function)
SAS	4.7	<0.0001	12q: amplified in osteosarcoma and GBM (member of transmembrane 4 superfamily)
METTL1	5.6	0.0004	12q: methyltransferase function
CDK4	5.8	<0.0001	12q: complexes with cyclin D1, promotes proliferation, commonly overexpressed in GBM
Cyclin D1	4.9	0.04	Complexes with cdk4, promotes proliferation, commonly overexpressed in GBM
CENTG1 (PIKE)	6.4	<0.0001	12q: PI3K enhancer regulates cyclin D1
CYP27B1	2.7	<0.0001	12q: GAS89, 25-hydroxyvitamin D3 1, alpha hydroxylase - increased in GBM
MAG	8.4	0.008	Oligodendroglial membrane protein important for myelination
MBP	6.4	0.02	Oligodendroglial protein involved in myelination
PLP1	4.5	0.01	Oligodendroglial protein involved in myelination
Nkx2.2	3.9	0.002	Oligodendroglial precursor differentiation homeodomain transcription factor
SOX10	7.4	0.006	Oligodendroglial differentiation transcription factor
MAL	4.4	0.03	Integral membrane protein in mature oligodendrocytes (MVP17, VIP17)
Scg10	3.9	0.06	Neuronal growth associated protein
Basp1	2.4	0.02	Intracellular signaling molecule with increased expression in cancer cell lines
Autotaxin	7.8	0.03	Potent cancer cell motility factor

OS: gene amplified in osteosarcoma, SAS: sarcoma amplified sequence, METTL1: methyltransferase-like 1, CDK4: cyclin dependent kinase 4, PIKE: PI3K enhancer, CYP27B1: 25-hydroxyvitamin D3 1 alpha-hydroxylase, GAS89: glioblastoma amplified sequence 89, MAG: myelin-associated glycoprotein, MBP: myelin basic protein, PLP1: proteolipid protein 1, SOX10: SRY-like HMG box family of transcription factors, MAL: myelin and lymphocyte protein, MVP17: myelin vesicular protein of 17 kDa, Scg10: superior cervical ganglia, neural specific 10, Basp1: brain acid-soluble protein 1.

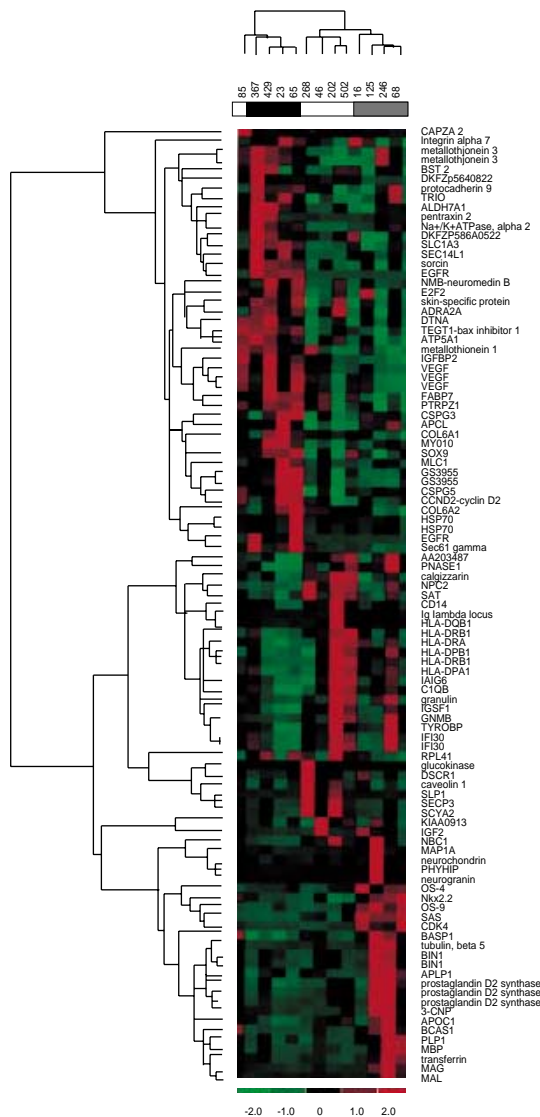


Fig. 2. One hundred and one probe sets (90 genes) are differentially expressed in epidermal growth factor receptor (EGFR) positive glioblastomas (GBMs) and EGFR negative GBMs. Hierarchical clustering dendrogram based on expression of these 90 genes is shown on the top. Black bar denotes tumors with EGFR overexpression, gray bar denotes GBMs with overexpression of genes on chromosome 12q13-15, and white bar denotes GBMs lacking either alteration.

