Comprehensive molecular characterization of pediatric treatment-induced high-grade glioma: A distinct entity despite disparate etiologies with defining molecular characteristics and potential therapeutic targets

John DeSisto[1]*, John T. Lucas Jr.[2]*, Ke Xu[3], Andrew Donson[1], Tong Lin[4], Bridget Sanford[1], Gang Wu[3], Quynh T. Tran[5], Dale Hedges[2], Chih-Yang Hsu[2], Gregory T. Armstrong[6,8], Michael Arnold[9], Smita Bhatia[8,10], Patrick Flannery[1], Rakeb Lemma[1], Lakotah Hardie[11], Ulrich Schüller[12], Lindsey M. Hoffman[1,11], Kathleen Dorris[1,11], Jean Mulcahy Levy[1,11], Todd C. Hankinson[1,11], Michael Handler[1,11], Arthur Liu[11], Nicholas Foreman[1,11], Rajeev Vibhakar[1,11], Kenneth Jones[1], Sariah Allen[5], Jinghui Zhang[3], Suzanne J. Baker[7], Thomas E. Merchant[2], Brent A. Orr[5#], Adam L. Green[1,11#]

*These authors contributed equally to this work

[#]Co-Senior Authors

[1]Morgan Adams Foundation Pediatric Brain Tumor Research Program, University of Colorado School of Medicine, Aurora, CO

Department of Radiation Oncology[2], Computational Biology[3], Biostatistics[4], Pathology[5], Epidemiology and Cancer Control[6], and Developmental Neurobiology[7], St. Jude Children's Research Hospital, Memphis, TN

[8]Childhood Cancer Survivor Study

[9]Department of Pathology, Nationwide Children's Hospital, Columbus, OH

[10]Division of Hematology and Oncology, University of Alabama at Birmingham

[11]Children's Hospital Colorado, Aurora, CO

[12]Institute of Neuropathology and Department of Pediatric Hematology and Oncology, University Medical Center; Children's Cancer Center, Hamburg, Germany

25 **Author Contributions**

26 ALG and JTL conceived the project and overall experimental design. AD assisted with sample
27 processing and data analysis. JTL and CH designed and analyzed *de novo* pHGG and TIHGG imaging data.
28 ALG and JTL compiled and analyzed TIHGG clinical and imaging data; LH and AL performed analysis of
29 radiation treatment for the Colorado cohort. MH and TCH procured tumor samples. US provided
30 methylation profiling data and clinical information for tumor samples. SA processed samples and
31 performed fluorescence in situ hybridization. TL, QT, and BAO analyzed 450/850K data from TIHGG
32 and pHGG cases. JTL, KX, GW, and BAO conceived of the relevant comparisons to *de novo* pHGG copy-
33 number and sequencing data. JD, BS, KJ, and ALG designed the relevant comparisons and performed
34 analyses of germline and somatic sequencing data, RNA-seq data, drug screening results. SJB provided
35 DNA methylation and sequencing data from PCGP cases with a history of therapeutic ionizing radiation.
36 JD performed all primary cell line experiments, assisted by PF and RL. LMH, KD, JML, NF, and RV assisted
37 with project planning and data analysis. GW, KX, KJ, and DH, performed all computational analyses and
38 interpretation of the germline and somatic sequencing data. JTL, JD, and KX performed statistical
39 analyses. MA, GA, BAO, NF, and SB contributed primary tumor samples and clinical data. KX completed
40 the reconstruction of episomes from WGS data. JTL, AG, BAO, TL, JD, KX, QT and GW contributed to the
41 interpretation of experiments. JTL and JD wrote the manuscript with input from all co-authors; all
42 coauthors were involved with manuscript revision, which was led by JTL, JD, BAO, and ALG.

