

Identification of molecular subtypes of glioblastoma by gene expression profiling

Paul S Mischel^{*1,6,7}, Ruty Shai^{2,7}, Tao Shi^{2,5}, Steve Horvath^{2,5}, Kan V Lu¹, Gheeyoung Choe¹, David Seligson¹, Thomas J Kremen³, Aarno Palotie¹, Linda M Liao^{3,6}, Timothy F Cloughesy^{4,6} and Stanley F Nelson^{2,6}

¹Department of Pathology, UCLA School of Medicine, Los Angeles, CA 90095, USA; ²Department of Human Genetics, UCLA School of Medicine, Los Angeles, CA 90095, USA; ³Department of Neurosurgery, UCLA School of Medicine, Los Angeles, CA 90095, USA; ⁴Department of Neurology, UCLA School of Medicine, Los Angeles, CA 90095, USA; ⁵Department of Biostatistics, UCLA School of Medicine, Los Angeles, CA 90095, USA; ⁶The Henry E Singleton Brain Tumor Program, UCLA School of Medicine, Los Angeles, CA 90095, USA

Epidermal growth factor receptor (EGFR) overexpression occurs in nearly 50% of cases of glioblastoma (GBM), but its clinical and biological implications are not well understood. We have used Affymetrix high-density oligonucleotide arrays to demonstrate that EGFR-overexpressing GBMs (EGFR+) have a distinct global gene transcriptional profile. We show that the expression of 90 genes can distinguish EGFR+ from EGFR nonexpressing (EGFR-) GBMs, including a number of genes known to act as growth/survival factors for GBMs. We have also uncovered two additional novel molecular subtypes of GBMs, one of which is characterized by coordinate upregulation of contiguous genes on chromosome 12q13–15 and expression of both astrocytic and oligodendroglial genes. These results define distinct molecular subtypes of GBMs that may be important in disease stratification, and in the discovery and assessment of GBM treatment strategies.

Oncogene (2003) 22, 2361–2373. doi:10.1038/sj.onc.1206344

Keywords: glioblastoma; EGFR; microarray; 12q13–15

Introduction

A new approach to cancer therapy focuses on inhibiting signal transduction pathways that are constitutively activated in the tumor cells (Griffin, 2001; Sawyers, 2002). Pharmacological agents that specifically target these signaling pathways demonstrate considerable promise (Kilic *et al.*, 2000; Druker *et al.*, 2001; Griffin, 2001; Neshat *et al.*, 2001; Sawyers, 2002). However, tools to identify which patients are most likely to respond to these targeted molecular therapies have lagged behind. Therefore, one of the critical challenges in cancer biology is to develop molecular classifications of tumors that reflect these underlying signal transduction

abnormalities and that can be used as the basis for patient stratification.

Glioblastoma (GBM), the most common malignant brain tumor of adults (and one of the most lethal of all cancers), may be an ideal candidate tumor for this kind of an approach. GBMs have a number of clearly defined signal transduction abnormalities, whose cooperative disruption appears to be a critical factor in regulating their biological and clinical behavior (Bachoo *et al.*, 2002). Further, the considerable molecular and clinical heterogeneity among GBMs suggest the presence of multiple molecular subtypes. One clear distinction among GBMs is based on clinical presentation and epidermal growth factor receptor (EGFR) status. Primary GBMs arise as *de novo* grade IV tumors and frequently contain EGFR overexpression/amplification (Watanabe *et al.*, 1996; Kleihues and Ohgaki, 1999; Kleihues *et al.*, 2000). In contrast, secondary GBMs progress from a lower grade and rarely overexpress EGFR glioma (Kleihues and Ohgaki, 1999; Kleihues *et al.*, 2000). Since EGFR signaling can be inhibited both at the level of the receptors and via downstream signaling intermediates, EGFR-overexpressing GBMs (EGFR+) provide an attractive target for therapy. However, the clinical and therapeutic implications of EGFR overexpression are far from clear. Further, the transcriptional consequences of EGFR overexpression on GBM cells have not been elucidated.

Large-scale gene expression profiling provides a powerful approach to identify transcriptional networks associated with upregulated signaling pathways, and to discover previously unrecognized tumor subtypes with distinct molecular and/or clinical phenotypes or responses to therapy (Golub *et al.*, 1999; Alizadeh *et al.*, 2000; Perou *et al.*, 2000; MacDonald *et al.*, 2001; Sorlie *et al.*, 2001; Chen *et al.*, 2002; Pomeroy *et al.*, 2002; Shipp *et al.*, 2002), including GBM (Sallinen *et al.*, 2000; Ljubimova *et al.*, 2001; Rickman *et al.*, 2001; Lal *et al.*, 2002). Gene expression profiling can also be used to detect distinct molecular signatures associated with specific mutations (Hedenfalk *et al.*, 2001). We hypothesized that overexpression of EGFR has a major effect on

*Correspondence: PS Mischel; E-mail: pmischel@mednet.ucla.edu

⁷These authors contributed equally

Received 19 August 2002; revised 12 December 2002; accepted 20 December 2002

the transcriptional profile of 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 (EGFR⁻) GBMs, characterized by their own transcriptional profiles, which may provide a different set of potential therapeutic targets. We demonstrate a global gene expression signature for EGFR-overexpressing GBMs. We show that it 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⁻ GBMs, based on their gene expression signatures. The transcriptional profiles of these EGFR⁻ GBMs suggest upregulation of different signaling pathways, which may require a distinct set of targeted molecular therapies.

Results

To exclude any potential effect of treatment on gene expression, we studied GBMs from untreated patients with *de novo* (primary GBMs) from which high-quality RNA could be obtained. A total of 13 cases were available for 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 (Li and Wong, 2001). First, we determined the EGFR status of each tumor on a transcript level (by microarray assay) and on a protein level (by immunohistochemistry). Eight of the 13 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 (Figure 1a, b). In contrast, four samples had very high levels of EGFR transcripts and were strongly immunoreactive for EGFR (Figure 1a, b). The EGFR transcript level was greater than 27-fold increased in the EGFR⁺ cases ($P < 0.0001$) (Table 1). Therefore, nine of the 13 test tumor samples were EGFR⁻ and four were strongly EGFR⁺.

To identify the genes that were most differentially expressed between EGFR⁺ and EGFR⁻ tumors, we used a thresholding approach, selecting for genes with a fold change ≥ 1.5 and an absolute difference of 50 in their model-based expression levels. These criteria resulted in 90 genes (101 probe sets) that were differentially expressed between the EGFR⁺ and EGFR⁻ GBMs (Figure 2a). To assess the validity of this approach, we performed a cross-validation analysis. Our gene selection procedure, coupled with weighted-gene voting prediction method (Golub *et al.*, 1999), led to a misclassification rate of zero, that is, all the 13 samples were correctly classified regarding EGFR status in a leave-one-out cross-validation analysis.

