Published OnlineFirst February 28, 2010; DOI: 10.1158/0008-5472.CAN-09-1707

Tumor and Stem Cell Biology

Cancer Research

Transcriptional Profiles of CD133$^+$ and CD133$^-$ Glioblastoma-Derived Cancer Stem Cell Lines Suggest Different Cells of Origin

Claudio Lottaz[1], Dagmar Beier[2], Katharina Meyer[1], Praveen Kumar[2], Andreas Hermann[3], Johannes Schwarz[4], Markus Junker[5], Peter J. Oefner[1], Ulrich Bogdahn[2], Jörg Wischhusen[5], Rainer Spang[1], Alexander Storch[3], and Christoph P. Beier[2,6]

Abstract

Glioblastoma multiforme (GBM) is paradigmatic for the investigation of cancer stem cells (CSC) in solid tumors. Growing evidence suggests that different types of CSC lead to the formation of GBM. This has prompted the present comparison of gene expression profiles between 17 GBM CSC lines and their different putative founder cells. Using a newly derived 24-gene signature, we can now distinguish two subgroups of GBM: Type I CSC lines display "proneural" signature genes and resemble fetal neural stem cell (fNSC) lines, whereas type II CSC lines show "mesenchymal" transcriptional profiles similar to adult NSC (aNSC) lines. Phenotypically, type I CSC lines are CD133 positive and grow as neurospheres. Type II CSC lines, in contrast, display (semi-)adherent growth and lack CD133 expression. Molecular differences between type I and type II CSC lines include the expression of extracellular matrix molecules and the transcriptional activity of the WNT and the transforming growth factor-β/bone morphogenetic protein signaling pathways. Importantly, these characteristics were not affected by induced adherence on laminin. Comparing CSC lines with their putative cells of origin, we observed greatly increased proliferation and impaired differentiation capacity in both types of CSC lines but no cancer-associated activation of otherwise silent signaling pathways. Thus, our data suggest that the heterogeneous tumor entity GBM may derive from cells that have preserved or acquired properties of either fNSC or aNSC but lost the corresponding differentiation potential. Moreover, we propose a gene signature that enables the subclassification of GBM according to their putative cells of origin. *Cancer Res; 70(5); 2030–40. ©2010 AACR.*

Introduction

Glioblastoma multiformes (GBM) are among the most deadly human cancers (1, 2). They contain a rare subpopulation of cells with stem cell–like properties, so-called cancer stem cells (CSC) or tumor-initiating cells. On implantation into nude mice, these CSC give rise to tumors that histologically mimic the original lesions, whereas other cells isolated from the same tumors are nontumorigenic *in vivo* (3–5).

Based on our recent observation that a subgroup of primary astrocytic GBM gave rise to CD133$^-$ CSC lines, which differed from previously described CD133$^+$ CSC lines with regard to transcriptional profiles and phenotypic properties, we already proposed that the histologically defined entity of GBM might comprise different GBM subtypes driven by different CSC (6). Further support came from an unrelated study that confirmed our findings in a second (larger) cohort of cell lines, proposing an additional array-based classification system of CSC lines based on their spontaneous clustering pattern (7). At present, it remains unclear if these two classification systems actually reflect similar biological entities and cover all relevant subgroups. In addition, neither of these studies addressed the question whether different CSC might be derived from different cells of origin (6, 7). Analysis of GBM according to their respective CSC may also shed light on the putative cells of origin of GBM. Postulated founder cell populations comprise mature astrocytes (8), unipotent "restricted" progenitor cells, or astrocytic progenitor cells fused with mesenchymal stem cells (MSC; refs. 9, 10). In addition, normally quiescent multipotent adult neural stem cells (aNSC) from the subventricular zone (SVZ) of the adult brain as well as fetal NSC (fNSC) may constitute founder cell populations (10–15). aNSC maintain

Authors' Affiliations: [1]Institute of Functional Genomics, University of Regensburg Medical School; [2]Department of Neurology, University of Regensburg, Regensburg, Germany; [3]Department of Neurology and Center for Regenerative Therapies Dresden, Dresden University of Technology, Dresden, Germany; [4]Department of Neurology, University of Leipzig, Leipzig, Germany; [5]Interdisciplinary Center for Clinical Research, Junior Research Group "Tumor progression and immune escape," University of Würzburg, Medical School, Clinics for Gynecology and Obstetrics, Würzburg, Germany; and [6]Department of Neurology, University of Aachen, Aachen, Germany

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

C. Lottaz and D. Beier contributed equally to this work.

Corresponding Author: Christoph P. Beier, Department of Neurology, University of Regensburg Medical School, Universitätsstrasse 84, 93053 Regensburg, Germany. Phone: 49-941-9410; Fax: 49-941-941-3205; E-mail: Christoph.Beier@gmx.de.

doi: 10.1158/0008-5472.CAN-09-1707

Published OnlineFirst February 28, 2010; DOI: 10.1158/0008-5472.CAN-09-1707

neurogenesis in the adult and share striking similarities with their malignant counterparts (3, 5, 14). However, aNSC do not express CD133, which is the only established surface marker for CSC in GBM. Conversely, fNSC express CD133 even in the early postnatal stage, but it is unclear if these cells persist until adulthood (16). Thus, the actual cells of origin of GBM CSC remain enigmatic.

Gene arrays have proven effective in establishing molecularly defined subgroups within histologically defined tumor entities (15, 17–20). Recently, Phillips and colleagues have described a 35-gene signature that divided "GBM" into three subtypes, which resembled different stages of neurogenesis, correlated with the patient's prognosis, and delineated a pattern of disease progression. According to the predominant physiologic function of the signature genes, the subtypes were named "proneural," "mesenchymal," and "proliferative" (15). Whether this classification reflects different types of tumors originating from different CSC remains an open question.

We addressed this question by analyzing the gene expression profiles of 17 GBM CSC lines, which had been shown to preserve the genetic and transcriptional alterations of the original tumor and hardly acquire new mutations when propagated *in vitro* (14), and a panel of cell types constituting possible founder cell populations. This revealed that proneural neurosphere-like growing $CD133^+$ GBM CSC lines share similarities with fNSC lines, whereas mesenchymal semiadherent/adherently growing $CD133^-$ GBM CSC lines resemble aNSC lines. Bioinformatic analysis yielded a 24-gene signature (based on 29 transcripts) that enabled a robust *in vitro* and *in vivo* class prediction for GBM CSC. The findings suggest two separate founder cell populations and allow the identification of molecular aberrations that distinguish CSC from healthy stem cells.