Acknowledgements

We thank members of the Zhang and Wu laboratories (St. Jude Children's Research Hospital) and members of the Foreman and Vibhakar laboratories (University of Colorado) for assistance and discussion. We also thank Matthew Lear (St. Jude Children's Research Hospital) for his help in acquiring and performing preliminary clinical analysis of primary tumor samples and Scott Olsen, Granger Ridout, and Emily Walker (Hartwell Center for Bioinformatics and Biotechnology) for their assistance in sequencing samples. RNA-seq library preparation and high-throughput sequencing were conducted at the University of Colorado Cancer Center Genomics Shared Resource; microscopy was performed at the University of Colorado Denver Advanced Light Microscopy Core Facility; and tumor sample preparation for imaging was performed by the University of Colorado Cancer Center Histology Shared Resource. This work was supported by the National Cancer Institute (CA55727, G.T. Armstrong, Principal Investigator), the Fördergemeinschaft Kinderkrebs-Zentrum Hamburg (US), St. Jude Children's Research Hospital Cancer Center Support (CORE) grant (CA21765, C. Roberts, Principal Investigator), ALSAC, the Morgan Adams Foundation (ALG), and a St. Baldrick's Foundation Scholarship (ALG).

Data Availability Statement

The datasets generated during and/or analyzed during the current study are currently available from the corresponding authors upon request and will be made publicly available. Our teams are current in the process of submitting all sequencing data to EMBL-EBI and will update the editors with the link(s) as soon as possible.

Abstract

Treatment-induced high-grade gliomas (TIHGGs) are an incurable late complication of cranial radiation therapy or combined radiation/chemotherapy used to treat pediatric cancer. We assembled a cohort of 33 TIHGGs from multiple institutions. The primary antecedent malignancies were medulloblastoma, acute lymphoblastic leukemia, astrocytoma, and ependymoma. We performed methylation profiling, RNA-seq, and genomic sequencing (whole-genome or whole-exome) on TIHGG samples. Methylation profiling revealed that TIHGGs cluster primarily with the pediatric receptor tyrosine kinase I subtype (26/31 samples). Common TIHGG copy-number alterations include Chromosome (Ch.) 1p loss/1q gain, Ch. 4 loss, Ch. 6q loss, and Ch. 13 and Ch. 14 loss; focal alterations include *PDGFRA* and *CDK4* gain and loss of *CDKN2A* and *BCOR*. Relative to *de novo* pediatric high-grade glioma (pHGG), *BCOR* loss ($p=0.004$) and *CDKN2A* loss ($p=0.005$) were significantly increased. Transcriptomic analysis identified two distinct TIHGG subgroups, one with a lesser mutation burden (0.12 mut/Mb), Ch. 1p loss/1q gain (5/6 samples), and stem cell characteristics, and one with a greater mutation burden (1.08 mut/Mb, $p<0.0002$), depletion of DNA repair pathways, and inflammatory characteristics. We observed increased chromothripsis in TIHGG versus pHGG (67% vs. 31%, $p=0.036$), which was associated with extrachromosomal circular DNA-mediated amplification of *PDGFRA* and *CDK4*. *In vitro* drug screening in one primary, patient-derived TIHGG cell line from each expression subgroup identified microtubule inhibitors/stabilizers, DNA-damaging agents, MEK inhibition, and, in the inflammatory subgroup, proteasome inhibitors as potentially effective therapies. This study provides a comprehensive molecular profile of TIHGG, including mechanistic insights to TIHGG oncogenesis, and identifies potentially effective therapeutic modalities for further investigation.

86 **Introduction**

87 Treatment-induced high-grade gliomas (TIHGGs) are pediatric high-grade astrocytomas with an
88 estimated incidence of 3% following cranial radiotherapy[1]. The highest cumulative incidence of TIHGG is
89 in survivors of medulloblastoma, ependymoma, and leukemia[1]. TIHGG represents a rare but significant
90 cause of late mortality in survivors of childhood cancer[2], as few therapeutic options are available at
91 diagnosis[3,4]. Genomic analysis of TIHGG has been restricted to only a few small cohorts, but early
92 evidence suggests that, in comparison with *de novo* pediatric high-grade gliomas (pHGG), TIHGG may be
93 enriched for recurrent copy-number alterations and for amplification of *PDGFRA* and other potential
94 activators of the MAP kinase pathway[5,6]. *TP53* mutations, gain of 1q, deletion of *CDKN2A*, and
95 overexpression of *ERBB3* and *SOX10* may also be more frequent[5,7]. Similarly, alterations in treatment-
96 related second cancers outside the brain show recurrent genomic copy-number variations, some of
97 which contain genes that modify tumor latency *(NF2*, *CHEK2*, and *RET* amplification)[8,9,10]. These
98 divergent oncogenic programs are spread over multiple histologies, limiting inference regarding
99 potential therapeutic options.

