Evidence for sequenced molecular evolution of *IDH1* mutant glioblastoma from a distinct cell of origin

Lai, et al

PATIENTS and METHODS

Cases and Specimens

IDH1 mutation status was determined from 604 adult and 14 pediatric treatment-naïve *de novo* GBM patient samples aggregated from 6 sample cohorts (cohorts A-F). In addition, two cohorts of newly-diagnosed lower grade glioma samples (cohorts G & H) and one cohort including both newly diagnosed and recurrent gliomas of grades II-IV (cohort I) were genotyped for $IDH1^{R132}$ and included in a limited number of analyses in the current study. Ethical review committee approval was obtained from each institutional source. All cases and samples included in this investigation are summarized in Table 1 and listed in Tables S1, S2, and S10. Identifiers include: Case ID (unique for each patient), Sample ID (unique for each surgical resection; and, in some instances, may also identify the tissue block), expression, array comparative genomic hybridization (aCGH), or methylation microarray ID (unique to each hybridization). Cohorts C-G have been previously described[1-4], with the exceptions that: (1) survival times were updated where available for cases in cohort C, and (2) that cohorts E, F & G have been expanded. Cohort C was originally created to study prognostic signatures[1], and this series is biased to over-represent long term survivors. With this exception for cohort C, and the confinement of cohort E to pediatric cases, all cases were selected solely on the basis of diagnosis and were unbiased for any other features including survival times or age.

All cases in cohorts A-F were GBMs (WHO Grade IV) that presented with no prior clinical history of brain tumor or treatment. All cases in cohorts B-F were verified to meet WHO GBM criteria[5] by a board certified neuropathologist who examined hematoxylin and eosin (H&E)-stained sections of formalin-fixed paraffin-embedded (FFPE) material. For cohort A, all *IDH1* mutant cases and the majority of *IDH1* WT cases were subjected to similar verification of WHO GBM diagnosis. Cohorts G and H consist of cases of grade II and III glioma samples. Some cases in cohort G were described in an earlier publication[1]. Frozen tissue was utilized for all tissue-based analyses related to cohorts C-G, while a combination of frozen and FFPE tissue was used for newly generated cohorts A, B, H, and I. All frozen and FFPE tissue specimens were examined by a board-certified neuropathologist and estimated to contain tumor content of >70% of viable cells. Samples with low tumor content, suboptimal DNA/RNA quality and yields, or high normal brain contamination index (as determined by expression profiling) were excluded from analysis. If an alternative tissue specimen from the same case was examined and found to be suitable for study, this replacement sample was included in the study and is designated “recut” in Table S1 or S2. In such instances, no record of the original sample appears in the current report. For analysis of expression profiles of matched pairs of primary and recurrent tumors, frozen tissue was utilized from 25 cases from cohorts A, C, and G (Table S8).

For hierarchical clustering of GBM gene expression profiles with those from human adult neural stem cells and brain tissue, samples were as follows: GBM samples annotated in Table S1 as profiled on U133 A&B arrays; 4 HANSE cell cultures prepared as previously described[6]; 4 post-mortem brain samples from 15-20 week

fetuses; 3 post-mortem samples and 4 surgical specimens from cortex of non-tumor bearing donors.

Sequencing

IDH1 Sequencing

Genomic DNA was isolated from frozen or formalin-fixed paraffin-embedded tissue using the DNeasy Blood and Tissue Kit (Qiagen). DNA and RNA were simultaneously isolated from a number of samples using the AllPrep DNA/RNA kit (Qiagen). For determination of sequence at *IDH1* residue 132, both Sanger and Sequenom methods were utilized. All $IDH1^{R132MUT}$ calls and nearly all (>95%) of $IDH1^{R132WT}$ calls were confirmed with both sequencing platforms with the exception that, for samples in cohort H, only Sanger sequencing was performed. Sanger sequencing primers were designed to include the R132 codon within exon 4. Some samples were analyzed using forward 5' – GCGTCAAATGTGCCACTATC – 3' and reverse 5' – GCAAAATCACATTATTGCCAAC – 3' to generate a 236 bp fragment. Other samples were analyzed using a nested PCR approach to generate a 562 bp and 513 bp fragment from the first and second round of PCR, respectively (1^{st} round: forward primer: 5' – GCACCCATCTTCTGTG – 3' reverse primer: 5' – GTGTAGATACCAAAAGATAAG – 3' and 2^{nd} round: forward primer: 5' – TGTGCCAGTGCTAAAACT – 3', reverse primer: 5' – AACACATACAAGTTGGAAATTTCT – 3'). The PCR products were sequenced using the BigDye Terminator v1.1 (Applied Biosystems) and analyzed on a 3730 sequencer (Applied Biosystems).

