SETA: A novel SH3 domain-containing adapter molecule associated with malignancy in astrocytes[1]

Oliver Bögler,[2] Frank B. Furnari, Andrea Kindler-Roehrborn, Virginia W. Sykes, Richard Yung, H.-J. Su Huang, and Webster K. Cavenee

Ludwig Institute for Cancer Research (O.B., F.B.F., H.-J.S.H., W.K.C.), Department of Medicine (H.-J.S.H., W.K.C.), Center for Molecular Genetics (W.K.C.), and Cancer Center (W.K.C.), University of California San Diego, La Jolla, CA 92093-0660; Institute of Cell Biology, University of Essen Medical School, and West German Cancer Center, Hufelandstrasse 55, D45122 Essen, Germany (A.K.-R); and Department of Anatomy and Division of Neurosurgery, Medical College of Virginia, Virginia Commonwealth University, Richmond VA 23298-0709 (O.B., V.W.S., R.Y.)

Differential display poymerase chain reaction analysis was used to compare five differentiation states of the O-2A progenitor-like cell line CG4: progenitor cells and cells at 12 h or 4 days after the induction of differentiation into oligodendrocytes or astrocytes. This led to the identification of 52 sequence tags that were expressed differentially with cellular phenotype. One sequence was upregulated during differentiation of CG4 cells and represented a novel gene that we named *SETA* (SH3 domain-containing gene expressed in tumorigenic astrocytes). This gene encodes an SH3 domain-containing adapter protein with sequence similarity to the *CD2AP* (CD2 adapter protein) and *CMS* (Cas ligand with multiple *Src* homology) genes. SETA mRNA was expressed at high levels in the developing rat brain but was barely detectable in the normal adult rat or human brain. However, SETA mRNA was found in approximately one half of the human gliomas tested, including astrocytomas grades II, III, and IV, as well as oligodendrogliomas, mixed oligoastrocytomas, and human glioma-derived cell lines. A rat glioma generated by treatment with the alkylating carcinogen ethylnitrosourea on postnatal day 1 and a derived cell line also expressed SETA mRNA. Furthermore, in an in vitro model of astrocytoma progression based on $p53^{-/-}$ astrocytes, expression of *SETA* was restricted to cells that are tumorigenic. *Neuro-Oncology 2, 6–15, 2000 (Posted to Neuro-Oncology [serial online], Doc. 99-31, December 13, 1999. URL <neuro-oncology. mc.duke.edu>)*

Introduction

Identifying molecules specifically associated with glial cell malignancy could potentially contribute to the development of new clinical approaches to gliomas. Genes expressed preferentially in glial tumors could serve as targets for new therapies or as markers to aid neuropathologic diagnoses (Bruner et al. 1997). One potential approach to identifying such molecules is to study the gene expression changes that accompany differentiation of glial progenitor cells whose transformation is thought

to give rise to gliomas. The recent isolation from a human glioblastoma multiforme of a cell with striking similarity to the O-2A progenitor cell described in the developing rat brain (Noble and Mayer-Proschel, 1997) suggests an analysis of gene expression in this cell type. The O-2A lineage provides a well-characterized model of cellular differentiation for which many of the factors controlling cell division, migration, survival, and differentiation are described. A small collection of proteins that are used as markers defining differentiation are differentially expressed. For example, O-2A progenitors express platelet-derived growth factor receptor-alpha (Hart et al., 1989), oligodendrocytes express MBP (Barbarese et al., 1981), and type-2 astrocytes are GFAP positive (Raff et al., 1983). However, little is known about differential gene expression at the mRNA level in this lineage. The description of the O-2A progenitor cell line central glia 4 (CG4; Louis et al., 1992) has facilitated the isolation of pure cultures of differentiated cells, and so has made it easier to perform molecular difference analyses. Here we have used pure populations of CG4 cells at various stages of differentiation for a comparative gene expression study.

The DD-PCR technique (Liang and Pardee, 1992) has brought several significant improvements to comparative gene expression analysis (Liang and Pardee, 1995; Wan et al., 1996). The most significant for the present study is the relative ease with which multiple samples can be compared in one experiment. Furthermore, quantitative increases and decreases in gene expression levels can be monitored more easily in the same experiment, allowing the isolation of genes with a larger variety in expression profiles. Here we performed an extensive DD-PCR analysis of CG4 mRNA and identified 52 differentially expressed sequences. To demonstrate the ability of this approach to lead to genes of potential interest to neuro-oncology, we then present data on a novel gene identified using one of these expressed sequence tags. This gene, *SETA,* was expressed in association with glial malignancy in the adult brain and is a member of a new subfamily of adapter molecules. The discovery of this gene suggests that the study of gene expression in differentiating glial cells may yield molecules of interest to glioma-biology.

Materials and Methods

RNA Extraction and DNase Treatment

CG4 cells at 50% to 80% confluence were cultured in conditions that either maintained the O-2A progenitor phenotype or caused differentiation into either oligodendrocytes or astrocytes. Total cellular RNA was extracted using Tryzol (Life Technologies, Grand Island, NY) according to the manufacturer's protocol. Aliquots of 50 µg of RNA were precipitated and resuspended in 90 µl Tris-EDTA buffer (10 mM Tris [pH 7.6], 1 mM EDTA), and traces of genomic DNA were removed by the addition of 4 µl 1 M Tris-Cl (pH 7.5), 1 µl 1 M MgCl$_2$, 1 µl 10 mM dithiothreitol (DTT), 3 µl RNasin (Promega, Madison, WI), 1 µl DNase I (Roche Molecular Biochemicals, Indianapolis, IN), and incubation at 37°C

for 60 min. The reaction was stopped by addition of 2 µl 0.5 M EDTA (pH 8.0) and 2 µl 10% sodium dodecyl sulfate, and extracted with an equal volume of phenol:chloroform:isoamyl, followed by choloroform. Back extractions of each organic phase with 100 µl diethyl pyrocarbonate (DEPC)-treated water were combined with the recovered aqueous phase. Then the RNA was precipitated and resuspended in Tris-EDTA buffer at 0.5 µg/µl.

Reverse Transcription

Aliquots of 0.2 µg DNase-treated RNA were subjected to reverse transcription in a volume of 11 µl, using 20 pmol of one of four primers: $T_{12}VA$, $T_{12}VT$, $T_{12}VG$, and $T_{12}VC$, where V is A, G, or C. After the reactions were incubated at 65°C for 5 min, they were snap cooled on ice. Then 4 µl 5 × RT buffer (Life Technologies), 2 µl 0.1 M DTT, 1.5 µl 250 µM dNTPs (Amersham-Pharmacia Biotech, Piscataway, NJ), and 0.5 µl RNasin were added, and the reactions were incubated at 37°C for 1–2 min to allow equilibration. Subsequently, 1 µl M-MLV reverse transcriptase (Life Technologies) was added, and the reaction was incubated for a further 60 min at 37°C, followed by heat inactivation at 90°C for 5 min to generate cDNA.