We then performed hierarchical clustering, a standard unsupervised learning method (Kaufman and Rousseeuw, 1990; Hastie *et al.*, 2001). The EGFR⁺ GBMs were clearly separable from the EGFR⁻ tumors; one of the clusters was significantly enriched for EGFR⁺ GBMs (Fisher's exact test; $P < 0.007$) (Figure 2a). Further, the dendrogram suggested the presence of two subtypes of EGFR⁻ GBMs (Figure 2a). 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

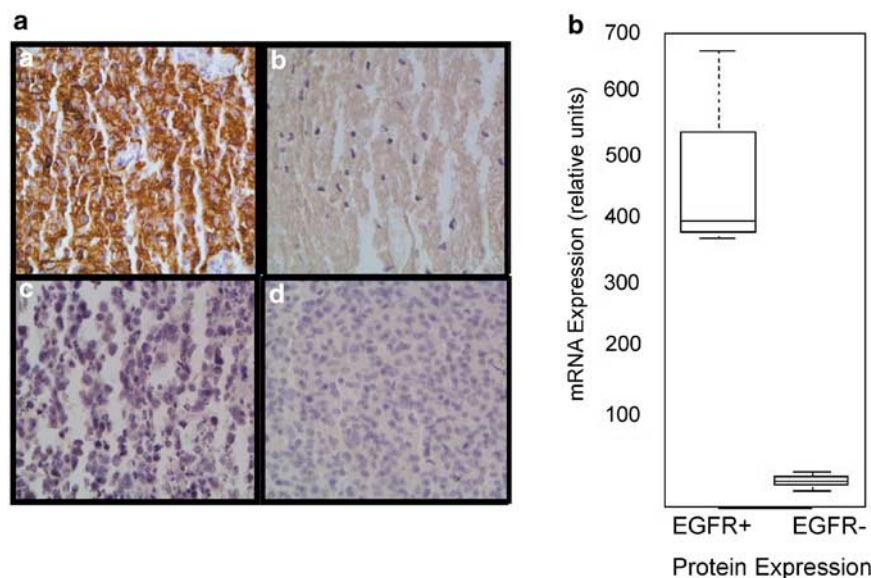


Figure 1 EGFR expression. (a) Immunohistochemical staining for EGFR. EGFR immunostaining on case 367 (a), the normal tissue core from that patient (b), an EGFR⁻ tumor 68 (c) and the negative control from case 367 (d). (b). Correlation between EGFR mRNA and protein. Relative EGFR mRNA expression, (y-axis) is highly correlated with EGFR protein expression (x-axis). EGFR mRNA and protein are either strongly expressed (EGFR⁺) or are barely detectable (EGFR⁻)

Table 1 Genes upregulated in EGFR-overexpressing GBMs

Gene	Ratio EGFR+ :EGFR-	P-value	Function
EGFR	27	0.0001	Growth factor receptor for GBM cells – signal transduction
Pleiotrophin (PTN) ^a	1.9	0.002	Growth factor for GBM cells – signal transduction
PTPRZ1	1.8	0.03	PTN receptor – signal transduction
VEGF	3.2	0.05	Angiogenesis in GBMs – signal transduction
Endothelin B receptor ^a	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	Antiapoptotic 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, α 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
CD99/MIC2 ^a	1.5	0.06	Cell-surface marker
Cyclin D2	2.0	0.03	Proliferation

Genes upregulated in 12q13-15-overexpressing GBMs

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, α 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
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

^aCoefficient of variation is between 0.4 and 0.5

osteosarcoma (OS-9) (12q13.2) and conserved gene amplified in osteosarcoma (OS-4) (12q13–15) (Fisher's exact test; $P < 0.0014$) (Figure 2a). These genes are overexpressed in approximately 10–20% of high-grade gliomas, usually in association with a DNA amplification event (Reifenberger *et al.*, 1994, 1995, 1996; Fischer *et al.*, 1996; Galanis *et al.*, 1998; Hui *et al.*, 2001). 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, 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) (Figure 2b, Table 1). We next determined whether this set of seven 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 seven genes (data not shown). 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.

As an independent test of the association of this gene expression profile with EGFR status, we analysed the gene expression pattern of an additional 16 GBMs. This included biopsies from seven untreated primary GBMs and four secondary GBMs, as well as five samples obtained from GBM autopsy patients. All of the secondary GBM and autopsy samples were derived from patients who had been treated with radiation and chemotherapy. Using hierarchical clustering of these additional 16 samples based on the 90 gene sets, we found that 4/5 tumors with EGFR transcripts clustered with the other EGFR+ tumors, and 9/11 GBMs lacking EGFR transcripts clustered with the other EGFR– tumors (Fisher's exact test; $P = 0.03$) (Figure 2c). Interestingly, the one EGFR+ tumor that did not cluster with the other EGFR+ tumors also contained upregulation of 12q13–15 (sample 459) and clustered with these samples. Therefore, for all 29 tumors tested, one cluster