For the Sequenom assay, primers (5' – ACGTTGGATGAAAATATCCCCCGGCTTGTG – 3' and 5' – ACGTTGGATGACATGACTTACTTGATCCCC – 3') were designed to amplify a 100 bp product with separate extension primers used to interrogate nucleotide position 394 (5' – GATCCCCATAAGCATGA – 3') and position 395 (5' – ATCCCCATAAGCATGAC – 3'). The resultant PCR products were analyzed by mass spectrometry (Sequenom, San Diego, CA). A more detailed Sequenom assay protocol is as follows: PCR reactions with the appropriate template DNA and 100nM of each PCR primer were set in a 5ul volume and cycled for 94°C for 2 min followed by 35 cycles at 94°C for 30 sec, 56°C for 30 sec and 72°C for 1 min. The reaction was terminated at a final extension for 8 min at 72°C, following which 0.3 units of Shrimp alkaline phosphatase (Sequenom) was added to each reaction mixture for 40 min at 37 °C to remove excess deoxynucleotides. The shrimp alkaline phosphatase was heat inactivated for 5 min at 85°C and a single base pair extension was performed using Thermosequenase and iPLEX nucleotides (Sequenom). Extension primers were used at a final concentration of 0.2μM in a 9μl reaction. Extension reactions were initially incubated at 94°C for 30 sec followed by 40 cycles of 94°C for 5 sec, 5 cycles of 52°C for 5 sec and 80°C for 5 sec. After this, primer extension reaction products were desalted with SpectroCLEAN resin (Sequenom). Ten nanoliters of the extension reaction was dispensed on a 384-format SpectroCHIP (Sequenom) prespotted with 3-hydroxypicolinic acid using a MassARRAY nanodispenser (Sequenom).

Sequencing of p53

To generate the data in Figure 4A on "any *p53* mutation", exons 4 - 10 were analyzed using Sanger sequencing on a series of selected GBMs and AAs (Tables S1 & S2). To query *p53* positions 1068 and 1069 (corresponding to 2nd and 3rd residues in codon 273), a sequenom assay was developed. For the data reported in Figure 4B, an expanded sample series including all available adult grade II-IV gliomas was examined (cohorts A-D, F-I; Table S10). A sequenom assay to sequence nucleotide position 1068, 1069 (corresponding to codon 273) of TP53 was developed using the following primer sets 5' – ACGTTGGATGGTAATCTACTGGGACGGAAC – 3' and 5' – ACGTTGGATGAGATTCTCTTCCTCTGTGCG – 3' for amplification and the extension primers for nucleotide position 1068 (5' – AGGACAGGCACAAACAC – 3') and for position 1069 (5' – CAGGACAGGCACAAACA – 3') (Sequenom).

Histological analysis

A set of pathologist-confirmed GBM cases for which FFPE sections were available was utilized for analysis (Table S3). Tissue selection and analyses were performed by pathologists blinded to the genotype of the cases examined. For MIB-1 counts, a pathologist (FP) examined sections stained by immunohistochemistry for Ki-67 and counterstained with a nuclear dye and scored a total of 1,000 tumor cell nuclei per specimen. This same pathologist scored each specimen for the occurrence of necrosis and vascular abnormalities on ordinal scales. For necrosis, the scale included: none, focal, moderate, or extensive. For vascular abnormalities, the presence of aberrations, apart from microvascular proliferation, was scored on an ordinal scale of absent, focal, or present. Vascular abnormalities scored included glomeruloid features, perivascular

cellular collars, and abnormal stromal elements, such as, excessive extracellular matrix. Assessments of percent oligodendroglial content were performed by a pair of neuropathologists (O.E.S., W.H.Y.) according to WHO criteria[5] and scored on an ordinal scale (<1% , 1-5%, 5-10%, 10-25%, 25-50%, 50-75%, >75%).