Differential Display PCR

Aliquots of cDNA were subjected to a matrix of PCR reactions using the same $T_{12}VN$ primer that was used for the reverse transcription in combination with the 10-mers described in Table 1. The following reagents were added to 2 µl of the completed first strand cDNA reaction: 2 µl 10X PCR buffer (Roche Molecular Biochemicals), 1.5 µl 25 µM dNTPs, 2 µl 10-mer (stock is 2 pmol/µl), 0.8 µl $T_{12}VN$ (stock is 10 pmol/µl), 0.5 µl ^{35}S-ATP, 10.9 µl water, and 0.3 µl *Taq* polymerase (Roche Molecular Biochemicals). The reactions were assembled in 96-well plates and subjected to the following temperature cycle: 94°C for 3 min; 40 cycles of 94°C for 30 s, 40°C for 2 min, a 70-s ramp to 72°C, 72°C for 30 s; 72°C for 10 min; 4°C until the reactions were recovered. Five microliters of each reaction was combined with 2 µl sequencing loading dye (Sequenase stop dye; Amersham-Pharmacia Biotech) and heated at 80°C for 3 min before being placed on ice. Then, 2.5 µl of each sample was loaded on a prerun 7% denaturing polyacrylamide 8 M urea gel and subjected to 85 W constant power at a temperature of approximately 50°C for 3 h and 15 min, or until the xylene cyanol dye reached the bottom of the gel. The gels were then dried and exposed to film.

Because optimizing the PCR conditions for each combination was impractical, all reactions were performed in the same buffer conditions and with the same temperature cycles. We observed that the $T_{12}VN$ primer, being longer than the 10-mers, sometimes dominated the reaction and generated identical sets of products regardless of which 10-mer was also present. This was apparent not only when the lanes generated from the 5 cDNA substrates with the same primer pair were compared but also when the sets of lanes generated by the same $T_{12}VN$ in combination with different 10-mers were compared. In

Table 1. Evaluation of combinations of 10-mers and $T_{12}VN$ primers in DD-PCR

Primer #	Sequence 5'– >3'	%GC	$T_{12}VT$	$T_{12}VA$	$T_{12}VG$	$T_{12}VC$
12	GAAAACAATA	20	–	–	–	–
7	CTTATCAATG	30	–	–	–	–
9	TAGATATACG	"	–	–	–	–
15	TTTCTGCATA	"	–	+	–	–
19	TATTCACATG	"	–	+	–	–
21	ATGATGATTC	"	–	+	–	–
13	CTTATAGCTC	40	–	+	–	–
4	AAGTTGCTAC	"	+	+	+	–
20	AAGATCGAAG	"	+	+	+	–
23	CTGTTTAGGT	"	+	+	+	–
6	ATGATACACG	"	+	+	+	+
11	CCAAATGATG	"	+	+	+	+
17	GAGATTGAAG	"	+	+	+	+
27	GTTGTTGAAG	"	+	+	+	+
5	ACACTTCACG	50	+	+	+	–
25	CGTGTGAATC	"	+	+	+	–
26	GGTAACTGTG	"	+	+	+	–
30	TGATCAACGG	"	+	+	+	–
31	AGACACTACC	"	+	+	+	–
32	CACCCTAAAG	"	+	+	+	–
1	CTGATCCATG	"	+	+	+	+
28	GCAATCTGTG	"	+	+	+	+
8	CTCGTCGAGA	60	–	–	+	–
14	GCAATCCAGG	"	+	+	+	–
16	TGGCAACCTC	"	+	+	+	–
24	CTAGTCCCTG	"	+	+	+	–
34	GTCCTACTCG	"	+	+	+	–
35	CTACACAGGC	"	+	+	+	–
36	GTCCTTAGCG	"	+	+	+	–
37	GTCCTCAACG	"	+	+	+	–
38	CTACTACCGC	"	+	+	+	–
39	GAGTCACTCG	"	+	+	+	–
40	GTCCTCAGTG	"	+	+	+	–
41	CGTCGTTACC	"	+	+	+	–
33	CGCAGTACTC	"	+	+	+	+
42	GCAGACTGAG	"	+	+	+	+
3	TCTAGCCCGC	70	+	+	+	+
10	CCGTTGGACC	"	+	+	+	+
18	GTGACCACGG	"	+	+	+	+
22	CTCAGCCTCG	"	+	+	+	+
43	GTCTCGTCGG	"	ND	ND	+	+
44	GGCCTACTCG	"	ND	ND	+	+
45	GTGTAGGGCG	"	ND	ND	+	+
46	CGGGTCGATC	"	ND	ND	+	+
47	GCCCTCTTCG	"	ND	ND	+	+
48	CAGGGGCATC	"	ND	ND	+	+
49	GAGACCTCCG	"	ND	ND	+	+
50	GGCCTTCAGG	"	ND	ND	+	+
51	GCTCTCACCG	"	ND	ND	+	+
52	TGCACGGACG	"	ND	ND	+	+
2	CGCGCCCAAG	80	+	+	+	+
29	CTGCCGGACG	"	+	+	+	+

Abbreviations: DD-PCR, differential display polymerase chain reaction; GC, guanine/cytosine.

The primers used in DD-PCR are listed according to their %GC content. In each case, the numerical designation of the primer, the sequence from 5' to 3', and the %GC content are indicated. In addition, the outcome of DD-PCR reactions of each 10-mer with each $T_{12}VN$ primer is shown, with + indicating a successful reaction, – indicating a reaction in which the 10-mer did not contribute to the reaction, and ND indicating reactions that were not done.

order to detect this problem, gels were run so that combinations of the same $T_{12}VN$ with various 10-mers were loaded side by side. This problem was most common with $T_{12}VC$, and then $T_{12}VG$, as would be expected because having either a G (guanine) or a C (cytosine) at the 3′ end increases the annealing temperature of the primer (Table 1). Similarly, the higher the GC (guanine, cytosine) content of the 10-mer used in combination with either $T_{12}VC$ or $T_{12}VG$, the less likely it was that the problem would occur (Table 1). However, the GC content of the 10-mer was not an absolute predictor of whether it would be able to participate in the reaction. For example, several 10-mers with 40% GC content worked in combination with $T_{12}VC$, although several with 60% GC content did not. When the GC content of the 10-mer was 70% or higher, every 10-mer used worked with both $T_{12}VC$ and $T_{12}VG$.

