Stem Cell Reports, Volume *8*

Supplemental Information

Prospective Isolation and Comparison of Human Germinal Matrix and Glioblastoma EGFR^{+} Populations with Stem Cell Properties

Jessica Tome-Garcia, Rut Tejero, German Nudelman, Raymund L. Yong, Robert Sebra, Huaien Wang, Mary Fowkes, Margret Magid, Martin Walsh, Violeta Silva-Vargas, Elena Zaslavsky, Roland H. Friedel, Fiona Doetsch, and Nadejda M. Tsankova

Figure S1 (Relates to Figure 1 and Figure 2)

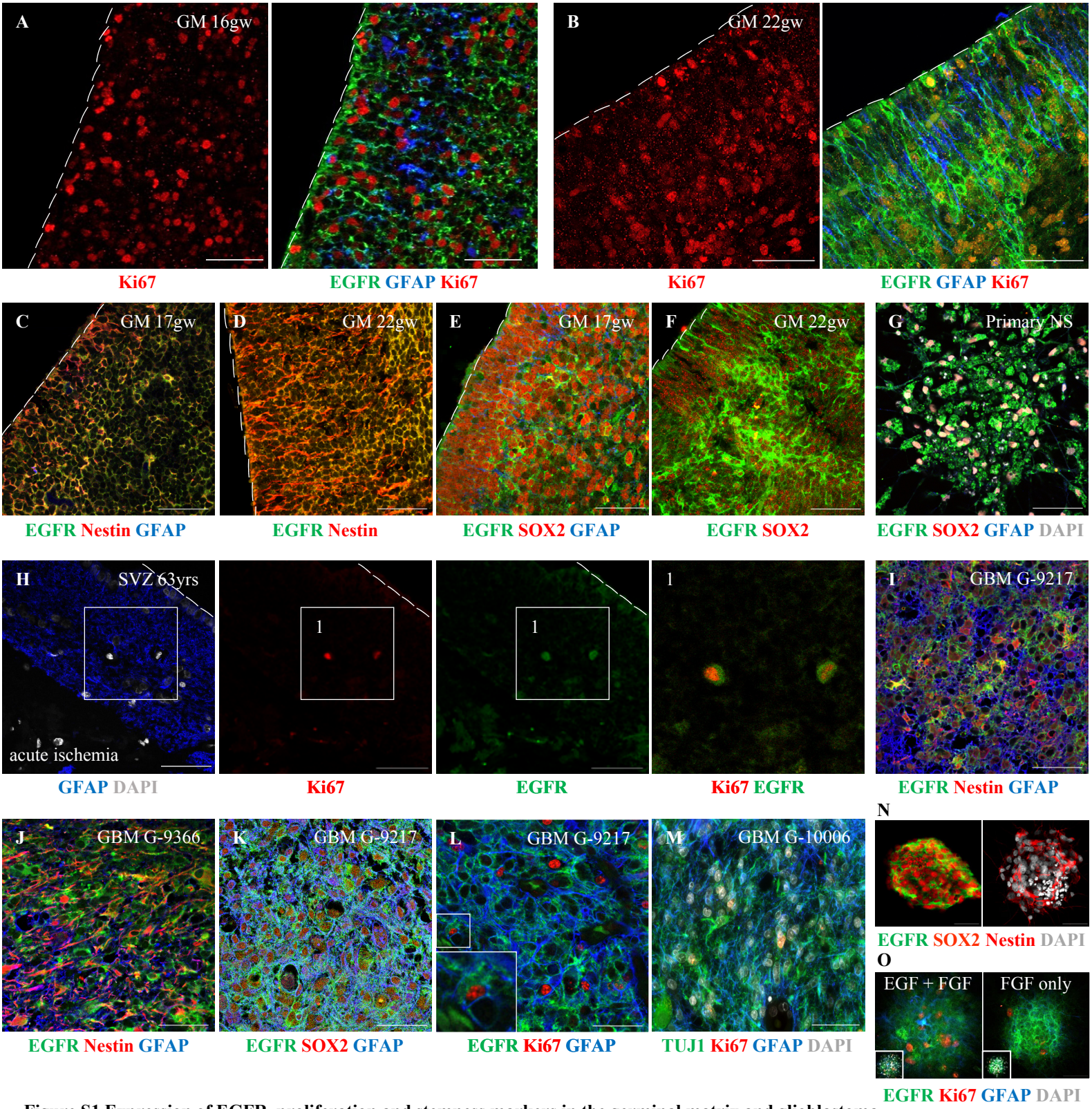


Figure S1 Expression of EGFR, proliferation and stemness markers in the germinal matrix and glioblastoma

Immunofluorescence analysis of EGFR expression in germinal matrix (GM) tissue at 16 and 22 gestational weeks (gw) reveals (**A-B**) active proliferation, as assessed by Ki67, in many EGFR+ but also in EGFR– cells (see also Figure 1A). Most EGFR+ cells in the GM co-express the stem cell markers Nestin (**C-D**) and SOX2 (**E-F**). NESTIN+EGFR– and SOX2+EGFR– cells are also present. (**G**) Single cell-derived clones (1 cell/well) from EGFR+ GM isolates are diffusely positive for EGFR and SOX2, and negative for GFAP. **(H)** Although normally quiescent, cells within the adult SVZ tissue can become reactivated to enter cell cycle after an acute ischemic injury, showing increased EGFR expression and occasional co-localization with Ki67. Immunofluorescence staining of glioblastoma (GBM) tissues (also used for FACS) show frequent co-expression of EGFR with Nestin **(I-J)**, SOX2 **(K),** Ki67 and GFAP **(L)**; as well as variable expression of the neuronal differentiating marker TUJ1 **(M)**. (**N**) Single-cell seeding (1cell/well) of ^{LB}EGFR+DAPI^{low} populations yields GS with diffuse expression of EGFR (ED), SOX2, and Nestin. (**O**) Low-passaged, self-renewing GS grow with and without EGF supplementation in culture, the latter displaying upregulated EGFR expression (10cells/ul, GBM G-12746, ^{LB}EGFR+DAPI^{low} population) . Dashed line indicates ependymal layer. Scale bar: 50µm

Figure S2 (Relates to Figure 1 and Figure 2)