100 In this study, we present an evaluation of the largest TIHGG cohort assembled to date, using
101 comprehensive clinical and molecular analyses. We conducted a clinical review of each case, including
102 the initial tumor location and type, treatment received, and the resulting TIHGG location and histology,
103 followed by DNA methylation analysis using 450/850k arrays, genomic sequencing of TIHGG and
104 matched blood samples, and RNA-seq. We assemble a comprehensive molecular profile of TIHGG,
105 enabling us to identify several previously unknown characteristics of the disease, and propose a method
106 of grouping TIHGG tumors phenotypically based upon a combination of genomic alterations and gene
107 expression characteristics. Our findings differentiate TIHGG from pHGG and provide rationale for
108 alternative therapeutic approaches, which we explore in *in silico* and *in vitro* pre-clinical therapeutic
109 screening studies.

110 **Results**

111 ***Patient characteristics***

112 The median age at diagnosis of first cancer within the TIHGG cohort was 6.5 years (range, 0.16-
113 19). The initial diagnoses included medulloblastoma (N=12; 36%), acute lymphoblastic leukemia (ALL)
114 (N=10; 30.3%), astrocytoma (N=3; 9%), ependymoma (N=2; 6%), germinoma (N=2; 6%), lymphoma (N=1;
115 3%), craniopharyngioma (N=1; 3%), neuroblastoma (N=1; 3%), and ganglioglioma (N=1; 3%) (Sup. Table
116 1). Two patients had known mismatch repair deficiencies resulting from a heterozygous germline loss of
117 *MSH2* in one case and *PMS2* in the other (Case 2, Case 32). Both were included in the study because
118 each had therapeutic radiation exposure to the cranium with sufficient latency (11 and 8 years) to meet
119 the TIHGG diagnostic criteria. All cases had ionizing radiation exposure, except for one case with a prior
120 history of high-dose chemotherapy exposure for treatment of high-risk neuroblastoma (further details
121 on clinical histories are discussed in Methods). The individual clinical histories for the TIHGG cohort with
122 extended clinical information are presented in Sup. Fig. 1.

123 Complete clinical-pathologic data regarding treatment history were available in 76% of cases
124 (Fig. 1A (showing analyses/information available for the entire cohort), Sup. Fig. 1, Sup. Table 1, 2).
125 Patients received total body irradiation (TBI) (N=1, 3%), cranial irradiation (N=6, 18%), craniospinal
126 irradiation (N=9, 24%), or focal brain irradiation (N=10, 29%) to treat the initial malignancy (Fig. 1B). One
127 patient received radiotherapy to the abdomen (Case 44). Radiotherapy field extent was unknown in 24%
128 (N=7) of cases. The indications for radiotherapy were adjuvant (N=12, 36%), definitive (N=1, 3%), and
129 salvage (N=9, 26.5%). Radiotherapy technique was 2D (N=13, 38.2%) or 3D conformal radiation therapy
130 (N=8, 23.5%). Technique was not known in 38% (N=13) of cases. The location of the TIHGG relative to
131 the initial radiotherapy field was in-field but outside the high-dose region in 24% (N=8) of evaluable
132 cases and in the high-dose (prescribed dose) region in 52% of cases (N=17).