Materials and Methods

Tissue culture and gene arrays. All experiments had been approved by the local ethics committee (University of Regensburg, No. 05/105, June 14, 2005). The CSC lines (R8, R18, R53, and R54 ± laminin) were established, propagated, and subjected to microarray analysis as previously described (6). R8-laminin, R18-laminin, R53-laminin, and R54-laminin were cultured on laminin for 7 d. The proliferation of CSC lines after transforming growth factor-β (TGF-β) treatment was determined using the AlamarBlue assay according to the manufacturer's instructions. fNSC and aNSC cultures have been described by Maisel and colleagues (21); two additional aNSC cultures were newly established as described by Maisel and colleagues (21). The following data sets downloaded from the National Center for Biotechnology Information–Gene Expression Omnibus (GEO) were used: GDS2728: CSC lines derived from primary astrocytic GBM (R8, R11, R18, R28, R43, and R46); GSE8049: CSC lines derived from newly diagnosed adult GBM (GS1 and GS3–GS9), as well as one CSC line derived from a long-term survivor with recurrent GBM with oligodendroglial features (GS2), all described in detail by Günther and colleagues (7); GSE15209: CSC lines derived from newly diagnosed GBM [these CSC lines were described by Pollard and colleagues (22)]; GSE9834: 3 samples of human astrocytes; GSE9835: 3 samples of human neurons derived from human cord blood stem cell culture via transdifferentiation; and GDS1816: 100 high-grade glioma samples described by Phillips and colleagues (15). R8-laminin, R53, R53-laminin, R54, R54-laminin, R18-laminin, GDS2728, GSE8049, GSE9834, GSE15209, and GSE9835 expression profiles were generated on the Affymetrix GeneChip HG-U133 Plus 2.0 microarray platform. In addition, 2 fNSC lines, 3 aNSC lines, 3 MSC lines, and 100 high-grade glioma samples (GDS1816) were analyzed by the HG-U133A platform.

Microarray preprocessing and quality control. The array data were merged by matching probe sets using the matchprobes extension from the Bioconductor suite of life science–related extensions (23) for the statistical computing environment R (24). From the publicly available data, we used raw expression values that were jointly normalized together with our own data. Gene expression profiles were background corrected and normalized on probe level using the variance stabilization method by Huber and colleagues (25). Normalized probe intensities were summarized into gene expression levels using the additive model described in Irizarry and colleagues (26) fitted by the median polish method (27). For quality control, we generated an overall box plot showing the distribution of gene expression levels for each array and one rank residual plot per sample to illustrate the variance stability across expression levels. Thus, we detected hybridization problems with one microarray and excluded this sample from the analysis.

Statistical analysis of microarray data. To group samples with similar expression profiles, we performed complete linkage hierarchical clustering. We selected 500 genes that were most variably expressed between all samples. Distances between samples were computed using Euclidean distance. We determined three groups of CSC lines by cutting the dendrogram accordingly. Stability of this clustering was evaluated using consensus clustering (23). We selected genes that were differentially expressed between the clusters using absolute correlation of gene expression values with cluster labels. We reported the top ranking genes only and no *P* values because the cluster labels were driven by the gene expression values, rendering *P* values meaningless. We calculated a list of differentially expressed genes between CSC lines and possible cells of origin using linear models implemented in the Bioconductor package limma. False discovery rates for lists of differentially expressed genes were calculated according to Benjamini-Hochberg (24). We tested for enrichment of Gene Ontology (GO) terms and pathways from the Kyoto Encyclopedia of Genes and Genomes using Fisher's exact test. Shrunken centroid classification implemented in the Bioconductor package limma prediction analysis of microarrays (PAM) was used (28) to learn a classifier for discriminating between cluster 2 (fNSC) and cluster 3 (aNSC). PAM identified a set of 29 transcripts with minimal cross-validation error. These 29 transcripts, which define 24-signature genes, were used to predict the stem cell phenotype of 6 independent samples. We also examined these 24-signature genes in the GBM data set GEO (GDS1815).

Published OnlineFirst February 28, 2010; DOI: 10.1158/0008-5472.CAN-09-1707

The signature genes were "averaged" using a standard additive model and fitted with a median polish procedure (27) to calculate a consensus expression index (referred to as CSC index).

Fluorescence-activated cell sorting. Cells were dissociated and resuspended in PBS containing 0.5% (w/v) nonspecific IgG and 2 mmol/L EDTA. Cells were stained with mouse anti-human CD44-FITC (clone MEM-85; ImmunoTools), mouse anti-human platelet-derived growth factor receptor α–phycoerythrin (PDGFRα-PE; clone 16A1; BioLegend), mouse anti-human CD133/2-allophycocyanin (clone 293C3; Miltenyi Biotech), rat anti-human epidermal growth factor receptor (EGFR)-PE (clone ICR10; AbD Serotec), or the corresponding isotype control antibody (mIgG1-PE, mIgG2b-FITC, mIgG2b-PE, and ratIgG2a-PE; Caltag Laboratories). Cells were analyzed on a BD FACSCalibur.

Results

Gene arrays from GBM CSC lines cluster with NSC lines. We compared gene expression profiles of 17 GBM CSC lines with those of different populations of putative founder cells comprising 3 samples of human astrocytes, 3 samples of human neurons, 2 fNSC lines, 3 aNSC lines, and 3 MSC lines.

Hierarchical clustering based on the 500 genes most variably expressed among all samples revealed three distinct clusters. Profiles of the CSC lines fell only into two clusters comprising either fNSC or aNSC lines but not in the third cluster containing neurons, astrocytes, and MSC lines (Fig. 1A). Note that no information on the cell type was used in the selection of genes used for clustering (i.e., the clustering was done unsupervised). To validate that our clustering was robust and reflected strong intracluster similarity and intercluster dissimilarity of expression profiles, consensus clustering was performed (29). We randomly selected 100 subsets of size 80 from the 500 top variance genes and repeated the clustering procedure for each subset of genes separately. For each pair of samples, we counted how often the two samples had been assigned to the same cluster. Stability and reproducibility of the clustering indicated strong similarities of the samples within each cluster (Fig. 1B).

We then selected genes that were differentially expressed between the three clusters by screening the complete set of genes on the array using the correlation of gene expression to the cluster labels. Because significance testing is not valid in a context where the groups are defined by the expression data, we do not report *P* values but selected a total of 300 genes, which were overexpressed in one of the three groups (Fig. 1C).

We have recently compared fNSC and aNSC lines in detail (21) and found that DCX, CD44, PDGFRα, EGFR, and CD133 were differentially expressed in aNSC lines ($CD44^{high}$, $CD133^{low}$, $PDGFR\alpha^{low}$, DCX^{neg}, and $EGFR^{low}$) and fNSC lines ($CD44^{low}$, $CD133^{high}$, $PDGFR\alpha^{high}$, DCX^{pos}, and $EGFR^{high}$). Nestin was expressed in both groups ($nestin^{high}$). We used these markers to validate the relationship between type I/II CSC lines and fNSC/aNSC lines. As previously described (6), both types of CSC lines expressed nestin. Similar to fNSC, type I CSC lines highly expressed CD133, PDGFRα, and EGFR and contained DCX-positive cells (Fig. 1D; data not shown). Conversely, type II CSC lines expressed CD44 but neither PDGFRα nor CD133 or DCX (Fig. 1D; data not shown). Interestingly, type II CSC lines also showed EGFR expression (Fig. 1D). Although the respective expression level may vary for specific markers (e.g., CD133) within a group, the qualitatively similar marker expression underscores the close relationship of type I/II CSC to their putative cells of origin suggested by the transcriptional profile. Of note, the expression of differentiation markers (βIII-tubulin, GFAP, GalC, and MBP) did not uniformly differ between both groups (data not shown).