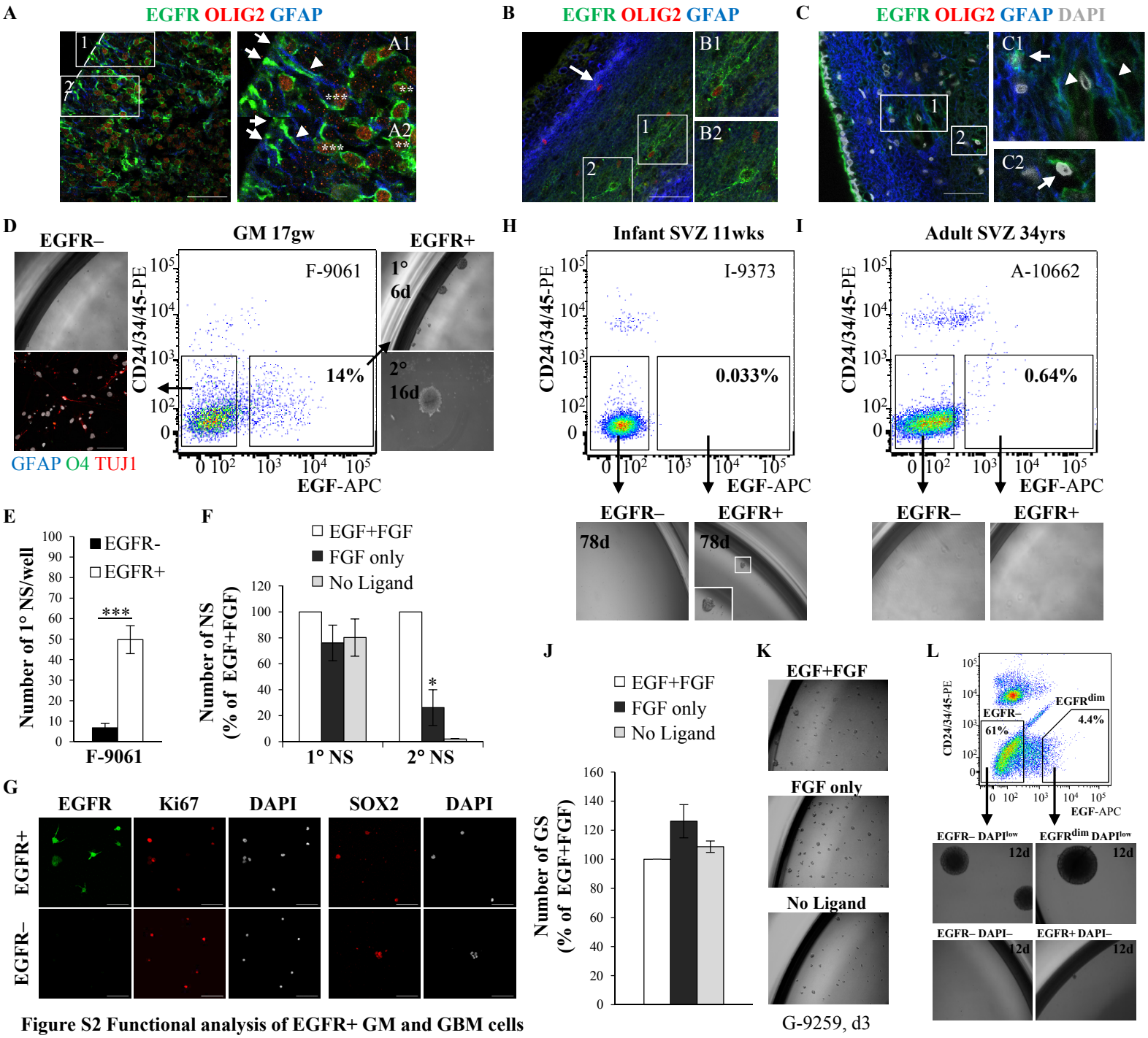


Figure S2 Functional analysis of EGFR+ GM and GBM cells

(A) Immunofluorescence (IF) of GM at 17 weeks of gestation reveals EGFR+ cells with radial morphology (arrows) next to the developing ependymal layer (dash lines) in the lateral ventricle, some of which co-localize with GFAP (arrowhead). Many EGFR+OLIG2+ (**) and scattered EGFR+GFAP+OLIG2+ (***) cells are seen in the deeper GM. **(B)** In the infant SVZ, the expression of EGFR diminishes, but some EGFR+GFAP+, EGFR+OLIG2+ (B1), as well as exclusively EGFR+ (B2) or OLIG2+ (arrow) cells are found. **(C)** Adult SVZ shows occasional EGFR+GFAP+ co-localization in the astrocyte ribbon (C1, arrow) and in the hypocellular layer (arrowheads); EGFR+GFAP– cells are also seen (C2, arrow). **(D-E)** FACS-isolated EGFR+ GM cells form primary (1°) and secondary (2°) neurospheres (NS). Rare 1° NS derived from EGFR– cells were noted in one case of early gestational age, F-9061, in much lower numbers than those derived from EGFR+ cells ($p=0.0009$) ($n=3$); these failed to form 2° NS or to differentiate effectively, undergoing apoptosis with rare TUJ1+ cells and no appreciable GFAP and O4 positivity. **(F)** While 1° NS formation from GM EGFR+ cells is not significantly affected by the withdrawal of EGF or FGF from the cell culture media, 2°NS are reduced ($p=0.03$) ($n=3$). **(G)** IF of acutely isolated EGFR+/– GM cells two hours after FACS shows predominant distribution of EGFR in the positive fraction, and comparable expression of Ki67 and SOX2 (F-9449, 21gw). **(H)** Postnatally, rare NS can be seen in EGFR+ cells up to 11 weeks of infancy, smaller in size, and unable to form 2° NS. **(I)** Occasional EGFR+ cells can be acutely isolated from adult SVZ, but they are not able to form NS, regardless of the age of the specimen ($n=10$). **(J-K)** In GBM, sphere-forming ability is retained in serially passaged GS after removal of EGF ligand ("FGF only") and after removal of all ligands ("No Ligand") from the culture medium ($n=5$). **(L)** GBM tumor, which lacked EGFR amplification, chromosome 7 polysomy, and EGFR expression, did not have a well-defined EGFR+DAPI^{low} population (EGFR^{dim}) and showed equivalent GS formations in the EGFR– and EGFR^{dim} fractions. gw=gestational weeks; wks=weeks, yrs=years; d=days. Percentages of populations shown in FACS plots are from total live events, as determined by DAPI staining. Scale bar=50μM.