Isolation of Differentially Displayed PCR Products

Using the autoradiograph as a guide, we excised areas of gels and soaked them in 50 µl water for 10 min at room temperature, and then boiled them for 15 min. The released DNA was EtOH-precipitated in the presence of glycogen. The DNA was resuspended in 10 µl TE Tris-EDTA buffer and amplified with the appropriate $T_{12}VN$ and 10-mer primers.

During secondary PCR amplification, multiple products may be amplified from an acrylamide gel slice, likely due to overlapping or very close migrating bands. To identify the majority product responsible for the band detected by autoradiography, we used a TA cloning kit (Invitrogen, Carlsbad, CA) to clone the secondary PCR products. Twenty to 30 white colonies, in addition to one blue colony from another ligation, were grown on LB/Amp plates, lifted onto nylon filters, and hybridized with a probe generated from one of the white colonies by a PCR reaction using TA-1 primer (5′ CCAGTGTGCTGGAATTCGGC) and TA-2 primer (5′ GGATATCTGCAGAATTCGGC), which flank the insert of the pCR-II plasmid from the TA cloning kit (Invitrogen) with minimal vector sequence. If the probe hybridized to a majority of streaked colonies, the colony that generated this probe was considered representative of the band originally excised from the acrylamide gel and was used on RNA dot blots containing 1–2 µg polyA$^+$ RNA/dot from the 5 differentiation time points originally used in the differential display. These blots were treated according to standard protocols and washed to a final stringency of 0.5× saline sodium citrate, 0.1% sodium dodecyl sulfate at 55°C. Conversely, if the probe only hybridized to a minority of streaked colonies, this colony was not considered representative of the cloned products, and a new probe was generated from a different colony.

RNA Dot Blot, Northern and Western Analysis

Total cellular RNA (10 µg) per lane was used in standard RNA dot blot and Northern blotting protocols. Equal RNA transfer was confirmed by methylene blue staining of the membranes. Northern blots were hybridized at 42°C in a formamide-containing hybridization buffer with a random primed ^{32}P-labeled probe derived from *SETA* cDNAs. Western blotting was performed on total cellular protein using an immunopurified rabbit antiserum raised against a histidine-tagged bacterial SETA protein encoding the longest protein. RNA was isolated from p53$^{-/-}$ astrocytes, derived tumors that were described previously (Bogler et al., 1999) and from human surgical biopsies using standard protocols.

Results

Differential Display PCR Analysis of the Oligodendrocyte Progenitor Cell Line CG4

To identify genes, the expression of which varied with differentiation state in CNS glial cells, we performed DD-PCR analysis on the O-2A progenitor cell-like cell line CG4 (Louis et al., 1992). These cells differentiate into oligodendrocytes or astrocytes in response to changes in their culture environment (see Fig. 1) in a fashion similar to primary cultures of rat O-2A progenitors, but are more amenable to long-term culture in vitro. Total cellular RNA from CG4 cells in five different states of differentiation was isolated for comparison. Cells were harvested either while in the progenitor cell-state or at 12 h and at 4 days after being transferred into culture conditions that caused them to differentiate into oligodendrocytes or astrocytes. These time points were chosen because they include the two extremes of the cellular differentiation process in this cell type. At 12 h after the induction of differentiation, there are no observable overt morphological or antigenic changes, whereas after 4 days, the cells have terminally differentiated (data not shown). Therefore, comparison of these RNAs would allow early changes in gene expression to be distinguished from later changes, potentially allowing the identification of genes with different roles in the process of differentiation.

DD-PCR was performed according to the method of Liang and Pardee (1992) with modifications (see Materials and Methods). A large number of different primer combinations were used, consisting of one $T_{12}VN$ primer (where V is A [adenine], C [cytosine], or G [guanine] and N is any base) and one random 10-mer. From the set of approximately 940 DD-PCR reactions, displaying approximately 100 bands per reaction, 268 differentially displayed bands were identified and isolated. Re-amplification, cloning, and unequivocal identification of the major molecular species in each band were achieved for 161 of these (see Table 2). Further analysis of these expressed sequence tags by RNA dot blot of the same total RNA samples used in the DD-PCR was performed to confirm differential expression patterns. This resulted in the identification of 52 differentially expressed sequences that were classified as shown in Table 2. Therefore, approximately one fifth of the bands that were initially identified on the DD-PCR gels resulted in expressed sequence tags that were differentially expressed. Further analysis of one of these fragments, #190, is the topic of the present paper.
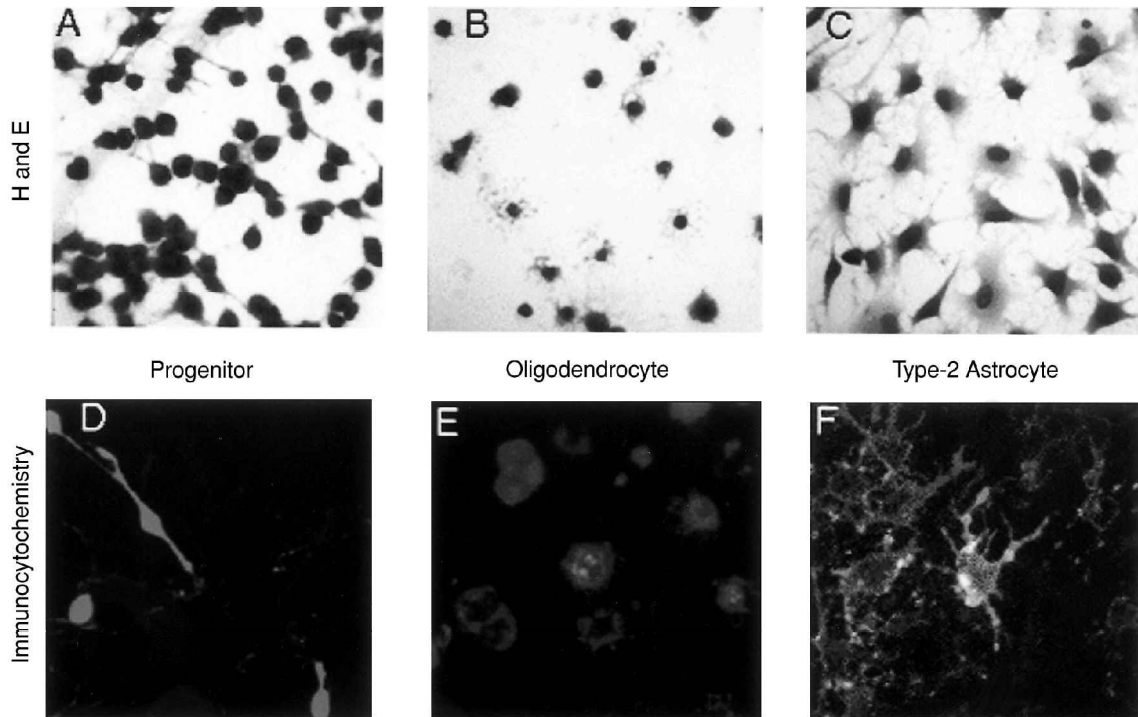
Fig. 1. Differentiation of CG4 cells. Shown are CG4 cells in their progenitor cell state (A and D) and 4 days after the initiation of differentiation into oligodendrocytes (B and E) or type-2 astroctyes (C and F). Cells were stained either with hemotoxylin and eosin (A–C) or with the differentiation state–specific antibodies A2B5 (D), anti-galactocerebroside (E), or anti-glial fibrillary acidic protein (F). Original magnification X2000.