CD133 expression correlates with transcriptional similarity to either aNSC or fNSC lines. Only two NSC clusters contained profiles of GBM CSC lines. Interestingly, CD133 expression of CSC lines correlated significantly with their classification into the two clusters ($P < 0.001$, χ^2 test). Eight of eight $CD133^{+}$ neurosphere-like growing GBM CSC lines clustered with fNSC lines. In contrast, both $CD133^{-/+}$ semiadherent and five of seven $CD133^{-}$ adherently growing GBM CSC lines clustered with aNSC lines (Fig. 1; Supplementary Table S1). In the following, we will refer to these clusters of CSC lines by type I and type II (Supplementary Table S1). The median age of patients with GBM driven by type I CSC was 70 years in contrast to 60 years in patients with GBM driven by type II CSC (Fig. 2A).

The assignment to a respective cell of origin is not driven by similar growth patterns. We observed that the clustering correlated significantly with the growth pattern of cell lines, suggesting that growth patterns might be a strong confounding factor driving the expression profiles. To test this hypothesis, we included four additional microarrays from type I (R18 and R54) and type II (R8 and R53) CSC lines now cultured as a semiconfluent monolayer on laminin (Fig. 2B) and CSC lines published by Pollard and colleagues (22), which had always been grown on laminin since their derivation from patient material. Irrespective of the growth conditions, gene expression profiling grouped these CSC lines together with type I/II cell lines (Fig. 2C; Supplementary Fig. S1), suggesting that the underlying biological differences between the CSC lines, but not the growth pattern, are driving the clustering. In addition, the growth pattern of fNSC and aNSC lines (Fig. 2D) was inversely correlated with the growth patterns of the associated CSC lines (Figs. 1A–C and 2C). Together, the different growth patterns (6, 7) seem to be the consequence of the underlying biological differences between type I and II CSC lines but not their cause.

Comparison of the two subtypes of GBM CSC lines. We then compared the transcriptional profiles of the 17 GBM CSC lines. Notably, transcripts indicating TGF-β activity as well as key proteins of the TGF-β/bone morphogenetic protein (BMP) signaling pathway were significantly upregulated in type II GBM CSC lines compared with type I GBM CSC lines (Table 1; Supplementary Table S3). In line with this observation, all type II GBM CSC lines were responsive to TGF-β and displayed reduced proliferation and migration after treatment with the cytokine (Fig. 3A). In contrast, TGF-β did not alter growth, clonogenicity, and migration of all type I GBM

Published OnlineFirst February 28, 2010; DOI: 10.1158/0008-5472.CAN-09-1707

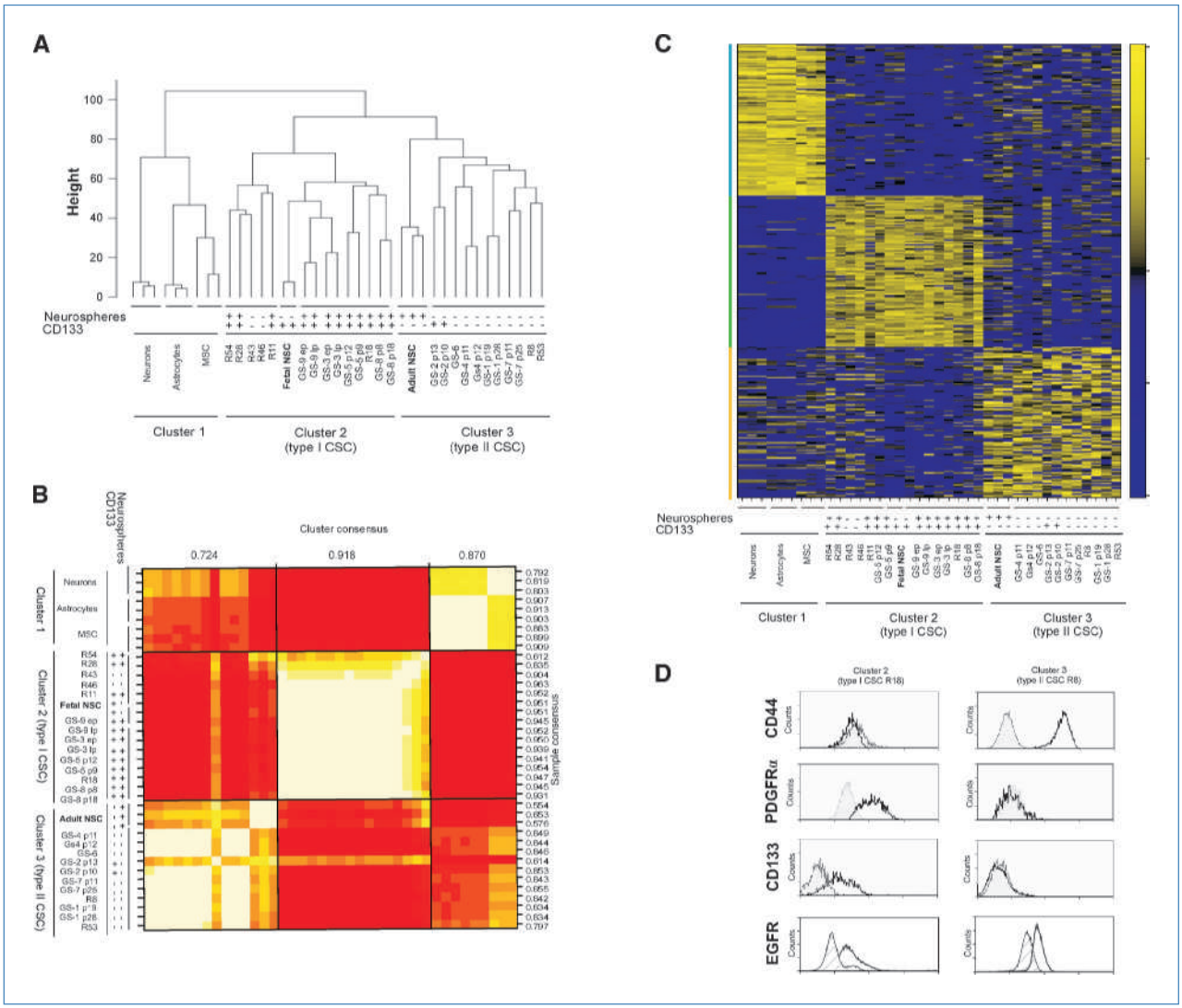

Figure 1. Relationship of CD133$^+$ and CD133$^-$ CSC lines and putative cells of origin. A, 500 genes were selected being most variably expressed between all samples. Distances between the samples were computed using Euclidean distance. Based on these distances, unsupervised clustering was performed using the complete linkage method. lp, late passage; ep, early passage; GS, glioblastoma sample; Rxy, cell line from Regensburg (± laminin); +, CD133 expression or neurosphere formation. B, consensus clustering was performed. We randomly selected 100 subsets of size 80 from the 500 top variance genes and computed a clustering for each subset of genes separately. For each pair of samples, we counted how often two samples fell in the same cluster. The counts are visualized in a color-coded consensus matrix. Dark red corresponds to pairs that were never clustered together, whereas beige corresponds to pairs that were always clustered together. The infrequent intermediate counts are represented by a color gradient. C, heat map of 300 genes selected based on the clustering results. Each cluster is represented by 100 genes that are overexpressed in this cluster. D, cell surface expression of markers for fNSC and aNSC was investigated on type I and II CSC lines using flow cytometry. Representative results of the CSC lines R18 (type I) and R8 (type II) are shown.