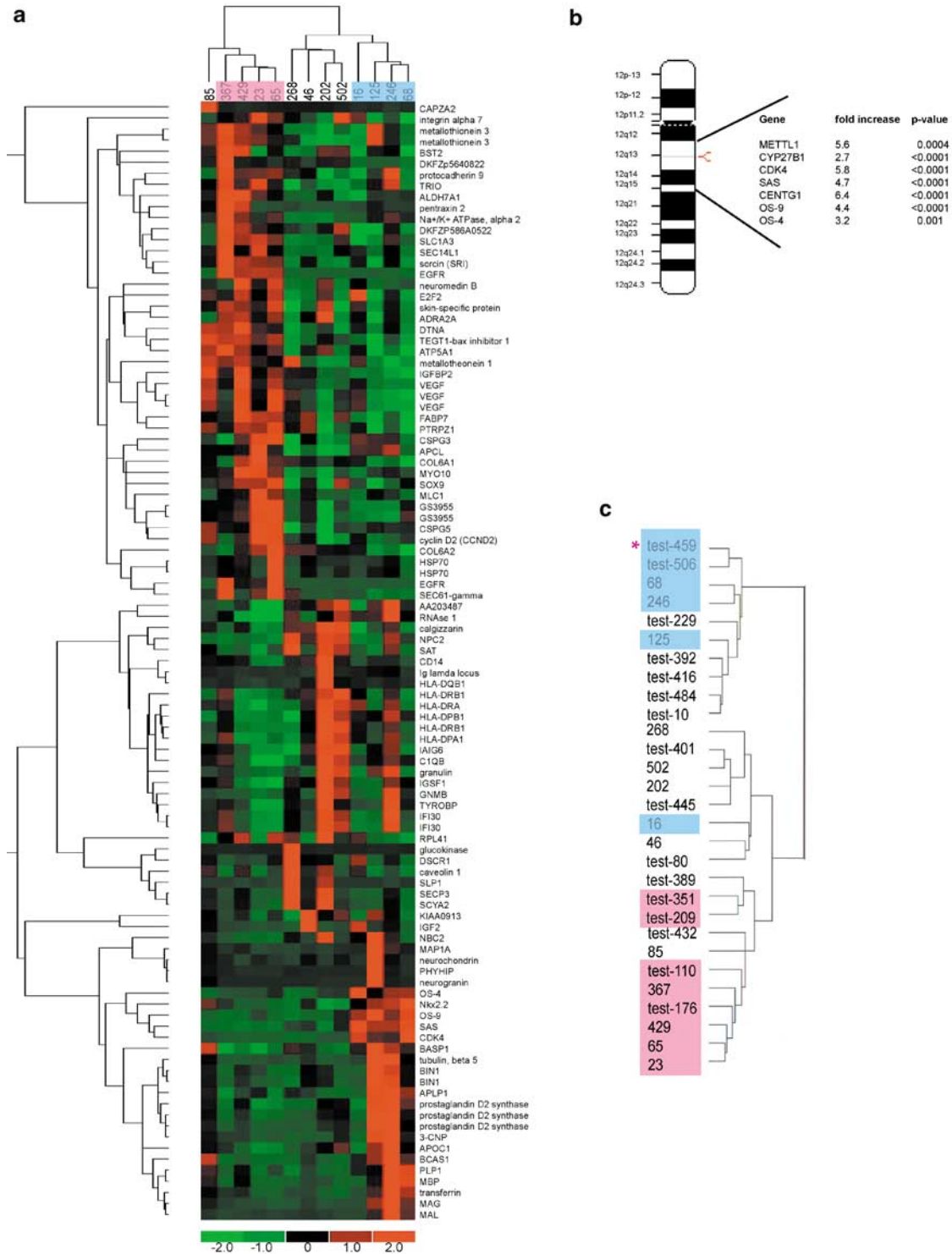


Figure 2 Genes differentially expressed between EGFR+ and EGFR- GBMs. (a) A total of 101 probe sets (90 genes) were differentially expressed in EGFR+ EGFR- GBMs, using our filtering criteria, and the tumors were hierarchically clustered based on expression of these 90 genes. Red bar denotes tumors with EGFR overexpression, blue bar denotes GBMs with overexpression of genes on chromosome 12q13-15, and white bar denotes GBMs lacking either alteration. (b) Genes with coordinate upregulation mapping to chromosome 12q13-15. (c) Enlargement of the hierarchical clustering dendrogram. Hierarchical clustering based on the expression of these 90 genes was performed on the 13 original tumor samples and an additional 16 independent test samples. EGFR+ GBMs (red) clustered together ($P=0.0003$), 12q13-15+ GBMs clustered ($P=0.01$) and EGFR-/12q13-15- GBMs clustered together ($P=0.02$). Sample 459 (blue bar and red asterisk) overexpressed both 12q13-15 and EGFR

was enriched for EGFR+ GBMs (Fisher's exact test; $P=0.0003$), one cluster was enriched for 12q13-15+ GBMs (Fisher's exact test; $P=0.01$) and one cluster was enriched for GBMs lacking either alteration (Fisher's exact test; $P=0.02$). These results further demonstrate the association of these 90 genes with EGFR+ expression, and indicate that simultaneous upregulation of EGFR and genes on 12q13-15 can occur.

To validate the presence of three distinct groups of GBMs, and to determine whether these molecular subsets had global transcriptional correlates, we analysed 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 (4255 genes), regardless of EGFR status. We then performed hierarchical clustering of the 13 GBMs based on expression of these 4255 most variable genes. The dendrogram indicated the presence of three global transcriptional groups, one of which was significantly enriched for EGFR+ GBMs (Fisher's exact test; $P=0.02$) (Figure 3a). The dendrogram also indicated the presence of two EGFR-

groups, one of which was significantly enriched for 12q13-15 upregulated GBMs (Fisher's exact test; $P=0.014$) and the other which was comprised of non-EGFR expressing, non-12q13-15 upregulated GBMs (Fisher's exact test; $P<0.007$). This hierarchical clustering pattern was observed over a range of thresholds (coefficient of variation between 0.4 and 0.6; 3000-7000 genes). An alternative view of tumor groupings was obtained by performing classical multidimensional scaling of the 4255 genes onto a three-dimensional plot (a form of principal component analysis). This is an unsupervised method of data reduction in which high-dimensional gene expression data are reduced to three viewable dimensions, representing linear combinations of genes that provide the most variation in the data set (Venables and Ripley, 1999). The EGFR+ tumors (red) were separable from the EGFR- tumors (Figure 3b), and the 12q13-15 upregulated EGFR- GBMs (blue) were distinct from the EGFR- non-12q13-15 upregulated GBMs (gray). These results lend support to the hypothesis that EGFR+ GBMs are a distinct subset, and further suggest that these

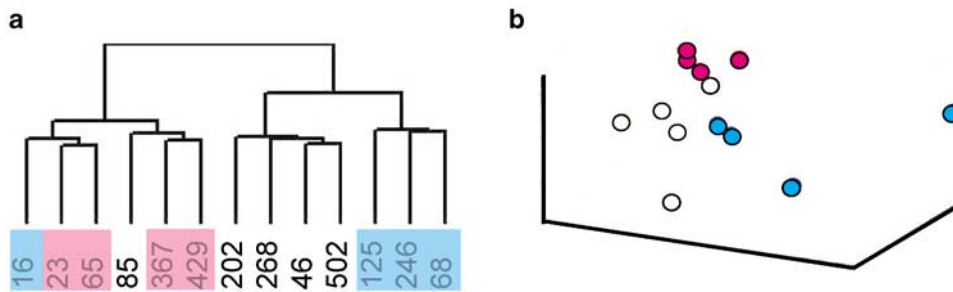


Figure 3 EGFR+ GBMs, 12q13-15+ GBMs and EGFR-/12q13-15- GBMs have globally distinct gene expression profiles. (a) Hierarchical clustering based on the 4255 most variable genes, independent of EGFR status. One cluster is enriched for EGFR+ tumors (red) (Fisher's exact test, $P=0.02$). Another cluster is enriched for GBMs with coordinate upregulation of 12q13-15 genes (blue) (Fisher's exact test, $P=0.014$), and the third cluster is enriched for GBMs lacking either alteration (Fisher's exact test, $P<0.007$). (b) Multidimensional scaling plot based on expression of the most variable 4255 genes, demonstrating that EGFR+ (red), 12q13-15+ (blue) and EGFR-/12q13-15- tumors (gray) have distinct global gene expression profiles

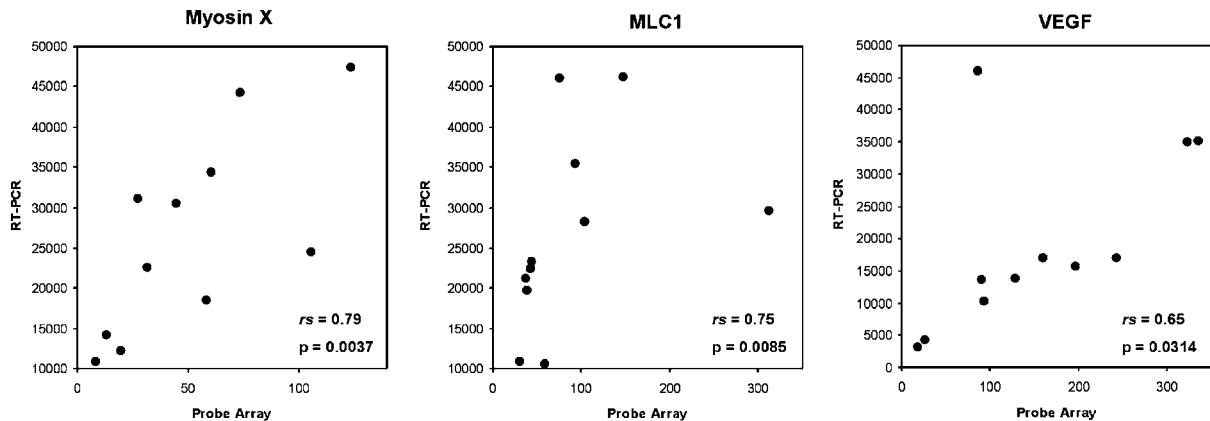


Figure 4 Comparison of RNA assays by microarray hybridization and RT-PCR. Representative examples of correlation between mRNA level as detected by microarray assay (x -axis) and RT-PCR (y -axis) for three genes: Myosin X (MYO10), MLC1 and VEGF. Each data point represents an individual tumor. Values on x - and y -axis are expressed in relative units. Spearman rank correlations (r 's) and P -values are listed

additional molecular subsets of EGFR– GBMs are robust.