Generation and Analysis of Microarray Data

Array CGH

Genomic DNA was extracted from selected GBM patient tumor specimens in cohort A (Table S4) using AllPrep DNA/RNA kit according to manufacturer's protocol (Qiagen). 500 ng of genomic DNA was labeled and hybridized to Human Genome 1 Million CGH microarrays following manufacturer's standard protocol (Agilent). Arrays were analyzed with Feature Extraction software version 10.7 (Agilent). Genome-wide analysis of copy number gains and losses were determined using Nexus Copy Number software (Biodiscovery). Analysis was performed using the FAAST rank segmentation algorithm with the following settings: threshold $\log_2$ value of +.2 for gain, and -.2 for loss. Amplification of EGFR was scored for all samples with focal $\log_2$ copy number gains that achieved ratios of >0.7 at the EGFR locus.

Gene expression profiling and analysis

Total RNA was extracted from frozen tissue using the AllPrep DNA/RNA kit and labeled cRNA was hybridized to U133A and B, U133P, or Agilent WHG chips according to standard manufacturer's protocol. A set of 35 signature genes, or, for TCGA dataset, a subset of 28 genes for which data was available (Table S6), were used in k means clustering to assign tumors into one of the 3 subtypes. Previously published

classification results for cohorts C, D, & G [1-2] are reported (Table S1). Data from cohorts A, B, and F (all profiled on Affymetrix U133P chips) were pooled for k means clustering (Tables S1 and S7). Similarity to each of the 3 centroids, as represented by Pearson's correlation coefficients, were then generated as described[1]. Tumors that mapped with a correlation coefficient of less than 0.2 to any of the three centroids were deemed 'poorly classified'.

Hierarchical clustering

Hierarchical clustering of gene expression data (Affy_mas5sig) generated on Affymetrix U133 A&B chips was performed using Spotfire Decision Site software. The genelist for clustering consisted of all probe sets with mean Affy_mas5sig values >200 across the GBM sample set for which a comparison between $IDH1^{R132MUT}$ vs $IDH1^{R132WT}$ samples showed 2 fold or greater difference with a t-test p value of $<1x10^{-5}$. For clustering, data was z-score normalized across all samples and agglomerative clustering was employed.

AGDEX

AGDEX methodology was employed as previously described[7] to compare expression profiles of human GBMs to that of a published dataset on neural precursors in mid-gestation mouse embryonic forebrain[8]. For each platform, the microarray probe for a given gene generating the most robust signal within the dataset under investigation was employed. In accordance with the conclusions of the study by Kawaguchi et al[8], samples from cluster 1 of the published mouse dataset were designated as neural stem cells and clusters 2, 3, & 4 were pooled to represent committed neural progenitors. Resulting AGDEX and Pearson coefficients, +0.147 and +0.144, respectively, were, as

anticipated, quite similar. The observed AGDEX value of +0.147 exceeds that produced by any of a series of 250 results generated with random permutations of samples, indicating that the probability of the observed result is <.004.

Methylation profiling and analysis

Genomic DNA isolated from frozen tissue was used to perform methylation sensitive restriction enzyme (MSRE) analysis[9] with Agilent CpG Island Microarray (Agilent Technologies) on selected GBM (cohort A, Table S1) and lower grade samples (cohort H, Table S2). Briefly, 750ng of isolated DNA was digested with BfaI, then ligated to oligonucleotide linkers. Following ligation, the product was divided into two equal portions, separately digested with HpaII (methylation sensitive) or MspI (methylation insensitive), amplified, labeled and co-hybridized to the microarray chip. Raw data was processed and normalized using R language (www.r-project.com)[10]. All log ratio values were converted to a standardized z-score (based on each experiment's internal control probes) and the CpG island fragment list (X and Y chromosome loci removed) with highest variability amongst all samples was subjected to unsupervised hierarchical clustering using R hclust package with distance=Euclidean and Linkage=Complete. Identical sample clustering results were obtained using CpG island fragment lists with coefficient of variation cutoffs at all values above 2. For figure 4D, all fragments with cv > 3 for which all samples yielded useable data were included (see Table S11).