CSC lines investigated (Fig. 3A; data not shown). Additional major differences included the expression of extracellular matrix (ECM) and focal adhesion proteins (7). Accordingly, the regulation of the TGF-β/BMP signaling cascade constitutes a major difference between the two subtypes of GBM CSC lines, which might be relevant for the selection of patients likely to benefit from experimental anti–TGF-β therapies. However, a selection of patients according to the GBM subtype would require a molecularly defined signature that permits the correct assignment of each patient to either group.

A 24-gene signature reliably discriminates between type I and type II GBM CSC lines. Although the expression of CD133 significantly correlated with the group assignment of type I and type II CSC lines, its expression did not define either group. To identify signature genes allowing a group assignment of new CSC lines, we used shrunken centroid classification and identified 29 transcripts encoding 24-signature genes, which discriminate between type I and type II GBM CSC lines (Supplementary Table S2). To test the signature on independent data, we classified microarrays of six additional, only recently characterized CSC and fNSC lines (17). In contrast to

Published OnlineFirst February 28, 2010; DOI: 10.1158/0008-5472.CAN-09-1707

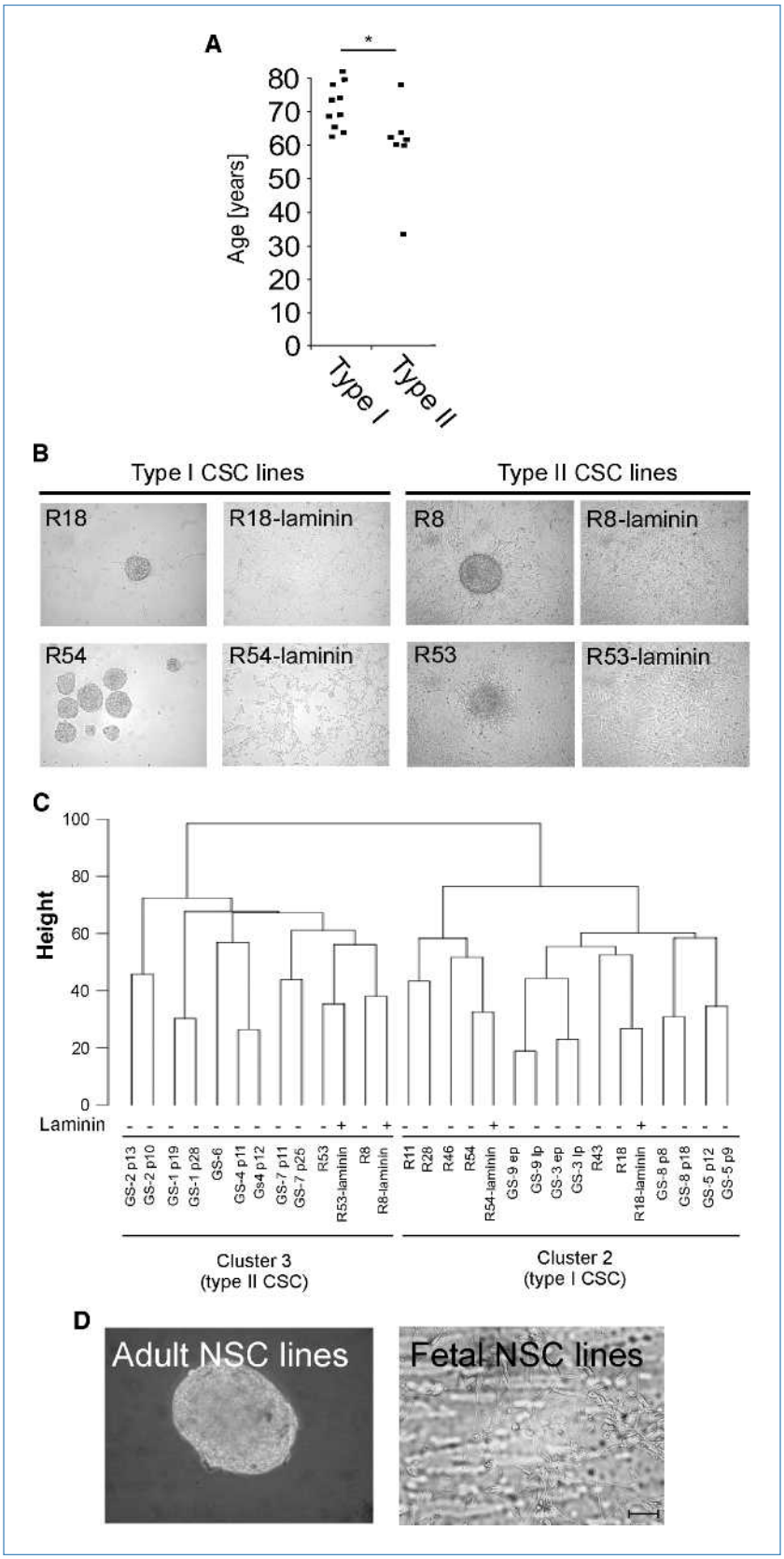

Figure 2. The group assignment is not driven by similar growth patterns. A, age distribution of patients diagnosed with GBM driven by either type I or type II CSC lines. *, $P = 0.03$, two-sided Student's t test assuming different variances. B, the CSC lines R8, R18, R53, and R54 were grown with or without laminin for 7 d. The resulting growth patterns are shown. C, 2,000 genes showing the strongest variance across all samples were selected. Distances between the samples were computed using Euclidean distance. Based on these distances, unsupervised clustering was performed using the complete linkage method. D, growth pattern of fNSC and aNSC lines (representative pictures are shown).

Published OnlineFirst February 28, 2010; DOI: 10.1158/0008-5472.CAN-09-1707

Table 1. Selection of differentially regulated genes between type I and type II GBM CSC lines

Category and gene symbol name		Probe set	Type I	Type II
BMP/TGF-β signaling				
SKIL	Ski-like oncogene	215889_at		+
SMAD3	SMAD family member 3	205397_x_at		+
SMAD7	SMAD family member 7	204790_at		+
BMP2	Bone morphogenetic protein 2	205289_at		+
TGFBR2	Transforming growth factor, β receptor II (70/80 kDa)	208944_at		+
TGFBI	Transforming growth factor, β-induced, 68 kDa	201506_at		+
ECM				
CD44	CD44 molecule (Indian blood group)	212014_x_at		+
ITGB5	Integrin, β_5	214021_x_at		+
ITGB3	Integrin, β_3 (platelet glycoprotein IIIa, antigen CD61)	204627_s_at		+
ITGB5	Integrin, β_5	214020_x_at		+
ITGB1	Integrin, β_1	211945_s_at		+
Signaling pathways				
VEGFC	Vascular endothelial growth factor C	209946_at		+
RAB27A	RAB27A, member RAS oncogene family	209514_s_at		+
FZD6	Frizzled homologue 6	203987_at		+
RAB32	RAB32, member RAS oncogene family	204214_s_at		+
DKK3	Dickkopf homologue 3	202196_s_at		+
MET	Met proto-oncogene (hepatocyte growth factor receptor)	213807_x_at		+
RASAL2	RAS protein activator like 2	219026_s_at		+
RAB22A	RAB22A, member RAS oncogene family	218360_at		+
AXL	AXL receptor tyrosine kinase	202686_s_at		+
SRC	v-src	213324_at		+
PML	Promyelocytic leukemia	211014_s_at		+
DLL3	Delta-like 3	219537_x_at	+	
MARK1	MAP/microtubule affinity-regulating kinase 1	221047_s_at	+	
RHOT2	Ras homologue gene family, member T2	221789_x_at	+	
EFNA3	Ephrin-A3	210132_at	+	
CTNND2	Catenin, δ2	209618_at	+	
Stem cell marker				
SOX2	SRY (sex determining region Y)-box 2	213722_at	+	
OLIG2	Oligodendrocyte lineage transcription factor 2	213825_at	+	
SOX11	SRY (sex determining region Y)-box 11	204915_s_at	+	