*CENTG1 (PIKE), SAS, CDK4, OS-4, CYP27B1-glioblastoma amplified sequence 89, and METTL1*) was significantly increased in this subset of GBMs ( $p < 0.001$  for each gene) (Fig. 3, Table 2). We next determined whether this set of 7 genes is coordinately regulated in other experiments. In 368 independent experiments performed at our cDNA microarray core facility, we found no evidence of coordinate transcriptional regulation of these 7 genes. The other subset of EGFR- GBMs showed an increased expression of neither EGFR gene nor the contiguous

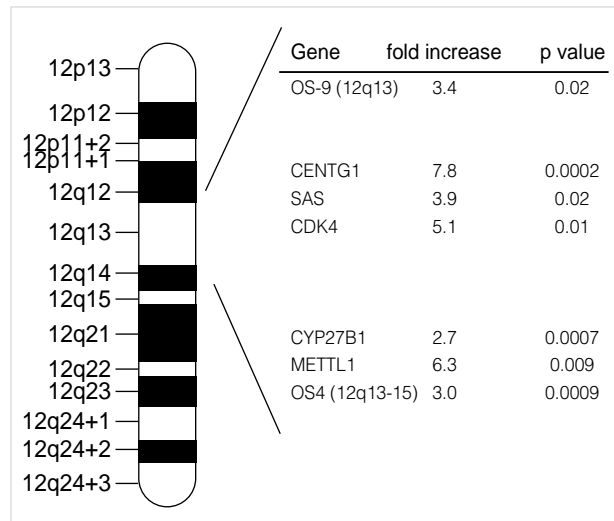


Fig. 3. Genes with coordinate upregulation mapping to chromosome 12q13-15.

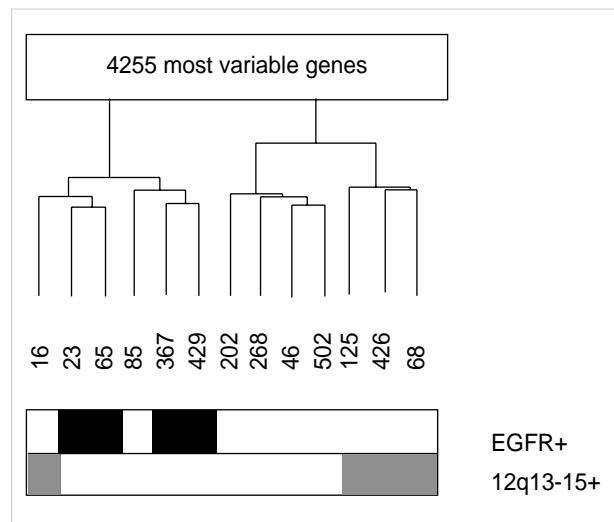


Fig. 4. Hierarchical clustering based on the most variable genes, independent of epidermal growth factor receptor (EGFR) status. One cluster is enriched for EGFR+ tumors (black bars) (Fisher's exact test,  $p = 0.02$ ). Another cluster is enriched for glioblastomas with coordinate upregulation of 12q13-15 genes (gray bars) (Fisher's exact test,  $p = 0.014$ ), and the third cluster is enriched for glioblastomas lacking either alteration (Fisher's exact test,  $p < 0.007$ ).

genes on chromosome 12q13-15. No genes were upregulated globally (in more than 3 out of 5 samples) in this EGFR-/non-12q13-15 upregulated subset. Taken together, these results suggest that there are three molecular subsets of GBMs: those with EGFR expression, those with contiguous upregulation of genes on 12q13-15, and those lacking either change.

