

## Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Eckel-Passow JE, Lachance DH, Molinaro AM, et al. Glioma groups based on 1p/19q, *IDH*, and *TERT* promoter mutations in tumors. *N Engl J Med* 2015;372:2499-508. DOI: 10.1056/NEJMoa1407279

<b>INDEX OF CONTENTS</b>	<b>PAGE</b>
<b>Supplementary Results</b>	
Association of glioma molecular groups with acquired genetic alterations	4
Association of glioma molecular groups with <i>MGMT</i> methylation	5
Association of glioma molecular groups with tumor location	5
Association of TCGA GBM RNA expression subtypes with germline variants	5
<b>Supplementary Methods</b>	
Mayo Clinic case-control study	7
UCSF adult glioma case-control study	7
The Cancer Genome Atlas (TCGA)	8
Mayo Genome Consortia (MayoGC) controls	8
Pathology review	8
<i>IDH1</i> and <i>IDH2</i> mutation	9
1p/19q codeletion	9
<i>TERT</i> promoter mutation	10
<i>ATRX</i> status	11
<i>MGMT</i> methylation	11
Additional molecular markers	11
DASL expression profiling	12
Genotyping	12
Statistical methods	12
<b>Supplementary Figures</b>	
Figure S1. Prevalence of <i>TERT</i> promoter mutation, <i>IDH</i> mutation, and 1p/19q codeletion in the Mayo Clinic, UCSF AGS, TCGA, and combined cases	15
Figure S2. Prevalence of the glioma molecular groups as defined by <i>TERT</i> promoter mutation, <i>IDH</i> mutation, and 1p/19q codeletion status in the Mayo Clinic, UCSF AGS, and TCGA cases	16
Figure S3. Pattern of <i>TERT</i> promoter mutation, <i>IDH</i> mutation, and 1p/19q codeletion status among 317 Mayo Clinic gliomas	17

Figure S4. Molecular characterization of the glioma molecular groups defined by <i>TERT</i> promoter mutation, <i>IDH</i> mutation, and 1p/19q codeletion status for (A) Mayo Clinic, (B) UCSF AGS, and (C) TCGA cases	19
Figure S5. Comparison of the glioma molecular groups defined by <i>TERT</i> promoter mutation, <i>IDH</i> mutation, and 1p/19q codeletion status versus the TCGA GBM expression subtypes for (A) 166 Mayo Clinic cases and (B) 62 UCSF AGS cases	23
Figure S6. Unadjusted Kaplan-Meier estimates of overall survival of the glioma molecular groups defined by <i>TERT</i> promoter mutation, <i>IDH</i> mutation, and 1p/19q codeletion status	24
Figure S7. Odds ratios for case-control SNP associations of 9 regions with development of the five glioma molecular groups as defined by <i>TERT</i> promoter mutation, <i>IDH</i> mutation, and 1p/19q codeletion status	25
<b>Supplementary Tables</b>	
Table S1. TCGA data downloaded for this project	27
Table S2A. Distribution of histologic type and grade in the Mayo Clinic, UCSF AGS, and TCGA cases stratified by molecular group	28
Table S2B. Age at diagnosis, gender, tumor location, and extent of resection for the Mayo Clinic, UCSF AGS, and TCGA cases stratified by molecular group	29
Table S2C. Association between molecular group and age at diagnosis for the Mayo Clinic, UCSF AGS, and TCGA cases	30
Table S3. Summary of common glioma alterations within molecular groups in the Mayo Clinic, UCSF AGS and TCGA cases	31
Table S4. Hazard rate (HR) and 95% confidence interval (95% CI) for all pairwise molecular group comparisons from a stratified (by dataset) Cox proportional hazards model	32
Table S5. Association of 22 SNPs in 9 independent regions known to be associated with glioma in the Mayo Clinic, UCSF AGS, and TCGA/MayoGC case-control studies, stratified by molecular group	33
Table S6. Association of 22 SNPs in 9 independent regions known to be associated with glioma in the Mayo Clinic cases, stratified by TCGA GBM RNA expression subtypes	34
<b>Acknowledgements</b>	35
<b>References</b>	36

## SUPPLEMENTARY RESULTS

### Association of glioma molecular groups with acquired genetic alterations

Table S3 provides a list of all copy number and point mutations that were available for the Mayo Clinic, UCSF AGS, and TCGA glioma cases. Figures S3 and Figures S4 compare the glioma groups with these acquired mutations and the TCGA GBM RNA expression subtypes.

Triple-positive gliomas nearly uniformly had a proneural RNA expression pattern with few, but noteworthy, mutations and copy-number alterations: monosomy 4 (31%), *CIC* mutations (mapped to 19q; 58%), *FUBP1* mutations (mapped to 1p; 27%), *NOTCH1* mutations (33%), *PIK3CA* mutations (19%), and *PIK3R1* mutations (9%). *CDKN2A/B* was often hemizygotously deleted in triple-positive gliomas (21%).

*IDH*-mutated-only gliomas frequently acquired *ATRX* alterations (86%) and *TP53* mutations (88%). These gliomas frequently had copy-number gain of 8q24 (which always included *MYC* and rs55705857; 31%), lost all or a portion of 19q alone (30%), gained all or portions of 7q (22%), and co-amplified *CDK4* and *GLI1* (7%) but rarely *MDM2* (<1%). *CDKN2A/B* was frequently hemizygotously deleted (30%) and had a low prevalence of homozygous deletion (7%). Gliomas in this group mainly had a proneural RNA expression pattern.

*TERT*-mutated-only gliomas harbored all of the known common alterations associated with primary GBM. These include gain of chromosome 7 (78%), amplification of *EGFR* (49%), presence of *EGFRvIII* (18%), loss of *CDKN2A/B* (73%) and *PTEN* (94%), and *RB1* mutations (7%). Unlike *IDH*-mutated-only gliomas, this group frequently co-amplified *CDK4* and *GLI1* (22%) as well as amplified *MDM2* (12%). Unlike triple-positive and *IDH*-mutated-only gliomas, *CDKN2A/B* was frequently homozygously deleted (54%). Similar to triple-positive gliomas, this group also acquired frequent mutations in *PIK3CA* (11%) and *PIK3R1* (7%). *TERT*-mutated-only gliomas generally had either a mesenchymal or classical RNA expression pattern.

In general, the patterns of alterations in triple-negative gliomas were similar to that of *TERT*-mutated-only gliomas. However, the prevalence of the alterations in triple-negative gliomas was usually lower than in *TERT*-

mutated-only gliomas. For example, chromosome 7 gain, *EGFR* amplification, *EGFRvIII*, *CDKN2A/B* loss, and *PTEN* loss were observed in 48%, 26%, 7%, 57%, and 60% of tumors, respectively. There were some differences compared to *TERT*-mutated-only gliomas. For example, *PIK3CA* and *PIKR1* mutations were notably absent in triple-negative gliomas. The prevalence of *MYC* gain was 15% (compared to 5% in the *TERT*-mutated-only gliomas). Triple-negative gliomas had all four of the RNA expression patterns (classical, mesenchymal, neural, and proneural).

*TERT*- and *IDH*-mutated gliomas were a small group. Like the *TERT*-mutated-only gliomas and the triple-negative gliomas, this group often homozygously deleted *CDKN2A/B* (24%), lost *PTEN* (40%), and amplified *PDGFRA/KIT* (20%). Like the *IDH*-mutated-only group, a significant proportion of this group gained all or part of 7q (20%), gained 8q24 (20%), and acquired *TP53* mutations (32%) and *ATRX* mutations (28%).

#### **Association of glioma molecular groups with *MGMT* methylation**

*MGMT* methylation was only available for TCGA cases, where it varied substantially among the groups. *MGMT* methylation was observed in TCGA cases at a frequency of 100% (64/64) in triple-positive, 91% (125/138) in *IDH*-mutated-only, 100% (6/6) in *TERT*- and *IDH*-mutated, 44% (56/128) in *TERT*-mutated-only, and 33% (13/39) in triple-negative gliomas.

#### **Association of glioma molecular groups with tumor location**

Tumor location data was only confirmed and analyzed in the Mayo Clinic cases (Table S2B). There was a significant association between group and tumor location ( $p=0.014$ ). While approximately 80% of triple-positive and *IDH*-mutated-only tumors occur in the frontal lobe, approximately 50% *TERT*-mutated-only and *TERT*- and *IDH*-mutated tumors and 60% of triple-negative tumors occur in the frontal lobe.

#### **Association of TCGA GBM RNA expression subtypes with germline variants**

Using the Mayo Clinic case-control study, we evaluated the association between risk of the TCGA GBM RNA expression subtypes and nine regions previously shown via GWAS to be associated with glioma risk: *TERC*

(3q26), *TERT* (5p15), *EGFR* (7p12; containing two independent regions), *CCDC26* (8q24), *CDKN2A/B* (9p21), *PHLDB1* (11q23), *TP53* (17p13), and *RTEL1* (20q13).<sup>1-5</sup> Of note, while there have been candidate genes/SNPs (e.g., *ERCC1*) prior to GWAS that have been published, these genes were not validated in our population and thus were not considered herein.<sup>6</sup> The *CCDC26* and *TERT* SNPs were associated with risk of developing proneural gliomas (Table S6).

## **SUPPLEMENTARY METHODS**

### **Mayo Clinic case-control study**

The Mayo Clinic glioma case-control study has been described previously.<sup>1,4</sup> This study was approved by the Mayo Clinic Office for Human Research Protection and informed written consent was obtained from all participants. Briefly, all cases were identified at diagnosis (diagnosed at Mayo Clinic) or at the time of pathologic confirmation (diagnosed elsewhere and treated at Mayo Clinic), were at least 18 years of age, and had a surgical resection or biopsy between 1989 and 2012. Patient clinical data were extracted from the electronic medical record. Pre-operative radiographic images were reviewed to confirm tumor size and location. Postoperative radiographic images were also reviewed to determine the extent of resection. Controls were at least 18 years of age, underwent a general medical examination at Mayo Clinic between 1989 and 2012, and had no previous history of a brain tumor. Controls were matched to cases by gender, age, ethnicity, and residence. Consenting participants provided blood, buccal, and/or saliva specimens and information during in-person or telephone interviews. A total of 317 cases and 789 controls were used as the discovery set.

### **UCSF adult glioma case-control study**

The UCSF case-control study includes participants of the San Francisco Bay Area Adult Glioma Study (AGS). This study was approved by the UCSF Committee on Human Research and informed written consent was obtained from all participants. Most details of subject recruitment for AGS have been reported previously.<sup>1,5,7-9</sup> Briefly, all cases were adults (>18 years of age) with newly diagnosed histologically confirmed glioma. Population-based cases diagnosed between 1991 and 2009 (Series 1-4) and residing in the six San Francisco Bay Area counties were ascertained using the Cancer Prevention Institute of California's early case ascertainment system. Clinic-based cases diagnosed between 2002 and 2012 (Series 3-5) were recruited from the UCSF Neuro-oncology Clinic, regardless of place of residence. From 1991 to 2010, population-based controls from the same residential area as the population-based cases were identified using random digit dialing and were frequency matched to population-based cases on age, gender, and ethnicity. Between 2010 and 2012, all controls were selected from the UCSF general medicine phlebotomy clinic. Clinic-based controls

were matched to clinic-based glioma cases on age, gender, and ethnicity. Consenting participants provided blood, buccal, and/or saliva specimens and information during in-person or telephone interviews. A total of 351 cases and up to 4504 controls (including 3390 iControls<sup>1</sup>) were used as the first replication set in this study. Extent of surgery was determined from SEER registry data for the population-based cases and from abstraction from medical records and pathology reports for cases not in the SEER registry. Extent of surgery was coded as either biopsy only or resection.

### **The Cancer Genome Atlas (TCGA)**

TCGA Glioblastoma Multiforme (GBM) and TCGA Brain Lower Grade Glioma (LGG) data were downloaded as detailed in Supplemental Table 1. Clinical and pathological data for TCGA GBM and LGG cases were downloaded from the corresponding TCGA Data Matrix. As detailed below, if available, *MGMT* methylation, gene expression subtypes, *IDH* mutation, and 1p/19q codeletion status were obtained from Supplementary Table S7 of Brennan et al.<sup>10</sup>

### **Mayo Genome Consortia (MayoGC) controls**

In order to perform a SNP association analysis using the TCGA GBM and LGG cases, the MayoGC Phase 1 controls were utilized as the corresponding control data.<sup>11</sup> Phase 1 included 6297 controls across three studies.

### **Pathology review**

Two pathologists (CG and TT) reviewed pathology as described previously for the Mayo Clinic and UCSF AGS cases, respectively.<sup>1</sup> Given historical practices at the Mayo Clinic, a few gliomas were classified as grade IV mixed oligoastrocytomas or oligodendrogliomas (Table S2A in Supplementary Appendix). Because these tumors behave as if they were lower grade, for the purpose of this paper they were grouped with grade II-III



mixed oligoastrocytomas or grade II-III oligodendrogliomas, respectively. Pathology for the TCGA cases was obtained from the clinical data available in the TCGA Data Matrix.