The Isolation of cDNAs Corresponding to Fragment #190

The differentially expressed sequence tag #190 obtained from DD-PCR was used to isolate cDNAs from a poly-dT primed CG4 cDNA library using conventional hybridization screening. This resulted in the isolation of two partially overlapping cDNAs encoding partial open reading frames with no start codons. The cDNAs were homologous at the 5′ end, but differed in their 3′ ends, predicting two alternative C-terminal protein sequences (see Fig. 2A). To extend the 5′ end of the cDNAs, two rounds of 5′ RACE were performed using CG4 and C6 glioma cell line RNA as substrate. The first round of 5′ RACE resulted in the isolation of two additional cDNAs, extending the known sequence and revealing a shared region that encoded a *Src* homology 3 (SH3) domain. The sequences diverged 21 codons 5′ of the N-terminal border of this SH3 domain, with the shorter sequence encoding only an additional 4 amino acids, including a methionine, predicted to be a translational start site due to the presence of an inframe TGA (thymine, guanine, adenine) stop codon 8 codons upstream. The longer sequence remained open in the absence of an ATG (adenine, thymine, guanine) codon, and so was used to design primers for an additional round of 5′ RACE. This in turn resulted in the isolation of sequence e ncoding an additional SH3 domain (Fig. 2A) and an ATG preceded by a TGA 14 codons upstream of it, suggesting this was the start methionine for the longer N-terminal SH3 domain-containing transcript.

Table 2. Categories of DD-PCR products that were successfully cloned and purified

	Astrocytic	O-2A progenitor	Oligodendrocytic	Bipotential	Total
Differentially expressed	23	2	6[a]	21	52
Not differentially expressed	40	23	35	11	109

Abbreviations: DD-PCR, differential display polymerase chain reaction.

The number of DD-PCR products that were either differentially expressed (upper row) or that were cloned because they appeared to be differentially expressed by DD-PCR, but were not by RNA dot-blot analysis (lower row), is shown. Each DD-PCR product is assigned to one lineage compartment category according to where expression was highest in RNA dot-blot analysis (upper row) or appeared to be highest in DD-PCR analysis (lower row). This assignment does not imply that expression was confined to that compartment.

[a]Three of the oligodendrocytic and 18 of the bipotential products were isolated because their DD-PCR products appeared to be specifically detectable in other categories but were reclassified after RNA dot-blot analysis.
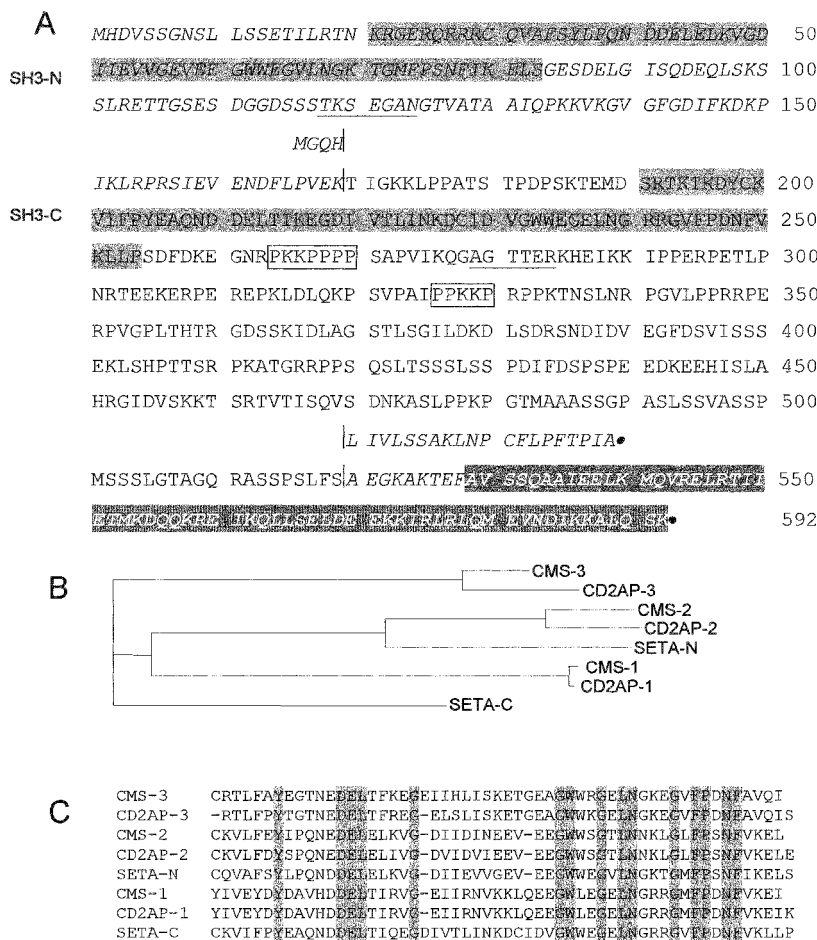
Fig. 2. Predicted protein sequence of SETA (SH3 domain-containing protein expressed in tumorigenic astrocytes). (A) Predicted protein sequence of the #190/*SETA* gene. The sequence common to both cDNAs and the first round of 5′ RACE is shown in roman type. The alternative sequences at the N- and C-termini are shown in italic, and the point of diversion is marked with vertical lines. The sites of the primers used in the two successive rounds of 5′ RACE are underlined. Proline-rich regions conserved with the CD2AP (CD2 adapter protein) and CMS (Cas ligand with multiple *Src* homology) proteins are boxed. SH3 domains are shaded with black type, and the C-terminal coiled-coil domain is shaded with white type. (B) A dendrogram showing the degree of sequence homology at the amino acid level between the SH3 domains encoded by the *SETA, CD2AP,* and *CMS* genes. SETA-N and SETA-C refer to the more N-terminal and more C-terminal SH3 domain of the SETA protein, respectively. (C) Alignment of the SH3 domain sequences of SETA, CD2AP, and CMS. SETA-N and SETA-C are defined in B above; CD2AP and CMS SH3 domains are numbered according to published reports (Dustin et al., 1998; Kirsch et al., 1999). The order of SH3 domains is the same as in the dendrogram. Conserved residues are shaded. (GenBank Accession Number AF131867)