the CSC lines analyzed thus far, Pollard and colleagues cultured all CSC lines on laminin, resulting in an adherent growth pattern of all CSC lines. Notably, published features of these six CSC lines agree with the classification (Fig. 3B). The $CD133^-$ CSC line GS166 was classified as type II. The $CD133^+$ type II line GS179 expressed GFAPδ filaments, which are specific for aNSC in the SVZ and were not detected in the analyzed type I CSC lines. Conversely, all CSC lines classified as type I CSC lines were $CD133^+$. In addition, the gene signature identified all fNSC lines analyzed by Pollard and colleagues and grouped them together with type I GBM CSC lines (Fig. 3B; Supplementary Table S1). Together, the signature genes discriminated between type I and type II CSC lines irrespective of the culture conditions and the growth pattern.

Relation of type I and II GBM CSC lines to the proneural or the mesenchymal subtype. Phillips and colleagues established an array-based classification, which uses a 35-gene signature to define three different GBM subtypes referred to as proneural, mesenchymal, and proliferative, respectively. We integrated our data derived from GBM CSC lines with the data of the 100 high-grade glioma samples of Phillips and colleagues and calculated a CSC index using the 24-signature genes. A low index indicated a type II–like expression profile, whereas a high index suggested a type I–like expression profile. Going back to the molecular subclasses of the samples published by Phillips and colleagues, we observed that the proneural phenotype was associated with a high CSC index, suggesting that these GBM CSC lines might be driven by type I GBM CSC, whereas the mesenchymal phenotype was associated with a low CSC index, suggesting that these GBMs might be related to type II GBM CSC (Fig. 3C and D).

Comparison of type I GBM CSC lines with fNSC lines. The cluster analysis suggested that type I GBM CSC lines had originated from cells that had acquired or preserved features

Published OnlineFirst February 28, 2010; DOI: 10.1158/0008-5472.CAN-09-1707

similar to those of fNSC. The most striking transcriptional differences suggested an impaired differentiation capacity of CSC lines compared with fNSC lines. fNSC lines expressed more transcripts characterizing mature neurons or astrocytes, including S100, receptors for neurotransmitters (e.g., AMPA4, nicotinergic acetylcholine, γ-aminobutyric acid A, catecholamine receptors, and serotonin 2A), and ion channels (e.g., KCND3 and CACNA1S). In addition, cytokines (e.g., CTNF, PDGFB, FGF18, TGFB2, BMP7, and FGF4) and genes indicating mesenchymal differentiation (e.g., myosin and troponin) were overexpressed in fNSC lines. Upregulated genes in type I CSC lines indicated rapidly proliferating and metabolically active cells (heat shock proteins, genes associated with RNA processing, and oxidative phosphorylation; Table 2; Supplementary Tables S3 and S4). Only few previously GBM-associated transcripts were differentially regulated in type I CSC lines. The overrepresented GO terms included atypical MHC-I proteins (e.g., HLA-G and HLA-E; ref. 30) and telomere-stabilizing genes (DKC1 and TERF1). Transcriptional profiling in CSC lines, however, revealed neither

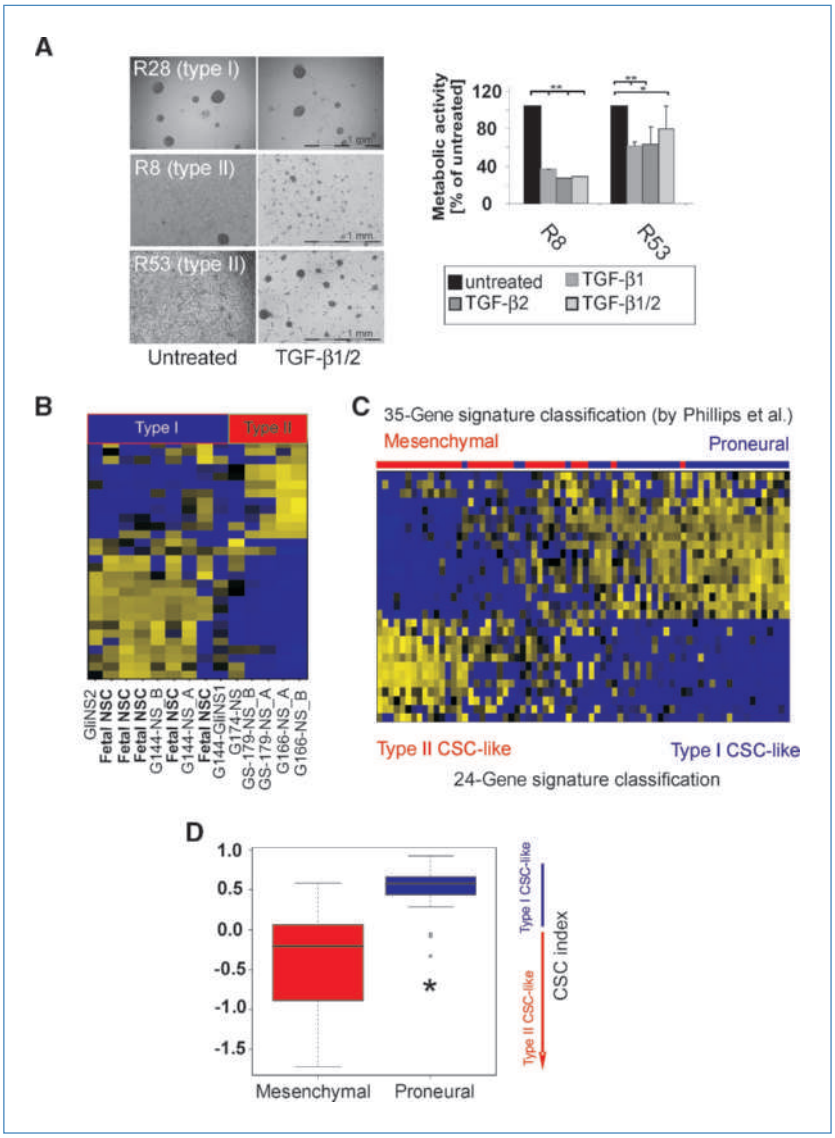

Figure 3. Relationship of CD133$^+$ and CD133$^-$ GBM CSC lines with mesenchymal and proneural GBM subtypes. A, type I (R11, R18, R28, R44, and R54) and type II (R8 and R53) CSC lines were treated with 10 ng/mL TGF-β for 7 d. Left, growth pattern of type I CSC lines (representative pictures of the type I CSC line R28) and of both type II CSC lines; right, the metabolic activity of both type II CSC lines was determined by the AlamarBlue assay. **, $P < 0.001$; *, $P < 0.01$, two-sided Student's t test. B, heat map of 6 different CSC lines and 4 fNSC lines described by Pollard and colleagues (22) clustered according to the 29 transcripts that encode the 24-signature genes. C, heat map of the 24-signature genes and the data on GBM subtypes of Phillips and colleagues (15). Only proneural and mesenchymal samples were included and sorted by the CSC index of signature genes. The color bar above the heat map encodes the original subclasses of Phillips and colleagues. Red, mesenchymal; blue, proneural. D, box plots of the CSC index in the two subgroups characterized by Phillips and colleagues as mesenchymal (red) and proneural (blue). *, $P < 10^{-10}$, two-sided Student's t test.