The gene expression signature of EGFR-overexpressing GBMs was notable for upregulation of growth factors, receptors and signal transduction molecules (over one-third of the probe sets), including vascular endothelial growth factor (VEGF), which plays a critical role in GBM angiogenesis and progression (Brat and van Meir, 2001). In addition, pleiotrophin (PTN) and its receptor PTRPZ1, and endothelin B receptor (ET(B)) were all upregulated in EGFR+ GBMs (Table 1), as was the antiapoptotic protein Bax inhibitor 1 (TEGT) (Table 1). The EGFR+ GBM gene expression signature was also notable for increased expression of plasma membrane-bound transporters and channels, including the multidrug chemoresistance gene SRI (sorcin) and MLC1, a recently cloned cell-surface protein whose mutation is associated with white matter brain defects (Hamada *et al.*, 1988; Leegwater *et al.*, 2002). In addition to their association with EGFR protein, we also detected a significant correlation between the transcript level of each of these genes and EGFR mRNA level, as expected ($r=0.65$). These data suggest upregulation of multiple growth factor-mediated signal transduction pathways in EGFR-overexpressing GBMs that can promote GBM cell proliferation, survival and angiogenesis. These results also suggest potential candidates for targeted molecular therapy, in this subset of GBMs. To independently validate this set of potential candidates, we performed RT–PCR analysis of nine of these differentially expressed genes. The transcript level of these target genes, as detected by RT–PCR, was highly correlated with the transcript level as detected by the microarray assay (mean r 's = 0.64, $P=0.04$) (Figure 4).

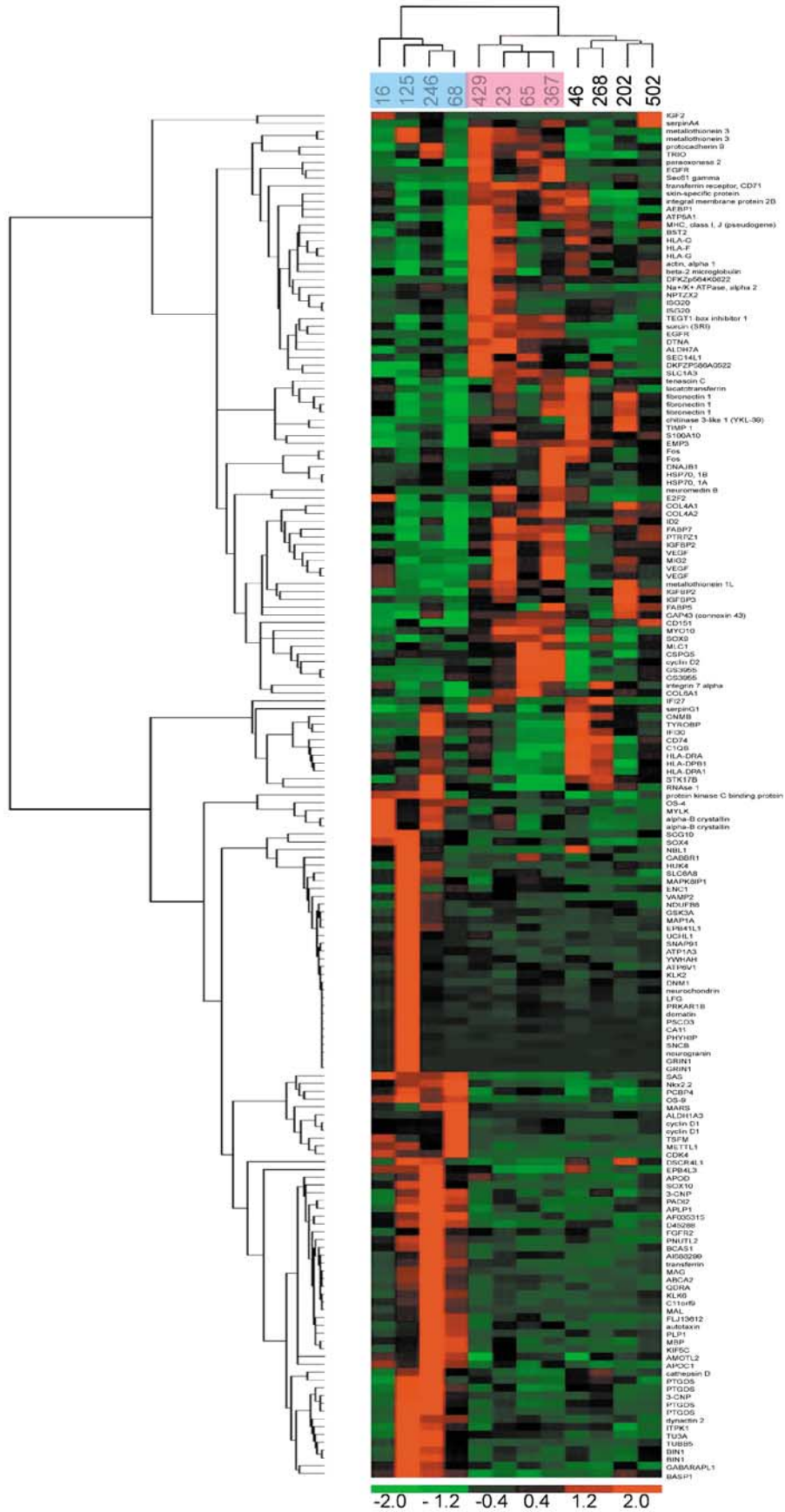
The gene expression signature of GBMs with overexpression of genes on 12q13–15 was notable for increased expression of CDK4 (located on 12q13–15), its binding partner cyclin D1 and CENTG1 (PIKE), a newly identified signal transduction molecule that enhances PI3K activity and increases cyclin D1 activity (Ye *et al.*, 2000). Taken together, this suggests coordinate upregulation of the cyclin D1 pathway in tumors with the 12q13–15 expression signature. In contrast to the 12q13–15+ GBMs, the EGFR-overexpressing GBMs expressed elevated levels of cyclin D2 (two fold increase, $P=0.03$; Table 1). The 12q13–15-upregulated GBMs were also significant for high levels of transcription of oligodendroglial genes (MBP, MAG, PLP1, Nkx2.2, Mal and Sox10) (Hajihosseini *et al.*, 1996; Landry *et al.*, 1997; Zhou *et al.*, 2001; Fu *et al.*, 2002). (Table 1). Morphologically, these tumors were indistinguishable from the other GBMs; they lacked oligodendroglial morphology and had equivalent levels of the astrocytic marker GFAP

(data not shown). To exclude the possibility that this gene expression signature was arising from entrapped normal brain tissue, we performed hierarchical clustering of the 13 tumor sample sets and 10 normal brain samples. The normal brain samples and the tumors clustered quite distinctly ($P<0.00001$) from each other (data not shown) for expression of the global gene signature (all 12533 probe sets) and the EGFR-specific gene set (90 genes). Thus, it is highly unlikely that entrapped normal brain tissue was contributing to the distinctive gene expression profiles we observed. These results raise a number of possibilities. GBMs with coordinated upregulation of genes on 12q13–15 may arise from a glial precursor cell capable of astrocytic or oligodendroglial differentiation (Holland *et al.*, 2000; Dai *et al.*, 2001; Holland, 2001). Alternatively, GBMs with this alteration may either dedifferentiate, or activate a transcriptional program of oligodendroglial genes (Holland *et al.*, 2000; Dai *et al.*, 2001; Bachoo *et al.*, 2002). 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 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 to identify genes that were differentially expressed between EGFR+ and 12q13–15+ GBMs. This resulted in 175 probe sets (157 genes) that were differentially expressed between EGFR+ and 12q13–15+ GBMs (Figure 5). Hierarchical clustering of these tumors based on these 157 genes continued to support the presence of three primary GBM subtypes (Figure 5). 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 (Ohnishi *et al.*, 1998; Herold-Mende *et al.*, 2002), while the 12q13–15+ group had very high transcript levels of autotaxin, a secreted motility factor that promotes tumor cell invasion and metastasis (Stracke *et al.*, 1992; Nam *et al.*, 2000). These results suggest that these GBM subtypes may differ in their invasion patterns, and suggest additional potential biological, and perhaps clinical differences.