Manual bisulfite sequencing validation

Two *IDH1*^{R132MUT} samples and two *IDH1*^{R132WT} were selected for validation on a panel of CpG island fragments selected from an expanded list of most variably methylated fragments (cv >2) (Table10 and Table S12). Bisulfite modified DNA was prepared as

described above. MethPrimer software was used to generate PCR products for sequencing[11]. Detection of a cytosine in the context of a CG dinucleoptide indicated a methylated cytosine.

EGFR and *PTEN* Fluorescent In-situ Hybridization (FISH) Analysis

5 micron thickness FFPE slices of a series of GBMs (cohorts A and B; Table S3) were subjected to FISH analysis as described previously[12] with modifications. Following an overnight incubation at 56ºC, the slides were deparaffinized in 3 washes of CitroSolv for 5 min each, followed by two washes in alcohol. After air-drying, the slides were incubated in a 1M solution of NaSCN for 30 min at 80ºC and then were treated with pepsin prior to additional washes in water and a series of ethanol. Dried slides were then co-denatured (76°C for 6 min) with the probe and were hybridized overnight at 37°C (ThermoBrite; Vysis, Downers Grove, IL). Post-hybridization washes and counter-staining were done in a manner similar to those previously described. For *EGFR*, commercially available probes for *EGFR*/CEP7 and CEP10 were utilized (Vysis/Abbott Laboratories). For *PTEN*, a probe consisting of a BAC contig that encompassed the *PTEN* loci and adjoining areas was labeled with Spectrum Orange using a nick translation kit (Vysis/Abbott Laboratories). Using FISHView software, a minimum of 70 and 150 non-overlapping nuclei per slide were analyzed for the *EGFR*/CEP7 and *PTEN* /CEP10 hybridization signals, respectively (Applied Spectral Imaging). For *EGFR*, a tumor was identified as amplified when >10% of the tumor cells either had an *EGFR*/CEP7 ratio of 2 or the presence of multiple *EGFR* signals in tight clusters. A tumor was categorized as deleted for the *PTEN* locus when >40% of the tumor cells

exhibited less than or equal to one copy of *PTEN*.

MGMT methylation analysis

MGMT methylation analysis was performed on GBM cases with available DNA in cohort A (Table S1) by methylation-specific PCR (MSP) according to a previously published protocol[13] with slight modifications. To generate bisulfite-modified DNA, genomic DNA isolated from FFPE using Recoverall Total Nucleic Acid Isolation Kit (Ambion) was modified using the EZ DNA Methylation-Gold Kit (Zymo Research) following the manufacturer's protocol.

For *MGMT*, samples were subjected to a two-stage nested PCR strategy using: first-stage primers (5'-GGATATGTTG GGATAGTT-3' and 5'-CCAAAAACCCCAAACCC-3') and second-stage primers (unmethylated reaction: 5'-TTTGTGTTTTGATGTTTGTAGGTTTTTGT-3' and 5'-AACTCCACACTCTTCCAAAAACAAAACA-3'; methylated reaction: 5'-TTTCGACGTTCGTAGGTTTTCGC-3' and 5'-GCACTCTTCCGAAAACGAAACG-3'). PCR products are analyzed on 3% agarose gels. Positive and negative control samples for the MSP reaction included U87MG DNA treated with SssI methyltransferase (NEB) and whole-genome amplification of U87MG DNA using the GenomiPhi V2 Amplification kit (Amersham Biosciences), respectively.

Analysis of MR Images

Available pretreatment (pre-initial operation) MRIs (T1 with contrast and T2 or FLAIR) from cohort A (Table S5), were analyzed by a neuroradiologist (W.B.P) blinded to

patient genotype for the tumor variables characterized according to previously reported criteria[14]. Tumor location was assessed for all patients from MRI scans, patient history or radiological report by a neurologist (J.C.). Each tumor was scored according to which cerebral lobe(s) involved and tumors spanning more than one lobe were counted towards both locations. In addition, a 'non-lobar' category was designated to include corpus callosum, thalamus, brainstem, and cerebellum. Tumor size (cm) was assessed by measuring the greatest tumor dimension (1-D) seen on axial imaging (J.C.)