### ***IDH1* and *IDH2* mutation**

*IDH1* and *IDH2* mutation analysis was performed as described previously for the Mayo Clinic and UCSF AGS cases.<sup>5,12,13</sup> *IDH1* and *IDH2* mutation status for TCGA cases was obtained from Brennan et al.,<sup>10</sup> for GBM subjects, when available, or from the somatic mutation data downloaded from the TCGA Data Matrix for TCGA LGG cases and the remaining GBM cases. *IDH* mutated denotes that the subject was *IDH1* or *IDH2* mutated.

### **1p/19q codeletion**

1p/19q codeletion testing was performed as described previously for the Mayo Clinic and UCSF AGS cases, unless noted below.<sup>12,13</sup> Briefly, 1p/19q codeletion status was determined in all Mayo Clinic cases either by FISH as a clinical test or by array comparative genomic hybridization (aCGH) utilizing an Agilent custom 8x60K array. 1p/19q codeletion was determined using a clinical FISH assay in the UCSF AGS cases. Because the rate of 1p/19q codeletion was rare in pure astrocytomas in both TCGA and Mayo Clinic cases, 1p/19q codeletion was not assessed in the UCSF AGS tumors classified as astrocytoma grades II-IV. Thus, UCSF AGS tumors classified as astrocytoma grades II-IV were inferred to be 1p/19q non-codeleted, which might lead to a very small misclassification of 1p/19q codeletion in the UCSF AGS data. 1p/19q codeletion status was determined for TCGA cases by evaluating the Affymetrix 6.0 Level 1 data using Genotyping Console (version 1.2.0.26) and the Affymetrix ChAS 2.1 software (Affymetrix, Santa Clara, CA) using previously-published methods.<sup>14-16</sup> The ChAS results were interpreted by four independent reviewers (RBJ, CEP, ARC, and TMK). 1p/19q codeletion for all Mayo Clinic and TCGA aCGH cases was defined as a translocation through the centromere, which results in whole arm deletion without whole chromosome loss. Evidence of mosaic low-level codeletion was considered positive.

## ***TERT* promoter mutation**

*TERT* promoter mutation for the Mayo Clinic and UCSF AGS cases was based upon a previously-published method<sup>17</sup> using reagents purchased from Life Technologies, Grand Island, NY, unless otherwise noted. A 244 base pair (bp) segment spanning the C228T and the C250T mutations in the *TERT* promoter was amplified from ~200ng genomic DNA using 2 pmol of the primers GCACAGACGCCAGGACCGCGCT and TTCCCACGTGCGCAGCAGGACGCA using 0.2 mM dNTPs, 0.5X PCR Enhancer and 1.5U of Platinum Taq DNA Polymerase in a total volume of 20 ul. Cycling conditions were set at 94°C for 1 minute and 72°C for 1 minute for 35 cycles. 1ul of the amplified DNA from the above PCR was then used as template for a second PCR in a volume of 20 ul with 2 pmol of the primers CAGGAAACAGCTATGACCATGATTACGGCACAGACGCCAGGACCGCGCT and CGTTGTAAAACGACGGCCAGTGAATTGTTCCCACGTGCGCAGCAGGACGCA, 0.5X PCR Enhancer and a 10XdNTP mix that contained 1.5 mM dGTP and 0.5mM deaza-GTP (Sigma-Aldrich, St. Louis, MO). The cycling condition was as described above. 12 ul of the amplified DNA from the 2nd PCR was mixed with 5.4 ul of ExoSAP-IT (Affymetrix, Santa Clara, CA), incubated at 37°C for 15 minutes, and then 7 minutes at 94°C. 5 ul of ExoSAP-treated DNA was then Sanger sequenced using 1 pmol of the primer CAGGAAACAGCTATGACCATGATTACG or CGTTGTAAAACGACGGCCAGTGAATTG.

For the TCGA GBM and LGG cases, *TERT* promoter mutation was obtained from the RNAseq data. As shown in Figure 4C of Brennan et al.,<sup>10</sup> *TERT* mRNA expression is highly correlated with *TERT* promoter mutation and thus we inferred *TERT* promoter mutation from the RNAseq data. Specifically, *TERT* was called mutated if the RSEM<sup>18</sup> normalized RNAseq value was larger than 5e-8. Of the 19 TCGA LGG subjects with known *TERT* promoter mutation data, this threshold resulted in 100% sensitivity (10 of the 10 known *TERT* mutated subjects were predicted to be mutated from the RNAseq data) and 89% specificity (1 of the 9 known *TERT* wild-type subjects was predicted to be mutated from the RNAseq data).

## ***ATRX* status**

*ATRX* immunostaining of the Mayo Clinic cases was performed at MSKCC (by JH) or at the Mayo Clinic (by CG) using previously-published methods.<sup>19</sup> *ATRX* immunostaining of the UCSF AGS cases was performed at the UCSF Brain Tumor Research Center Tissue Core (by MP, using the same methods as JH). *ATRX* status in TCGA cases was obtained from the sequencing-based somatic mutations downloaded from TCGA Data Matrix.

## ***MGMT* methylation**

Only TCGA glioma cases had *MGMT* methylation data available. *MGMT* methylation status was obtained for 387 (125 GBM and 262 LGG) TCGA cases that were assigned to one of the five glioma molecular groups. All 262 LGG cases and 54 of the 125 GBM cases had data from the Illumina 450K methylation array, and 74 of the 125 GBM cases had data from the Illumina 27K methylation array data (3 GBM cases had data on both arrays). For both the 450K and 27K platform we used a two-probe model (probes cg12434587 and cg12981137)<sup>20</sup> to call *MGMT* methylation. Specifically, the non-normalized methylated values (M-values) were extracted using Bead Studio and the probability of being methylated was determined from the two-probe model described by Brady et al.<sup>20</sup> Subjects with a probability larger than 0.358 were classified as being *MGMT* methylated.

## **Additional molecular markers**

In the Mayo Clinic cases, aCGH (Agilent custom 8x60K array) was used to ascertain commonly-acquired copy-number alterations. Mutations in *TP53* and amplification of *EGFR* were performed in the UCSF AGS cases as previously described<sup>21</sup> and *p16* analyses were performed using *CDKN2A* FISH with Spectrum probes (Abbott Laboratories). For TCGA cases, Affymetrix 6.0 Level 1 data were examined using Affymetrix ChAS 2.1 software and interpreted by four independent reviewers (RBJ, CEP, ARC, and TMK) for chromosomes 1, 4, 7, 8, 9, 10, 12, and 19. The deletion, gain, duplication, and amplification status of *PDGFRA*, *KIT*, *EGFR*, *MYC*,

*CDKN2A/B*, *PTEN*, *GLI*, *CDK4*, and *MDM2* was specifically examined. The Level 2 TCGA data were used to ascertain the presence of somatic mutations in *TP53*, *ATRX*, *CIC*, *FUBP1*, *NOTCH1*, *PIK3CA*, *PIK3R1*, *PTEN*, *EGFR*, *NF1*, *PDGFRA*, and *RB1*. TCGA *EGFRvIII* calls were obtained from Supplemental Table S5 in Brennan et al.<sup>10</sup>; *EGFRvIII* allele fractions (delta 2-7) > 0.01 were called *EGFRvIII* positive.

### **DASL expression profiling**

For the Mayo Clinic and UCSF AGS cases, RNA was extracted from formalin-fixed, paraffin-embedded blocks using the AllPrep DNA/RNA Mini Kit (Qiagen, Valencia, CA). The expression profiles of the resulting RNAs used the Illumina DASL method (HumanRef-8 v3 BeadChip, Illumina, San Diego, CA), as previously described.<sup>22</sup> Quality-control metrics were evaluated<sup>23</sup> and subsequently, the data were normalized for the Mayo Clinic and UCSF AGS cases separately using quantile normalization.

### **Genotyping**

Custom genotyping on the Mayo Clinic cases and controls, as well as on the UCSF cases and controls, was done using the Illumina GoldenGate assay (San Diego, CA), as described previously.<sup>4,5</sup> For the analyses described herein, we were particularly focused on evaluating germline associations of nine regions that have been previously shown to be associated with glioma. Specifically, we evaluated 22 SNPs within or near *TERC* (3q26), *TERT* (5p15), *EGFR* (7p12 – two regions), *CCDC26* (8q24), *CDKN2A/B* (9p21), *PHLDB1* (11q23), *TP53* (17p13), and *RTEL1* (20q13). Quality control of the Mayo Clinic and UCSF case-control custom SNP data were performed as described previously.<sup>4,5</sup>

### **Statistical methods**

Five glioma molecular groups were defined based on *TERT* promoter mutation, *IDH* mutation, and 1p/19q codeletion. Molecular groups with a prevalence <4% were not analyzed for associations with age at diagnosis,

outcome, or germline variants due to the small sample size associated with these groups (lack of statistical power to detect associations). As such, groups with a prevalence <4% were grouped together and classified as “other”. Age at diagnosis was compared across the five groups using contrast estimates from an analysis of variance model. Comparisons were made both within each dataset separately (Mayo Clinic, UCSF AGS, and TCGA) as well as for the combined dataset. Both unadjusted and adjusted Kaplan-Meier survival curves were used to estimate overall survival for each of the five groups. The adjusted survival curves adjusted for gender and age at diagnosis (based on the 2010 US white population) using the reweighted (direct adjustment) method.<sup>24</sup> For grade II-III and grade IV gliomas separately, comparisons were made both within each dataset (Mayo Clinic, UCSF AGS, and TCGA) as well as for the combined (Mayo Clinic + UCSF AGS + TCGA) dataset. A stratified (by dataset) Cox proportional hazards model was used to determine if the molecular groups were associated with overall patient survival after adjustment for gender, age at diagnosis, histology, and grade. The stratified Cox model was first applied to the combined grade II-IV data with the following independent variables: gender, age at diagnosis, grade, molecular group, and a grade-by-molecular group interaction. Since the grade-by-molecular group interaction was statistically significant, all subsequent Cox models were generated separately within grade II-III and within grade IV gliomas. Hazard rates (HRs) and 95% confidence intervals (95% CIs) were obtained for both unadjusted and age-, gender-, and grade-adjusted stratified Cox proportional hazard models.

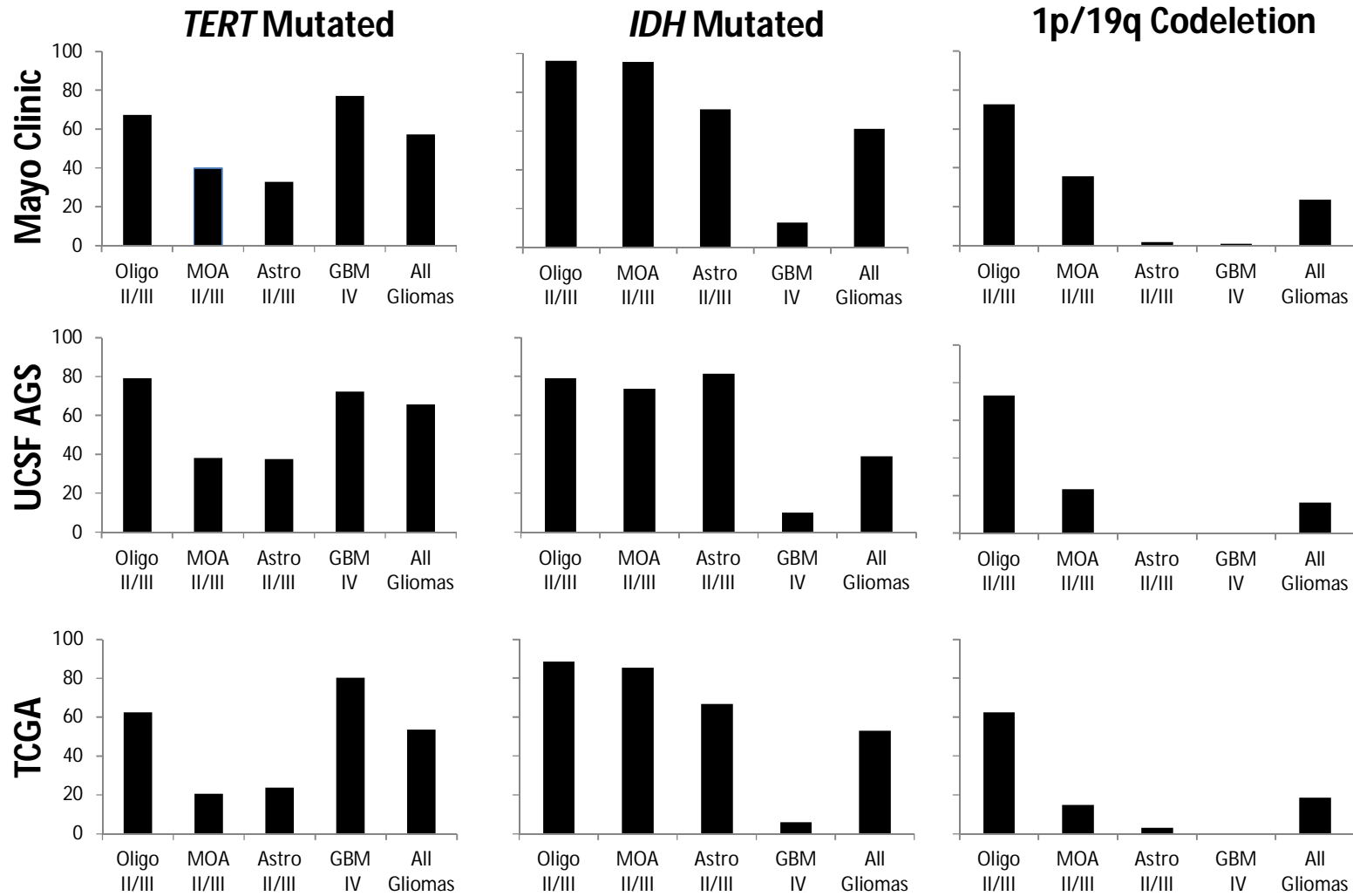
TCGA and MayoGC SNP data were phased using shapeit2<sup>25</sup> and imputed using Impute2<sup>26</sup> with 1000 Genomes<sup>27</sup> as the reference population. Prior to imputation, duplicates were removed and SNPs with more than 5% missing data were removed. Genotypes were forward-strand aligned to the 1000 genome reference and for ambiguous SNPs the Browning strand checking utility was used ([http://faculty.washington.edu/sguy/beagle/strand\\_switching/strand\\_switching.html](http://faculty.washington.edu/sguy/beagle/strand_switching/strand_switching.html)). For the SNP association analyses, we utilized the 207 Mayo Clinic cases that had *TERT* mutation, *IDH* mutation, 1p/19q codeletion, and SNP data available and these cases were compared to 789 controls. Similarly, 351 UCSF AGS cases had all required data available and these were compared with up to 4504 controls. The TCGA/MayoGC case-control study contained 402 TCGA cases that had all required data available and these were compared to the 6297