To validate the presence of the three distinct groups of GBMs, and to determine whether these molecular subsets had global

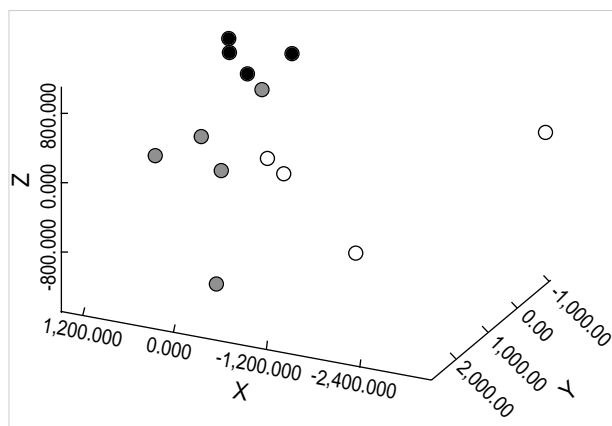


Fig. 5. Multidimensional scaling plot demonstrates that epidermal growth factor receptor (EGFR)+ (black), 12q13-15+ (gray) and EGFR-/12q13-15- (white) primary glioblastomas have distinct gene expression profiles.

transcriptional correlates, we analyzed a less selected set of genes from the microarray data. To obtain genes with a sufficiently strong signal, we restricted the analysis to those probe sets with a coefficient of variation  $>0.5$  (4,255 genes), regardless of EGFR status. We then performed hierarchical clustering of the 13 GBMs based on expression of these 4,255 most variable genes. The dendrogram indicated the presence of three global transcriptional groups, one of which was significantly enriched for EGFR+ GBMs (Fishers exact test;  $p=0.02$ ) (Fig. 4). The dendrogram also indicated the presence of two EGFR- groups, one of which was significantly enriched for 12q13-15 upregulated GBMs (Fishers exact test;  $p=0.014$ ). The other group was comprised of non-EGFR expressing, non-12q13-15 upregulated GBMs (Fishers exact test;  $p<0.007$ ). This hierarchical clustering pattern was observed over a range of thresholds (coefficients of variation between 0.4 and 0.6; 3,000-7,000 genes). An alternative view of tumor groupings was obtained by performing multidimensional scaling of the 4,255 genes onto a three dimensional plot (a form of principal component analysis). The EGFR+ tumors were separable from the EGFR- tumors (Fig. 5), and the 12q13-15 upregulated EGFR- GBMs were distinct from the EGFR-/non-12q13-15 upregulated GBMs. These results lend support to the hypothesis that EGFR+ GBMs are a distinct subset and further suggest that these additional molecular subsets of EGFR negative GBMs are robust.

The gene expression signatures of EGFR overexpressing GBMs were notable for upregulation of growth factors, receptors and signal transduction molecules (over 1/3 of the probe sets), including vascular endothelial growth factor (VEGF), which plays a critical role in GBM angiogenesis and progression.<sup>16</sup> In addition,