The Predicted Sequence of Proteins Encoded by #190 cDNAs

Protein sequences predicted by #190 cDNAs share a central common region, and two alternative N and C termini (Fig. 2A). The central common region encoded one SH3 domain, termed SH3-C because of its more C-terminal position, and two proline-rich motifs (boxed in Fig. 2A). At the N terminus, the longer combined cDNA sequence predicted 169 amino acids encoding an additional SH3-domain, termed SH3-N because of its more N-terminal position. The shorter N-terminal alternative sequence predicted only 4 amino acids. At the C terminus, the longer sequence encoded 73 amino acids and is predicted to adopt a coiled-coil motif (Lupas et al., 1991), while the shorter sequence substituted 21 amino acids for this region.

Comparison of #190 sequences with the public gene databases by using the basic local alignment search tool (BLAST) did not reveal any identical sequences, suggesting a novel gene. However, it did reveal homology to two other genes that have been recently described: *CD2AP,* the CD2 adapter protein, which plays a role in T-cell receptor clustering (Dustin et al., 1998), and *CMS*, Cas ligand with multiple *Src* homology, which binds to p130[Cas], a tyrosine phosphorylated docking protein (Kirsch et al., 1999). Alignment of the SH3 domains from these proteins reveals that *CD2AP* and *CMS* are more closely related to each other than either is to the #190 cDNA (Fig. 2B). However, the SH3 domains from all three genes share invariant amino acids (Fig. 2C), suggesting that they may be functionally related.

In addition to the similarities in their SH3 domains, an alignment of the predicted protein sequence of the longer form of #190 cDNA, including the SH3-N domain and the C-terminal coiled coil, with the coding regions of CD2AP and CMS, revealed that the proteins share approximately 34% identities over 614 amino acids. The degree of homology was highest within the SH3-domains, which collectively share 60% identities at the amino acid level but remain at 28% outside the SH3 domains, indicating that these genes are related. This analysis also showed that *CD2AP* and *CMS* are more closely related to each other than either is to the #190 gene. Furthermore, all three proteins share the two P-x-x-P motifs boxed in Fig. 2A, which have been shown in CMS to bind to the SH3 domains of other proteins (Kirsch et al., 1999). The relatively high homology between these genes and lack of similarity with other sequences in the database suggest that they represent a new subfamily of adapter proteins.

Expression of #190 in Differentiating Glia and Developing Brain

Dot-blot analysis of CG4 RNA with the original DD-PCR-generated fragment #190 as a probe revealed an increase in expression in differentiated CG4 cells relative to the progenitor state, with more expression in the astrocyte pathway (not shown). Confirmation of this differential expression pattern was sought by Northern analysis using cDNA isolated from the CG4 cDNA library as probe (Fig. 3A).

CG4 cell differentiation was monitored by examining mRNA levels for the astrocytic marker GFAP and the oligodendrocyte marker MBP, as well as by immunohistochemistry with differentiation stage-specific antibodies (Fig. 1). GFAP and MBP mRNA were expressed in CG4 cells that were morphologically bipolar and antigenically A2B5+, GFAP−, GalC−, characteristics of undifferentiated O-2A progenitors (Fig. 3A; Louis et al., 1992). As the cells differentiated into oligodendrocytes, GFAP mRNA was downregulated, and MBP mRNA expression was maintained. Conversely, as CG4 cells differentiated into type-2 astrocytes, MBP expression was downregulated and GFAP expression was maintained.

On CG4 Northern blots, mRNA for the #190 gene appeared as two bands running between the 28S and 18S ribosomal RNA, with approximate sizes of 2.5 and 4 kb. Although #190 mRNA was detected in all CG4 differentiation states, an increase in signal was observed in differentiating astrocytes at days 2 and 3 and in differentiating oligodendrocytes on day 4, relative to the levels of the constitutively expressed Pan1/2 mRNA (Fig. 3A; Jacobs et al., 1994). Interestingly, this increase in #190 mRNA was similar to that seen for GFAP, which also showed a stronger signal on days 1 through 3 as compared with day 4 after the induction of astrocyte differentiation and on day 4 of oligodendrocyte differentiation.

Examination of #190 mRNA expression in developing rat brain showed a strong signal in late embryonic and early postnatal samples, while expression in adult rat brain was barely detectable by Northern analysis (Fig. 3B). On this blot, #190 mRNA appeared as a major band
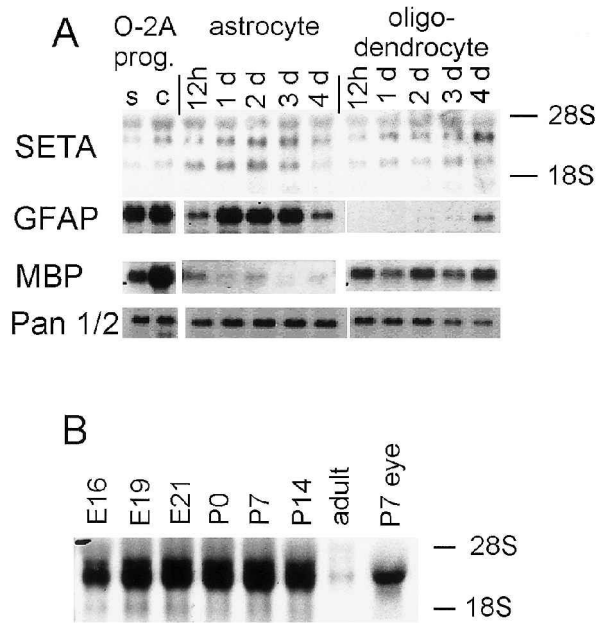