MayoGC controls. An additive logistic regression model was used to assess the association between each SNP and disease status, with genotype coded as 0, 1, or 2 copies of the minor allele. For the primary analysis, the three case-control studies were combined and analyzed adjusted for case-control study and associations were stratified by glioma molecular group. The combined analysis used the data across all three case-control studies. The exception was for the chromosome 8 and 17 SNPs, where only the Mayo Clinic and UCSF AGS case-control studies were combined (these SNPs have low minor allele frequency and thus we did not trust the estimated odds ratios that were obtained from the imputed results in the TCGA/MayoGC data). For the combined analysis of chromosome 8 and 17 SNPs, the logistic models were adjusted for age, gender, and case-control study; age and gender were available for the Mayo Clinic and UCSF AGS case-control studies but were not for the MayoGC control subjects. Overall, the primary analysis entailed evaluation of 22 SNPs that represented nine regions across eight genes that have been previously shown via GWAS to be associated with glioma risk. These nine regions were evaluated for association with risk of each of the five molecular groups and therefore a Bonferroni correction ( $\alpha=.05/45=0.0011$ ) was used to determine statistical significance in the combined analysis. Secondly, analyses were performed separately for each case-control study.

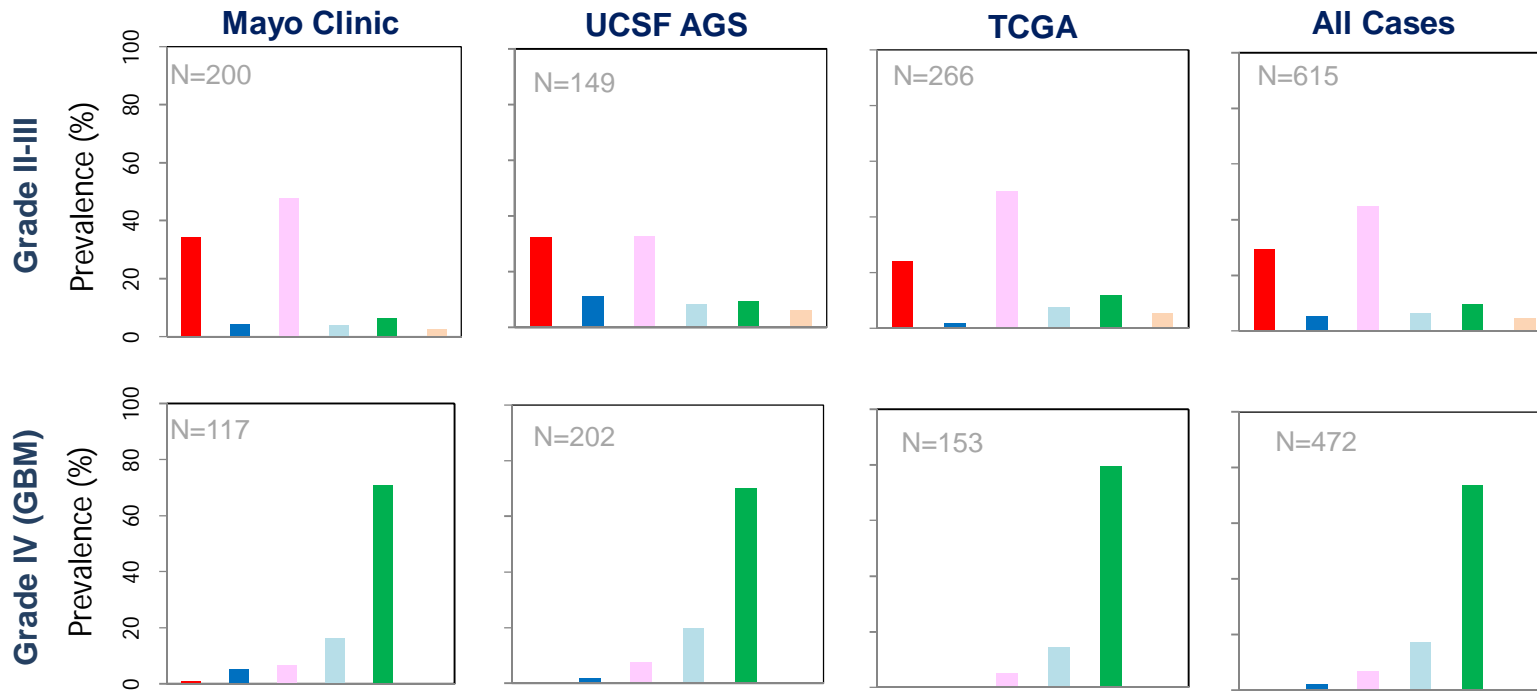
One-hundred-sixty-six Mayo Clinic and 62 UCSF AGS cases were assigned to the four TCGA GBM RNA expression subtypes using ClaNC.<sup>28</sup> To verify that the ClaNC algorithm had suitable classification accuracy, using the TCGA GBM expression data we developed a classification model using 80% of the TCGA GBM data and then applied this model to the remaining 20%, which resulted in 97% classification accuracy. The final classification model was built using all of the TCGA GBM expression data. For the SNP associations stratified by the TCGA GBM RNA expression subtypes, we utilized the 149 glioma cases from the Mayo Clinic case-control study that had both SNP data as well as DASL gene expression data available and these were compared to 789 controls.

## SUPPLEMENTARY FIGURES

**Figure S1:** Prevalence of *TERT* promoter mutation, *IDH* mutation, and 1p/19q codeletion in the Mayo Clinic, UCSF AGS, TCGA, and combined cases. Data are stratified by histologic type for the oligodendrogliomas and mixed oligoastrocytomas, and for astrocytomas by grade (grade II-III and grade IV). Abbreviations: Oligo=oligodendrogloma, MOA=mixed oligoastrocytoma, Astro=astrocytoma, GBM=glioblastoma.



**Figure S2:** Prevalence of the glioma molecular groups as defined by *TERT* promoter mutation, *IDH* mutation, and 1p/19q codeletion status in the Mayo Clinic, UCSF AGS, and TCGA cases. The prevalence of the molecular groups for grade II-III (astrocytomas, mixed oligoastrocytomas and oligodendrogliomas) and grade IV (glioblastoma multiforme or GBM) gliomas are shown. The key indicates how to interpret the different colored bars that represent the results for the different groups.

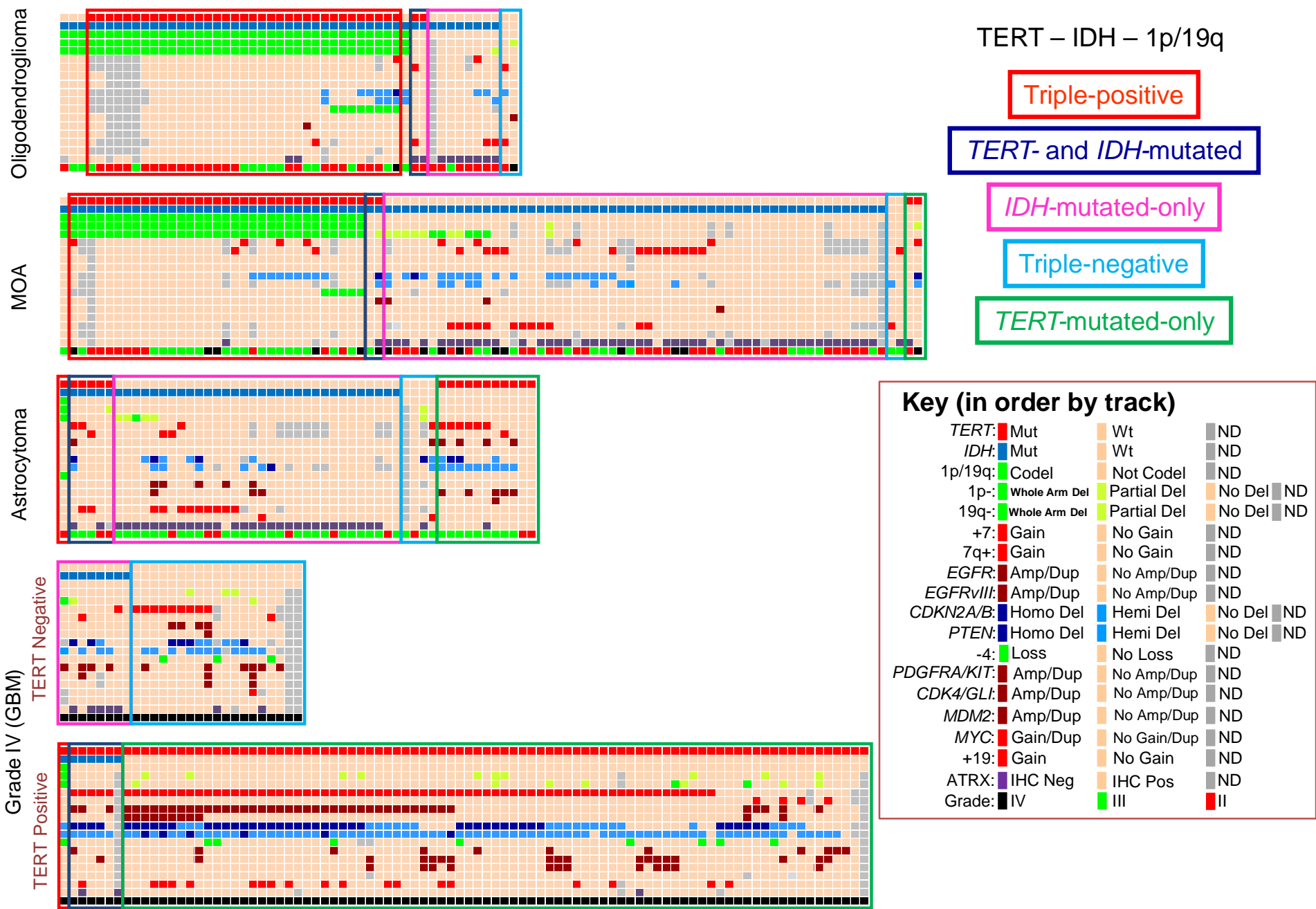


**KEY**

Bar	TERT mutated	IDH mutated	1p/19q codelet	Molecular group
Red	X	X	X	Triple-positive
Blue	X	X		<i>TERT</i> - and <i>IDH</i> -mutated
Pink		X		<i>IDH</i> -mutated-only
Light Blue				Triple-negative
Green	X			<i>TERT</i> -mutated-only
Orange		(other combinations)		Other