Figure S3 (Relates to Figure 2)

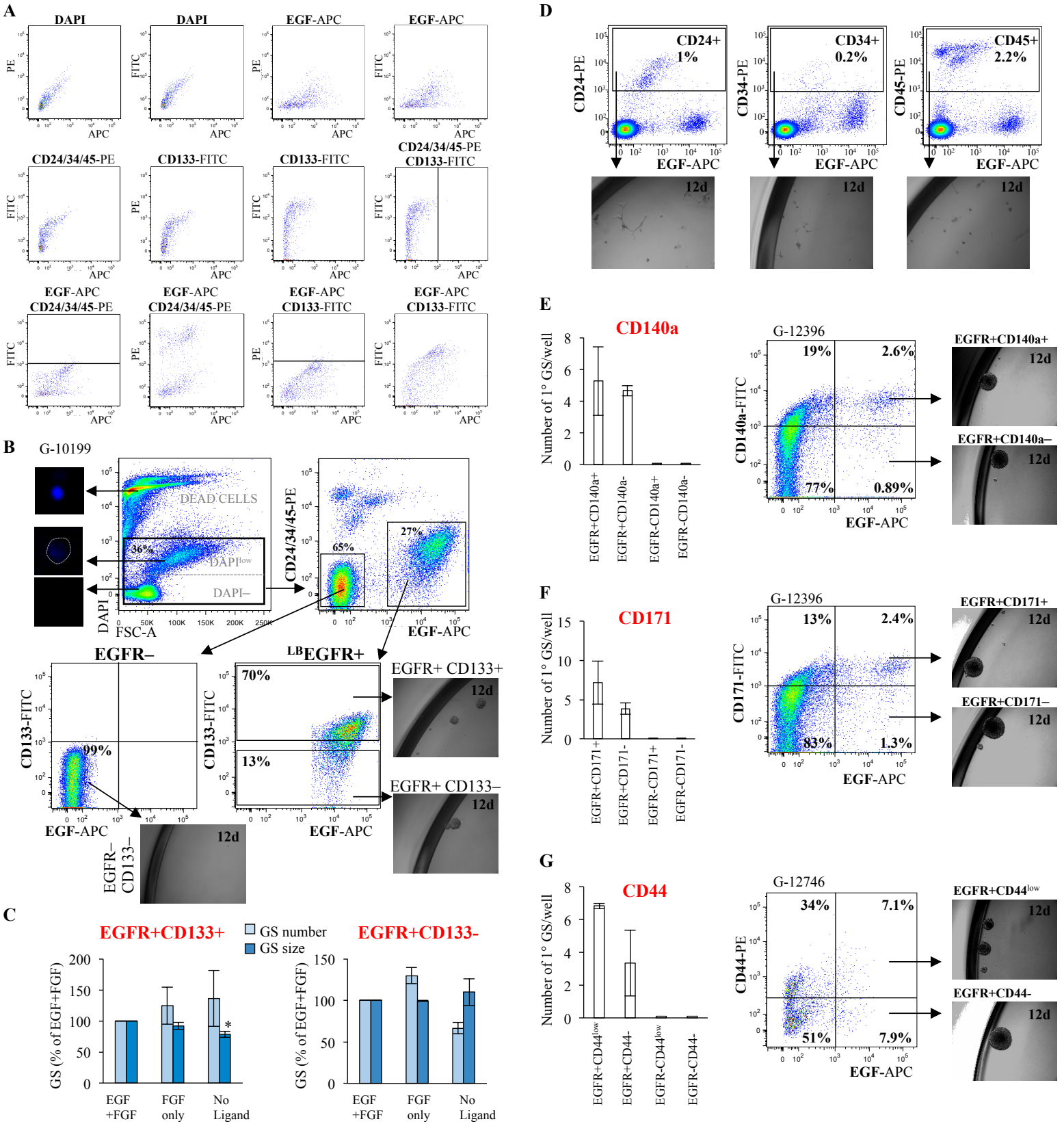
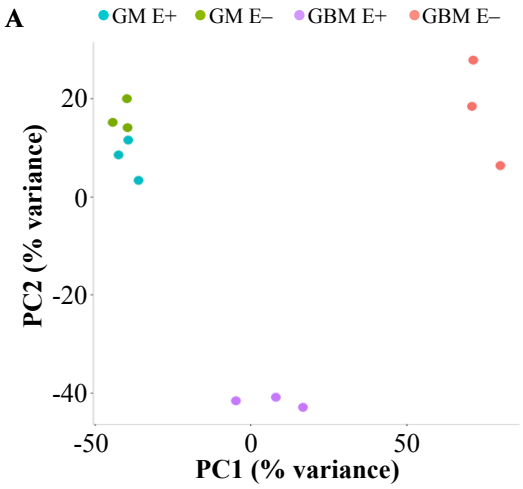
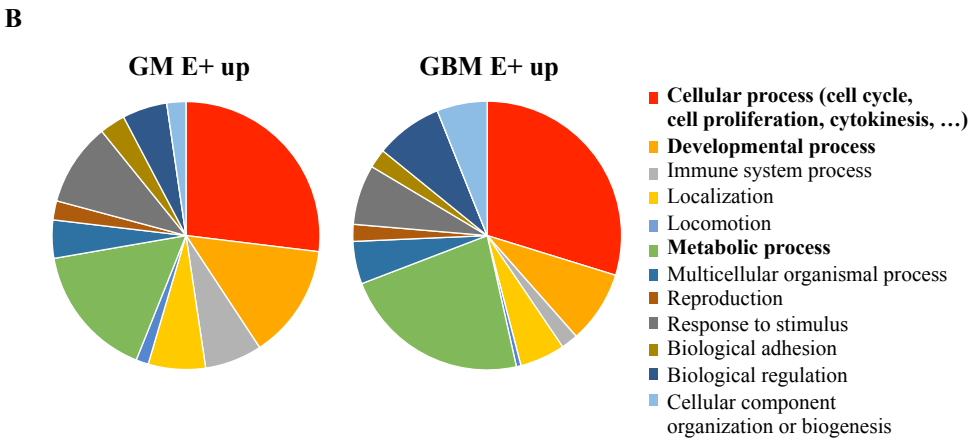


Figure S3 Comparison of EGF ligand and antibody-based isolation of sphere-forming populations in GBM