Fig. 3. Northern analysis of SETA in differentiation of glial cells and CNS development. (A) Northern analysis of total RNA isolated from CG4 cells that we maintained in the O-2A-progenitor-like differentiation state (O-2A), either at subconfluent density(ies) or at confluence (c), and 12 h, 1, 2, 3, or 4 days after induction of differentiation into astrocytes or oligodendrocytes, as indicated. The Northern blots were analyzed sequentially with probes for the #190/*SETA*, GFAP (glial fibrillary acidic protein), MBP (myelin basic protein), or *Pan1/2* genes as indicated. (B) Northern analysis of total RNA derived from rat brain or eye with a probe for the *SETA* gene. E16, E19, and E21 refer to rat brain at embryonic days 16, 19, and 21, respectively. P0, P7, and P14 are rat brain at postnatal days 0, 7, and 14, respectively. P7 eye is rat eye at postnatal day 7.

at approximately 4 kb, while the smaller mRNA species was not detected. As the ratio of the two species of mRNA in CG4 cells appeared to be close to equal, this would suggest that cells in the CNS outside the oligodendrocyte lineage express SETA mRNA during development. Further evidence for this comes from the observation that SETA mRNA is present in the developing eye (Fig. 3B), where the retina is part of the CNS. However, oligodendrocyte lineage cells are excluded from the eye by the lamina cribrosa, while type-1 astrocytes, which are likely derived from a separate lineage, migrate into the eye from the optic nerve (ffrench-Constant et al., 1988). Furthermore, the possibility that #190 mRNA in the developing CNS is found in neurons cannot be excluded.

Expression in a Glioma Model Leads to Renaming #190 as SETA

As a first step toward examining the role of the #190 gene in glial cell transformation, we analyzed expression in p53−/− astrocytes. In vitro, these cells rapidly became transformed and with time could gain the ability to form tumors in nude mouse flank or brain, depending on the growth conditions (Bogler et al., 1995; Bogler et al.,
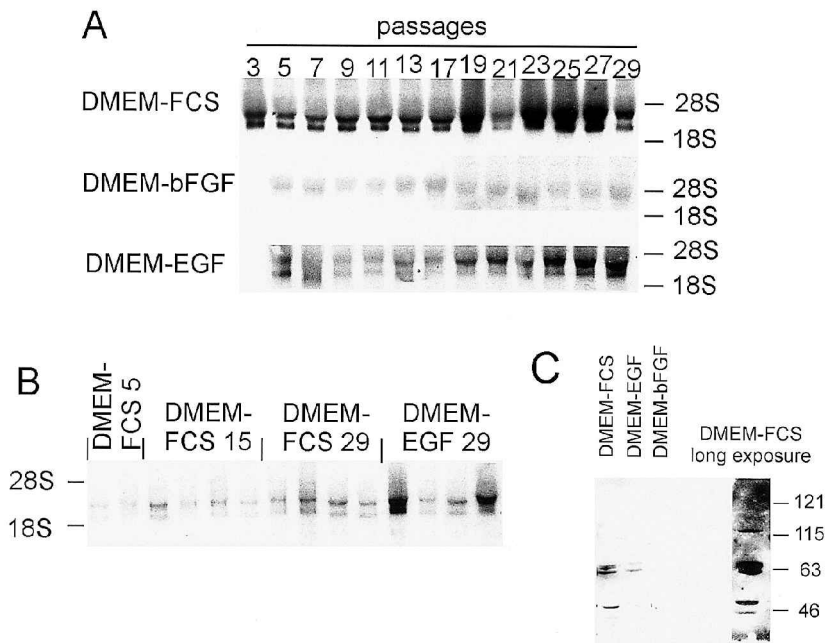
Fig. 4. Analysis of SETA in p53[−/−] astrocytes in culture and in tumors. (A) Northern analysis with a *SETA* cDNA probe of total RNA of p53[−/−] astrocytes grown in DMEM-FCS, DMEM-bFGF, or DMEM-EGF for a varying number of passages, as indicated. (B) Northern analysis with a *SETA* cDNA probe of total RNA of s.c. tumors derived from p53[−/−] astrocytes grown in DMEM-FCS or DMEM-EGF for the number of passages indicated. (C) Western analysis of total cellular protein extracted from p53[−/−] astrocytes grown in DMEM-FCS, DMEM-bFGF, or DMEM-EGF for 29 passages, as indicated, with an immunopurified polyclonal antiserum raised against bacterially expressed SETA protein. Equal protein loading was confirmed by Coomassie staining of a similarly loaded protein gel.

1999; Yahanda et al., 1995). p53[−/−] astrocytes grown in DMEM-FCS or DMEM-EGF became tumorigenic, while those in DMEM-bFGF did not (Bogler et al., 1999). As shown previously, the expression of a group of glioma-associated genes, including those for EGFR, PDGFRs, and VEGF, associates with the tumorigenic phenotype in this system (Bogler et al., 1999).

As shown in Fig. 4A, #190 mRNA was detected only in p53[−/−] astrocytes that had the ability to form tumors: those grown in DMEM-FCS or DMEM-EGF (Bogler et al., 1999). In contrast, cells grown in DMEM-bFGF and not tumorigenic did not express #190 mRNA. The level of SETA mRNA showed some variability over different samples, but overall appeared to increase with higher passage numbers in DMEM-FCS and DMEM-EGF cultures, potentially correlating with increased malignancy (as shown in (Bogler et al., 1999). Northern analysis of a series of s.c. tumors generated by p53[−/−] astrocytes grown in DMEM-FCS or DMEM-EGF for different periods of time revealed that #190 gene expression was maintained in tumors (Fig. 4B). All the tumors analyzed expressed #190 mRNA, although at varying levels, suggesting that its expression is maintained in an in vivo tumor and so is not simply the result of in vitro growth conditions. This led us to rename the #190 gene *SETA* for SH3 domain-containing gene expressed in tumorigenic astrocytes.

In order to investigate whether SETA protein synthesis correlated with the presence of mRNA, protein extracts of p53[−/−] astrocytes at passage 29 in the three conditions were subjected to Western blotting and probed with an affinity-purified polyclonal antiserum against bacterially expressed SETA protein (Fig. 4C). SETA protein was detected in p53[−/−] astrocytes grown in DMEM-FCS and DMEM-EGF but not in DMEM-bFGF cultures as expected from Northern analysis. Six protein bands were revealed by Western analysis, with bands at 45, 48, and 116 kDa and a predominant triplet around 63 kDa. The predicted molecular mass of the longest SETA isoform predicted by the sequences shown in Fig. 1 is 64.8 kDa, suggesting that this may correspond to the triplet detected in Western blots. The sequence variations that gave rise to the smaller and longer proteins detected are not clear at present.