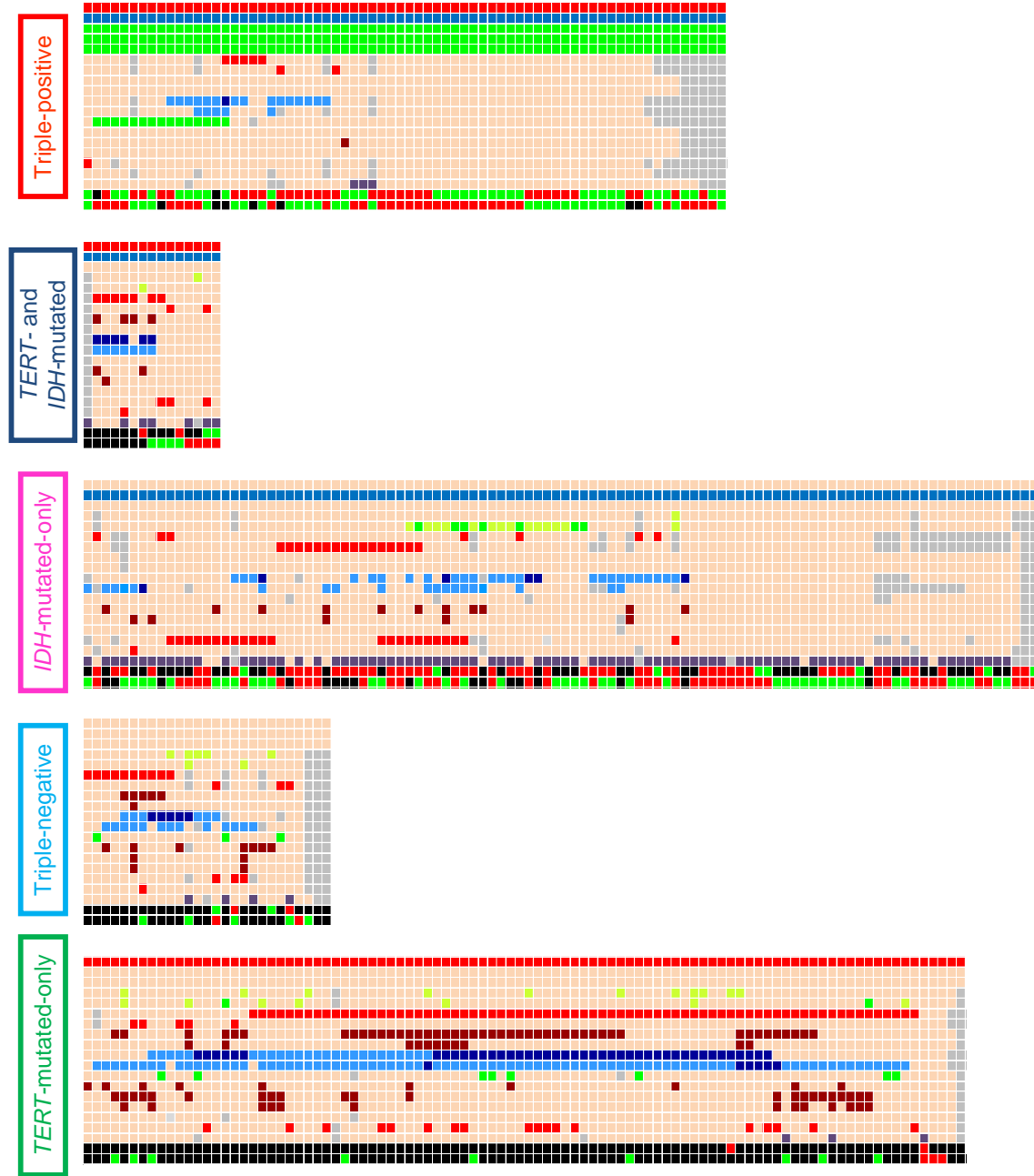


**Figure S3:** Pattern of *TERT* promoter mutation, *IDH* mutation, and 1p/19q codeletion status among 317 Mayo Clinic gliomas. The data are organized by histologic type. Displayed are track summaries of selected copy-number results as well as the presence or absence of ATRX expression by immunohistochemistry. The inset key indicates how to interpret the tracks. Colored boxes indicate the glioma molecular group: Red=triple-positive (*TERT* promoter mutated, *IDH* mutated, and 1p/19q codeleted) gliomas; Blue=*TERT*- and *IDH*-mutated gliomas; Pink=*IDH*-mutated-only gliomas; Aqua=triple-negative gliomas; and Green=*TERT*-mutated-only gliomas. Abbreviations: MOA=mixed oligoastrocytoma, GBM=glioblastoma, Mut=mutated; Wt=wild type; Amp=gene amplification; Dup=gene duplication; Codel=codeletion; Del=deletion; hemi=hemizygous; homo=homozygous; IHC=immunohistochemistry; ND=not done/failed/equivocal.



**Figure S4:** Molecular characterization of the glioma molecular groups defined by *TERT* promoter mutation, *IDH* mutation, and 1p/19q codeletion status for (A) Mayo Clinic, (B) UCSF AGS, and (C) TCGA cases. The data are organized by glioma molecular group: Red=triple-positive (*TERT* promoter mutated, *IDH* mutated, and 1p/19q codeleted) gliomas; Blue=*TERT*- and *IDH*-mutated gliomas; Pink=*IDH*-mutated-only gliomas; Aqua=triple-negative gliomas; and Green=*TERT*-mutated-only gliomas. Also displayed are track summaries of selected copy-number results, point mutations, and the presence or absence of ATRX expression by immunohistochemistry (Mayo Clinic and UCSF AGS) or ATRX mutation by sequencing (TCGA), as available for each case series. The inset key indicates how to interpret the tracks. Abbreviations: MOA=mixed oligoastrocytoma, GBM=glioblastoma, Mut=mutated; Wt=wild type; Amp=gene amplification; Dup=gene duplication; Codel=codeletion; Del=deletion; hemi=hemizygous; homo=homozygous; IHC=immunohistochemistry; ND=not done/failed/equivocal.

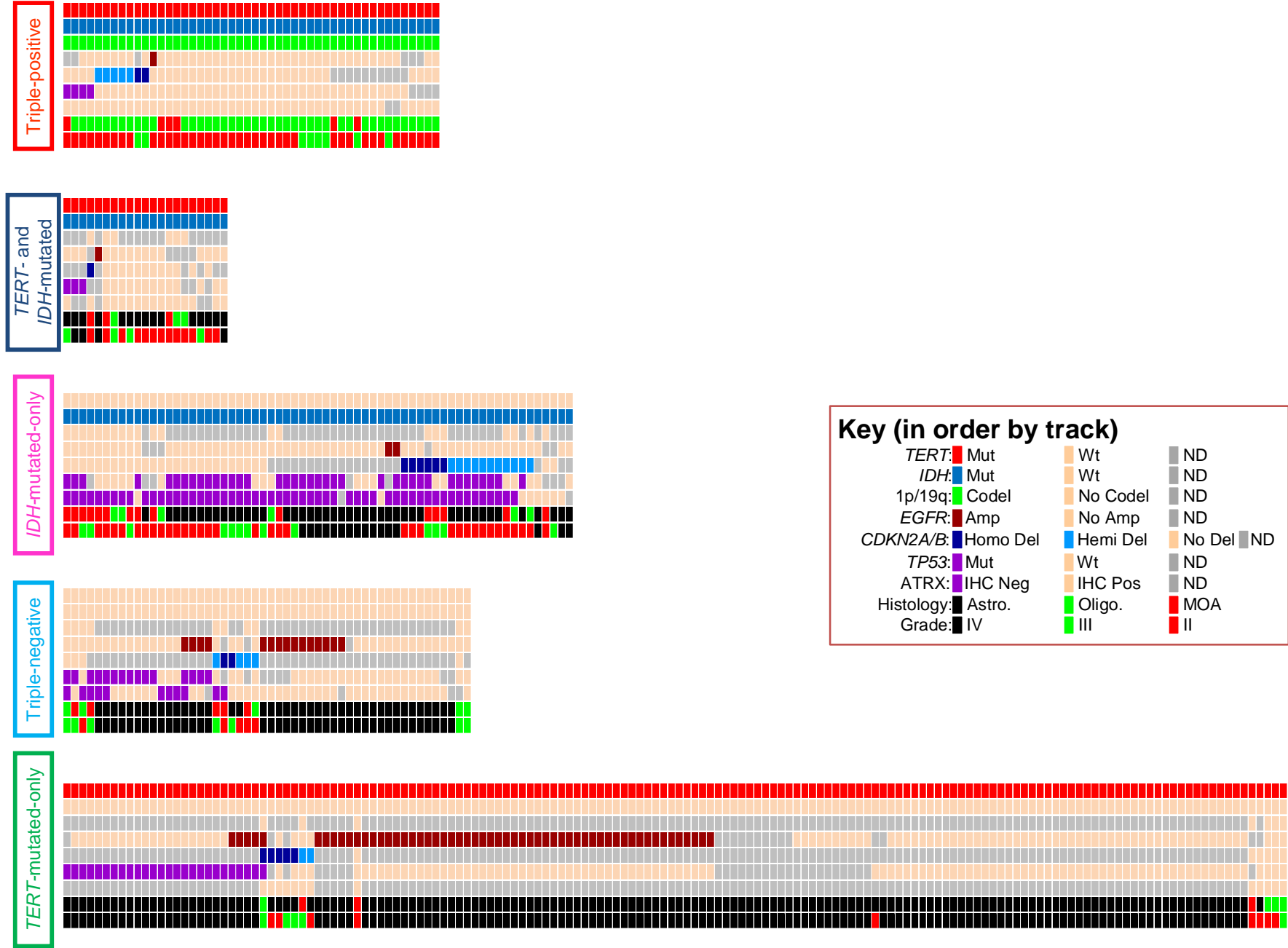
A



**Key (in order by track)**

<i>TERT</i> : Mut	Wt	ND
<i>IDH</i> : Mut	Wt	ND
1p/19q: Codel	Not Codel	ND
1p-: Whole Arm Del	Partial Del	No Del ND
19q-: Whole Arm Del	Partial Del	No Del ND
+7: Gain	No Gain	ND
7q+: Gain	No Gain	ND
<i>EGFR</i> : Amp/Dup	No Amp/Dup	ND
<i>EGFRvIII</i> : Amp/Dup	No Amp/Dup	ND
<i>CDKN2A/B</i> : Homo Del	Hemi Del	No Del ND
<i>PTEN</i> : Homo Del	Hemi Del	No Del ND
-4: Loss	No Loss	ND
<i>PDGFRA/KIT</i> : Amp/Dup	No Amp/Dup	ND
<i>CDK4/GLI</i> : Amp/Dup	No Amp/Dup	ND
<i>MDM2</i> : Amp/Dup	No Amp/Dup	ND
<i>MYC</i> : Gain/Dup	No Gain/Dup	ND
+19: Gain	No Gain	ND
<i>ATRX</i> : IHC Neg	IHC Pos	ND
Histology: Astro	Oligo	MOA
Grade: IV	III	II