Imaging overlay/tumor probability analysis

All *IDH1*^{R132MUT} cases from cohort A with available digital imaging and a random sampling of *IDH1*^{R132WT} cases from this same cohort were subjected to overlay analysis (Table S5). Images were registered to a high-resolution (1.0mm isotropic), T1-weighted brain atlas (MNI152; Montreal Neurological Institute) using a mutual information algorithm and a 12-degree of freedom transformation using FSL (FMRIB, Oxford, UK; http://www.fmrib.ox.ac.uk/fsl/)[15]. Fine registration (1-2 degrees & 1-2 voxels) was then performed using a Fourier transform-based, 6-degree of freedom, rigid body registration algorithm[15] followed by visual inspection to ensure adequate alignment. Following registration with the atlas, the skull was removed from each post-contrast T1-weighted and T2-weighted/FLAIR volume dataset using the Brain Extraction Tool (BET; FMRIB, Oxford, UK; http://www.fmrib.ox.ac.uk/fsl/)[16]. Contrast-enhancing and T2 hyperintense lesions were segmented from each volume dataset, excluding regions of edema outside the primary tumor site while retaining cystic/necrotic regions within contrast-enhancing tumor. The segmented contrast-enhancing and T2 hyperintense regions of interest (ROIs) were combined and used as an estimate of the tumor burden for each patient.

The probability of a patient in a particular group having tumor burden in a particular voxel location was calculated by dividing the number of times a patient's tumor was located within a voxel of interest by the total number of patients per group. Image maps containing voxel-wise tumor probability were then exported to ANALYZE format and imported into the open source 3D Slicer software package (http://www.slicer.org/). A 3D cortical surface model was generated in 3D Slicer from the T1-weighted MNI152 brain atlas and 3D surface models were then generated for tumor probabilities of >0% through >90% at increments of 10% tumor probability.

Determination of area of differential involvement (ADIFFI)

Scans were assumed to be independent across patients. Each voxel was analyzed separately. For each patient, each voxel within the brain was classified by a pair of two-level factors: (1) whether the patient with that voxel has an $IDH1^{R132MUT}$ tumor or a $IDH1^{R132WT}$ tumor (2) whether that voxel in that patient represents a region of the brain containing tumor, or a region containing healthy brain. For each voxel, a 2x2 contingency table was created enumerating across patients according to these two factors. For each such table, the hypothesis that patient IDH1 mutation status was independent of whether that voxel contained tumor was tested via Fisher's Exact Test [17], and the two-sided p-values recorded.

Statistical analysis

Survival analysis used the survival package [18] in the R language and estimated Kaplan-Meier curves were compared via the log-rank test. For parametric analyses, two-tailed t-tests were employed and for analysis of frequencies of nominal data, two-tailed

Fisher's Exact Test was employed. Analysis of frequencies of ordinal data included the following procedure: for each ordinal end point, association between *IDH1* mutational status and the end point's levels was assessed via a generalized linear model [19] for an ordinal response as a function of *IDH1* status. As an alternative to a standard contingency table method (e.g., Fisher's Exact Test), this incorporates the knowledge about the ordering of the ordinal levels to achieve more power to detect a difference based on mutational status. Specifically, a proportional odds linear regression model with a logistic link was implemented via the polr() function from the MASS library [20] in R. In estimating the probability of *IDH1* mutation given patient tumor grade (AA or GBM) and age, the procedure was as follows: Separately for each grade, the probability that a tumor harbored the *IDH1* mutation for patients of a give age was estimated by a logistic regression that modeled the binary outcome (i.e., $IDH1^{R132WT}$ or $IDH1^{R132MUT}$) as a smooth function of age, where the 'smooth function' was estimated using a cubic B-spline [20] of patient age.