SETA is Expressed in Human and Rat Gliomas

The association of *SETA* gene expression with the ability of p53[−/−] astrocytes to form tumors suggested that it may also be present in human gliomas. Northern analysis of primary human brain tumors using the rat #190 cDNA as a probe revealed that SETA mRNA is detectable in most of these tumors, and at higher levels than in normal brain (Fig. 5A). Specifically, this analysis revealed that SETA mRNA was present in 5 of 10 grade II astrocytomas, 2 of 5 grade III astrocytomas, and 7 of 11 grade IV astrocytomas (Fig. 5A). In addition, SETA mRNA was detected in 2 out of 2 oligodendroglioma and oligoastrocytoma tumors analyzed. In contrast, SETA mRNA was barely detected in two samples of normal human brain. It should be noted that the detection of SETA mRNA in these human tumors was difficult because the rat gene-
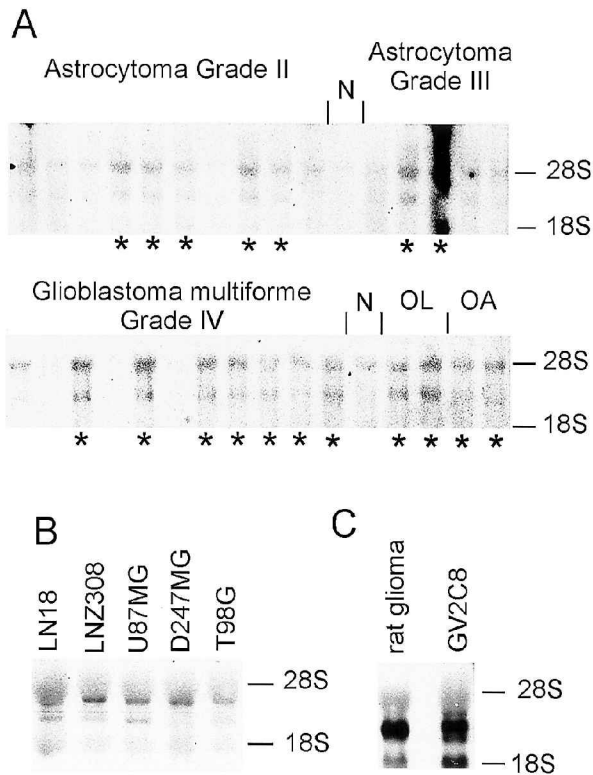
Fig. 5. Northern analysis of human and rat gliomas. (A) Northern analysis of primary human brain tumors. Total RNA isolated from fresh frozen biopsy materials of human brain tumors was probed with *SETA* rat cDNA at low stringency to allow hybridization of the rat probe to human RNA. Furthermore, limiting amounts of total RNA available from these samples required longer exposure times to detect hybridization. As a result, some background signal was observed, particularly in the region of the 28S rRNA. Therefore, only samples that revealed a SETA-specific band midway between the 28S and 18S rRNA were considered to contain SETA mRNA and are marked with an *. Equal RNA transfer was confirmed by methylene blue staining of the membrane (not shown). (B) Northern analysis with a *SETA* cDNA probe of total RNA from human glioma-derived cell lines, as indicated. Equal RNA transfer was confirmed by methylene blue staining of the membrane (not shown). (C) Northern analysis with a *SETA* cDNA probe of total RNA from a rat glioma and the rat glioma-derived cell line GV2C8, as indicated. Equal RNA transfer was confirmed by methylene blue staining of the membrane (not shown).

derived probe did not perform well with human RNA and because large amounts of tumor RNA are difficult to obtain. Therefore, we also examined *SETA* expression in human glioma cell lines, where large amounts of RNA can easily be generated. In agreement with the detection of SETA mRNA in human gliomas, Northern analysis of 5 glioma-derived cell lines also showed the presence of SETA mRNA (Fig. 5B), with LNZ308 and U-87 MG having highest levels of expression. Lastly, because analysis of human RNA with the *SETA* cDNA probe used in these experiments resulted in lower signals than those obtained from rodent material, we analyzed *SETA* expression in an experimentally induced rat glioma. The glioma was generated in a BDIX rat by a single s.c. injec-

tion of N-ethyl-N-nitrosourea on postnatal day 1 and was harvested from the adult rat. In addition, a cell line, GV2C8, derived from such a glioma was also analyzed. As shown in Fig. 5C, both the glioma and cell line showed strong SETA mRNA expression, suggesting that the lower levels seen in human gliomas were a result of reduced cross-species hybridization and not inherently lower *SETA* expression in i.c. tumors.

Discussion

As a first step toward the identification of molecules that may be of use in the treatment or diagnosis of glioma, we have isolated genes that are differentially expressed in the course of glial cell differentiation. The application of DD-PCR to the CG4 cell line model of O-2A lineage differentiation resulted in the identification of 52 sequences that changed expression with differentiation, as assessed by RNA dot-blot analysis.

Most of the differentially expressed genes identified in this study showed significant differences in the mRNA levels between the progenitor state and one or both differentiated states. Approximately half of these sequences, 23 of 52, showed elevated levels of mRNA in the course of astrocytic differentiation, and a similar number, 21, showed an increase in the course of differentiation down both pathways. In contrast, we identified only 6 sequences whose transcription was elevated as cells progressed down the default oligodendrocytic differentiation pathway only, and only 2 sequences that were progenitor specific. It is interesting in this context that oligodendrocyte differentiation requires no instructive factor and occurs when growth factors are withdrawn, while differentiation into type-2 astrocytes requires the presence of instructive signals, such as those present in serum (Louis et al., 1992; Raff et al., 1983; Raff, 1989). Therefore, it is possible that the differential isolation of genes specifically associated with these two pathways reflects the relative importance of transcription as a regulatory mechanism in their execution. According to this hypothesis, many oligodendrocyte genes may already be transcriptionally active in the O-2A progenitor, and so enable the cell to rapidly differentiate in the absence of inducing factors. In contrast, the induced pathway of type-2 astrocyte differentiation may be mediated via the de novo transcriptional activation of genes in response to the appearance of instructive factors in the cell's environment.

In the course of this exhaustive DD-PCR screen, we came across the technical problem of primer competition not previously described in the literature, to our knowledge, concerning this technique. When reactions were displayed on gels in such a way that neighboring sets shared the same $T_{12}VN$ primer, it often became clear that the identity of the 10-mer was of no consequence to the observed pattern of bands. This was more often the case when $T_{12}VG$ or $T_{12}VC$ was combined with 10-mers of low GC content, presumably because the melting temperatures of the primers in these combinations were most unequal. As the admission of such primer combinations would lead to an overestimate of the true number of cDNAs displayed, we excluded them from the data set.