(A) Representative single and FMO color controls in combined CD133/EGF FACS. **(B)** In GBM, FACS isolation using EGF demonstrates a unique live cell population with forward scatter shift, which is able to incorporate some DAPI (termed DAPI^{low}; dashes outline nucleus). Combined CD133 and EGF FACS analysis demonstrates three defined populations: EGFR+CD133+, EGFR+CD133–, and EGFR–CD133–. Both EGFR+CD133+ and EGFR+CD133– cells display 1° GS formations *in vitro* (day 12, EGF+bFGF) (see also Figure 2I-K). **(C)** EGFR+CD133+ and EGFR+CD133– sorted populations show similar number of GS when grown with or without EGF supplementation, and even with no ligand in the culture medium. EGFR+CD133+ GS formed in no ligand conditions have smaller diameter compared to EGF+FGF (low-passaged, patient-derived GS, n=3 tumors; 6 days). **(D)** CD24+, CD34+, and CD45+ single populations do not form GS (3 tumors, 12 days). **(E-G)** Functional analysis of populations derived from combined FACS using EGF and either CD140a (PDGFRalpha) (n=3) (E), CD171 (n=3) (F), or CD44 (n=2) (G) shows that in each case the ^{LB}EGFR+ populations encompass all sphere-forming cells, including those that lack CD140a, CD171, and CD44 (12 days).

Figure S4 (Relates to Figure 4)





C

GM EGFR+ *vs* EGFR– up (GM E+ up)

TermID	Term	logP
GO:0045880	Positive regulation of smoothened signaling p.	-10.4
GO:0048568	Embryonic organ development	-9.5
GO:0007346	Regulation of mitotic cell cycle	-5.8
GO:0008284	Positive regulation of cell proliferation	-5.8
GO:0005006	EGF-activated receptor activity	-4.9
GO:2000648	Positive regulation of stem cell proliferation	-4.5
GO:0090068	Positive regulation of cell cycle process	-3.9
WP47	Hedgehog signaling pathway	-6.2
WP61	Delta-Notch signaling pathway	-3.5

GM EGFR– *vs* EGFR+ up (GM E+ down)

TermID	Term	logP
GO:0007399	Nervous system development	-17
GO:0007275	Multicellular organismal development	-15.6
GO:0001764	Neuron migration	-14.3
GO:0048856	Anatomical structure development	-14.1
GO:0050808	Synapse organization	-14.0
GO:0048468	Cell development	-12.9
GO:0030182	Neuron differentiation	-12.5
WP2267	Synaptic vesicle pathway	-8.8

D

GBM EGFR+ *vs* EGFR– up (GBM E+ up)

TermID	Term	logP
GO:0000278	Mitotic cell cycle	-67.6
GO:0007059	Chromosome segregation	-32.0
GO:0051726	Regulation of cell cycle	-25.4
GO:0044843	Cell cycle G1/S phase transition	-18.2
GO:0060828	Regulation of canonical Wnt pathway	-8.4
GO:2000648	Positive regulation of stem cell proliferation	-7.9
GO:0005006	EGF-activated receptor activity	-5.5
GO:0031572	G2 DNA damage checkpoint	-4.9
WP47	Hedgehog signaling pathway	-10.7

GBM EGFR– vs EGFR+ up (GBM E+ down)

TermID	Term	logP
GO:0007154	Cell communication	-30.2
GO:0002376	Immune system process	-28.5
GO:0032879	Regulation of localization	-26.1
GO:0030334	Regulation of cell migration	-19.1
GO:0021782	Glial cell development	-8.9
GO:0045597	Positive regulation of cell differentiation	-7.1
WP306	Focal Adhesion	-9.4
WP453	Inflammatory Response Pathway	-6.0

Differential gene expression and gene set enrichment analyses in GM and GBM

(A) Principal component analysis and hierarchical clustering (see Figure 4A) of all RNAseq samples confirms reproducibility, showing a tight clustering of biological replicates. While developing GM EGFR+ and EGFR– transcriptomes are similar to one another, GBM EGFR+ transcriptomes are more similar to GMs than to their EGFR– neoplastic counterpart. **(B-D)** Gene enrichment analyses are performed on the sets of most highly differentially regulated genes (FDR adjusted p ≤0.05, $|\text{Log}_2(\text{fold change})| \geq 2$) in GM EGFR+ *vs*. EGFR– and in GBM EGFR+ *vs*. EGFR–. **(B)** Analysis of differentially upregulated gene sets in EGFR+ GM (GM E+ up) and in EGFR+ GBM (GBM E+ up) using the PANTHER tool depicts prominent representation of top biological processes related to cellular process (cell cycle, cell proliferation, cytokinesis, chromosome segregation, cellular component movement, and cell communication), metabolic process, and developmental process in both data sets. **(C-D)** Analysis for gene set enrichment using the HOMER tool (p ≤0.05, $|\text{Log}_2(\text{fold change})| \geq 2$) in both differentially upregulated (E+ up / E– down) and downregulated (E+ down / E– up) GM and GBM gene sets. Representative top biological and molecular processes (GO) and Wikipathways (WP) are depicted, showing enrichment for cell division, cell cycle regulation, stem cell proliferation, and EGFR signaling in EGFR+ populations of both GM (C) and GBM (D) (top), and more variable processes related to cell differentiation, migration, and inflammation in EGFR– populations (bottom).

Supplemental Table 1 (Relates to Figures 1-4)