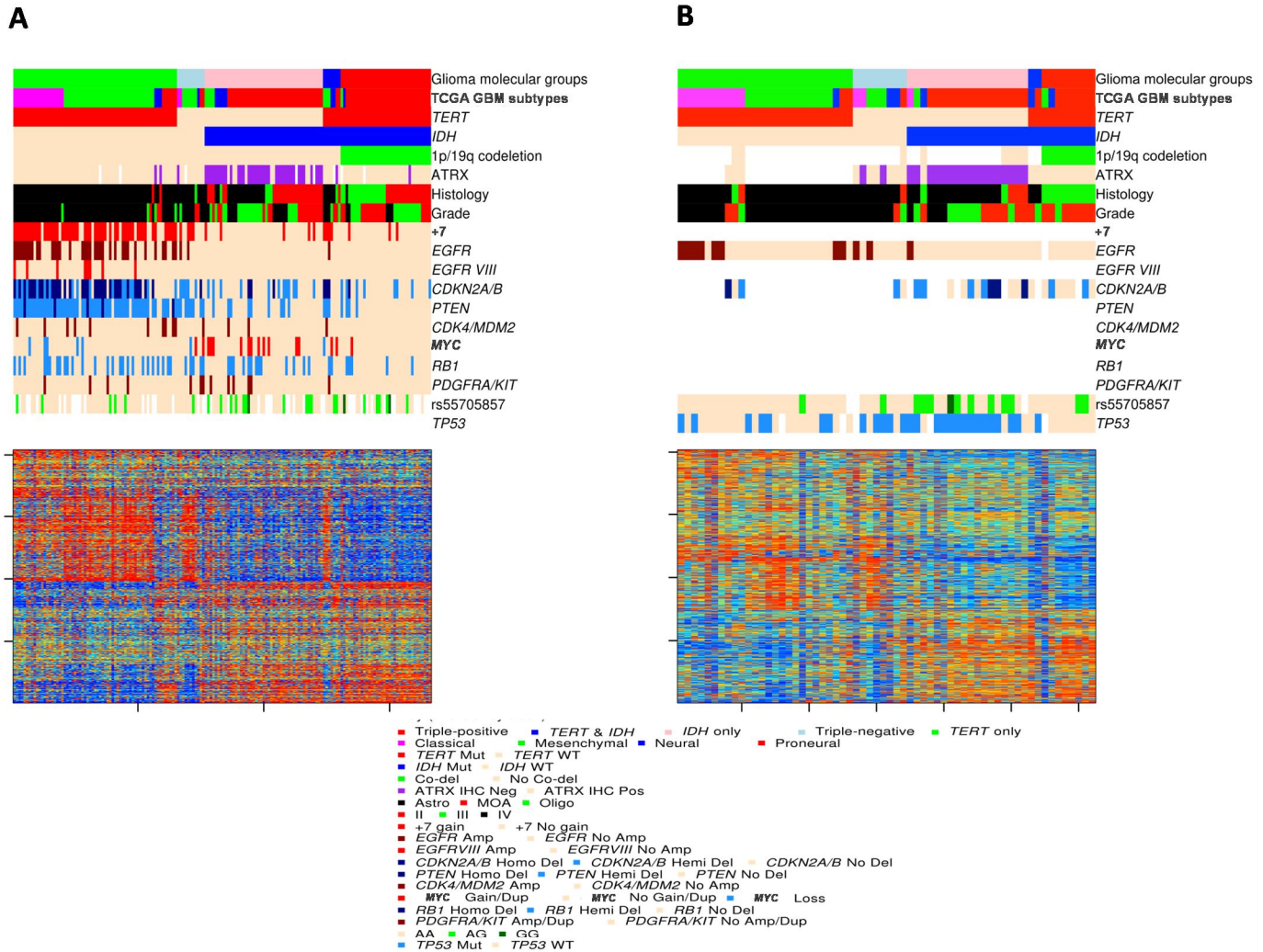
**B**



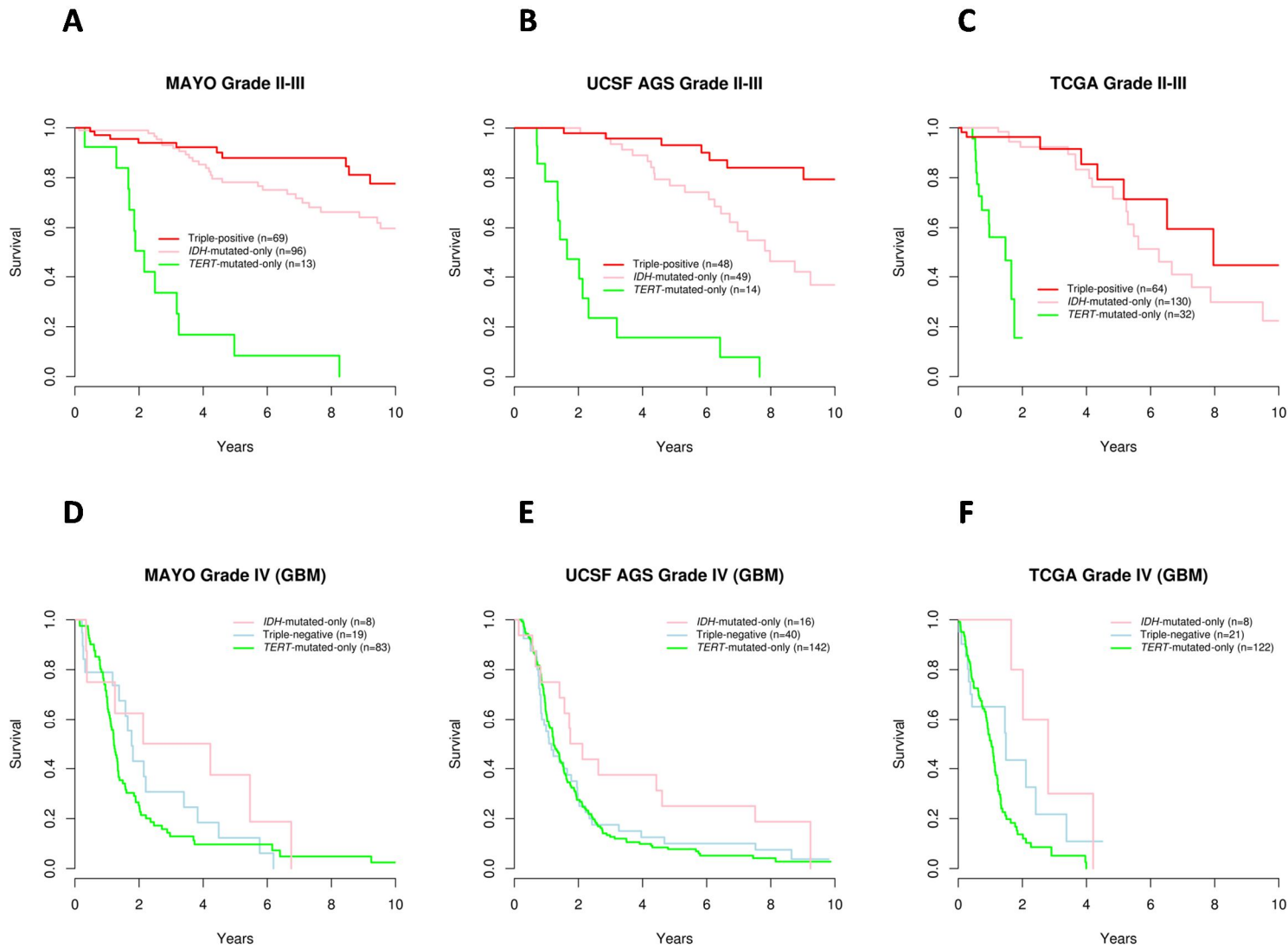
C



**Figure S5:** Comparison of the glioma molecular groups defined by *TERT* promoter mutation, *IDH* mutation, and 1p/19q codeletion status versus the TCGA GBM RNA expression subtypes for (A) 166 Mayo Clinic cases and (B) 62 UCSF AGS cases. The cases were assigned to TCGA GBM RNA expression subtypes as described in the Supplementary Methods. The top track illustrates the glioma molecular group as defined by *TERT* promoter mutation, *IDH* mutation, and 1p/19q codeletion status: Red=triple-positive (*TERT* promoter mutated, *IDH* mutated, and 1p/19q codeleted) gliomas; Blue=*TERT*- and *IDH*-mutated gliomas; Pink=*IDH*-mutated-only gliomas; Aqua=triple-negative gliomas; and Green=*TERT*-mutated-only gliomas. The second track provides the TCGA GBM expression subtypes: Pink=Classical; Green=Mesenchymal; Blue=Neural; and Red=Proneural. The inset Key defines the remaining tracks. The heat map was created using the genes TCGA used to define the GBM RNA expression subtypes.

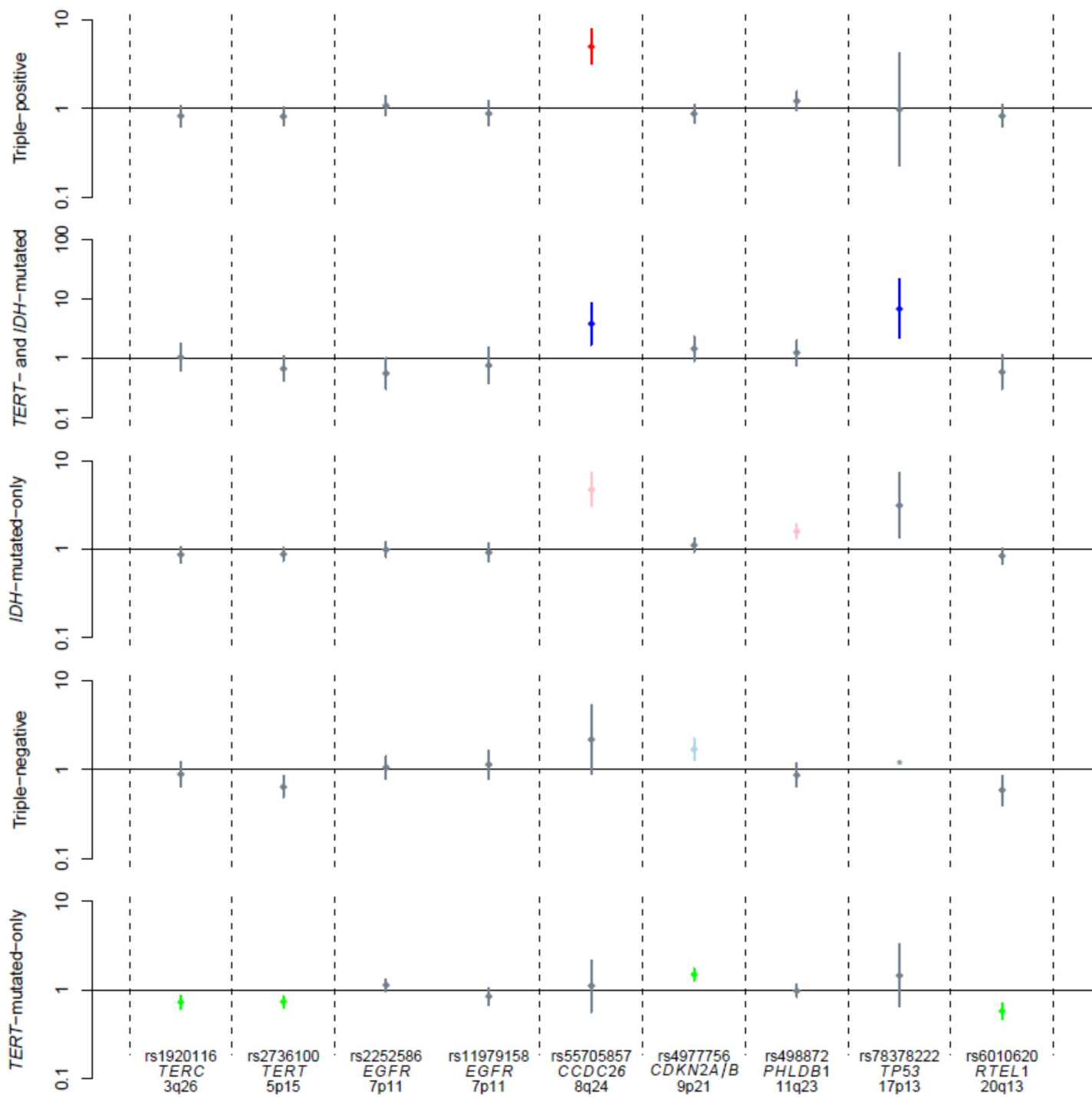


**Figure S6:** Unadjusted Kaplan-Meier estimates of overall survival of the glioma molecular groups defined by *TERT* promoter mutation, *IDH* mutation, and 1p/19q codeletion status. Survival curves are provided for groups with a prevalence of 8 or more glioma cases. Overall survival for the triple-positive, *IDH*-mutated-only, and *TERT*-mutated-only grade II-III gliomas in the (A) Mayo Clinic, (B) UCSF AGS, and (C) TCGA cases. Overall survival for the *IDH*-mutated-only, triple-negative, and *TERT*-mutated-only grade IV gliomas in the (D) Mayo Clinic, (E) UCSF AGS, and (F) TCGA cases.





**Figure S7:** Odds ratios for case-control SNP associations of nine regions with development of the five glioma molecular groups as defined by *TERT* promoter mutation, *IDH* mutation, and 1p/19q codeletion status. Each subfigure (from top-to-bottom) indicates one molecular group: Red=triple-positive (*TERT* promoter mutated, *IDH* mutated, and 1p/19q codeleted) gliomas; Blue=*TERT*- and *IDH*-mutated gliomas; Pink=*IDH*-mutated-only gliomas; Aqua=triple-negative gliomas; and Green=*TERT*-mutated-only gliomas. Vertical solid colored lines indicate 95% confidence intervals associated with the odds ratios. Between each pair of dashed lines are the results for each associated gene region and associated chromosomal band. SNPs with p-value  $\leq 0.0011$  (Bonferroni corrected p-value for testing nine regions in each of the five molecular groups) are denoted in color whereas SNPs with p-value  $> 0.0011$  are denoted in grey. The asterisk indicates that there were too few cases in the triple-negative group to assess the rs78378222 variant. For this figure the cases and controls from Mayo Clinic, UCSF AGS, and TCGA/MayoGC were combined. See Table S6 in Supplementary Appendix for a complete summary of the data for all 22 SNPs evaluated.



**Table S1:** TCGA data downloaded for this project

Data Type	Source	Platform	TCGA Level*	# GBM**	# LGG**	GBM Download Date	LGG Download Date
<b>Germline SNPs</b>							
Imputed germline variants (LGG & GBM)	Mayo Clinic repository (BORA <sup>11</sup> )	Affymetrix 6.0		512	297	28-Feb-13	23-Dec-13
<b>Copy Number and EGFRvIII</b>							
Probe level copy number on tumors	TCGA Data Matrix <sup>29</sup>	Affymetrix 6.0	Levels 1 & 2	538	387	13-Nov-13	13-Nov-13
<i>EGFRvIII</i>	Brennan et al. <sup>10</sup>			164	0		
<b>MGMT Methylation Status</b>							
<i>MGMT</i> methylation	Brennan et al. <sup>10</sup>			351			
Methylation	Broad firehose sttdata <sup>30</sup>	Illumina 450K	Level 3	135	408	10-Dec-13	10-Dec-13
Methylation	Broad firehose sttdata <sup>30</sup>	Illumina 27K	Level 3	285		10-Dec-13	10-Dec-13
<b>TERT Promoter Mutation Status</b>							
RSEM <sup>18</sup> normalized & raw gene counts***	Broad firehose sttdata <sup>30</sup>	RNAseq	Level 3	166	275	15-Oct-13	15-Oct-13
<b>Somatic Mutations</b>							
Somatic mutations	TCGA Data Matrix <sup>29</sup>	Sequencing	Level 2	291	296	18-Feb-14	18-Feb-14
<b>Clinical Data</b>	TCGA Data Matrix <sup>29</sup>			583	421	28-08-14	28-08-14

\* TCGA levels range from Level 1 (raw data) to Level 3 (highest platform-specific preprocessed data). <https://tcga-data.nci.nih.gov/tcga/tcgaDataType.jsp>

\*\* The number of samples does not include duplicates, i.e., it denotes the number of independent subjects. The numbers provided above denote the number of samples that were downloaded. However, only 153 GBM and 266 LGG cases could be assigned to one of the five molecular groups and thus were analyzed in this manuscript.

\*\*\* RSEM: RNA-Seq by Expectation Maximization<sup>18</sup>, is an algorithms adopted by the Broad TCGA group to quantify RNA-Seq transcript counts using reference transcript. This method accounts and corrects for uncertainties in mapping due to highly homologous sequences in the reference transcript set.