133 The median time between initial treatment and TIHGG diagnosis was 8.0 years (95% CI 7.3-13.2
134 years) (Fig. 1C). The median age at the time of diagnosis of TIHGG was 13.6 years (9-26 years) (Sup.
135 Table 1). TIHGG tumor grades were WHO grade III (N=3, 9%), grade IV (N=12, 55%), and pHGG not
136 otherwise specified (N=12, 36%) (Sup. Table 1). Assessment by methylation array identified *MGMT*
137 promoter methylation in 8 of 31 (25.8%) cases in which methylation status was evaluated during the
138 clinical course of treatment.

139 Salvage therapy characteristics for TIHGG were available in 38% of cases (N=13) and included
140 resection (N=3, 23.1%), repeat radiotherapy (N=12, 92.3%), and/or chemotherapy (N=4, 30.7%) (Sup.
141 Table 2). The median radiotherapy dose for the second course was 54 Gy (range 30-59.4 Gy). At the time
142 of this analysis, all TIHGG patients with known outcome were deceased (N=26), and 7 had unknown
143 clinical status (Fig. 1D). The median survival in the TIHGG cohort was 9 months (95% CI 5-24.7 months)
144 (Fig. 1D, Sup. Table 2).

145 The use of radiotherapy, chemotherapy, and/or resection did not substantially improve
146 treatment outcomes, although a paucity of details available regarding therapy limited power for this
147 analysis (Sup. Table 2). These outcomes are consistent with prior published clinical series of TIHGG[4,8].

148

149 ***Voxel-wise lesion mapping of TIHGG relative to* de novo *pHGG***

150 An MRI was available at diagnosis in 21 TIHGG cases (Sup. Fig. 2A). In cases for which location
151 information was available, TIHGGs were located in the infratentorial brain (posterior fossa) (N=11,
152 32.3%), supratentorial brain (N=12, 35.3%), supra- and infratentorial brain (N=1, 3%), and spine (N=1,
153 3%) (Fig. 1E, Sup. Fig. 2B). An additional cohort of 84 *de novo* pHGG was used for comparison. Compared
154 to a cohort of 84 de novo pHGG, TIHGG tumors were more likely to involve the frontal and occipital
155 lobe, and posterior fossa (cerebellum and inferior brainstem) regions relative to primary pHGG (Fig. 1E,
156 Sup. Fig. 2B).

157

158 ***TIHGGs cluster into a specific methylation group***

159 We began our molecular analysis using DNA methylation profiling because methylation data (or
160 a DNA sample sufficient to conduct methylation profiling) were available for the greatest number of
161 TIHGG samples in our cohort. We first analyzed the methylation data using unsupervised clustering of 34
162 TIHGG samples from 31 tumors against a combined reference cohort containing pediatric and adult CNS
163 malignancies[11,12]. Twenty-five of 31 cases clustered among a group of epigenetically similar tumors
164 consisting of *H3K27M*-negative midline gliomas and gliomas in the pediatric receptor tyrosine kinase I
165 (pedRTK I) subgroup (Fig. 2A-C and Sup. Fig. 3A-C)[13]. TIHGGs were not readily distinguishable by
166 methylation from IDH wt, GBM-MID, and pedRTK I cases from the Capper and Korshunov reference
167 cohorts[11], so the combined group of TIHGG cases is designated as "pedRTK I" going forward, reflecting
168 their closer overall proximity to the centroids of the pedRTK I reference cases (Sup. Table 3). The
169 remaining six TIHGGs clustered with the GBM K27M (N=1), control cerebellum (N=2), anaplastic pilocytic
170 astrocytoma (N=1), HGNET-MN1 (N=1), and pleomorphic xanthoastrocytoma (N=1) methylation groups.
171 We confirmed the initial classification results using hierarchical and supervised methods (Sup. Table 4
172 and Sup. Table 5). No trend was noted associating initial cancer diagnosis with subsequent TIHGG
173 methylation class (Sup. Fig. 4A)[14].The latency and survival of TIHGG tumors classified as pedRTK I did not
174 differ significantly from that of those classified with other subclasses (Sup. Fig. 4B-C). The odds of being
175 assigned to the pedRTK I subgroup in the TIHGG cohort were 29.2 times the odds of being assigned to
176 that subgroup in the *de novo* pHGG cohort (95% CI 9.15-92.9, $p<0.0001$) (Sup. Fig. 5A, Sup. Table 6).
177 Survival of the TIHGG-pedRTK I was inferior to pHGG pedRTK I (median survival 9 months, [95% CI, 5-26]
178 vs. 16 months [95% CI, 10-not reached, $p=0.00021$] (Sup. Fig. 5B).