To further assess the clinical utility of this approach, we used immunohistochemistry to validate the differential expression of a couple of the genes. For this purpose, we generated a tissue microarray of 48 GBM specimens (including nine from patients used in this study). We chose to analyse cyclin D1, whose

Figure 5 Genes differentially expressed between EGFR+ and 12q13–15+ GBMs. In total, 175 probe sets (157 genes) were differentially expressed between EGFR+ and 12q13–15-overexpressing GBMs, using our filtering criteria, and the tumors were hierarchically clustered based on expression of these 157 genes. Red bar denotes tumors with EGFR overexpression; blue bar denotes GBMs with overexpression of genes on chromosome 12q13–15



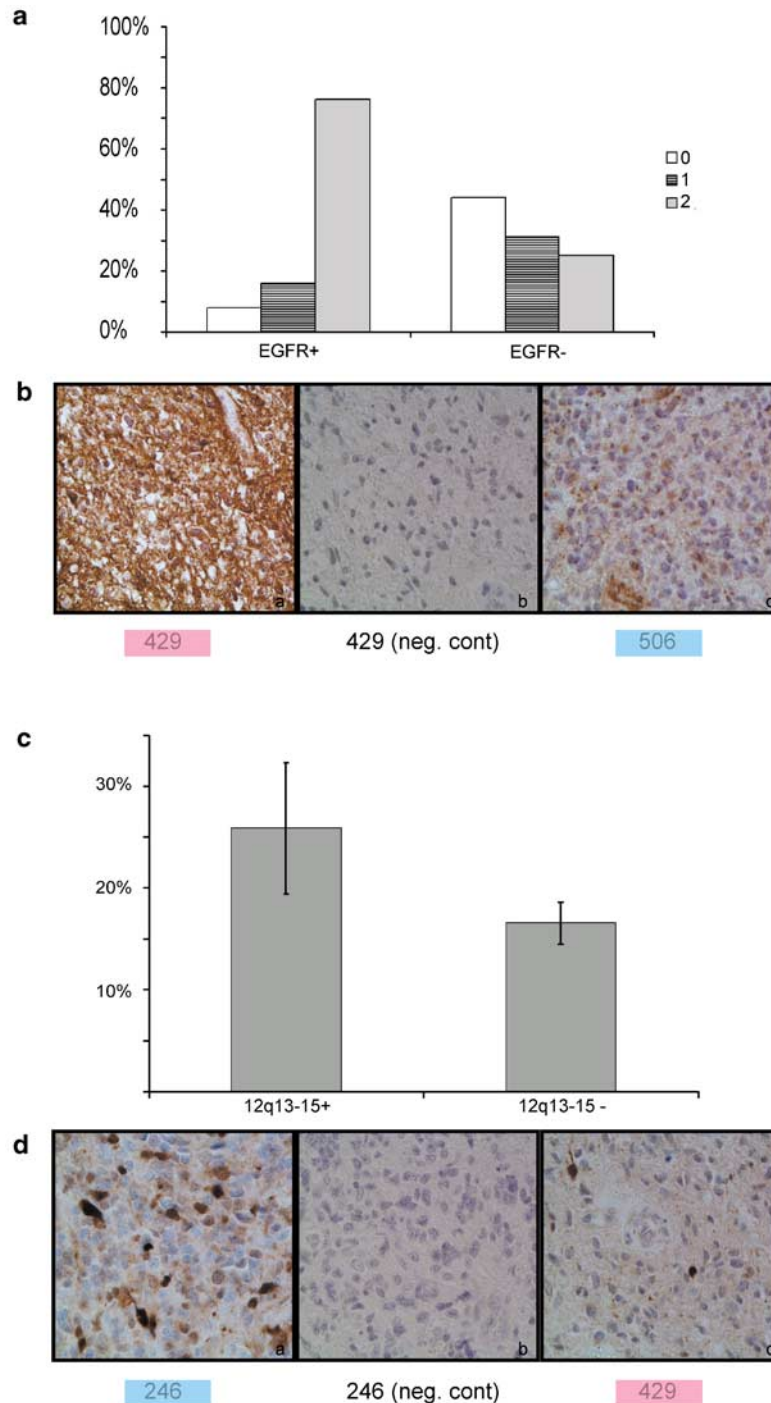


Figure 6 Immunohistochemical staining for CD99/MIC2 and cyclin D1. (a) CD99/MIC2 staining intensity is significantly correlated with EGFR expression (χ^2 test, $P=0.004$) in a tissue array of 47 primary GBMs. CD99/MIC2 staining intensity was scored on a 0–2 scale of increasing intensity. (b) Representative CD99/MIC2 immunohistochemical staining in a EGFR+ case 429 (a), its negative control (b) and an EGFR–/12q13–15+ case 506 (c). (c) Percentage of cyclin-D1-positive cells in six GBMs with 12q13–15 overexpression (as assessed by the microarray assay) and 10 non-12q13–15-overexpressing GBMs (also assessed by microarray assay). Percentage of cyclin-D1-positive cells was assessed in three representative fields from each tumor. Cyclin D1 staining is associated with 12q13–15-overexpressing GBMs (t -test, $P=0.07$). (d) Representative cyclin D1 staining in a 12q13–15+ case 246 (a), its negative control (b) and in an EGFR+/non-12q13–15-overexpressing GBM (c)

transcript level was correlated with 12q13–15-overexpressing GBMs (4.9-fold increase, $P=0.04$), and CD99/MIC2, whose expression was correlated with EGFR-overexpressing GBMs (1.5-fold increase, $P=0.06$), since

these antibodies were available in our laboratory and have been optimized for use with paraffin-embedded biopsy tissues. CD99/MIC2 protein expression was highly correlated with EGFR overexpression ($P=0.004$), while

cyclin D1 protein expression was correlated with 12q13–15-overexpressing GBMs ($P=0.07$) (Figure 6). Thus, our findings may be extended using standard immunohistochemical analysis.

Discussion

The new challenge in cancer biology is to move from purely morphological classification of tumors to one that is based on molecular criteria. In the light of the development of new pharmacologic pathway inhibitors for cancer therapy, this goal is now even more important for the purposes of treatment discovery and selection of susceptible patient subsets. GBM may be an ideal tumor for this approach for the following reasons: (1) they have a number of clear-cut signal transduction abnormalities that influence their biological behavior and that are potentially targetable; (2) inhibition of these pathways with small molecules has yielded promising results in GBM preclinical models (Kilic *et al.*, 2000; Neshat *et al.*, 2001) and (3) none of the current therapies are highly effective (Preston-Martin, 1999). 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 molecular subtypes of GBMs, and suggests possible new therapeutic targets for each of these molecular subtypes.

Our observation that EGFR expression confers a distinct transcriptional phenotype to primary GBMs is a novel finding. Previous work from our own laboratory (Choe *et al.*, 2002) and the laboratories of other investigators (O'Rourke *et al.*, 1998; Barker *et al.*, 2001) suggest that EGFR signaling alters the biological behavior of GBMs. Our finding that EGFR expression globally impacts the transcriptional program (Figure 2), 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 (Figure 1b), 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 (Maity *et al.*, 2000). VEGF plays a major role in promoting angiogenesis and tumor growth of GBMs *in vivo* (Berkman *et al.*, 1993; Cheng *et al.*, 1996; Yuan *et al.*, 1996; Chan *et al.*, 1998;

Brat and van Meir, 2001; Chaudhry *et al.*, 2001), 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 (Egidy *et al.*, 2000). Since activation of this receptor promotes a prosurvival/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 antiapoptotic protein (Xu and Reed, 1998), 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, were upregulated in EGFR+ GBMs. PTN is a potent growth factor for GBM cells in culture (Powers *et al.*, 2002). PTRPZ1 expression, one of the receptors for PTN (Meng *et al.*, 2000), was also elevated in the EGFR+ samples, suggesting coordinate upregulation of this pathway.