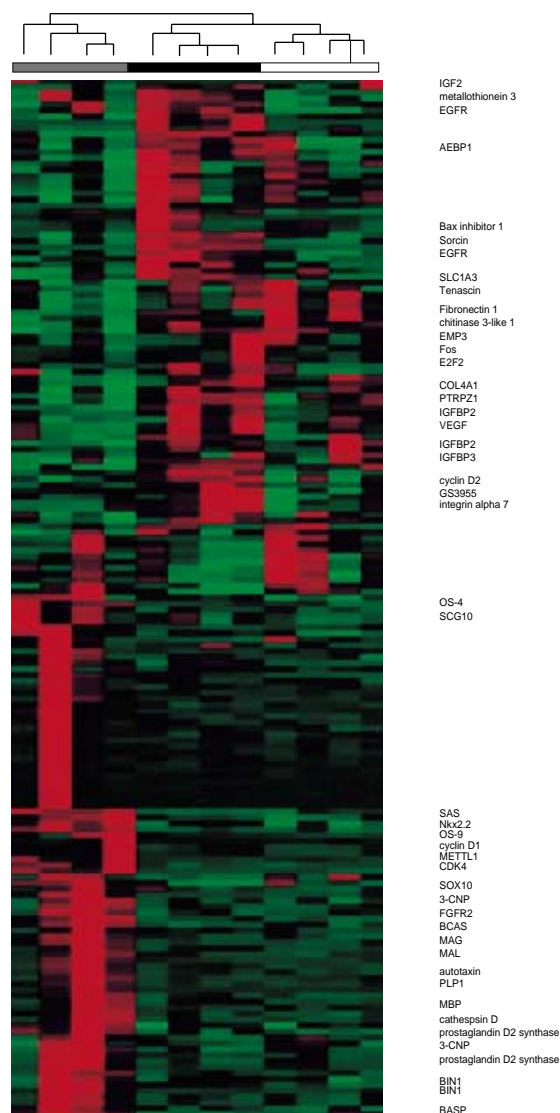


Fig. 6. One hundred and seventy five probe sets (157 genes) are differentially expressed between epidermal growth factor receptor (EGFR) + and 12q13-15 overexpressing glioblastomas (GBMs). Hierarchical clustering dendrogram based on expression of these 157 genes is shown on the top. Black bar denotes tumors with EGFR overexpression, gray bar denotes GBMs with overexpression of genes on chromosome 12q13-15, and white bar denotes GBMs lacking either alteration. Due to limited space, only selected gene names are shown.

pleiotrophin (PTN) and its receptor PTRPZ1, and endothelin B receptor [ET(B)] were all upregulated in EGFR+ GBMs, as was the anti-apoptotic protein BAX inhibitor 1 (TEGT) (Table 1). The EGFR+ GBM gene signature was also notable for increased expression of plasma membrane-bound transporters and channels, including the multi-drug chemo-resistance gene SRI (sorcin) and MLC1, a recently cloned cell surface protein whose mutation is associated with white matter brain defects.<sup>17,18</sup> These data suggest

the upregulation of multiple growth factor-mediated signal transduction pathways in EGFR expressing GBMs that can promote GBM cell proliferation, survival and angiogenesis; and they also suggest potential candidates for targeted molecular therapy.

The gene expression signature of GBMs with overexpression of genes on 12q13-15 was notable for the increased expression of CDK4 (located on 12q13-15), its binding partner cyclin D1 and CENTG1 (PIKE), a newly identified signal transduction molecule that enhances PI3K activity and increases cyclin D1 activity.<sup>19</sup> Taken together, these suggest the coordinate upregulation of the cyclin D1 pathway in tumors with the 12q13-15 expression signature. In contrast, the EGFR overexpressing GBMs expressed elevated levels of cyclin D2 (2 fold increase,  $p=0.03$ ). The 12q13-15 upregulated GBMs were also significant for high levels of transcription of oligodendroglial genes (*MBP*, *MAG*, *PLP1*, *Nkx2.2*, *Sox10* and *MAL*) (Table 2). Morphologically, these tumors were indistinguishable from the other GBMs; they lacked oligodendroglial morphology and had equivalent levels of the astrocytic marker GFAP. These results raise two possibilities: either GBMs with coordinated upregulation of genes on 12q13-15 may arise from a less committed glial precursor cell, or they may activate a pattern of gene transcription characteristic of dedifferentiation.<sup>7</sup> Other researchers have observed that these genes may be coordinately elevated in some GBMs, but the observation that coordinated transcriptional upregulation of genes over this region confers a distinctive global gene expression profile and alters cell commitment phenotype is a novel finding.