**Table S2A:** Distribution of histologic type and grade in the Mayo Clinic, UCSF AGS, and TCGA cases stratified by molecular group

	All Cases	Triple-positive	TERT- and IDH-mutated	IDH-mutated-only	Triple-negative	TERT-mutated-only	Other
<b>TERT Promoter Mutated</b>		Yes	Yes	No	No	Yes	
<b>IDH Mutated</b>		Yes	Yes	Yes	No	No	
<b>1p/19q Codeleted</b>		Yes	No	No	No	No	
<b>Mayo Clinic Cases</b>							
N	317	70	15	104	27	96	5
AII	12 (4%)	1 (1%)	2 (13%)	6 (6%)	1 (4%)	2 (2%)	0 (0%)
AIII	40 (13%)	0 (0%)	3 (20%)	25 (24%)	3 (11%)	9 (9%)	0 (0%)
AIV	117 (37%)	1 (1%)	6 (40%)	8 (8%)	19 (70%)	83 (86%)	0 (0%)
MOAII	39 (12%)	10 (14%)	0 (0%)	28 (27%)	0 (0%)	1 (1%)	0 (0%)
MOAIII	38 (12%)	18 (26%)	1 (7%)	17 (16%)	2 (7%)	0 (0%)	0 (0%)
MOAIV	16 (5%)	5 (7%)	1 (7%)	9 (9%)	0 (0%)	1 (1%)	0 (0%)
OII	37 (12%)	23 (33%)	2 (13%)	10 (10%)	1 (4%)	0 (0%)	1 (20%)
OIII	16 (5%)	11 (16%)	0 (0%)	1 (1%)	0 (0%)	0 (0%)	4 (80%)
OIV	2 (1%)	1 (1%)	0 (0%)	0 (0%)	1 (4%)	0 (0%)	0 (0%)
<b>UCSF AGS Cases</b>							
N	351	48	21	65	52	156	9
AII	32 (9%)	0 (0%)	8 (38%)	18 (28%)	1 (2%)	5 (3%)	0 (0%)
AIII	16 (5%)	0 (0%)	3 (14%)	10 (15%)	1 (2%)	2 (1%)	0 (0%)
AIV	202 (58%)	0 (0%)	4 (19%)	16 (25%)	40 (77%)	142 (91%)	0 (0%)
MOAII	26 (7%)	5 (10%)	3 (14%)	12 (18%)	2 (4%)	2 (1%)	2 (22%)
MOAIII	8 (2%)	1 (2%)	0 (0%)	3 (5%)	3 (6%)	1 (1%)	0 (0%)
MOAIV	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
OII	50 (14%)	35 (73%)	2 (10%)	6 (9%)	2 (4%)	2 (1%)	3 (33%)
OIII	17 (5%)	7 (15%)	1 (5%)	0 (0%)	3 (6%)	2 (1%)	4 (44%)
OIV	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
<b>TCGA Cases</b>							
N	419	64	6	139	42	154	14
AII	30 (7%)	1 (2%)	0 (0%)	24 (17%)	5 (12%)	0 (0%)	0 (0%)
AIII	63 (15%)	0 (0%)	0 (0%)	36 (26%)	4 (10%)	21 (14%)	2 (14%)
AIV	153 (37%)	0 (0%)	1 (17%)	8 (6%)	22 (52%)	122 (79%)	0 (0%)
MOAII	36 (9%)	5 (8%)	1 (17%)	27 (19%)	1 (2%)	0 (0%)	2 (14%)
MOAIII	32 (8%)	2 (3%)	1 (17%)	19 (14%)	4 (10%)	5 (3%)	1 (7%)
MOAIV	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
OII	61 (15%)	31 (48%)	2 (33%)	19 (14%)	3 (7%)	2 (1%)	4 (29%)
OIII	43 (10%)	25 (39%)	1 (17%)	5 (4%)	3 (7%)	4 (3%)	5 (36%)
OIV	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Missing	1 (0.2%)	0 (0%)	0 (0%)	1 (1%)	0 (0%)	0 (0%)	0 (0%)
<b>All Cases</b>							
N	1087	182	42	308	121	406	28
AII	74 (7%)	2 (1%)	10 (24%)	48 (16%)	7 (6%)	7 (2%)	0 (0%)
AIII	119 (11%)	0 (0%)	6 (14%)	71 (23%)	8 (7%)	32 (8%)	2 (7%)
AIV	472 (43%)	1 (1%)	11 (26%)	32 (10%)	81 (67%)	347 (85%)	0 (0%)
MOAII	101 (9%)	20 (11%)	4 (10%)	67 (22%)	3 (2%)	3 (1%)	4 (14%)
MOAIII	78 (7%)	21 (12%)	2 (5%)	39 (13%)	9 (7%)	6 (1%)	1 (4%)
MOAIV	16 (1%)	5 (3%)	1 (2%)	9 (3%)	0 (0%)	1 (0.2%)	0 (0%)
OII	148 (14%)	89 (49%)	6 (15%)	35 (11%)	6 (5%)	4 (1%)	8 (29%)
OIII	76 (7%)	43 (24%)	2 (5%)	6 (2%)	6 (5%)	6 (1%)	13 (46%)
OIV	2 (0.2%)	1 (0.5%)	0 (0%)	0 (0%)	1 (1%)	0 (0%)	0 (0%)
Missing	1 (0.1%)	0 (0%)	0 (0%)	1 (0.3%)	0 (0%)	0 (0%)	0 (0%)

Abbreviations: A = Astrocytoma, MOA = Mixed Oligoastrocytoma; O = Oligodendrogliomas

**Table S2B:** Age at diagnosis, gender, tumor location, and extent of resection for the Mayo Clinic, UCSF AGS, and TCGA cases stratified by molecular group

	All Cases	Triple-positive	TERT- and IDH-mutated	IDH-mutated-only	Triple-negative	TERT-mutated-only	Other
<b>TERT Promoter Mutated</b>		Yes	Yes	No	No	Yes	
<b>IDH Mutated</b>		Yes	Yes	Yes	No	No	
<b>1p/19q Codeleted</b>		Yes	No	No	No	No	
<b>Mayo Clinic Cases</b>							
N	317	70	15	104	27	96	5
Median age	45	41	39	35.5	48	59	38
Mean age ± SD	46±14	43±11	48±17	38±10	46±15	58±11	42±13
% Male	63	56	73	68	56	65	40
Extent of resection							
% Biopsy	5	14	7	1	7	2	0
% STR/GTR*	95	86	93	99	93	98	100
Tumor location**							
% Frontal lobe	68	83	53	76	63	52	80
% Parietal lobe	11	7	20	9	11	16	20
% Temporal lobe	19	9	27	14	26	30	0
% Other	1	1	0	1	0	2	0
<b>UCSF AGS Cases</b>							
N	351	48	21	65	52	156	9
Median age	51	43.5	44	35	52	57	48
Mean age ± SD	50±13	44±10	44±11	37±11	52±14	57±10	49±13
% Male	64	69	62	63	62	65	56
Extent of resection							
% Biopsy	3	4	10	5	2	2	0
% STR/GTR*	97	96	90	95	98	98	100
<b>TCGA Cases</b>							
N	419	64	6	139	42	154	14
Median age	49	46	48	36	49	61.5	45.5
Mean age ± SD	49±16	45±13	49±6	37±11	49±18	61±11	47±15
% Male	60	56	50	62	48	64	64
Extent of resection							
% Biopsy	6	2	17	3	5	10	0
% STR/GTR*	94	98	83	97	95	90	100
<b>All Cases</b>							
N	1087	182	42	308	121	406	28
Median age	48.5	42.5	44.5	35	51	59.5	45.5
Mean age ± SD	49±15	44±11	46±13	37±11	50±16	59±11	47±14
% Male	62	59	64	65	55	64	57
Extent of resection							
% Biopsy	5	7	10	3	4	5	0
% STR/GTR*	95	93	90	97	96	95	100

\*STR denotes subtotal resection and GTR denotes gross total resection.

\*\*Tumor Location data were consistently available and quality checked only for the Mayo Clinic gliomas

**Table S2C:** Association between molecular group and age at diagnosis for the Mayo Clinic, UCSF AGS, and TCGA cases\*

Pairwise Comparison		Site			
		Mayo Clinic	UCSF AGS	TCGA	All Cases
Triple-positive	<i>TERT-</i> and <i>IDH</i> -mutated	0.1223	0.8474	0.5436	0.2043
Triple-positive	<i>IDH</i> -mutated-only	0.0014	0.003	<0.0001	<0.0001
Triple-positive	Triple-negative	0.3037	0.0003	0.1165	<0.0001
Triple-positive	<i>TERT</i> -mutated-only	<0.0001	<0.0001	<0.0001	<0.0001
<i>TERT-</i> and <i>IDH</i> -mutated	<i>IDH</i> -mutated-only	0.0008	0.0142	0.0295	<0.0001
<i>TERT-</i> and <i>IDH</i> -mutated	Triple-negative	0.5199	0.008	0.9037	0.1626
<i>TERT-</i> and <i>IDH</i> -mutated	<i>TERT</i> -mutated-only	0.003	<0.0001	0.0132	<0.0001
<i>IDH</i> -mutated-only	Triple-negative	0.0008	<0.0001	<0.0001	<0.0001
<i>IDH</i> -mutated-only	<i>TERT</i> -mutated-only	<0.0001	<0.0001	<0.0001	<0.0001
Triple-negative	<i>TERT</i> -mutated-only	<0.0001	0.0017	<0.0001	<0.0001

\* P-values are reported for each pairwise comparison. Mean ages are provided in Supplemental Table S2B. P-values were obtained from contrast statements created from an ANOVA model that was run for each site separately. P-values < 0.005 are highlighted in yellow and 0.005<p-values<0.05 are highlighted in orange.



**Table S4:** Hazard rate (HR) and 95% confidence interval (95% CI) for all pairwise molecular group comparisons from a stratified (by dataset) Cox proportional hazards model

Molecular Group 1	Molecular Group 2	Unadjusted HR* (95% CI)	Age-Adjusted HR* (95% CI)	Age- and Grade- Adjusted HR* (95% CI)
<b>Grade II-III</b>				
<i>TERT</i> -mutated-only	Triple-negative	5.41 (2.88 - 10.18)	3.17 (1.64 - 6.18)	3.64 (1.85 - 7.17)
<i>TERT</i> -mutated-only	<i>TERT</i> - and <i>IDH</i> -mutated	13.36 (6.01 - 29.67)	9.50 (4.24 - 21.28)	9.06 (4.02 - 20.43)
<i>TERT</i> -mutated-only	<i>IDH</i> -mutated-only	10.34 (6.64 - 16.1)	6.12 (3.73 - 10.02)	5.83 (3.56 - 9.57)
<i>TERT</i> -mutated-only	Triple-positive	21.92 (12.75 - 37.70)	15.18 (8.66 - 26.59)	14.26 (8.12 - 25.07)
Triple-negative	<i>TERT</i> - and <i>IDH</i> -mutated	2.47 (1.04 - 5.88)	2.99 (1.25 - 7.13)	2.49 (1.03 - 6.02)
Triple-negative	<i>IDH</i> -mutated-only	1.91 (1.08 - 3.39)	1.92 (1.07 - 3.44)	1.60 (0.88 - 2.92)
Triple-negative	Triple-positive	4.05 (2.13 - 7.71)	4.77 (2.48 - 9.17)	3.92 (2.0 - 7.67)
<i>IDH</i> -mutated-only	<i>TERT</i> - and <i>IDH</i> -mutated	1.29 (0.6 - 2.76)	1.55 (0.72 - 3.36)	1.55 (0.72 - 3.37)
<i>IDH</i> -mutated-only	Triple-positive	2.12 (1.33 - 3.37)	2.48 (1.56 - 3.96)	2.44 (1.53 - 3.90)
<i>TERT</i> - and <i>IDH</i> -mutated	Triple-positive	1.64 (0.73 - 3.69)	1.60 (0.71 - 3.61)	1.57 (0.70 - 3.56)
<b>Grade IV</b>				
<i>TERT</i> -mutated-only	Triple-negative	1.25 (0.95 - 1.63)	0.99 (0.75 - 1.31)	
<i>TERT</i> -mutated-only	<i>TERT</i> - and <i>IDH</i> -mutated	2.19 (1.03 - 4.67)	1.59 (0.74 - 3.41)	
<i>TERT</i> -mutated-only	<i>IDH</i> -mutated-only	2.12 (1.40 - 3.21)	1.27 (0.82 - 1.97)	
Triple-negative	<i>TERT</i> - and <i>IDH</i> -mutated	1.76 (0.80 - 3.85)	1.60 (0.73 - 3.51)	
Triple-negative	<i>IDH</i> -mutated-only	1.7 (1.07 - 2.70)	1.28 (0.80 - 2.05)	
<i>IDH</i> -mutated-only	<i>TERT</i> - and <i>IDH</i> -mutated	1.03 (0.44 - 2.40)	1.25 (0.54 - 2.91)	

\*Hazard Rate (HR) denotes the hazard of molecular group 1 relative to molecular group 2. A HR with a 95% CI that does not include one is denoted in bold font.