By analysing 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 (Reifenberger *et al.*, 1994, 1995, 1996; Fischer *et al.*, 1996; Galanis *et al.*, 1998; Hui *et al.*, 2001). 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 analysed 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. This 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) (Reifenberger *et al.*, 1996). Similarly, we found no evidence of MDM2 upregulation in these tumors (data not shown), confirming these previous findings.

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 appeared to have an impact on cell-fate specification. 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 (Hajhosseini *et al.*, 1996; Landry *et al.*, 1997; Zhou *et al.*, 2001; Fu *et al.*, 2002). Morphologically, these tumors were GBMs, not oligodendrogliomas or mixed gliomas, and they contained equal level of GFAP transcripts relative to the non-12q13–15+ tumors. Further, previous studies have suggested that two of these markers, PLP1 and MBP, can be expressed by some astrocytomas (Landry *et al.*, 1997). This raises two important

possibilities: development from a multipotent precursor cell or dedifferentiation of an astrocytic cell. Experimental evidence suggests that GBMs may arise from precursor cells (Holland *et al.*, 2000; Dai *et al.*, 2001; Holland, 2001). Alternatively, terminally differentiated astrocytic tumor cells could be dedifferentiated into neural stem cells by activation/disruption of specific signal transduction pathways (Holland *et al.*, 2000; Dai *et al.*, 2001; Bachoo *et al.*, 2002).

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 a global genomic strategy 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 a similar response 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, cyclin D1 and CDK4 in 12q13–15), it is likely that this approach may be fruitful. Verification awaits future preclinical studies and clinical trials. In the future, it will be important to determine whether the EGFR+ gene expression signature seen in GBMs is also characteristic of other cancer types bearing EGFR overexpression/amplification.

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 (Kleihues *et al.*, 2000; Smith *et al.*, 2001), are also commonly mutated in GBMs. This may have 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– sample 85 has a global transcriptional pattern similar to the EGFR+ tumors (Figure 2a, b), may suggest that a different upstream lesion is activating similar signaling pathways. Third, in nearly 50% of GBMs with high-level EGFR expression, there is coexpression of mutant EGFRs, most commonly the EGFRvIII variant (Kleihues *et al.*, 2000). EGFRvIII results from an in-frame genomic deletion of exons 2–7, producing an EGFR that lacks its ligand-binding domain (Kuan *et al.*, 2000; Nagane *et al.*, 2001). EGFRvIII is constitutively active, oncogenic (Batra *et al.*, 1995; Huang *et al.*, 1997) and may have profound transcriptional consequences (Lal *et al.*, 2002). Unfortunately, EGFRvIII is not detectable by the current Affymetrix oligonucleotide arrays used in

this study. The recent finding that EGFRvIII expression imparts a distinctive transcriptional signature (Lal *et al.*, 2002) further indicates the importance of determining 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 (Kurten *et al.*, 1996; Chin *et al.*, 2001). Similar to the related Erb-family receptor her2/neu, it appears that overexpression vs amplification may have profound impact on the tumor's biologic behavior and response to treatment (Pauletti *et al.*, 2001). In the future, it will be important to determine whether the mechanism of EGFR overexpression impacts upon the transcriptional program of GBMs. Finally, our study cannot determine whether the unique transcriptional profile of EGFR+ GBMs is because of EGFR-mediated signaling.

In summary, we have used screened patient biopsy samples using a gene expression profiling approach to identify novel subsets of primary GBMs. We have identified a molecular signature associated with EGFR overexpression and we have uncovered two additional subsets of GBMs, each with a distinct transcriptional profile. This provides an important first step in developing a molecular taxonomy of GBMs and provides a potential approach by which to begin to stratify GBM patients for targeted molecular therapy.

Materials and methods

Tissue and RNA isolation

All patients participating in this study gave informed consent prior to surgery. At the time of resection, the tumor was 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 min of surgical resection. The portion for RNA extraction was snap frozen in liquid nitrogen and stored at –80°C. Total RNA was extracted from 100 to 150 mg of frozen tissue by using Trizol (Gibco BRL) and one round of cleanup by the Qiagen Rneasy total RNA isolation kit. Total RNA (10 µg) was used to generate double-stranded cDNA. Biotin-labeled antisense cRNA was produced by *in vitro* transcription using the ENZO BioArray HighYield kit. cRNA (15 µg) was fragmented and hybridized to Affymetrix U95Av2 GeneChip. 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).

Preprocessing and statistical analysis

The CEL files for all the microarray hybridizations generated by Affymetrix Microarray Suite Software were imported into the software dChip (Li and Wong, 2001) to compute the model-based expression index for each gene. All arrays were normalized against the array with median overall intensity. To reduce the noise in our data, we eliminated genes with a coefficient of variation smaller than 0.5 across all samples from the analysis, and ended up with 4255 genes. 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 two prediction methods that used the same gene-filtering criteria in their construction. To be specific, we trained (constructed) a *k*-nearest neighbor (Hastie *et al.*, 2001) and a gene voting (Golub *et al.*, 1999) predictor on all but one sample by using the same gene-filtering criterion. The resulting predictors were applied to the left-out observation, 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 hierarchical clustering of the samples or genes using Euclidean distance (Kaufman and Rousseeuw, 1990). Since Euclidean distances between samples were used, the MDS plot produced here is equivalent to plotting the samples with their first three principal components (Venables and Ripley, 1999). To evaluate the significance of differentially expressed genes, two group unpaired *t*-tests were used. Genes with a *p*-value less than 0.05 were considered significant. Analyses were performed using the statistical software S-Plus (www.insightful.com).

RT-PCR confirmation

Total RNA was available from nine of the 13 test samples (cRNA was available for an additional two samples). Total RNA (500 ng) was reverse transcribed using SuperScript First-Strand Synthesis System (Invitrogen, Carlsbad, CA, USA). The following genes were analysed by RT-PCR: VEGF, MYO10, MLC1, GS3955, TEGT, OS-9, OS-4, SAS, CDK-4, Cyclin D1 and MAG. Primers were designed to amplify 3' mRNA regions that did not overlap with probe sequences represented by Affymetrix U95av2 oligonucleotide array. The following forward and reverse primers were used for the representative examples shown in Figure 4: VEGF – 5'-GTC TTG GGT GCA TTG GAG CCT T-3' and 5'-ACA GGG ATT TTC TTG TCT TGC T-3' (414 bp product); MYO10 – 5'-CTG GCT GCC ACA TCC GAG GTT-3' and 5'-CTC TTC TCC AGC CGT TCA CAA-3' (328 bp product); and MLC1 – 5'-CCT GCT CGG GTC CTG AAA TCT-3' and 5'-CAC TTG CTG GGA CAC TCT GCT-3' (196 bp product). Primer sequences for the other genes tested are available upon request. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used to normalize amplification between samples (5'-TTC CAT GGC ACC GTC AAG GCT GAG A-3' and 5'-CAC GTT GGC AGT GGG GAC ACG GAA G-3' – 555 bp fragment). First-strand cDNA (2 μ l) was amplified in a 50 μ l PCR reaction volume containing 1 \times PCR

buffer (20 mM Tris-HCl (pH 8.4), 50 mM KCl), 2 mM MgCl₂, 0.2 mM each dNTP, 0.2 μ M each primer and 2 U of Platinum Taq Polymerase (Invitrogen) under the following conditions: initial denaturation at 94°C for 2 min, followed by 20, 25 or 30 amplification cycles with denaturation at 95°C for 20 s, annealing at 59°C for 20 s and extension at 72°C for 30 s. A 5 μ l volume of each PCR reaction was then loaded onto 1.5% agarose gels, stained with ethidium bromide and the band intensities analysed by densitometry using AlphaEase software version 5.04 (Alpha Innotech, San Leandro, CA, USA). Gene expression was normalized for GAPDH and statistical analysis was performed using Analyse-It software version 1.65 (Leeds, UK).