To further characterize the biological differences between the EGFR+ and 12q13-15 GBMs and to identify any additional signal transduction pathways that might be targeted for future therapy, we used the same filtering procedure for twelve cases consisting of four cases per each molecular subtype. Because EGFR negative sample #85 had a global transcriptional pattern similar to the EGFR+ tumors, this exceptional case was excluded from the procedure. This resulted in 175 probe sets (157 genes) that were differentially expressed between EGFR+ and 12q13-15+ GBMs (Fig. 6). Hierarchical clustering of these tumors based on these 157 genes continued to support the presence of three primary GBM subtypes (Fig. 6). In addition to the previously mentioned genes, the EGFR+ tumors had increased expression of extracellular matrix proteins including tenascin C and fibronectin, which play a role in GBM cell invasion,<sup>20,21</sup> while the 12q13-15+ group had very high transcript levels of autotaxin, a secreted motility factor that promotes tumor cell invasion and metastasis.<sup>22</sup> These results suggest that these GBM subtypes may differ in their invasion patterns. The results also suggest the pres-

ence of additional potential biological, and perhaps clinical differences.

## DISCUSSION

The new challenge in cancer biology is to move from purely morphological classification of tumors to a classification that is based on molecular criteria. In light of the development of new pharmacologic pathway inhibitors for cancer therapy, this goal is now even more important. GBMs may be an ideal tumor type for this approach for the following reasons: 1) they have a number of clear-cut, potentially targetable signal transduction abnormalities that influence their biological behavior, 2) inhibition of these pathways with small molecules has yielded promising results in GBM pre-clinical models,<sup>5</sup> and 3) none of the current therapies are highly effective. To approach this goal, we used gene expression profiling to uncover three novel molecular subsets of primary GBMs: EGFR+ GBMs, GBMs with upregulation of genes on chromosome 12q13-15, and GBMs lacking either of these changes. These molecular subsets have previously been indistinguishable by current histopathological criteria, but here we show that they have distinct transcriptional profiles. We show that these tumor types can be distinguished by a relatively small number of differentially expressed genes (90 genes), many of which are themselves signal transduction molecules that promote the growth and survival of GBMs. This work demonstrates the value of using a genomic approach to identify different molecular subtypes of GBMs and suggests possible new therapeutic targets for each of these molecular subtypes.

Our observation that EGFR expression confers a distinctive transcriptional phenotype to primary GBMs is a novel finding. Previous works from our own laboratory and from the laboratories of other investigators suggest that EGFR signaling alters the biological behavior of GBMs.<sup>23</sup> The fact that EGFR expression globally impacts the transcriptional program lends support to the hypothesis that EGFR+ GBMs are biologically distinct from other histologically similar GBMs. Many of the genes that are differentially upregulated in EGFR overexpressing GBMs are themselves signal transduction molecules, some of which promote growth, survival and angiogenesis. VEGF, ET (B), PTN and its receptor PTRPZ1, and BAX inhibitor 1 were all upregulated in EGFR+ GBMs. This is consistent with previous work demonstrating that EGFR activation transcriptionally upregulates VEGF expression in GBM cells.<sup>24</sup> VEGF plays a major role in promoting angiogenesis and tumor growth of

GBMs *in vivo*,<sup>16</sup> and may therefore be critical for EGFR-mediated pathogenesis. This also suggests an important molecular basis on which to select GBM patients for anti-angiogenic therapy. ET(B) is expressed by GBM cells *in vivo*; its ligand ET-1 is secreted by the GBM tumor vasculature. Since activation of this receptor promotes a pro-survival/antiapoptotic cascade in GBM cells, our results suggest that this pathway may play a role in protecting EGFR+ cells from apoptosis. Similarly, Bax inhibitor 1, a recently cloned anti-apoptotic protein, is also upregulated in EGFR+ GBMs suggesting another mechanism by which EGFR+ cells may escape apoptosis. We also found that PTN, and its receptor PTRPZ1,<sup>25</sup> were upregulated in EGFR+ GBMs. PTN is a potent growth factor for GBM cells in culture. Interestingly, transcription of PTRPZ1, one of the receptors for PTN, was also elevated in the EGFR+ samples, suggesting coordinate upregulation of this pathway.