SUPPLEMENTAL TABLE LEGENDS

Table S1: Patient information for GBM cohorts A-F; Relates to Figures 1-4
Includes: Sample identifiers for cohorts A-F, *IDH1* mutational status, survival data (Figure 1A), *MGMT* promoter methylation status (Figure 1B), gene expression subtype (Figure 2 A-C), tumor location (Figure 3A, E), age at diagnosis (Figure 3F), *p53* mutational status (Figure 4A), identification of samples profiled on methylation arrays (Figure 4D).

Table S2: Patient information for lower grade glioma cohorts G and H; Relates to Figures 2A-C, 3B-D & F, 4A and 4D.
Includes: Sample identifiers for cohorts G & H, survival data for cohort G & H, *IDH1* mutational status, gene expression subtype (Figure 2 A-C), tumor location (Figure 3B-D), age at diagnosis (Figure 3F), *p53* mutational status (Figure 4A), identification of samples profiled on methylation arrays (Figure 4D)

Table S3: Histology and FISH results of GBMs (sample series from cohorts A and B); Relates to Figures 1B and 4E.
Includes: Results of individual samples for histological examination (Figure 1B) and FISH results (Figure 4E)

Table S4: Tabulation of selected DNA copy number alterations detected by CGH (sample series from cohort A); Relates to Figure 1C
Includes: Tabulation of DNA copy number alterations in individual samples reported in Figure 1C.

Table S5: Scoring of radiological features of GBMs (cohort A cases with available pre-operative imaging); Relates to Figures 1E and 3E
Includes: MRI imaging scores summarized in Figure 1E and list of cases used in composite MRI imaging overlay in Figure 3E.

Table S6: Probesets used for subtyping samples on different gene expression microarray platforms; Relates to Figure 2A-C
Includes: Probesets used in gene expression subtyping analysis (Figure 2A-C)

Table S7: Data used for subtyping GBM samples in cohorts A, B, F on Affymetrix U133P platform; Relates to Figure 2A-C
Includes: Microarray expression values for 35 signature genes used for k-means classification of samples profiled on U133P platform (Figure 2A-C).

Table S8: Gene expression subtypes of matched pairs of primary and recurrent tumors (cases from cohorts A, C, and G); Relates to Figure 2C
Includes: Sample identifiers and subtype classification results for matched specimen pairs depicted in Figure 2C.

Table S9: Microarray data for hierarchical clustering of GBMs (in cohort C) cwith human neural stem cells and brain samples; Relates to Figure 2D
Includes: Raw microarray data utilized to generate heatmap in Figure 2D.

Table S10: Sequencing results for *IDH1*^{R132} and *p53*^{R273} (all adult cases in cohorts A-H, and additional glioma samples from cohort I); Relates to Figure 4B
Includes: Sequencing results for individual samples in Figure 4B

Table S11: Methylation array data (from Cohorts A & H) used to generate Figure 4D
Includes: Normalized log ratio data from CpG island fragments (annotated by gene name and location) found to display coefficient of variation (cv) >2.0 across the dataset. Data included in Figure 4D (all genes for which cv>3.0 and data for all samples met quality control criteria) is highlighted in orange.

Table S12: Bisulfite sequencing validation of methylation calls; Relates to Figure 4D
Each CG within a methylation sensitive enzyme site (CCGG) is depicted as M (methylated), U (unmethylated), X (not covered) or B (equal methylated and unmethylated peaks). The methylation status of CG sites outside of CCGGs were also tabulated. Highlighted in yellow are the regions where the CG methylation status is patchy.