Sample information

Sample ID	Sample Type, Age, Sex (Ordered by age)	% Live Cells	% EGFR+ cells	PMT	*In vitro* NS assays	Acute Immunofluorescence	EGF / FGF exclusion	ELDA (stem cell frequency)	RNA-seq
Fetal (F)									
F-9061	GM, 17gw, U	45.3	14	12 hrs	✓	✓			
F-10302	GM, 17gw, U	45	1.9	24 hrs	✓			✓	
F-12437	GM, 19gw, F	77	22	72 hrs	✓				✓ (GM2)
F-13759	GM, 18gw, M	52	9	3 hrs	✓		✓	✓	
F-13970	GM, 19gw, U	66	14.5	6 hrs	✓		✓	✓	
F-10929	GM, 20gw, F	96	1	6 hrs	✓		✓		
F-13969	GM, 21gw, U	77	9.8	6 hrs	✓		✓	✓	
F-9449	GM, 21gw, U	8.9	6.1	48 hrs	✓	✓	✓		✓ (GM1)
F-9062	GM, 21gw, F	38.8	2.1	12 hrs	✓	✓			
F-12266	GM, 22gw, F	73	25	<12 hrs	✓				✓ (GM3)
F-11045	GM, 24gw, F	30.8	8.5	<7 hrs	✓		✓	✓	
F-9491	GM, 25gw, M	11.9	1.5	24 hrs	✓				
Infant (I)									
I-8561	SVZ, 3days, M	66.9	2.72	4 hrs	✓				
I-9063	SVZ, 1wk, F	77	0.3	36 hrs	✓				
I-9373	SVZ, 11wks, F	92.4	0.03	24 hrs	✓				
Adult (A)									
A-10661	SVZ, 9yrs, F	78.9	0.1	12 hrs	✓				
A-10662	SVZ, 34yrs, M	90.8	0.64	12 hrs	✓				
A-10663	SVZ, 44yrs, F	78.1	<0.1	14 hrs	✓				
A-8503	SVZ, 57yrs, M	61	<0.1 (600 c)	21 hrs	✓				
A-10664	SVZ, 71yrs, M	72	<0.1	24 hrs	✓				
A-10665	SVZ, 75yrs, F	64.8	0.1	12 hrs	✓				
A-8155	SVZ, 80yrs, M	93	0.14	23 hrs	✓				
A-8751	SVZ, 80yrs, M	86	0.39	7 hrs	✓				
A-10659	SVZ, 84yrs, M	89.9	0.3	24 hrs	✓				
A-10658	SVZ, 94yrs, F	63	<0.1 (600 c)	14 hrs	✓				

Supplemental Table 1 (Continuation)

Sample ID	Sample Type, Age, Sex	% Live Cells	*EGFR* amp / Chr. 7 status	*EGFRvIII* detection	IDH1 status	EGFR expression (IHC)	*In vitro* GS assays	*In vivo* assays	FACS_Other surface markers	EGF / FGF exclusion	ELDA (stem cell frequency)	RNA-seq
G-9217	GBM, 76yrs, M	93	Amp+ diffuse	Y	WT	Y	✓					
G-9259	GBM, 67yrs, F	90	Amp+ focal	Y	WT	Y	✓	✓		✓	✓	
G-9366	GBM, 88yrs, M	83	Amp+ focal	Y	WT	Y	✓					
G-10006	GBM, 76yrs, M	76	Amp+ focal	Y	WT	Y	✓					✓ (GBM2)
G-10199	GBM, 55yrs, M	36	Amp+ diffuse	Y	WT	Y	✓		✓			
G-10719	GBM, 55yrs, F	84	Amp+ diffuse	Y	WT	Y	✓					
G-11039	GBM, 69yrs, M	90	Amp+ multifocal / Chr.7 Polysomy	Y	WT	Y	✓			✓	✓	
G-11074	GBM, 77yrs, F	75	Amp+ diffuse	Y	WT	Y	✓	✓				
G-11075	GBM, 62yrs, F	96	Chr.7 Polysomy	N	WT	Y	✓	✓	✓	✓	✓	
G-11314	GBM, 58yrs, M	76	Amp+ multifocal	Y	WT	Y	✓	✓	✓	✓		
G-11592	GBM, 39yrs, F	67	Chr.7 Polysomy	N	Mutant	Y	✓				✓	
G-11846	GBM, 49yrs, M	99	Amp+ diffuse	Y	WT	Y	✓					✓ (GBM1)
G-12022	GBM, 81yrs, M	86	Chr.7 Polysomy	N	WT	Y	✓					
G-12048	GBM, 64yrs, F	98	Amp+ diffuse	Y	WT	Y	✓					✓ (GBM3)
G-12396	GBM, 72yrs, M	69	Amp+ diffuse	N	WT	Y	✓		✓			
G-12746	GBM, 60yrs, M	60	None	N	WT	Y	✓		✓	✓	✓	
G-13063	GBM, 69yrs, F	50	None	N	WT	Y			✓	✓	✓	
G-13181	GBM, 73yrs, F	76	None	N	WT	Y	✓’			✓	✓	
G-13306	GBM, 71yrs, M	71	Chr.7 Polysomy	N	WT	Y	✓’			✓	✓	
G-13629	GBM, 69yrs, M	80	Chr.7 Polysomy	N	WT	Y	✓’		✓			
G-10416	GBM, 78yrs, F	81	None	N	WT	N	✓’					
G-11702	GBM, 67yrs, F	80	None	N	WT	N	✓’					

De-identified information for postmortem fetal (F), infant (I), and adult (A) samples, and for glioblastomas (G), used in the various experiments, including code, age, gender, % live cells isolated by fluorescence-activated cell sorting (FACS) assessed by DAPI, % EGFR+ live cells and postmortem time (PMT) (for non-neoplastic only), and EGFR gene amplification (CISH), EGFRvIII mutation (RT-qPCR), IDH1-R132H mutation (IHC), and EGFR expression status (IHC, for tumors only). FACS analysis was performed in all samples, but due to limited surgical and viable postmortem tissue availability, other experiments had more restricted sample allocation. (M = male; F = female; U = unknown gender; gw = gestational weeks; yrs = years; wks = weeks; GM = germinal matrix; SVZ = subventricular zone; c = cells; WT = wild type; Amp = amplification; Chr = chromosome; CISH = Chromogenic In Situ Hybridization; IHC = Immunohistochemistry; ’not included in GS quantification; Y = Yes, N = No).

Supplemental Table 2 (Relates to Figures 1-2)

Analysis of stem cell frequency in GM and GBM

SAMPLE ID	TISSUE TYPE	FACS EGFR+	STEM CELL FREQUENCY (%)		
			EGF+FGF (E+F)	FGF only (F)	No ligands (NL)
F-10302	GM	1° GS	4.0		
F-11045	GM	1° GS	16		
F-13759	GM	1° GS	4.5	4.0	3.0
F-13969	GM	1° GS	22.6	20.3	10.4
F-13970	GM	1° GS	37.7	32.7	25.9
G-11039	GBM	1° GS	3.6		
G-11075	GBM	1° GS	5.5		
G-11347	GBM	1° GS	5.7		
G-9259	GBM	lp GS	30	46	32.7
G-12746	GBM	lp GS	1.9	1.9	
G-13063	GBM	lp GS	22.9	13.7	9.3
G-13181	GBM	lp GS	9	4.3	2.0
G-13306	GBM	lp GS	6.8	7.8	8.2
			Un-paired T-test (p-value) GBM vs. GM		
			E+F	**F**	**NL**
			0.36	**0.74**	**0.99**

TTEST		
Un-paired T-test (p-value) GM		
E+F vs. F	**E+F vs. NL**	**F vs. NL**
0.85	**0.70**	**0.61**
Un-paired T-test (p-value) GBM		
E+F vs. F	**E+F vs. NL**	**F vs. NL**
0.61	**0.74**	**0.88**

Extreme limited dilution assay (ELDA) analysis of primary (1°) and low-passaged (lp) spheres reveals comparable stem cell frequency in EGFR+DAPI– germinal matrix (GM) and ^{LB}EGFR+DAPI^{low} GBM cells, with the ability of both to generate clonal NS/GS from single cells. Un-paired t-test analysis shows no significant difference in stem cell frequency between GBM and GM EGFR+ populations, including under different ligand supplementation conditions: EGF+FGF (E+F), FGF only (FGF), or in absence of both ligands (No ligands, NL).