By analyzing the genes that most distinguish EGFR+ from EGFR-GBMs, we uncovered a novel subtype of GBMs characterized by coordinate upregulation of genes on chromosome 12q13-15. Other investigators have previously shown amplification or upregulation of a set of genes on 12q13-15 in approximately 10-20% of GBMs.<sup>14,15</sup> However, our finding that this subset of GBMs has a distinctive global gene expression pattern is highly novel, and it suggests that these tumors are biologically distinct. We analyzed the expression of all probe sets in this locus and found that their expression was upregulated in this subset of GBMs. We found no evidence of coordinated transcriptional regulation of these genes across a wide variety of experiments with different tumor tissues, normal tissues and cell lines. Our finding strongly suggests the presence of a genomic amplification event. Previous refined mapping studies showed that the amplicon commonly associated with 12q13-15 in GBMs did not include MDM2 (12q13.5-15), although it can also be amplified in GBMs.<sup>15</sup> Similarly, we found no evidence of MDM2 upregulation in these tumors, confirming these previous findings.

Further analysis of the most differentially expressed genes clearly indicated that GBMs with 12q13-15 upregulation had a markedly different set of potentially targetable pathways, and the 12q13-15+ GBMs appeared to have a potentially important cell-fate specification difference. 12q13-15+ tumors were remarkable for upregulation of oligodendroglial genes, including (*MAG*, *MBP*, *PLP1*, *Nkx2.2*, *Sox10*, *MAL* and 2',3'-cyclic nucleotide 3' phosphodiesterase-3-CNP).<sup>26</sup> Morphologically, these tumors were GBMs, not oligodendrogliomas or mixed gliomas, and they contained equal levels of GFAP transcripts relative to the non-12q13-15+ tumors. Furthermore, previous studies have suggest-

ed that two of these markers, PLP1 and MBP, can be expressed by some astrocytomas.<sup>26</sup> This raises two important possibilities: development from a multi-potent precursor cell or development from dedifferentiation of an astrocytic cell. Researchers have long suspected that GBMs may arise from precursor cells.<sup>27</sup> Alternatively, terminally differentiated astrocytic tumor cells can be dedifferentiated into neural stem cells by activation/disruption of specific signal transduction pathways.<sup>7,28</sup> Most importantly, our finding that coordinated transcriptional upregulation of the 12q13-15 region confers a distinctive global gene expression profile and alters cell commitment phenotype is novel and potentially important.

Are the subsets we identified biologically distinct? There are multiple approaches to assessing and developing molecular subsets. Since gene expression probably has a powerful effect on overall tumor behavior, using global genomic strategies such as this is a legitimate way to dissect out biologically distinct tumors. However, the bottom line of these analyses will be the identification of distinct subsets that have similar responses to therapy. On initial analysis, the molecular subsets identified here were not associated with clear survival differences. However, given the extremely short median survival of GBM patients, and the lack of consistent response to any of the current therapies, this result is not surprising. The real utility of this approach lies in its ability to generate potentially targetable genes and pathways. Considering the potential to target genes and pathways upregulated in these specific tumor subsets, (e.g. VEGF and PTN in EGFR+ GBMs; cyclinD1 and CDK4 in 12q13-15+ GBMs), we think it is likely that this approach may be fruitful. Verification awaits future pre-clinical studies and clinical trials.

Our study has a number of limitations. First, a number of other genes that regulate signal transduction, such as PTEN, p53, p16/Ink4a, and PDGFR are also commonly mutated in GBMs,<sup>1</sup> which may have a profound impact on transcriptional profiles. In the future, it will be important to assess the effect of these mutations on GBM gene expression profiles, and to assess their potential interaction with the molecular lesions identified here (EGFR+, 12q13-15+). Second, it is also important to consider that different upstream lesions may have similar downstream signal transduction consequences, and thus, may promote similar transcriptional profiles. Our observation that EGFR negative sample #85 has a global transcriptional pattern similar to the EGFR+ tumors, may suggest that similar signaling pathways are being activated by a different upstream lesion. Third, in nearly 50% of GBMs with high-level EGFR expression, there is co-expression of mutant EGFRs, most commonly the EGFR variant III.