Tissue microarray generation and immunohistochemistry

Three representative 0.6 mm cores (two tumors, 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. Sections (5 μ m) were cut from the tissue array and immunohistochemistry was performed (Hoos *et al.*, 2002). For CD99/MIC2, staining intensity was scored on a scale of 0–2. For cyclin D1 immunostaining, three regions from each tumor were photographed and positive nuclei were counted using the Alphasizer with Alpha ease software. χ^2 test was used to assess correlation between CD99/MIC22 staining intensity and EGFR status. Student's *t*-test was used to assess the correlation between per cent of cyclin-D1-positive cells and 12q13–15 status.

Acknowledgements

We are grateful to Drs Charles Sawyers, Daniel Geschwind, William McBride, Harry Vinters and Jonathan Braun for their helpful comments on this manuscript. This work was supported by U01 CA88127 from the National Cancer Institute (SFN) and K08NS43147-01 from the National Institute of Neurological Disorders and Stroke, NIH (to PSM). PSM was also supported by an Accelerate Brain Cancer Cure Award, a Henry E Singleton Brain Tumor Fellowship, a generous donation from the Kevin Riley family to UCLA Comprehensive Brain Tumor Program and the Harry Allgauer Foundation through The Doris R Ullmann Fund for Brain Tumor Research Technologies. GC was supported by a postdoctoral fellowship from Korea Science & Engineering Foundation (KOSEF) (to GC). Tao Shi is a predoctoral trainee supported by the UCLA IGERT Bioinformatics Program funded by NSF DGE 9987641. Kan V Lu is supported by USHHS Institutional National Research Service Award #T32 CA09056.

References

- Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson Jr J, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Levy R, Wilson K, Grever MR, Byrd JC, Botstein D, Braun PO and Staudt LM. (2000). *Nature*, **403**, 503–511.
- Bachoo RM, Maher EA, Ligon KL, Sharpless NE, Chan SS, You MJ, Tang Y, DeFrances J, Stover E, Weissleder R, Rowitch DH, Louis DN and DePinho RA (2002). *Cancer Cell*, **1**, 269–277.
- Barker II FG, Simmons ML, Chang SM, Prados MD, Larson DA, Sneed PK, Wara WM, Berger MS, Chen P, Israel MA and Aldape KD. (2001). *Int. J. Radiat. Oncol. Biol. Phys.*, **51**, 410–418.
- Batra SK, Castelino-Prabhu S, Wikstrand CJ, Zhu X, Humphrey PA, Friedman HS and Bigner DD. (1995). *Cell Growth Differ.*, **6**, 1251–1259.