Supplemental Table 3 (Relates to Figure 4)

Gene enrichment analysis of differentially expressed EGFR+ GM and GBM genes

Whole transcriptome analysis via RNAseq was performed on 3 GBM ^{LB}EGFR+/EGFR– and 3 GM EGFR+/EGFR– acutely sorted populations. After mapping data to the hg38 genome, differential expression analysis was performed between ^{LB}EGFR+ and EGFR– GBM samples, and between EGFR+ and EGFR– GM samples (see also Figure S4 and supplemental experimental procedures). The data was analyzed for gene set enrichment using the HOMER tool (Heinz et al., 2010) (FDR adjusted $p \leq 0.05$, $Log_2(2)$), including analysis of biological processes, molecular functions, wikipathways, and the COSMIC mutation (Forbes et al., 2015) database gene sets. Also included is the list of 50 differentially upregulated genes shared between GM and GBM EGFR+ populations (see also Figure 4C).

Supplemental Experimental Procedures

Sample collection

All specimen collection was performed in accordance with the policies and regulation of the Icahn School of Medicine at Mount Sinai and its institutional review board. For non-neoplastic studies (IRB exemption HS#14-

01007), GM or SVZ were carefully dissected from fresh de-identified pediatric or adult postmortem brains, with comparable representation of either sex, post mortem time < 24 h for adult and < 48 h for pediatric, and without gross brain pathology. One portion of the tissue was placed in fresh PIPES solution for subsequent FACS analysis and cell culture; another portion was frozen on dry ice for downstream molecular analyses; and the remainder was fixed in 4% paraformaldehyde (PFA) for 24-72hrs. For glioma studies (IRB AAAJ9652-Y1M00, HS14#01007), de-identified GBM samples were obtained from the tissue biorepository, within 0.5-3 hours of neurosurgical resection. Only tumor specimens with histological features of GBM on frozen section, evaluated by another neuropathologist, were used for downstream molecular analyses in this study. Refer to Table S1 for additional sample information.

Fluorescence-activated cell sorting

Fresh tissue (~500mg of dissected GM/SVZ and 50-500mg GBM) was first finely minced mechanically, then enzymatically dissociated with papain (Worthington, LS003119, 0.3mg/ml, 45 units/sample) at 37°C for 13 min in rotation, and finally triturated in DMEM media with DNase (Worthington, LS002139, 0.1mg/ml) and ovomucoid protease inhibitor (Sigma Aldrich, 0.7mg/ml). Extracellular debris and myelin were removed by passage through 22% Percoll; red blood cells were lysed for 10min at RT (Ebioscience, 00-4300-54); then cells were washed once with HBSS+1% BSA+0.1% glucose. Live cell number was calculated via trypan blue incorporation using an automated cell counter (Countess, Invitrogen). Cells were incubated with Alexa Fluor® 647 (APC) EGF complex (Life Technologies, E-35351) (1μg/10^6 live cells) for 30 min on ice. In some experiments, antibodies against CD133-FITC, CD171-FITC, CD140a-FITC, and CD44-PE were used in combination with EGF. Exclusion of ependyma/neuroblasts, endothelium, and inflammatory cells were accomplished with CD24-PE, CD34-PE, CD45-PE, respectively (Codega et al., 2014). DAPI was used to discern between live and dead cells. No color, single color, and fluorescence minus one (FMO) color controls were done to establish negative/positive cut-off values for the populations. Cells were sorted on a Becton Dickinson FACSAria™ III sorter. See antibody summary below.

Self-renewal, differentiation, and ELDA protocols

For primary NS/GS formation analysis, immediately after sorting, cells were seeded on 96-well low-adherence plates at a density of 10c/μl, in triplicates, in NS media (1X N2, 1X B27, 20μM glutamine, 1X Insulin/Transferrin/Selenium, 15mM HEPES, 0.6% glucose, 1X Antibiotic/Antimycotic, in DMEM/F12 media) supplemented with EGF (20ng/ml) and bFGF (20ng/ml) (FGF). In some cases, EGF and/or bFGF was omitted. Cells were maintained at 37°C and 5% CO_2 for up to 4 weeks, changing 2/3 of media after day 6^{th} and every 3-4 days thereafter. For acute staining analysis, 2000 EGFR– and EGFR+ cells were seeded immediately after FACS isolation on PDL-laminin-coated coverslips, incubated at 37°C for 2hr in NS media without growth factors, then gently centrifuged for 1min at 1000rpm, fixed for 10min with 4% PFA, washed with 1X PBS, and stored at 4°C for IF. The rest of the sorted cells were pelleted at 500g for 5min at 4°C and immediately frozen at -80°C for subsequent DNA/RNA analyses. Images of NS/GS formation were captured with a light inverted microscope (Motic AE31) at Day6, Day12, and Day25 after seeding. Pictures covering the entire surface of the wells were taken at 10X and were used for subsequent counting. NS/GS diameter size was measured in all clones of all wells, 3 replicates/sample, and analyzed using ImageJ 1.43u software. For secondary/tertiary sphere passaging, single NS/GS reaching minimum size of 40μM were picked, pelleted at 100g for 1min, dissociated in 1ml of Accumax solution (Innovative Cell Technologies) for 10mins at RT with trituration, and incubated in NS media with growth factors (see above) at a clonal density of 1c/μl ((Pastrana et al., 2011). For differentiation, NS/GS were picked up and seeded on PDL-laminin coated glass coverslips or on 16-well chamber-slides (Thermofisher, 12-565) in differentiating NS media without B27, EGF, and bFGF, incubated for 14 days, and then fixed in 4% PFA for 10mins at RT. For some cases of differentiation O4 analysis, T3 hormone (1:1000) was added every 3 days. In all the experiments cells were seeded in triplicates.