EGFRvIII results from an in-frame genomic deletion of exons 2-7, producing an EGFR that lacks its ligand binding domain.<sup>29</sup> EGFRvIII is constitutively active and oncogenic, and may have profound transcriptional consequences. Unfortunately, EGFRvIII is not detectable by the current Affymetrix oligonucleotide arrays used in this study. In the future, it will be important to use other approaches to determine the effect of EGFRvIII expression on the transcriptional program of EGFR+ GBMs. Fourth, EGFR overexpression is usually the result of an amplification, but increased EGFR protein can be detected in the absence of such a genomic change.<sup>30</sup> Taking an analogy from a related Erb-family receptor *her2/neu*, it appears that overexpression vs. amplification may have a profound impact on the tumor's biologic behavior and response to treatment. In the future, it will be important to determine whether the mechanism of EGFR overexpression has any impacts upon the transcriptional program of GBMs. Finally, our study cannot determine whether the unique transcriptional profile of EGFR+ GBMs is due to EGFR-mediated signaling. In the future, it will be important to determine whether EGFR-activated signaling is required for upregulation of these genes.

In summary, we have used screened patient biopsy samples and used a gene expression profiling approach to identify novel subsets of primary GBMs. We have shown that there are EGFR+ primary GBMs, GBMs with coordinate upregulation of genes on chromosome 12q13-15, and primary GBMs lacking either alteration that have distinct transcriptional profiles. This strongly suggests that the three are biologically different tumor types, despite their identical microscopic appearance, and provides an important first step in developing a molecular taxonomy of GBMs.

## REFERENCES

1. Kleihues P, Burger PC, Collins VP, Newcomb EW, Ohgaki H, Cavane WK. Glioblastoma. In: P. Kleihues P, Cavane WK, eds. *Tumors of the Nervous System*. Lyon: IARC Press, 2000; 29-39.
2. Kleihues P, Ohgaki H. Primary and secondary glioblastomas: from concept to clinical diagnosis. *Neurooncology* 1999; 1: 44-51.
3. Watanabe K, Tachibana O, Sata K, Yonekawa Y, Kleihues P, Ohgaki H. Overexpression of the EGF receptor and p53 mutations are mutually exclusive in the evolution of primary and secondary glioblastomas. *Brain Pathol* 1996; 6: 217-23.
4. Griffin J. The biology of signal transduction inhibition: basic science to novel therapies. *Semin Oncol* 2001; 28: 3-8.
5. Kilic T, Alberta JA, Zdunek PR, *et al*. Intracranial inhibition of platelet-derived growth factor-mediated glioblastoma cell growth by an orally active kinase inhibitor of the 2-phenylaminopyrimidine class. *Cancer Res* 2000; 60: 5143-50.
6. Druker BJ, Sawyers CL, Kantarjian H, *et al*. Activity of a specific inhibitor of the BCR-ABL tyrosine kinase in the blast crisis of chronic myeloid leukemia and acute lymphoblastic leukemia with the Philadelphia chromosome. *N Engl J Med* 2001; 344: 1038-42.
7. Bachoo RM, Maher EA, Ligon KL, *et al*. Epidermal growth factor receptor and *Ink4a/Arf*. Convergent mechanisms governing terminal differentiation and transformation along the neural stem cell to astrocyte axis. *Cancer Cell* 2002; 1, 269-77.
8. Golub TR, Slonim DK, Tamayo P, *et al*. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999; 286: 531-7.
9. Pomeroy SL, Tamayo P, Gaasenbeek M, *et al*. Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature* 2002; 415: 436-42.
10. Hedenfalk I, Duggan D, Chen Y, *et al*. Gene-expression profiles in hereditary breast cancer. *N Engl J Med* 2001; 344: 539-48.
11. Alizadeh AA, Eisen MB, Davis RE, *et al*. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000; 403: 503-11.
12. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc Natl Acad Sci USA* 2001; 98: 31-6.
13. Kaufman L, Rousseeuw PJ. *Finding groups in data: an introduction to cluster analysis*. New York: Wiley, 1990; 24-89.
14. Hui AB, Lo KW, Yin XL, Poon WS, Ng HK. Detection of multiple gene amplifications in glioblastoma multiforme using array-based comparative genomic hybridization. *Lab Invest* 2001; 81: 717-23.
15. Reifenberger G, Ichimura K, Reifenberger J, Elkahlon AG, Meltzer PS, Collins VP. Refined mapping of 12q13-q15 amplicons in human malignant gliomas suggests CDK4/SAS and MDM2 as independent amplification targets. *Cancer Res* 1996; 56: 5141-5.
16. Brat DJ, Van Meir EG. Glomeruloid microvascular proliferation orchestrated by VPF/VEGF: a new world of angiogenesis research. *Am J Pathol* 2001; 158: 789-96.
17. Hamada H, Okochi E, Oh-hara T, Tsuruo T. Purification of the Mr 22,000 calcium-binding protein (sorcini) associated with multidrug resistance and its detection with monoclonal antibodies. *Cancer Res* 1988; 48: 3173-8.
18. Leegwater PA, Boor PK, Yuan BQ, *et al*. Identification of novel mutations in MLC1 responsible for megalencephalic leukoencephalopathy with subcortical cysts. *Hum Genet* 2002; 110: 279-83.
19. Ye K, Hurt KJ, Wu FY, *et al*. Pike. A nuclear gtpase that enhances