REFERENCES

1. 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 9:157-73, 2006
2. Chen R, Nishimura MC, Bumbaca SM, et al: A Hierarchy of Self-Renewing Tumor-Initiating Cell Types in Glioblastoma. Cancer Cell 17:362-375, 2010
3. Schiffman JD, Hodgson JG, VandenBerg SR, et al: Oncogenic BRAF mutation with CDKN2A inactivation is characteristic of a subset of pediatric malignant astrocytomas. Cancer Res 70:512-9, 2010
4. Gunther HS, Schmidt NO, Phillips HS, et al: Glioblastoma-derived stem cell-enriched cultures form distinct subgroups according to molecular and phenotypic criteria. Oncogene 27:2897-909, 2008
5. Louis DN, International Agency for Research on Cancer.: WHO classification of tumours of the central nervous system (ed 4th). Lyon, International Agency for Research on Cancer, 2007
6. Muller F-J, Laurent LC, Kostka D, et al: Regulatory networks define phenotypic classes of human stem cell lines. Nature 455:401-405, 2008
7. Johnson RA, Wright KD, Poppleton H, et al: Cross-species genomics matches driver mutations and cell compartments to model ependymoma. Nature 466:632-636, 2010
8. Kawaguchi A, Ikawa T, Kasukawa T, et al: Single-cell gene profiling defines differential progenitor subclasses in mammalian neurogenesis. Development 135:3113-3124, 2008
9. Tran A, Escovedo C, Migdall-Wilson J, et al: In silico enhanced restriction enzyme based methylation analysis of the human glioblastoma genome using Agilent 244K CpG Island microarrays. Frontiers in Neurogenomics 1, 2009
10. Team RDC: R: A Language and Environment for Statistical Computing, 2009
11. Li LC, Dahiya R: MethPrimer: designing primers for methylation PCRs. Bioinformatics 18:1427-31, 2002
12. Pandita A, Aldape KD, Zadeh G, et al: Contrasting in vivo and in vitro fates of glioblastoma cell subpopulations with amplified EGFR. Genes Chromosomes Cancer 39:29-36, 2004
13. Hegi ME, Diserens AC, Gorlia T, et al: MGMT gene silencing and benefit from temozolomide in glioblastoma. N Engl J Med 352:997-1003, 2005
14. Pope WB, Sayre J, Perlina A, et al: MR imaging correlates of survival in patients with high-grade gliomas. AJNR Am J Neuroradiol 26:2466-74, 2005
15. Cox RW, Jesmanowicz A: Real-time 3D image registration for functional MRI. Magn Reson Med 42:1014-8, 1999
16. Smith SM: Fast robust automated brain extraction. Hum Brain Mapp 17:143-55, 2002
17. Agresti A: Categorical data analysis (ed 2nd). New York, Wiley-Interscience, 2002