Stem cell (self-renewal) frequency was calculated by using Extreme Dilution Limited Assay (ELDA) (Hu and Smyth, 2009). Briefly, FACS-sorted EGFR+ cells were plated in decreasing numbers from 100 cells/well to 1 cell/well in 200μl in similar conditions as for NS/GS formation analysis. Cultures were maintained until day 25 as described above. Then, the number of wells containing spheres for each cell plating density (number of positive cultures) was recorded, and plotted using online ELDA analysis program (http://bioinf.wehi.edu.au/software/elda).

RNAseq data analysis

Illumina sequencer output FASTQ files for each sample were subjected to a quality control assessment using the FASTQC package. RNAseq reads were aligned to the human genome (GENCODE GRCh38) using STAR with default settings (Dobin et al., 2013). Gene counts were obtained using the featureCount utility of the subread

package (Liao et al., 2014). Differential expression analysis on the RNAseq data was performed using the DESeq2 R package (Love et al., 2014), which models the data with a negative binomial distribution and uses Empirical Bayes shrinkage for dispersion and fold-change estimation. Gene set enrichment analysis was done utilizing the HOMER tool (Heinz et al., 2010). The gene background set for the enrichment analysis was defined as the set of genes passing the independent filtering low expression threshold by DESeq2. All RNAseq tests were FDR adjusted for multiple testing correction.

Gene expression analysis

For bulk tumor, total RNA was isolated using RNA easy mini kit (Qiagen, 71404) with DNase treatment (Qiagen, 79254). For sorted cells, total RNA was isolated via Trizol /chloroform /isopropanol extraction method following manufacturer's instructions (Life Technologies), followed by RNA cleanup-and-concentrator kit with DNase treatment (Zymo Research, R1013). cDNA was generated using the High-Capacity RNA-to-cDNA Kit (Life Technologies, 4387406). qPCR reactions were run in duplicates using the SYBR-Green based system (Quanta Biosciences). Each sample Ct value was normalized against the expression of the 40S Ribosomal Protein S11 (*RPS11*) as a housekeeping gene (HKG). Data was expressed either as the exponential fraction of HKG expression levels $2^{-(CtSample - CtHKG)}$ ($2^{-\Delta Ct}$) or as a fold change relative to the EGFR– fractions ($2^{-\Delta\Delta Ct}$). All primers used for the reactions (see table below) were designed to span exon-exon junctions in order to minimize genomic contamination except for those ones with no introns in their DNA sequence. Melting curves were analyzed to ensure specificity of the primers for a single product, and product size was confirmed by 2% agarose gel electrophoresis.

Table summary of primers used for RT-qPCR analysis

GENE	**Forward (5'-3')**	**Reverse (5'-3')**
ASCL1	GGACTTTGGAAGCAGGGTG	CGCCACTGACAAGAAAGCA
CENPE	TTCAGCCTGATAGGATGGCG	TTGGGCAGTTTCTCCAAGTGA
DBF4	GGACATTAAGGATCTGGGAGGG	TTGTGCAAATTTAGCTTCCTTCT
EGFR exons 17-18	GATCGGCCTCTTCATGCGA	GTAAGAGGCTCCACAAGCTCC
EGFR exons 3-4	TCCAGGAGGTGGCTGGTTAT	ACGGCGCCATGCAGGAT
EGFRvIII exons 1/8	GGCTCTGGAGGAAAAGAAAGGTAA	CAAGGCCCTTCGCACTTCTTA
KNSTRN	GAGCTGCTGGAGAAGTTTCG	CCAGGGTCTCACTGCCTAAAG
KPNA2	AGCTCCAAGCTACTCAAGCTG	CCAGCCCGGATTATGTTGTCT
HKG	**Forward (5'-3')**	**Reverse (5'-3')**
RPS11	ATGTCCAGCCTCAGAACTTC	GCCGAGACTATCTGCACTAC

Orthotopic transplantation

All procedures performed in studies involving animals were in accordance with the ethical standards of the Icahn School of Medicine at Mount Sinai (ISMMS). FACS-sorted cell populations from GBM samples (^{LB}EGFR+DAPI^{low}, ^{LB}EGFR+DAPI–, EGFR–DAPI^{low}, or EGFR–DAPI–) without prior culturing were injected into the striatum of 2-month-old male mice with B & T cell ICR-Severe Combined Immunodeficiency (SCID) (IcrTac:ICR-Prkdc^{SCID} strain, Taconic). Microinjections were performed using a stereotactic apparatus (Stoelting) and microsyringe (Hamilton). For each mouse, $1x10^5$ cells were injected in a volume of 2 µl (Opti-MEM, Thermofisher) at the following stereotactic coordinates: 2mm right lateral to Bregma and 3mm deep. Note: for ^{LB}EGFR+DAPI– mice only, $0.5x10^5$ cells were available for injection. After 2 months, mice were sacrificed and tumor initiation was analyzed by histological examination. At later time points, additional mice were sacrificed if they became clinically symptomatic.

Immunofluorescence/Immunohistochemistry

Specimens were fixed in 4% paraformaldehyde (PFA) in phosphate buffered saline (1X PBS) for 24hr (SVZ) to 72hr (GM) at 4°C, or in 10% formalin/1X PBS for 1-15 days (GBM and some GM and SVZ cases). Floating