Published OnlineFirst February 28, 2010; DOI: 10.1158/0008-5472.CAN-09-1707

Table 2. Selection of differentially regulated genes between fNSC lines and type I GBM CSC lines

Category and gene symbol name		Probe set	fNSC lines	Type I CSC lines
Transcription and translation				
EIF1	Eukaryotic translation initiation factor 1	211956_s_at		+
EIF2AK1	Eukaryotic translation initiation factor 2-α kinase 1	217736_s_at		+
EIF4H	Eukaryotic translation initiation factor 4H	206621_s_at		+
TCEA1	Transcription elongation factor A (SII), 1	216241_s_at		+
TSEN34	tRNA splicing endonuclease 34 homologue (*S. cerevisiae*)	218132_s_at		+
Heat shock proteins				
HSPD1	Heat shock 60 kDa protein 1 (chaperonin)	200807_s_at		+
HSPE1	Heat shock 10 kDa protein 1 (chaperonin 10)	205133_s_at		+
HSPB11	Heat shock protein family B (small), member 11	203960_s_at		+
HSPA8	Heat shock 70 kDa protein 8	221891_x_at		+
CCT8	Chaperonin containing TCP1, subunit 8 (θ)	200873_s_at		+
HSP90AA1	Heat shock protein 90 kDa α (cytosolic), class A member 1	214328_s_at		+
HSP90B1	Heat shock protein 90 kDa β (Grp94), member 1	216449_x_at		+
Oxidative phosphorylation				
NDUFA7	NADH dehydrogenase (ubiquinone) 1 α subcomplex, 7, 14.5 kDa	202785_at		+
ATP5E	ATP synthase, H+ transporting, mitochondrial F1 complex, ε subunit	217801_at		+
IDH3A	Isocitrate dehydrogenase 3 (NAD+) α	202069_s_at		+
SDHB	Succinate dehydrogenase complex, subunit B, iron sulfur (Ip)	202675_at		+
NQO2	NAD(P)H dehydrogenase, quinone 2	203814_s_at		+
Neuronal differentiation				
HTR4	5-Hydroxytryptamine (serotonin) receptor 4	207578_s_at	+	
CHRNG	Cholinergic receptor, nicotinic, γ	221355_at	+	
GRIA4	Glutamate receptor, ionotrophic, AMPA 4	208464_at	+	
GABRA5	γ-Aminobutyric acid (GABA) A receptor, α5	206456_at	+	
GFRA2	GDNF family receptor α2	205721_at	+	
S100A14	S100 calcium binding protein A14	218677_at	+	
Cytokines and miscellaneous				
TP63	Tumor protein p63	211195_s_at	+	
EGFR	Epidermal growth factor receptor	211551_at	+	
TGFB2	Transforming growth factor, β2	220407_s_at	+	
Telomere regulation				
DKC1	Dyskeratosis congenita 1, dyskerin	201479_at		+
TERF1	Telomeric repeat binding factor (NIMA-interacting) 1	203448_s_at		+

aberrant activation of signaling pathways (e.g., Shh, WNT, and Notch) nor upregulation of antiapoptotic proteins. In summary, type I GBM CSC and fNSC lines showed surprisingly few differences and fairly similar transcription profiles. The few major transcriptional differences suggest that CSC lines proliferated faster than their nonmalignant counterparts possibly favored by impaired differentiation capacity and stabilization of telomeres.

Comparison of type II GBM CSC lines with aNSC lines. Type II GBM CSC lines might originate from cells acquiring or preserving features similar to those of aNSC. Again, the comparison of type II CSC and aNSC lines revealed only a very limited set of transcriptional differences (Table 3; Supplementary Tables S3 and S5). Type II GBM CSC lines showed an impaired spontaneous differentiation pattern compared with aNSC lines. Transcripts indicating oligodendroglial differentiation (e.g., MBP, MOG, and MAG) and proteins typically expressed in neurons (receptors for neurotransmitters and voltage-gated ion channels) were downregulated. Conversely, transcripts associated with increased metabolic activity or mRNA and DNA synthesis indicated a higher proliferation rate of GBM CSC lines compared with aNSC lines. The gene expression profiles did not indicate a differential activation of intracellular signaling pathways associated with "stemness" or antiapoptotic genes, suggesting that aNSC and CSC lines use similar signaling cascades to maintain pluripotency. Taken together, the differences between type II GBM CSC lines and aNSC lines comprised impaired differentiation combined with accelerated cell cycle and increased metabolic activity.

Discussion

Using Affymetrix Human Genome U133A arrays, Phillips and colleagues described three molecular subclasses of GBM seemingly reflecting different stages of neurogenesis (15, 18). Conversely, our group (6) and Günther and colleagues (7)

Published OnlineFirst February 28, 2010; DOI: 10.1158/0008-5472.CAN-09-1707

reported two different types of CSC lines derived from newly diagnosed GBM and cultured in medium, supporting the growth of NSC and GBM CSC. This study now integrates these three reports and suggests that different cells of origin give rise to different types of GBM CSC that account for the heterogeneity of GBM (Supplementary Table S1).

The array-based classification published by Günther and colleagues (7) was controversial because the distinct growth patterns of cluster 1 and 2 GBM CSC lines might have introduced a bias into the analysis. In addition, Pollard and colleagues (22) cultured six different CSC lines on laminin and did not report spontaneous clustering based on the transcriptional patterns, although our 24-gene signature classified them unambiguously. This suggests that differences in the growth pattern may superimpose the biological differences and make them undetectable if only few CSC lines are analyzed. However, our clustering analysis, which is based on the collected data available now, shows that, in line with a recent study by Kenny and colleagues (31), neither the culture conditions (± ECM proteins) nor the growth patterns have a dominant effect on the group assignment of CSC lines based on the transcriptional profile. In fact, the assignment to a given group reflects more profound biological differences as indicated by the different age of patients (Fig. 2A) and the 24-gene signature derived from *in vitro* CSC lines that could also be used to classify gene expression data obtained from primary GBM samples used directly *ex vivo* without culturing (Fig. 3).