- PI3kinase activity and is regulated by protein 4.1N. *Cell* 2000; 103: 919-30.
20. Herold-Mende C, Mueller MM, Bonsanto MM, Schmitt HP, Kunze S, Steiner HH. Clinical impact and functional aspects of tenascin-C expression during glioma progression. *Int J Cancer* 2002; 98: 362-9.
  21. Ohnishi T, Hiraga S, Izumoto S, *et al.* Role of fibronectin-stimulated tumor cell migration in glioma invasion in vivo: clinical significance of fibronectin and fibronectin receptor expressed in human glioma tissues. *Clin Exp Metastasis* 1998; 16: 729-41.
  22. Nam SW, Clair T, Campo CK, Lee HY, Liotta LA, Stracke ML. Autotaxin (ATX), a potent tumor motogen, augments invasive and metastatic potential of ras-transformed cells. *Oncogene* 2000; 19: 241-7.
  23. Choe G, Park JK, Jouben-Steele L, *et al.* Active MMP-9 expression is associated with primary glioblastoma subtype. *Clin Cancer Res* 2002; 8: 2894-901.
  24. Maity A, Pore N, Lee J, Solomon D, O'Rourke DM. Epidermal growth factor receptor transcriptionally up-regulates vascular endothelial growth factor expression in human glioblastoma cells via a pathway involving phosphatidylinositol 3'-kinase and distinct from that induced by hypoxia. *Cancer Res* 2000; 60: 5879-86.
  25. Meng K, Rodriguez-Pena A, Dimitrov T, *et al.* Pleiotrophin signals increased tyrosine phosphorylation of beta beta-catenin through inactivation of the intrinsic catalytic activity of the receptor-type protein tyrosine phosphatase beta/zeta. *Proc Natl Acad Sci USA* 2000; 97: 2603-8.
  26. Landry CF, Verity MA, Cherman L, *et al.* Expression of oligodendrocytic mRNAs in glial tumors: changes associated with tumor grade and extent of neoplastic infiltration. *Cancer Res* 1997; 57: 4098-104.
  27. Holland EC. Progenitor cells and glioma formation. *Curr Opin Neurol* 2001; 14: 683-8.
  28. Dai C, Celestino JC, Okada Y, Louis DN, Fuller GN, Holland EC. PDGF autocrine stimulation dedifferentiates cultured astrocytes and induces oligodendrogliomas and oligoastrocytomas from neural progenitors and astrocytes in vivo. *Genes Dev* 2001; 15: 1913-25.
  29. Batra SK, Castelino-Prabhu S, Wikstrand CJ, *et al.* Epidermal growth factor ligand-independent, unregulated, cell-transforming potential of a naturally occurring human mutant EGFRvIII gene. *Cell Growth Differ* 1995; 6: 1251-9.
  30. Chin LS, Raynor MC, Wei X, Chen HQ, Li L. Hrs interacts with sorting nexin 1 and regulates degradation of epidermal growth factor receptor. *J Biol Chem* 2001; 276: 7069-78.