vibratome sections (40-60µm, stored in 1X PBS /0.01% NaN3) and formalin-fixed paraffin-embedded (FFPE) sections (4µm) were used. Mice were anesthetized with pentobarbital (i.p. injection of 650 mg/kg) and transcardially perfused with 1X PBS first and then with 4% PFA/1X PBS. Brains were removed and post-fixed in 4% PFA/1X PBS overnight, cryoprotected in sucrose gradient, and then embedded in O.C.T. Brains were serially cut with a cryostat (25µM). For immunofluorescence (IF) experiments, FFPE and vibrotome sections underwent antigen retrieval in boiling citrate buffer pH6 (Vector labs), blocking (10% normal donkey serum (NDS)/0.5% triton-X (TX), 1hr at RT), primary antibody incubation (1% NDS/0.25% TX overnight at 4°C), and species-appropriate fluorochrome-conjugated secondary antibody incubation (1% NDS/0.25% TX for 4hr at RT). Nuclear counterstain was with DAPI. For EGFR IF of adult SVZ, the primary antibody signal was amplified with TSA amplification system (PerkinElmer Labs, NEL700A001KT). Briefly, enzymatic digestion of the tissue was performed with protease enzyme (10µg/ml) at 37°C for 10min followed by 10min at RT. After blocking in TNB buffer for 30min at RT, sections were incubated with primary monoclonal mouse EGFR antibody (Invitrogen) overnight at 4°C. Secondary anti-mouse HRP-conjugated antibody was incubated for 4h at RT and one cycle of amplification was performed with tyramide and FITC-conjugated Streptavidin following manufacturer's instructions. Pertinent negative controls were simultaneously performed to ensure specific signal. FFPE sections were processed similarly to vibratome sections, with the exception of 1hr deparaffinization with subsequent rehydration in decreasing gradient of ethanol, antigen retrieval for 20min, blocking in 5% NDS/0.5% TX for 30min. IF on cells, mouse brain tissue, or differentiated NS/GS were performed similarly to human vibrotome tissues, except for omitting antigen retrieval and signal amplification. Specifically, O4 IF was performed on live unfixed cells, incubated with primary and secondary antibodies for 40min at 4°C and later fixed with 4% PFA. Zeiss LSM710 confocal microscope was used for visualization. Acute IF staining was performed similarly to NS/GS, except with 0.5hr blocking and 2hrs secondary antibody incubation. All immunohistochemical and EGFR chromogenic in situ hybridization (CISH) experiments were performed on automated Ventana XT/ULTRA machines following standard protocols (Ventana). Immunoreactivity was detected by means of the ultraview universal DAB detection kit (Ventana, 760-500) or ultraview SISH DNP detection kit (Ventana, 760-098).

Table summary of antibodies used for immunofluorescence and FACS

ANTIBODY	COMPANY	CATALOG NUMBER	DILUTION
Mouse anti-EGFR ED	INVITROGEN	280005	1:50
Rabbit anti-EGFR ID	MILLIPORE	06-847	1:50
Rat anti-GFAP	LIFE TECHNOLOGIES	13-0300	1:1000
Mouse anti-human GFAP	CELLARTIS	Y40420	1:500
Mouse anti-HNA	MILLIPORE	MAB1281	1:400
Rabbit anti-Ki67	ABCAM	Ab15580	1:250
Mouse anti-Nestin	MILLIPORE	MAB5326	1:250
Mouse anti-O4	MILLIPORE	MAB345	1:200
Rabbit anti-OLIG2	MILLIPORE	AB9610	1:250
Rabbit anti-SOX2	STEMCELL TECHNOLOGIES	60055	1:100
Mouse anti-TUJ1	BIOLEGEND	MMS-435P	1:500
Rabbit anti-TUJ1	COVANCE	PRB-435P	1:1000
EGF ligand-AF647 conjugated	LIFE TECHNOLOGIES	E-35351	1:200
Anti-CD24 PE-conjugated	BD BIOSCIENCES	560991	1:10
Anti-CD34 PE-conjugated	BD BIOSCIENCES	550619	1:10
Anti-CD45 PE-conjugated	BD BIOSCIENCES	555483	1:10
Anti-CD24 FITC-conjugated	BD BIOSCIENCES	560992	1:10
Anti-CD34 FITC-conjugated	BD BIOSCIENCES	560942	1:10
Anti-CD45 FITC-conjugated	BD BIOSCIENCES	560976	1:10
Anti-CD133 Biotin-conjugated	MILTENYI BIOTEC	130-090-664	1:10
Anti-CD44 PE-conjugated	BD BIOSCIENCES	550989	1:10
Anti-CD140a Biotin-conjugated	BIOLEGEND	323503	2µg/10^6 cells
Anti-CD171 Biotin-conjugated	MILTENYI BIOTEC	130-100-703	1:11
Donkey anti-Mouse AF488 conjugated	JACKSON IMMUNORESEARCH	715-545-151	1:250

Donkey anti-Rabbit Cy3 conjugated	JACKSON IMMUNORESEARCH	711-165-152	1:250
Donkey anti-Rat AF647 conjugated	JACKSON IMMUNORESEARCH	712-605-153	1:500
Donkey anti-Mouse Cy3 conjugated	JACKSON IMMUNORESEARCH	715-165-151	1:500
Donkey anti-Rabbit AF488 conjugated	JACKSON IMMUNORESEARCH	711545152	1:500
IgM anti-Mouse-AF488	JACKSON IMMUNORESEARCH	115-545-075	1:500
IgM anti-Mouse-AF568	MOLECULAR PROBES	A-21043	1:500
DyLight 488 Streptavidin	VECTOR LABS	SA-5488-1	1:250
EGFR-CISH probe	VENTANA	760-1216	-
Chromosome 7-CISH probe	VENTANA	760-1219	-

Statistical analysis

Two-tailed unpaired Students' t-test was used to calculate significance (*$p \leq 0.05$, **$p \leq 0.01$, ***$p < 0.001$), which was corrected for non-homogeneous variances. For non-parametric analysis, we used the Mann-Whitney U test. Bar graph data is represented as mean ± SEM except for bar graphs in Figure 4E where plots show the median (black horizontal line), 25th and 75th percentiles (boxes) and range (whiskers) of the data. All pooled data in the main and supplementary figures comes from independent experiments, except for Figure S2E (3 wells from one sample).

References

Codega, P., Silva-Vargas, V., Paul, A., Maldonado-Soto, A.R., Deleo, A.M., Pastrana, E., and Doetsch, F. (2014). Prospective identification and purification of quiescent adult neural stem cells from their in vivo niche. Neuron *82*, 545-559.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Forbes, S.A., Beare, D., Gunasekaran, P., Leung, K., Bindal, N., Boutselakis, H., Ding, M., Bamford, S., Cole, C., Ward, S.*, et al.* (2015). COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic acids research *43*, D805-811.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell *38*, 576-589.

Hu, Y., and Smyth, G.K. (2009). ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. Journal of immunological methods *347*, 70-78.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923-930.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology *15*, 550.

Pastrana, E., Silva-Vargas, V., and Doetsch, F. (2011). Eyes wide open: a critical review of sphere-formation as an assay for stem cells. Cell stem cell *8*, 486-498.