- Berkman RA, Merrill MJ, Reinhold WC, Monacci WT, Saxena A, Clark WC, Robertson JT., Ali IU and Oldfield EH. (1993). *J. Clin. Invest.*, **91**, 153–159.
- Brat DJ and Van Meir EG. (2001). *Am. J. Pathol.*, **158**, 789–796.
- Chan AS, Leung SY, Wong MP, Yuen ST, Cheung N, Fan YW and Chung LP. (1998). *Am. J. Surg. Pathol.*, **22**, 816–826.
- Chaudhry IH, O'Donovan DG, Brenchley PE, Reid H and Roberts IS. (2001). *Histopathology*, **39**, 409–415.
- Chen X, Cheung ST, So S, Fan ST, Barry C, Higgins J, Lai KM, Ji J, Dudoit S, Ng IO, Van De Rijn M, Botstein D and Brown PO. (2002). *Mol. Biol. Cell*, **13**, 1929–1939.
- Cheng SY, Huang HJ, Nagane M, Ji XD, Wang D, Shih CC, Arap W, Huang CM and Cavenee WK. (1996). *Proc. Natl. Acad. Sci. USA*, **93**, 8502–8507.
- Chin LS, Raynor MC, Wei X, Chen HQ and Li L. (2001). *J. Biol. Chem.*, **276**, 7069–7078.
- Choe G, Jouben-Steele L, Park JK, Vinters HV, Liau LM, Cloughesy TF and Mischel PS. (2002). Active MMP-9 expression is associated with primary glioblastoma subtype. *Clin. Cancer Res.*, **8**, 2894–2901.
- Dai C, Celestino JC, Okada Y, Louis DN, Fuller GN and Holland EC. (2001). *Genes Dev.*, **15**, 1913–1925.
- Druker BJ, Sawyers CL, Kantarjian H, Resta DJ, Reese SF, Ford JM, Capdeville R and Talpaz M. (2001). *N. Engl. J. Med.*, **344**, 1038–1042.
- Egidy G, Eberl LP, Valdenaire O, Irmeler M, Majdi R, Diserens AC, Fontana A, Janzer RC, Pinet F and Juillerat-Jeanneret L. (2000). *Lab. Invest.*, **80**, 1681–1689.
- Fischer U, Meltzer P and Meese E. (1996). *Hum. Genet.*, **98**, 625–628.
- Fu H, Qi Y, Tan M, Cai J, Takebayashi H, Nakafuku M, Richardson W and Qiu M. (2002). *Development*, **129**, 681–693.
- Galanis E, Buckner J, Kimmel D, Jenkins R, Alderete B, O'Fallon J, Wang CH, Scheithauer BW and James CD. (1998). *Int. J. Oncol.*, **13**, 717–724.
- Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD and Lander ES. (1999). *Science*, **286**, 531–537.
- Griffin J. (2001). *Semin. Oncol.*, **28**, 3–8.
- Hajihosseini M, Tham TN and Dubois-Dalcq M. (1996). *J. Neurosci.*, **16**, 7981–7994.
- Hamada H, Okochi E, Oh-hara T and Tsuruo T. (1988). *Cancer Res.*, **48**, 3173–3178.
- Hastie T, Tibshirani R and Friedman J. (2001) *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Springer: New York.
- Hedenfalk I, Duggan D, Chen Y, Radmacher M, Bittner M, Simon R, Meltzer P, Gusterson B, Esteller M, Kallioniemi OP, Wilfond B, Borg A and Trent J. (2001). *N. Engl. J. Med.*, **344**, 539–548.
- Herold-Mende C, Mueller MM, Bonsanto MM, Schmitt HP, Kunze S and Steiner HH. (2002). *Int. J. Cancer*, **98**, 362–369.
- Holland EC. (2001). *Curr. Opin. Neurol.*, **14**, 683–688.
- Holland EC, Li Y, Celestino J, Dai C, Schaefer L, Sawaya RA and Fuller GN. (2000). *Am. J. Pathol.*, **157**, 1031–1037.
- Hoos A, Stojadinovic A, Singh B, Dudas ME, Leung DH, Shaha AR, Shah JP, Brennan MF, Cordon-Cardo C and Ghossein R. (2002). *Am. J. Pathol.*, **160**, 175–183.
- Huang HS, Nagane M, Klingbeil CK, Lin H, Nishikawa R, Ji XD, Huang CM, Gill GN, Wiley HS and Cavenee WK. (1997). *J. Biol. Chem.*, **272**, 2927–2935.
- Hui AB, Lo KW, Yin XL, Poon WS and Ng HK. (2001). *Lab. Invest.*, **81**, 717–723.
- Kaufman L and Rousseeuw PJ. (1990) *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley, Inc.: New York.
- Kilic T, Alberta JA, Zdunek PR, Acar M, Iannarelli P, O'Reilly T, Buchdunger E, Black PM and Stiles CD. (2000). *Cancer Res.*, **60**, 5143–5150.
- Kuan CT, Wikstrand CJ and Bigner DD. (2000). *Brain Tumor Pathol.*, **17**, 71–78.
- Kleihues P, Burger PC, Collins VP, Newcomb EW, Ohgaki H and Cavenee WK. (2000). Glioblastoma. *Tumors of the Nervous System*. Kleihues P and Cavenee WK (eds). IARC Press: Lyon, pp. 29–39.
- Kleihues P and Ohgaki H. (1999). *Neuro-oncology*, **1**, 4–51.
- Kurten RC, Cadena DL and Gill GN. (1996). *Science*, **272**, 1008–1010.
- Lal A, Glazer CA, Martinson HM, Friedman HS, Archer GE, Sampson JH and Riggins GJ. (2002). *Cancer Res.*, **62**, 3335–3339.
- Landry CF, Verity MA, Cherman L, Kashima T, Black K, Yates A and Campagnoni AT. (1997). *Cancer Res.*, **57**, 4098–4104.
- Leegwater PA, Boor PK, Yuan BQ, van der Steen J, Visser A, Konst AA, Oudejans CB, Schutgens RB, Pronk JC and van der Knaap MS. (2002). *Hum. Genet.*, **110**, 279–283.
- Li C and Wong WH. (2001). *Proc. Natl. Acad. Sci., USA*, **98**, 31–36.
- Ljubimova JY, Lakhter AJ, Loksh A, Yong WH, Reidinger MS, Miner JH, Sorokin LM, Ljubimov AV and Black KL. (2001). *Cancer Res.*, **61**, 5601–5610.
- MacDonald TJ, Brown KM, LaFleur B, Peterson K, Lawlor C, Chen Y, Packer RJ, Cogen P and Stephan DA. (2001). *Nat. Genet.*, **29**, 143–152.
- Maity A, Pore N, Lee J, Solomon D and O'Rourke DM. (2000). *Cancer Res.*, **60**, 5879–5886.
- Meng K, Rodriguez-Pena A, Dimitrov T, Chen W, Yamin M, Noda M and Deuel TF. (2000). *Proc. Natl. Acad. Sci. USA*, **97**, 2603–2608.
- Nagane M, Lin H, Cavenee WK and Huang HJ. (2001). *Cancer Lett.*, **161**, S17–S21.
- Nam SW, Clair T, Campo CK, Lee HY, Liotta LA and Stracke ML. (2000). *Oncogene*, **19**, 241–247.
- Neshat MS, Mellinghoff IK, Tran C, Stiles B, Thomas G, Petersen R, Frost P, Gibbons JJ, Wu H and Sawyers CL. (2001). *Proc. Natl. Acad. Sci. USA*, **98**, 10314–10319.
- Ohnishi T, Hiraga S, Izumoto S, Matsumura H, Kanemura Y, Arita N and Hayakawa T. (1998). *Clin. Exp. Metast.*, **16**, 729–741.
- O'Rourke DM, Kao GD, Singh N, Park BW, Muschel RJ, Wu CJ and Greene MI. (1998). *Proc. Natl. Acad. Sci. USA*, **95**, 10842–10847.
- Pauletti G, Dandekar S, Rong H, Ramos L, Peng H, Seshadri R and Slamon DJ. (2001). *J. Clin. Oncol.*, **18**, 3651–3664.
- Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO and Botstein D. (2000). *Nature*, **406**, 747–752.
- Pomeroy SL, Tamayo P, Gaasenbeek M, Sturla LM, Angelo M, McLaughlin ME, Kim JY, Goumnerova LC, Black PM, Lau C, Allen JC, Zagzag D, Olson JM, Curran T, Wetmore C, Biegel JA, Poggio T, Mukherjee S, Rifkin R, Califano A, Stolovitzky G, Louis DN, Mesirov JP, Lander ES and Golub TR (2002). *Nature*, **415**, 436–442.
- Powers C, Aigner A, Stoica GE, McDonnell K and Wellstein A. (2002). *J. Biol. Chem.*, **277**, 14153–14158.
- Preston-Martin S. (1999). *The Gliomas*, Berger MS and Wilson CB (eds). W.B. Saunders Company: Philadelphia, pp. 2–11.

- Reifenberger G, Ichimura K, Reifenberger J, Elkahoulou AG, Meltzer PS and Collins VP. (1996). *Cancer Res.*, **56**, 5141–5145.
- Reifenberger G, Reifenberger J, Ichimura K and Collins VP. (1995). *Cancer Res.*, **55**, 731–734.
- Reifenberger J, Ichimura K, Meltzer PS and Collins VP. (1994). *Cancer Res.*, **54**, 4299–4303.
- Rickman DS, Bobek MP, Misek DE, Kuick R, Blaivas M, Kurnit DM, Taylor J and Hanash SM. (2001). *Cancer Res.*, **61**, 6885–6891.
- Sallinen SL, Sallinen PK, Haapasalo HK, Helin HJ, Helen PT, Schraml P, Kallioniemi OP and Kononen J. (2000). *Cancer Res.*, **60**, 6617–6622.
- Sawyers CL. (2002). *Curr. Opin. Genet. Dev.*, **12**, 111–115.
- Shipp MA, Ross KN, Tamayo P, Weng AP, Kutok JL, Aguiar RC, Gaasenbeek M, Angelo M, Reich M, Pinkus GS, Ray TS, Koval MA, Last KW, Norton A, Lister TA, Mesirov J, Neuberg DS, Lander ES, Aster JC and Golub TR. (2002). *Nat. Med.*, **8**, 68–74.
- Smith JS, Tachibana I, Passe SM, Huntley BK, Borell TJ, Iturria N, O'Fallon JR, Schaefer PL, Scheithauer BW, James CD, Buckner JC and Jenkins RB (2001). *J. Natl. Cancer Inst.*, **93**, 1246–1256.
- Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein Lonning P and Borresen-Dale AL. (2001). *Proc. Natl. Acad. Sci. USA*, **98**, 10869–10874.
- Stracke ML, Krutzsch HC, Unsworth EJ, Arestad A, Cioce V, Schiffmann E and Liotta LA. (1992). *J. Biol. Chem.*, **267**, 2524–2529.
- Venables WN and Ripley BD. (1999) *Modern Applied Statistics With S-Plus*. Springer: New York.
- Watanabe K, Tachibana O, Sata K, Yonekawa Y, Kleihues P and Ohgaki H. (1996). *Brain Pathol.*, **6**, 217–223.
- Xu Q and Reed JC. (1998). *Mol. Cell*, **1**, 337–346.
- Ye K, Hurt KJ, Wu FY, Fang M, Luo HR, Hong JJ, Blackshaw S, Ferris CD and Snyder SH. (2000). *Cell*, **103**, 919–930.
- Yuan F, Chen Y, Dellian M, Safabakhsh N, Ferrara N and Jain RK. (1996). *Proc. Natl. Acad. Sci. USA*, **93**, 14765–14770.
- Zhou Q, Choi G and Anderson DJ. (2001). *Neuron*, **31**, 791–807.