Further analysis of one of these differentially expressed sequences led to the identification of the gene *SETA*, the expression of which correlates with the malignant state in adult brain and astrocytes in culture. Its expression at high levels in the brain during development, but only in glial tumors in the adult, appears to imply that it has a role common to both states, as has been suggested for genes such as the *EGFR* (epidermal growth factor receptor) and *PDGFR* (platelet-derived growth factor receptor) (reviewed in, for example, Westermark et al., 1995). In support of this, *SETA* expression, like the expression for the *EGFR* and *PDGFR* genes, correlates with the ability of p53$^{-/-}$ astrocytes to form tumors in nude mice (Bogler et al., 1999). The absence of SETA mRNA in p53$^{-/-}$ astrocytes grown in DMEM-bFGF, which are highly mitotic, aneuploid, and transformed, but unable to form tumors (Bogler et al., 1999), clearly indicates that a proliferative, transformed, or p53-null state is not sufficient to upregulate the *SETA* gene.

The predicted protein sequences of SETA reveal the presence of SH3 domains and P-x-x-P motifs, known to be involved in specific protein-protein interactions (Mayer et al., 1995). Although their presence does not indicate a specific function for the SETA protein, it suggests that it is a member of the adapter protein family and binds to other proteins, either to modulate their function or to bind them into a complex. The possibility that the two mRNA species observed by Northern blot-ting and the two different 5′ ends isolated for *SETA* cDNAs represent splice variants allows that different forms of *SETA* may function differently. Comparison with *SETA*'s closest homologs, *CD2AP* (Dustin et al., 1998) and *CMS* (Kirsch et al., 1999), suggests that a more N-terminal SH3 domain may yet be missing, particularly as *SETA*'s N-terminal SH3 domain was most similar to the middle SH3 domain of *CD2AP* and *CMS*, and further analyses to address this are underway. Furthermore, the observation that only the most N-terminal SH3 domain of *CD2AP* interacts with CD2 (Dustin et al., 1998) indicates that different SH3 domains in this new subfamily of SH3 domain-containing adapters may have different binding partner preferences and, thus, that different isoforms need to be studied independently.

SETA's tumor-associated high-level expression pattern in the brain, combined with the possibility that its product may function in signal transduction in relation to the organization of the cytoskeleton as has been suggested for *CMS* and *CD2AP*, suggests that further study of this molecule in the context of brain tumors may lead to an increased understanding of the malignant transformation of glia.

Acknowledgment

References

Barbarese, E., and Pfeiffer, S.E. (1981) Developmental regulation of myelin basic protein in dispersed cultures. *Proc. Natl. Acad. Sci. U.S.A.* **78**, 1953–1957.

Bögler, O., Huang, S.H.-J., and Cavenee, W.K. (1995) Loss of wild-type p53 bestows a growth advantage on primary cortical astrocytes and facilitates their in vitro transformation. *Cancer Res.* **55**, 2746–2751.

Bögler, O., Nagane, M., Gillis, J., Huang, S.H.-J., and Cavenee, W.K. (1999) Malignant transformation of p53-deficient astrocytes is modulated by environmental cues in vitro. *Cell Growth Differ.* **10**, 73–86.

Bruner, J.M., Inouye, L., Fuller, G.N., and Langford, L.A. (1997) Diagnostic discrepancies and their clinical impact in a neuropathology referral practice. *Cancer* **79**, 796–803.

Dustin, M.L., Olszowy, M.W., Holdorf, A.D., Li, J., Bromley, S., Desai, N., Widder, P., Rosenberger, F., van der Merwe, P.A., Allen, P.M., and Shaw, A.S. (1998) A novel adaptor protein orchestrates receptor patterning and cytoskeletal polarity in T-cell contacts. *Cell* **94**, 667–677.

ffrench-Constant, C., Miller, R.H., Burne, J.F., and Raff, M.C. (1988) Evidence that migratory oligodendrocyte-type-2 astrocyte (O-2A) progenitor cells are kept out of the rat retina by a barrier at the eye-end of the optic nerve. *J. Neurocytol.* **17**, 13–25.

Hart, I.K., Richardson, W.D., Bolsover, S.R., and Raff, M.C. (1989) PDGF and intracellular signaling in the timing of oligodendrocyte differentiation. *J. Cell Biol.* **109**, 3411–3417.

Jacobs, Y., Xin, X.Q., Dorshkind, K., and Nelson, C. (1994) Pan/E2A expression precedes immunoglobulin heavy-chain expression during B lymphopoiesis in nontransformed cells, and Pan/E2A proteins are not detected in myeloid cells. *Mol. Cell Biol.* **14**, 4087–4096.

Kirsch, K.H., Georgescu, M.-M., Ishimura, S., and Hanafusa, H. (1999) CMS: An adapter molecule involved in cytoskeletal rearrangements. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 6211–6216.

Liang, P., and Pardee, A.B. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* **257**, 967–971.

Liang, P., and Pardee, A.B. (1995) Recent advances in differential display. *Curr. Opin. Immunol.* **7**, 274–280.

Louis, J.C., Magal, E., Muir, D., Manthorpe, M., and Varon, S. (1992) CG-4, a new bipotential glial cell line from rat brain, is capable of differentiating in vitro into either mature oligodendrocytes or type-2 astrocytes. *J. Neurosci. Res.* **31**, 193–204.

Lupas, A., Van Dyke, M., and Stock, J. (1991) Predicting coiled coils from protein sequences. *Science* **252**, 1162–1164.

Mayer, B.J., and Eck, M.J. (1995) SH3 domains. Minding your p's and q's. *Curr. Biol.* **5**, 364–367.

Noble, M., and Mayer-Proschel, M. (1997) Growth factors, glia and gliomas. *J. Neurooncol.* **35**, 193–209.

Raff, M.C. (1989) Glial cell diversification in the rat optic nerve. *Science* **243**, 1450–1455.

Raff, M.C., Miller, R.H., and Noble, M. (1983) A glial progenitor cell that develops in vitro into an astrocyte or an oligodendrocyte depending on culture medium. *Nature* **303**, 390–396.

Wan, J.S., Sharp, S.J., Poirier, G.M., Wagaman, P.C., Chambers, J., Pyati, J., Hom, Y.L., Galindo, J.E., Huvar, A., Peterson, P.A., Jackson, M.R., and Erlander, M.G. (1996) Cloning differentially expressed mRNAs. *Nat. Biotechnol.* **14**, 1685–1691.

Westermark, B., Heldin, C.-H., and Nister, M. (1995) Platelet-derived growth factor in human glioma. *Glia* **15**, 257–263.

Yahanda, A.M., Bruner, J.M., Donehower, L.A., and Morrison, R.S. (1995) Astrocytes derived from p53-deficient mice provide a multistep in vitro model for development of malignant gliomas. *Mol. Cell Biol.* **15**, 4249–4259.