18. Lumley TTaT: survival: Survival analysis, including penalised likelihood. 2009

19. McCullagh P, and Nelder, J. A.: Generalized linear models, (ed 2nd). Chapman and Hall, 1989

20. Venables WN, Ripley BD: Modern applied statistics with S (ed 4th). New York, Springer, 2002

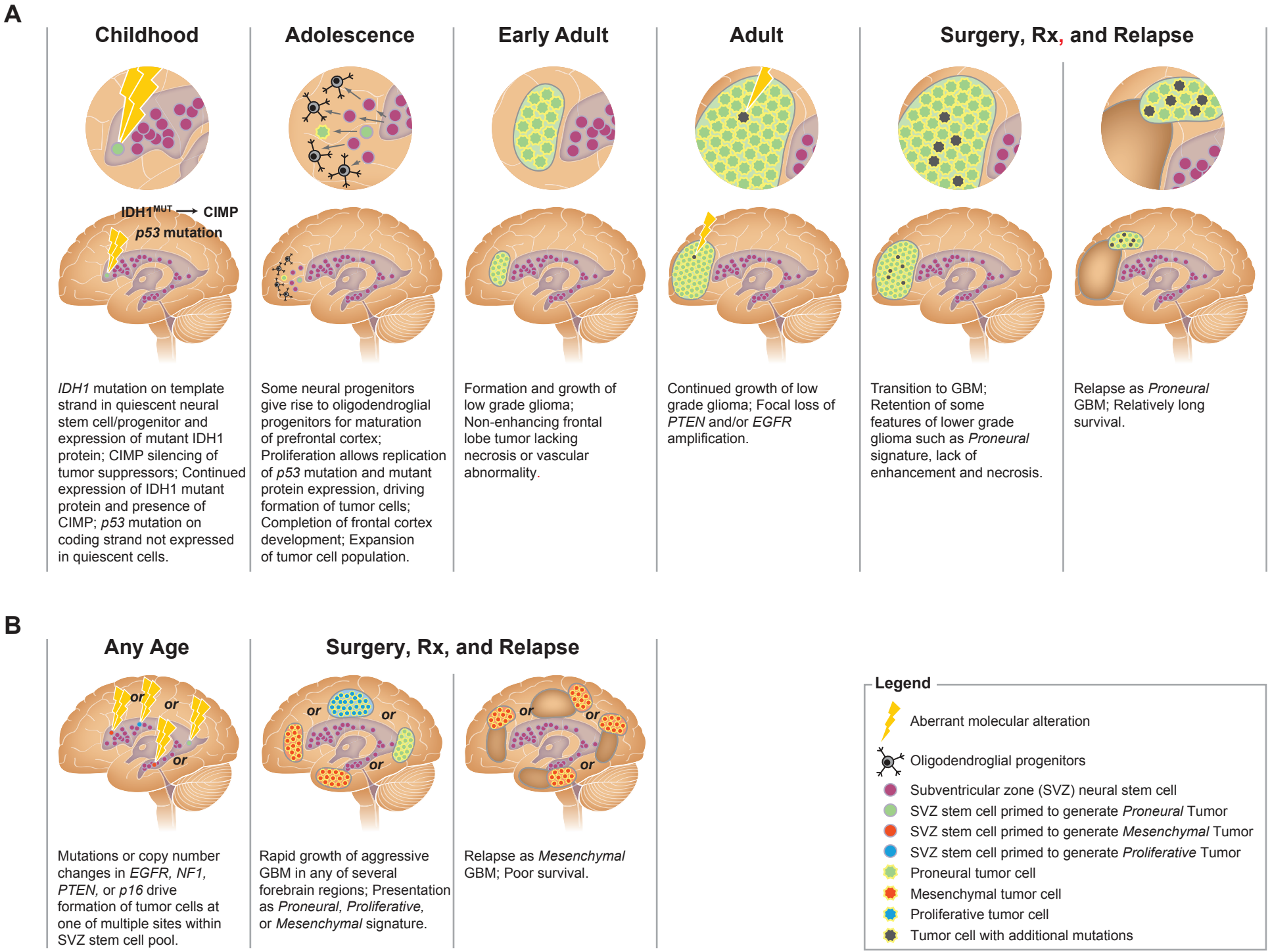
A
Childhood
Adolescence
Early Adult
Adult
Surgery, Rx, and Relapse
IDH1MUT → CIMP
p53 mutation
IDH1 mutation on template strand in quiescent neural stem cell/progenitor and expression of mutant IDH1 protein; CIMP silencing of tumor suppressors; Continued expression of IDH1 mutant protein and presence of CIMP; p53 mutation on coding strand not expressed in quiescent cells.
Some neural progenitors give rise to oligodendroglial progenitors for maturation of prefrontal cortex; Proliferation allows replication of p53 mutation and mutant protein expression, driving formation of tumor cells; Completion of frontal cortex development; Expansion of tumor cell population.
Formation and growth of low grade glioma; Non-enhancing frontal lobe tumor lacking necrosis or vascular abnormality.
Continued growth of low grade glioma; Focal loss of PTEN and/or EGFR amplification.
Transition to GBM; Retention of some features of lower grade glioma such as Proneural signature, lack of enhancement and necrosis.
Relapse as Proneural GBM; Relatively long survival.
B
Any Age
Surgery, Rx, and Relapse
or
Mutations or copy number changes in EGFR, NF1, PTEN, or p16 drive formation of tumor cells at one of multiple sites within SVZ stem cell pool.
Rapid growth of aggressive GBM in any of several forebrain regions; Presentation as Proneural, Proliferative, or Mesenchymal signature.
Relapse as Mesenchymal GBM; Poor survival.
Legend
Aberrant molecular alteration
Oligodendroglial progenitors
Subventricular zone (SVZ) neural stem cell
SVZ stem cell primed to generate Proneural Tumor
SVZ stem cell primed to generate Mesenchymal Tumor
SVZ stem cell primed to generate Proliferative Tumor
Proneural tumor cell
Mesenchymal tumor cell
Proliferative tumor cell
Tumor cell with additional mutations

Figure 5