**Table S5: Association of 22 SNPs in 9 independent regions known to be associated with glioma in the Mayo Clinic, UCSF AGS, and TCGA/MayoGC case-control studies, stratified by molecular group\***

Mayo Clinic (controls n=789)		Triple-positive (n=43)				TERT- and IDH-mutated (n=9)				IDH-mutated-only (n=74)				Triple-negative (n=16)				TERT-mutated-only (n=65)				All Mayo Gliomas (n=207)												
chr	snp	bp	a1	MAF_co	OR	IS	US	p	MAF_ca	OR	IS	US	p	MAF_ca	OR	IS	US	p	MAF_ca	OR	IS	US	p	MAF_ca	OR	IS	US	p	MAF_ca					
3	rs1920116	169579971	A	0.298	0.52	0.30	0.92	0.0250	0.179	0.70	0.24	2.10	0.5278	0.222	1.18	0.80	1.72	0.4062	0.321	0.72	0.31	1.66	0.4365	0.233	0.86	0.57	1.30	0.4753	0.262	0.83	0.65	1.06	0.1399	0.261
5	rs2736100	1286516	A	0.518	0.75	0.47	1.18	0.2069	0.465	0.33	0.01	1.98	0.0469	0.278	0.59	0.40	0.88	0.0084	0.432	0.68	0.32	1.41	0.2977	0.438	0.81	0.56	1.16	0.2469	0.454	0.70	0.56	0.88	0.0026	0.440
7	rs2252586	54978924	A	0.258	0.16	0.65	1.74	0.8132	0.286	0.30	0.07	1.34	0.1160	0.111	1.18	0.80	1.74	0.4164	0.311	0.66	0.27	1.65	0.3775	0.200	1.23	0.83	1.83	0.2939	0.323	1.09	0.85	1.40	0.0025	0.293
7	rs6969537	55082418	A	0.145	0.99	0.55	1.79	0.9704	0.151	1.19	0.35	4.02	0.7848	0.167	1.04	0.65	1.64	0.8793	0.169	0.81	0.29	2.28	0.6846	0.125	1.10	0.66	1.83	0.7125	0.146	1.05	0.78	1.41	0.7603	0.155
7	rs1015793	55114316	G	0.151	0.64	0.32	1.30	0.2186	0.105	1.70	0.55	2.26	0.3545	0.222	0.94	0.57	1.53	0.7886	0.149	0.56	0.17	1.86	0.3432	0.094	0.87	0.51	1.50	0.6225	0.131	0.87	0.63	1.19	0.3732	0.133
7	rs11971958	55159349	G	0.174	0.62	0.32	1.20	0.1543	0.116	0.98	0.28	3.41	0.9737	0.167	0.84	0.52	1.36	0.4816	0.162	0.66	0.23	1.88	0.4392	0.125	0.94	0.58	1.53	0.7933	0.162	0.84	0.62	1.13	0.2429	0.150
8	rs72714236	13048065	A	0.051	<b>2.83</b>	<b>1.47</b>	<b>5.47</b>	<b>0.0020</b>	0.140	1.07	0.14	7.91	0.9512	0.056	3.99	2.24	<b>7.12</b>	<b>2.68E-06</b>	0.155	0.59	0.08	4.37	0.6054	0.031	0.80	0.32	1.97	0.6226	<b>0.038</b>	<b>2.05</b>	<b>1.37</b>	<b>3.05</b>	<b>0.00042</b>	<b>0.101</b>
8	rs891835	130491752	C	0.208	<b>1.92</b>	<b>1.21</b>	<b>3.06</b>	<b>0.0059</b>	0.349	0.78	0.22	2.73	0.6988	0.167	1.69	1.13	2.52	0.0105	0.314	1.02	0.43	2.40	0.9666	0.219	0.89	0.55	1.42	0.6174	0.185	1.35	1.04	1.74	0.0222	0.266
8	rs72714295	130569398	A	0.066	<b>2.26</b>	<b>1.22</b>	<b>4.21</b>	<b>0.0098</b>	0.151	1.63	0.36	7.29	0.5264	0.111	2.92	1.72	<b>4.98</b>	<b>8.04E-05</b>	0.169	0.92	0.22	3.89	0.9117	0.063	0.83	0.38	1.83	0.6483	0.055	1.79	1.25	<b>2.58</b>	<b>0.0017</b>	<b>0.119</b>
8	rs72714302	130588045	C	0.049	<b>2.74</b>	<b>1.44</b>	<b>5.21</b>	<b>0.0021</b>	0.140	1.01	0.14	7.59	0.9890	0.056	3.35	1.89	<b>5.96</b>	<b>3.62E-05</b>	0.142	0.62	0.08	4.61	0.6421	0.031	0.63	0.23	1.73	0.3661	0.031	<b>1.85</b>	<b>1.24</b>	<b>2.77</b>	<b>0.0027</b>	<b>0.094</b>
8	rs72716319	130599332	G	0.050	<b>2.69</b>	<b>1.41</b>	<b>5.12</b>	<b>0.0026</b>	0.140	0.99	0.13	7.44	0.9927	0.056	3.18	1.79	<b>5.65</b>	<b>5.58E-05</b>	0.142	0.61	0.08	4.50	0.6243	0.031	0.78	0.31	1.96	0.5969	0.038	<b>1.88</b>	<b>1.26</b>	<b>2.80</b>	<b>0.0020</b>	<b>0.097</b>
8	rs72716328	130609332	A	0.046	<b>2.74</b>	<b>1.40</b>	<b>5.38</b>	<b>0.0034</b>	0.128	1.07	0.14	7.98	0.9447	0.056	3.43	1.93	<b>6.10</b>	<b>2.63E-05</b>	0.142	0.66	0.09	4.59	0.6947	0.031	0.72	0.26	1.99	0.5322	0.031	<b>1.95</b>	<b>1.29</b>	<b>2.95</b>	<b>0.0016</b>	<b>0.092</b>
8	rs147958197	130631395	G	0.041	<b>4.11</b>	<b>2.10</b>	<b>8.07</b>	<b>3.95E-05</b>	0.151	4.35	1.16	16.31	0.0291	0.167	3.11	1.71	<b>5.64</b>	<b>0.00019</b>	0.122	0.75	0.10	5.70	0.7605	0.031	1.05	0.41	2.70	0.9172	0.038	<b>2.34</b>	<b>1.54</b>	<b>3.51</b>	<b>6.47E-05</b>	<b>0.097</b>
8	rs55708587	130645692	C	0.052	<b>3.35</b>	<b>1.82</b>	<b>6.15</b>	<b>0.00010</b>	0.174	3.20	0.90	11.37	0.0725	0.167	3.42	2.02	<b>5.78</b>	<b>4.68E-06</b>	0.176	0.56	0.07	4.18	0.5706	0.031	1.19	0.54	2.62	0.6884	0.055	<b>2.36</b>	<b>1.63</b>	<b>3.41</b>	<b>4.98E-06</b>	<b>0.126</b>
8	rs4295627	130685457	C	0.176	<b>2.37</b>	<b>1.46</b>	<b>3.85</b>	<b>0.0005</b>	0.337	0.89	0.26	3.05	0.8474	0.167	1.56	1.04	<b>2.35</b>	<b>0.0337</b>	0.264	0.17	0.43	2.65	0.8906	0.188	0.89	0.54	1.47	0.6445	0.154	1.39	1.06	1.81	0.0165	0.234
8	rs1063192	22003366	G	0.427	<b>1.18</b>	<b>0.76</b>	<b>1.84</b>	0.4613	0.465	2.20	0.84	5.74	0.1071	0.611	1.51	1.05	2.17	0.0269	0.514	0.75	0.86	3.56	0.1236	<b>0.563</b>	<b>1.61</b>	<b>1.11</b>	<b>2.34</b>	<b>0.0114</b>	<b>0.547</b>	<b>1.50</b>	<b>1.20</b>	<b>1.87</b>	<b>0.0042</b>	<b>0.522</b>
9	rs2157719	22033366	G	0.419	1.17	0.75	1.82	0.4998	0.454	2.32	0.88	6.10	0.0875	0.611	1.54	1.07	2.22	0.0217	0.507	1.81	0.89	3.70	0.1029	<b>0.563</b>	<b>1.59</b>	<b>1.10</b>	<b>2.30</b>	<b>0.0146</b>	<b>0.531</b>	<b>1.49</b>	<b>1.19</b>	<b>1.87</b>	<b>0.0046</b>	<b>0.512</b>
9	rs4977756	22068652	G	0.392	1.08	0.68	1.71	0.7462	0.405	3.30	1.21	9.10	0.0197	0.667	1.64	1.14	2.37	0.0084	0.493	1.80	0.89	3.68	0.1043	<b>0.511</b>	<b>1.47</b>	<b>1.02</b>	<b>2.13</b>	<b>0.0397</b>	<b>0.492</b>	<b>1.49</b>	<b>1.19</b>	<b>1.87</b>	<b>0.0044</b>	<b>0.485</b>
11	rs498872	11847367	A	0.329	0.99	0.63	1.56	0.9691	0.326	1.06	0.40	2.77	0.9101	0.333	1.70	<b>1.20</b>	<b>2.42</b>	<b>0.0029</b>	0.453	0.81	0.38	1.75	0.5939	0.281	1.06	0.73	1.54	0.7653	0.346	1.21	0.97	1.52	0.0886	0.374
17	rs78378222	7571752	C	0.106	0.78	0.10	5.94	0.8085	0.012	4.85	0.56	41.66	0.1504	0.056	3.93	<b>1.50</b>	<b>10.29</b>	<b>0.0054</b>	0.047	0.00	0.00	inf	0.9971	0.000	0.45	0.06	3.41	0.4367	0.008	1.64	0.77	3.50	0.2035	0.024
20	rs6010620	62309839	A	0.247	0.92	0.55	1.54	0.7543	0.233	0.59	0.17	2.08	0.4101	0.167	0.95	0.62	1.45	0.0149	0.243	0.57	0.22	1.49	0.2488	0.156	<b>0.51</b>	<b>0.31</b>	<b>0.86</b>	<b>0.0105</b>	<b>0.146</b>	<b>0.74</b>	<b>0.57</b>	<b>0.98</b>	<b>0.0334</b>	<b>0.201</b>
20	rs2297440	62312299	A	0.243	0.87	0.52	1.48	0.6161	0.221	0.60	0.17	2.10	0.4229	0.167	0.85	0.55	1.30	0.0476	0.223	0.45	0.16	1.28	0.1344	0.125	<b>0.50</b>	<b>0.30</b>	<b>0.84</b>	<b>0.0087</b>	<b>0.139</b>	<b>0.69</b>	<b>0.52</b>	<b>0.91</b>	<b>0.0087</b>	<b>0.186</b>

For the data displayed for the individual Mayo Clinic, UCSF AGS, and TCGA/MayoGC case-control studies, the bold cells indicate SNPs that are significant in at least 2 of the 3 individual case-control studies at the 0.05 level (i.e., Mayo + UCSF + TCGA; Mayo + UCSF + MayoGC; Mayo + TCGA + MayoGC; Mayo + UCSF + MayoGC + TCGA). For the data displayed in the combined analysis, the bold cells indicate p-values < 0.001 (Bonferroni correction 0.05/45). The combined analysis used the data for all 3 case-control studies except for the chromosome 8 and 17 SNPs, where only the Mayo Clinic and UCSF AGS case-control studies were combined (these SNPs have low minor allele frequency and thus we did not trust the estimated odds ratios that were obtained from the imputed results in the TCGA/MayoGC data). Because the TCGA/MayoGC case-control data were imputed and the resultant gene dosage was analyzed, the TCGA/MayoGC MAF is not reported. OR, IS and US denote odds ratio and the lower and upper 95% confidence limits for each comparison; inf indicates that the upper 95% CI was infinity. MAF\_ca and MAF\_co denote the minor allele frequencies for the cases and controls, respectively, in the Mayo and UCSF AGS case-control studies.