Table 3. Selection of differentially regulated genes between aNSC lines and type II GBM CSC lines

Category and gene symbol name		Probe set	aNSC lines	Type II CSC lines
Ribosome, transcription, and translation				
RPS21	Ribosomal protein S21	200834_s_at		+
RPL36	Ribosomal protein L36	219762_s_at		+
RPL18A	Ribosomal protein L18a	200869_at		+
SFRS10	Splicing factor, arginine/serine-rich 10	200892_s_at		+
EEF2	Eukaryotic translation elongation factor 2	204102_s_at		+
EIF2AK1	Eukaryotic translation initiation factor 2-α kinase 1	217736_s_at		+
EIF4A1	Eukaryotic translation initiation factor 4A, 1	211787_s_at		+
EIF3M	Eukaryotic translation initiation factor 3, M	202232_s_at		+
Metabolism				
CYCS	Cytochrome *c*, somatic	208905_at		+
ATP6V1C2	ATPase, H+ transporting, lysosomal 42 kDa, V1 subunit C2	208638_at		+
OAZ1	Ornithine decarboxylase antizyme 1	215952_s_at		+
NDUFA4	NADH dehydrogenase 1α subcomplex, 4, 9 kDa	217773_s_at		+
COX7B	Cytochrome *c* oxidase subunit VIIb	202110_at		+
SDHD	Succinate dehydrogenase complex, subunit D, integral membrane protein	202026_at		+
IDH3A	Isocitrate dehydrogenase 3 (NAD+) α	202069_s_at		+
Cell cycle				
CDK4	Cyclin-dependent kinase 4	202246_s_at		+
MYCBP	c-myc binding protein	203359_s_at		+
CDK2AP1	CDK2-associated protein 1	201938_at		+
CKS2	CDC28 protein kinase regulatory subunit 2	204170_s_at		+
Markers for differentiation				
MBP	Myelin basic protein	209072_at	+	
MOG	Myelin oligodendrocyte glycoprotein	214650_x_at	+	
MAG	Myelin associated glycoprotein	216617_s_at	+	
SNCA	Synuclein, α (non-A4 component of amyloid precursor)	204467_s_at	+	
FAIM2	Fas apoptotic inhibitory molecule 2	203619_s_at	+	
NCAM1	Neural cell adhesion molecule 1	214952_at	+	
GABBR1	γ-Aminobutyric acid (GABA) B receptor, 1	203146_s_at	+	
Miscellaneous				
TGFB2	Transforming growth factor, β2	220407_s_at		+
OLIG2	Oligodendrocyte lineage transcription factor 2	213825_at	+	
FGFR2	Fibroblast growth factor receptor 2	208228_s_at	+	
SUMO2	SMT3 suppressor of mif two 3 homologue 2	215452_x_at		+
SUMO1	SMT3 suppressor of mif two 3 homologue 1	211069_s_at		+

Published OnlineFirst February 28, 2010; DOI: 10.1158/0008-5472.CAN-09-1707

The cell of origin of GBM is still vague. In line with a previous report (14), we show that all GBM CSC lines cluster with NSC lines but not with astrocytes. Thus, GBMs are likely to originate from cells that have acquired (e.g., via dedifferentiation) or preserved features of NSC.

However, no such statement can be made for those ~60% of GBM that did not give rise to CSC lines on propagation in NSC medium (5, 6, 32). A limitation of our microarray-based study is the inability to actually observe the transformation of a putative founder cell into a GBM CSC. Although type I and II CSC and their putative cells of origin share similar marker expression patterns, our approach can only generate hypotheses but not prove them. Only new conditional *in vivo* experiments, which extend recently established animal models (33, 34) to the point where the putative founder populations can be specifically transformed by the selective induction of the respective genetic changes, will allow final conclusions.

Nevertheless, our data imply that $CD133^-$ type II GBM CSC may originate from $CD133^-$ aNSC located at the SVZ/hippocampus (16), which concords with several recent reports on the development of GBM (8, 34, 35). Considering that there is heterogeneity among the type II CSC lines, different molecular mechanisms may be involved in the genesis of type II CSC. Larger series will be needed to identify possible subgroups. $CD133^+$ type I GBM CSC, in contrast, cannot be traced back to a specific cell of origin at this moment. The respective cell has to preserve or acquire a fNSC-like phenotype. Putative candidate cells comprise among others $CD133^+$ radial glia-like ependymal cells published by Coskun and colleagues (36). As aNSC can be reprogrammed into pluripotent embryonal stem cells by activation of the transcription factor Oct4 (37), the similarities between type I CSC lines and fNSC lines could also be due to the reacquisition of a fNSC-like phenotype by aNSC.

The finding that both type I and the molecularly distinct type II GBM CSC cannot be distinguished from their putative cells of origin by the activation of aberrant signaling pathways but rather by a massively impaired differentiation capacity is quite striking: This implies that signaling pathways, which are constitutively active in NSC, are fully sufficient to drive tumor growth (38, 39). Accordingly, new therapies for GBM should not only focus on the inhibition of growth factor signaling but also consider strategies to differentiate or eliminate CSC (40, 41).

However, although impaired differentiation potential seems to be a common characteristic of both type I and type II CSC lines, gene expression profiling also revealed marked biological differences between the two GBM subgroups. This includes the differential expression of integrins β_3, β_5, and α_V (among other ECM-related proteins), which might predict the responsiveness to new integrin inhibitors such as cilengitide (42). Therefore, a correct assignment of a tumor to the respective subtype may have profound clinical implications. This may be achieved in part by the assessment of CD133 expression, which is of high clinical importance *in vivo* (43, 44) and shows a significant correlation with the molecular array-based assignment. However, we here provide a new 24-gene signature, which seems to be more specific than CD133 expression and should be further evaluated for its potential to reliably discriminate between the two tumor subtypes and the respective CSC.

In conclusion, we confirmed the existence of two distinct types of GBM CSC. They develop differently and derive from different cells of origin or via different molecular steps. During transition from the respective NSC to the related CSC, the loss of the appropriate differentiation potential may be the most critical event, although additional alterations leading to accelerated proliferation are also required for the malignant phenotype (particularly in type II GBM CSC). Our data thus help to better understand the development of GBM, the heterogeneity of these tumors, and the accordingly variable prognosis of patients with GBM and may lead to the development of more personalized and improved therapies targeting GBM CSC.

Disclosure of Potential Conflicts of Interest

C.P. Beier: commercial research grant, Merck Serono. The other authors disclosed no potential conflicts of interest.

Acknowledgments

We thank Ludwig Aigner for the support of the study and Birgit Jachnik for excellent technical support.

Grant Support

BayGene and NGFNplus Brain Tumor Network Subproject 7 No. 01GS0887.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received 05/11/2009; revised 11/24/2009; accepted 12/10/2009; published OnlineFirst 02/09/2010.

References

1. Stupp R, Mason WP, van den Bent MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. N Engl J Med 2005;352:987–96.
2. Beier CP, Schmid C, Gorlia T, et al. RNOP-09: pegylated liposomal doxorubicine and prolonged temozolomide in addition to radiotherapy in newly diagnosed glioblastoma—a phase II study. BMC Cancer 2009;9:308.
3. Singh SK, Hawkins C, Clarke ID, et al. Identification of human brain tumour initiating cells. Nature 2004;432:396–401.
4. Ignatova TN, Kukekov VG, Laywell ED, et al. Human cortical glial tumors contain neural stem-like cells expressing astroglial and neuronal markers *in vitro*. Glia 2002;39:193–206.
5. Galli R, Binda E, Orfanelli U, et al. Isolation and characterization of tumorigenic, stem-like neural precursors from human glioblastoma. Cancer Res 2004;64:7011–21.
6. Beier D, Hau P, Proescholdt M, et al. $CD133^+$ and $CD133^-$ glioblastoma-derived cancer stem cells show differential growth characteristics and molecular profiles. Cancer Res 2007;67:4010–5.
7. Günther HS, Schmidt NO, Phillips HS, et al. Glioblastoma-derived stem cell-enriched cultures form distinct subgroups according to molecular and phenotypic criteria. Oncogene 2008;27:2897–909.
8. Uhrbom L, Dai C, Celestino JC, et al. Ink4a-Arf loss cooperates with KRas activation in astrocytes and neural progenitors to generate