**Table S6:** Association of 22 SNPs in 9 independent regions known to be associated with glioma in the Mayo Clinic cases, stratified by TCGA GBM RNA expression subtypes\*

Mayo Clinic (controls n=789)					Classical (n=22)					Mesenchymal (n=52)					Neural (n=13)					Proneural (n=62)				
chr	snp	bp	a1	MAF_co	OR	l95	u95	p	MAF_ca	OR	l95	u95	p	MAF_ca	OR	l95	u95	p	MAF_ca	OR	l95	u95	p	MAF_ca
3	rs1920116	169579971	A	0.298	0.52	0.24	1.13	0.09934	0.182	1.21	0.79	1.85	0.3906	0.337	1.00	0.41	2.39	0.9933	0.292	0.73	0.48	1.13	0.1568	0.234
5	rs2736100	1286516	A	0.518	1.01	0.55	1.85	0.9819	0.523	0.93	0.62	1.40	0.7233	0.490	1.40	0.60	3.26	0.4318	0.615	<b>0.47</b>	<b>0.32</b>	<b>0.71</b>	<b>0.0003166</b>	<b>0.363</b>
7	rs2252586	54978924	A	0.276	1.00	0.51	1.94	0.9914	0.273	0.93	0.59	1.47	0.7517	0.265	0.80	0.32	2.01	0.6309	0.231	1.29	0.86	1.93	0.2169	0.328
7	rs6969537	55082418	A	0.145	1.81	0.88	3.72	0.1061	0.227	0.53	0.26	1.10	0.08882	0.077	0.70	0.21	2.28	0.5542	0.115	1.46	0.93	2.29	0.0984	0.210
7	rs1015793	55114316	G	0.151	1.45	0.68	3.08	0.3374	0.205	0.49	0.24	1.03	0.0601	0.077	0.00				0.000	1.34	0.83	2.14	0.2293	0.194
7	rs11979158	55159349	G	0.174	1.39	0.69	2.81	0.3593	0.227	0.59	0.31	1.11	0.09897	0.106	0.18	0.02	1.36	0.09755	0.038	1.17	0.74	1.85	0.5049	0.202
8	rs72714236	130468065	A	0.051	1.45	0.46	4.58	0.5292	0.068	0.95	0.38	2.38	0.9191	0.048	2.29	0.66	7.89	0.1907	0.115	<b>3.10</b>	<b>1.76</b>	<b>5.45</b>	<b>0.0008439</b>	<b>0.145</b>
8	rs891835	130491752	C	0.208	1.47	0.74	2.93	0.2767	0.273	1.25	0.78	2.02	0.3494	0.240	0.92	0.34	2.50	0.8698	0.208	1.80	1.19	2.71	0.005207	0.325
8	rs72714295	130569398	A	0.066	0.79	0.19	3.28	0.7447	0.048	1.31	0.64	2.65	0.4583	0.087	2.24	0.77	6.50	0.1375	0.154	<b>2.46</b>	<b>1.44</b>	<b>4.23</b>	<b>0.00106</b>	<b>0.153</b>
8	rs72714302	130588045	C	0.049	1.02	0.25	4.13	0.9812	0.045	1.55	0.73	3.29	0.2517	0.077	2.27	0.68	7.55	0.18	0.115	<b>3.09</b>	<b>1.76</b>	<b>5.42</b>	<b>0.0008043</b>	<b>0.145</b>
8	rs72716319	130599332	G	0.050	1.02	0.25	4.13	0.9826	0.045	1.54	0.73	3.27	0.2595	0.077	2.20	0.66	7.34	0.202	0.115	<b>3.02</b>	<b>1.73</b>	<b>5.30</b>	<b>0.0001101</b>	<b>0.145</b>
8	rs72716328	130606932	A	0.046	0.60	0.08	4.19	0.6034	0.023	1.71	0.81	3.62	0.1607	0.077	2.34	0.71	7.76	0.1638	0.115	<b>2.91</b>	<b>1.61</b>	<b>5.25</b>	<b>0.000384</b>	<b>0.129</b>
8	rs147958197	130631395	G	0.041	1.38	0.32	5.93	0.667	0.045	1.73	0.76	3.93	0.192	0.067	3.98	1.30	12.20	0.01568	0.154	<b>3.12</b>	<b>1.69</b>	<b>5.75</b>	<b>0.0002768</b>	<b>0.121</b>
8	rs55705857	130645692	G	0.052	1.04	0.26	4.20	0.9604	0.045	1.54	0.73	3.27	0.2608	0.077	2.85	0.95	8.54	0.06109	0.154	<b>3.57</b>	<b>2.11</b>	<b>6.05</b>	<b>2.31E-06</b>	<b>0.177</b>
8	rs4295627	130685457	C	0.176	0.79	0.33	1.91	0.6041	0.136	1.05	0.62	1.76	0.8589	0.183	1.35	0.53	3.44	0.5327	0.231	1.64	1.07	2.50	0.02258	0.266
9	rs1063192	22003367	G	0.427	2.66	1.37	5.16	0.003997	0.667	1.66	1.10	2.50	0.01476	0.548	2.22	0.99	4.98	0.05335	0.615	1.03	0.71	1.50	0.8677	0.427
9	rs2157719	22033366	G	0.419	2.41	1.27	4.57	0.00708	0.636	1.82	1.20	2.74	0.004524	0.558	2.74	1.19	6.29	0.01747	0.654	1.07	0.73	1.56	0.727	0.427
9	rs4977756	22068652	G	0.392	2.39	1.28	4.49	0.006501	0.614	1.83	1.21	2.75	0.003899	0.539	2.24	1.00	5.02	0.04906	0.577	0.96	0.65	1.43	0.8502	0.375
11	rs498872	118477367	A	0.329	1.24	0.67	2.29	0.4885	0.386	1.03	0.68	1.55	0.8968	0.337	1.28	0.58	2.79	0.5406	0.385	1.33	0.92	1.92	0.13	0.395
17	rs78378222	7571752	C	0.016	0.00				0.000	0.64	0.08	4.86	0.6662	0.010	2.76	0.34	22.67	0.3438	0.038	1.68	0.49	5.81	0.4129	0.024
20	rs6010620	62309839	A	0.247	0.31	0.11	0.88	0.02808	0.091	0.58	0.34	1.00	0.04937	0.164	1.12	0.46	2.72	0.8063	0.269	1.10	0.72	1.67	0.6735	0.266
20	rs2297440	62312299	A	0.243	0.33	0.12	0.92	0.03438	0.091	0.56	0.32	0.96	0.03674	0.154	0.92	0.36	2.33	0.8563	0.231	1.02	0.66	1.56	0.9419	0.250

\*Colored cells indicate SNPs that had a p-value < 0.0014 (= 0.05/36; Bonferroni correction for testing 9 regions across 4 groups). OR, l95 and u95 denote odds ratio and the lower and upper 95% confidence limits for each comparison. MAF\_ca and MAF\_co denote the minor allele frequencies for the cases and controls, respectively.

## ACKNOWLEDGEMENTS

Work at the Mayo Clinic was supported by the National Institutes of Health (grant numbers P50CA108961 and P30 CA15083), National Institute of Neurological Disorders and Stroke (grant number RC1NS068222Z), the Bernie and Edith Waterman Foundation, and the Ting Tsung and Wei Fong Chao Family Foundation. Work at University of California, San Francisco, was supported by the National Institutes of Health (grant numbers R01CA52689, P50CA097257, R25CA112355, R01CA139020, R01CA126831, and UL1 RR024131 ), as well as the National Center for Research Resources, National Center for Advancing Translational Sciences, National Brain Tumor Foundation, the Stanley D. Lewis and Virginia S. Lewis Endowed Chair in Brain Tumor Research, the UCSF Robert Magnin Newman Endowed Chair in Neuro-oncology, and by donations from families and friends of John Berardi, Helen Glaser, Elvera Olsen, Raymond E. Cooper, and William Martinusen. The collection of UCSF AGS cancer incidence data used in this study was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885; the National Cancer Institute's Surveillance, Epidemiology and End Results Program under contract HHSN261201000140C awarded to the Cancer Prevention Institute of California, contract HHSN261201000035C awarded to the University of Southern California, and contract HHSN261201000034C awarded to the Public Health Institute; and the Centers for Disease Control and Prevention's National Program of Cancer Registries, under agreement # U58DP003862-01 awarded to the California Department of Public Health. The ideas and opinions expressed herein are those of the author(s) and endorsement by the State of California Department of Public Health, the National Cancer Institute, and the Centers for Disease Control and Prevention or their Contractors and Subcontractors is not intended nor should be inferred. This publication was supported by the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through UCSF-CTSI Grant Number UL1 RR024131. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. The authors wish to acknowledge study participants, the clinicians and research staff at the participating medical centers, Katherine Cornelius, the late Dr. Bernd Scheithauer, the Mayo Clinic Center for Individualized Medicine, the Mayo Clinic Comprehensive Cancer Center Biospecimens and Processing and Genotyping Shared Resources, and the UCSF Diller Cancer Center Genomics Core. The

results published here are in part based upon data generated by The Cancer Genome Atlas managed by the NCI and NHGRI. Information about TCGA can be found at <http://cancergenome.nih.gov>.

## REFERENCES

1. Wrensch M, Jenkins RB, Chang JS, et al. Variants in the *CDKN2B* and *RTEL1* regions are associated with high-grade glioma susceptibility. *Nat Genet* 2009;41:905-8.
2. Shete S, Hosking FJ, Robertson LB, et al. Genome-wide association study identifies five susceptibility loci for glioma. *Nat Genet* 2009;41:899-904.
3. Stacey SN, Sulem P, Jonasdottir A, et al. A germline variant in the *TP53* polyadenylation signal confers cancer susceptibility. *Nat Genet* 2011;43:1098-103.
4. Jenkins RB, Wrensch MR, Johnson D, et al. Distinct germ line polymorphisms underlie glioma morphologic heterogeneity. *Cancer Genet* 2011;204:13-8.
5. Jenkins RB, Xiao Y, Sicotte H, et al. A low-frequency variant at 8q24.21 is strongly associated with risk of oligodendroglial tumors and astrocytomas with *IDH1* or *IDH2* mutation. *Nat Genet* 2012;44:1122-5.
6. Walsh KM, Anderson E, Hansen HM, et al. Analysis of 60 reported glioma risk SNPs replicates published GWAS findings but fails to replicate associations from published candidate-gene studies. *Genet Epidemiol* 2013;37:222-8.
7. Felini MJ, Olshan AF, Schroeder JC, et al. Reproductive factors and hormone use and risk of adult gliomas. *Cancer Causes Control* 2009;20:87-96.
8. Wrensch M, Lee M, Miike R, et al. Familial and personal medical history of cancer and nervous system conditions among adults with glioma and controls. *Am J Epidemiol* 1997;145:581-93.
9. Wiemels JL, Wiencke JK, Sison JD, Miike R, McMillan A, Wrensch M. History of allergies among adults with glioma and controls. *Int J Cancer* 2002;98:609-15.
10. Brennan CW, Verhaak RG, McKenna A, et al. The somatic genomic landscape of glioblastoma. *Cell* 2013;155:462-77.
11. Bielinski SJ, Chai HS, Pathak J, et al. Mayo Genome Consortia: a genotype-phenotype resource for genome-wide association studies with an application to the analysis of circulating bilirubin levels. *Mayo Clin Proc* 2011;86:606-14.
12. Christensen BC, Smith AA, Zheng S, et al. DNA methylation, isocitrate dehydrogenase mutation, and survival in glioma. *J Natl Cancer Inst* 2011;103:143-53.
13. Smith JS, Perry A, Borell TJ, et al. Alterations of chromosome arms 1p and 19q as predictors of survival in oligodendrogliomas, astrocytomas, and mixed oligoastrocytomas. *J Clin Oncol* 2000;18:636-45.
14. Chen W, Yuan L, Cai Y, et al. Identification of chromosomal copy number variations and novel candidate loci in hereditary nonpolyposis colorectal cancer with mismatch repair proficiency. *Genomics* 2013;102:27-34.
15. Pei J, Robu V, Feder M, et al. Copy neutral loss of heterozygosity in 20q in chronic lymphocytic leukemia/small lymphocytic lymphoma. *Cancer Genet* 2014;207:98-102.
16. Solomon DA, Kim JS, Bondaruk J, et al. Frequent truncating mutations of *STAG2* in bladder cancer. *Nat Genet* 2013;45:1428-30.
17. Killela PJ, Reitman ZJ, Jiao Y, et al. *TERT* promoter mutations occur frequently in gliomas and a subset of tumors derived from cells with low rates of self-renewal. *Proc Natl Acad Sci U S A* 2013;110:6021-6.
18. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 2011;12:323.
19. Jiao Y, Killela PJ, Reitman ZJ, et al. Frequent *ATRX*, *CIC*, *FUBP1* and *IDH1* mutations refine the classification of malignant gliomas. *Oncotarget* 2012;3:709-22.
20. Bady P, Sciuscio D, Diserens AC, et al. *MGMT* methylation analysis of glioblastoma on the Infinium methylation BeadChip identifies two distinct CpG regions associated with gene silencing and outcome, yielding a prediction model for comparisons across datasets, tumor grades, and CIMP-status. *Acta Neuropathol* 2012;124:547-60.
21. Wiencke JK, Aldape K, McMillan A, et al. Molecular features of adult glioma associated with patient race/ethnicity, age, and a polymorphism in O<sup>6</sup>-methylguanine-DNA-methyltransferase. *Cancer Epidemiol Biomarkers Prev* 2005;14:1774-83.

22. Nakagawa T, Kollmeyer TM, Morlan BW, et al. A tissue biomarker panel predicting systemic progression after PSA recurrence post-definitive prostate cancer therapy. *PLoS One* 2008;3:e2318.
23. Mahoney DW, Therneau TM, Anderson SK, et al. Quality assessment metrics for whole genome gene expression profiling of paraffin embedded samples. *BMC Res Notes* 2013;6:33.
24. Therneau TM, Crowson CS, Atkinson EJ. Adjusted Survival Curves. In: *R Code*.
25. Delaneau O, Zagury JF, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods* 2013;10:5-6.
26. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* 2012;44:955-9.
27. Abecasis GR, Auton A, Brooks LD, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012;491:56-65.
28. Dabney AR. ClaNC: point-and-click software for classifying microarrays to nearest centroids. *Bioinformatics* 2006;22:122-3.
29. The Cancer Genome Atlas (TCGA) Data Portal (Accessed at <https://tcga-data.nci.nih.gov/tcga/dataAccessMatrix.htm>.)
30. Broad Institute's Genome Data Analysis Center (GDAC). (Accessed at <https://confluence.broadinstitute.org/display/GDAC/Home>.)