Published OnlineFirst February 28, 2010; DOI: 10.1158/0008-5472.CAN-09-1707

glioblastomas of various morphologies depending on activated Akt. Cancer Res 2002;62:5551–8.
9. Bjerkvig R, Tysnes BB, Aboody KS, Najbauer J, Terzis AJ. Opinion: the origin of the cancer stem cell: current controversies and new insights. Nat Rev Cancer 2005;5:899–904.
10. Stiles CD, Rowitch DH. Glioma stem cells: a midterm exam. Neuron 2008;58:832–46.
11. Shiras A, Chettiar S, Shepal V, et al. Spontaneous transformation of human adult non-tumorigenic stem cells to cancer stem cells is driven by genomic instability in a human model of glioblastoma. Stem Cells 2007;25:1478–89.
12. Gil-Perotin S, Marin-Husstege M, Li J, et al. Loss of p53 induces changes in the behavior of subventricular zone cells: implication for the genesis of glial tumors. J Neurosci 2006;26:1107–16.
13. Jackson EL, Garcia-Verdugo JM, Gil-Perotin S, et al. PDGFRα-positive B cells are neural stem cells in the adult SVZ that form glioma-like growths in response to increased PDGF signaling. Neuron 2006;51:187–99.
14. Lee J, Kotliarova S, Kotliarov Y, et al. Tumor stem cells derived from glioblastomas cultured in bFGF and EGF more closely mirror the phenotype and genotype of primary tumors than do serum-cultured cell lines. Cancer Cell 2006;9:391–403.
15. Phillips HS, Kharbanda S, Chen R, et al. Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. Cancer Cell 2006;9:157–73.
16. Pfenninger CV, Roschupkina T, Hertwig F, et al. CD133 is not present on neurogenic astrocytes in the adult subventricular zone, but on embryonic neural stem cells, ependymal cells, and glioblastoma cells. Cancer Res 2007;67:5727–36.
17. Li A, Walling J, Ahn S, et al. Unsupervised analysis of transcriptomic profiles reveals six glioma subtypes. Cancer Res 2009;69:2091–9.
18. Lee Y, Scheck AC, Cloughesy TF, et al. Gene expression analysis of glioblastomas identifies the major molecular basis for the prognostic benefit of younger age. BMC Med Genomics 2008;1:52.
19. Freije WA, Castro-Vargas FE, Fang Z, et al. Gene expression profiling of gliomas strongly predicts survival. Cancer Res 2004;64:6503–10.
20. Nutt CL, Mani DR, Betensky RA, et al. Gene expression-based classification of malignant gliomas correlates better with survival than histological classification. Cancer Res 2003;63:1602–7.
21. Maisel M, Herr A, Milosevic J, et al. Transcription profiling of adult and fetal human neuroprogenitors identifies divergent paths to maintain the neuroprogenitor cell state. Stem Cells 2007;25:1231–40.
22. Pollard SM, Yoshikawa K, Clarke ID, et al. Glioma stem cell lines expanded in adherent culture have tumor-specific phenotypes and are suitable for chemical and genetic screens. Cell Stem Cell 2009;4:568–80.
23. Gentleman RC, Carey VJ, Bates DM, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 2004;5:R80.
24. R Development Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2008.
25. Huber W, von Heydebreck A, Sultmann H, Poustka A, Vingron M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. Bioinformatics 2002;18 Suppl 1:S96–104.
26. Irizarry RA, Bolstad BM, Collin F, et al. Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 2003;31:e15.
27. Tukey JW. Exploratory data analysis. Reading (MA): Addison-Wesley; 1977.
28. Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci U S A 2002;99:6567–72.
29. Monti S, Tamayo P, Mesirov JP, Golub TR. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. Mach Learn 2003;52:91–118.
30. Wischhusen J, Waschbisch A, Wiendl H. Immune-refractory cancers and their little helpers—an extended role for immunetolerogenic MHC molecules HLA-G and HLA-E? Semin Cancer Biol 2007;17:459–68.
31. Kenny PA, Lee GY, Myers CA, et al. The morphologies of breast cancer cell lines in three-dimensional assays correlate with their profiles of gene expression. Mol Oncol 2007;1:84–96.
32. Pallini R, Ricci-Vitiani L, Banna GL, et al. Cancer stem cell analysis and clinical outcome in patients with glioblastoma multiforme. Clin Cancer Res 2008;14:8205–12.
33. Alcantara Llaguno S, Chen J, Kwon CH, et al. Malignant astrocytomas originate from neural stem/progenitor cells in a somatic tumor suppressor mouse model. Cancer Cell 2009;15:45–56.
34. Zheng H, Ying H, Yan H, et al. p53 and Pten control neural and glioma stem/progenitor cell renewal and differentiation. Nature 2008;455:1129–33.
35. Zhu Y, Guignard F, Zhao D, et al. Early inactivation of p53 tumor suppressor gene cooperating with NF1 loss induces malignant astrocytoma. Cancer Cell 2005;8:119–30.
36. Coskun V, Wu H, Blanchi B, et al. CD133$^+$ neural stem cells in the ependyma of mammalian postnatal forebrain. Proc Natl Acad Sci U S A 2008;105:1026–31.
37. Kim JB, Sebastiano V, Wu G, Arauzo-Bravo MJ, Schöler HR. Oct4-induced pluripotency in adult neural stem cells. Cell 2009;136:411–9.
38. Li ZZ, Wang H, Eyler CE, Hjeimeiand AB, Rich JN. Turning cancer stem cells inside out: an exploration of glioma stem cell signaling pathways. J Biol Chem 2009;284:16705–1609.
39. Bleau AM, Hambardzumyan D, Ozawa T, et al. PTEN/PI3K/Akt pathway regulates the side population phenotype and ABCG2 activity in glioma tumor stem-like cells. Cell Stem Cell 2009;4:226–35.
40. Beier D, Rohrl S, Pillai DR, et al. Temozolomide preferentially depletes cancer stem cells in glioblastoma. Cancer Res 2008;68:5706–15.
41. Piccirillo SG, Reynolds BA, Zanetti N, et al. Bone morphogenetic proteins inhibit the tumorigenic potential of human brain tumour-initiating cells. Nature 2006;444:761–5.
42. Reardon DA, Nabors LB, Stupp R, Mikkelsen T. Cilengitide: an integrin-targeting arginine-glycine-aspartic acid peptide with promising activity for glioblastoma multiforme. Expert Opin Investig Drugs 2008;17:1225–35.
43. Joo KM, Kim SY, Jin X, et al. Clinical and biological implications of CD133-positive and CD133-negative cells in glioblastomas. Lab Invest 2008;88:808–15.
44. Zeppernick F, Ahmadi R, Campos B, et al. Stem cell marker CD133 affects clinical outcome in glioma patients. Clin Cancer Res 2008;14:123–9.

AACR American Association for Cancer Research

Cancer Research

The Journal of Cancer Research (1916–1930) | The American Journal of Cancer (1931–1940)

Transcriptional Profiles of CD133$^+$ and CD133$^-$ Glioblastoma-Derived Cancer Stem Cell Lines Suggest Different Cells of Origin

Claudio Lottaz, Dagmar Beier, Katharina Meyer, et al.

Cancer Res 2010;70:2030-2040. Published OnlineFirst February 28, 2010.