179 Alignment of TIHGGs with pedRTK I suggested the possibility that TIHGGs, despite their
180 disparate clinical precursors, might share common molecular alterations, which we investigated by
181 comparing copy-number alteration and transcriptomic profiles discussed in the following sections.
182
183 ***TIHGGs show recurrent copy-number abnormalities***
184 We next reviewed the methylation profiling data for large segment and focal copy-number
185 alterations. Total copy-number variation was significantly increased in TIHGG when comparing *de novo*
186 pHGG and TIHGG in the pedRTK I subgroup (Sup. Fig. 5C). Recurrent large segment copy-number
187 alterations identified from methylation profiling included Ch.1p loss (3/31, 9.7%), Ch.1q gain (3/31,
188 9.7%), Ch.6q loss (5/31, 16.1%), Ch.13 loss (8/31, 25.8%), and Ch.14 loss (3/31, 9.7%) (Fig. 3A, Sup. Table
189 5, Sup. Fig. 6). Copy-number alterations in *PDGFRA* gain/amplification (15/31, 48.4%), *CDK4*
190 amplification (5/31, 12.9%), *CDKN2A* loss (12/31, 38.7%), and *BCOR* loss (7/31, 22.6%) were also noted.
191 These were increased relative to *de novo* pHGG for *PDGFRA* (p=0.0055) and *BCOR* (p=0.0043) (Sup. Table
192 7, Sup. Fig. 7). Changes in copy number of *PDGFRA, CDKN2A,* and *BCOR* were associated with an
193 accompanying change in gene expression in cases with available RNA-Seq data (N=12, 34.3%) (Fig. 3B).
194 When comparing total genomic copy-number variations by case, TIHGG favored loss over gain (Fig. 3C).
195 Overall, TIHGGs showed large segment genomic alterations in Ch.1 (1p loss/1q gain) and Ch.13
196 (loss), and focal alterations in *PDGFRA* (amplification)*, CDKN2A* (loss) and *BCOR* (loss). As further
197 discussed below, where alterations were identified in both WES/WGS and methylation data, we
198 obtained consistent results between the two methods. The *PDGFRA* amplification that we observed is
199 consistent with the alignment of most TIHGGs with the pedRTK I methylation subclass[11].
200

201 ***TIHGG has a high frequency of chromothripsis and episomal amplification of oncogenes***

202 To further investigate the potential sources of copy-number amplification, we evaluated the
203 rate of chromothripsis in the TIHGG samples relative to that in primary non-brainstem and brainstem
204 pHGG from Wu *et al.*[15]. We observed an increase in the frequency of chromothripsis in the TIHGG cohort
205 (8/12 cases) relative to that in diffuse intrinsic pontine glioma (DIPG) (66.7% vs. 30.0%, p=0.048), non-
206 brainstem pHGG (66.7% vs. 33.3%, p=0.091), and all primary pHGG combined (66.7% vs. 31.4%, p=0.036)
207 (Fig. 4A, Sup. Table 8). Both intra-chromosomal (Case 29-dm2) rearrangements leading to
208 extrachromosomal circular DNA (eccDNA) structures, as well as inter-chromosomal rearrangements
209 (Case 29-dm1) leading to eccDNA structures, were observed in case 29 (Fig. 4B, D). *PDGFRA* and *CDK4*
210 appeared to be amplified separately on the two eccDNA structures (Fig. 4B, D). Amplification was
211 confirmed in selected cases by fluorescence *in situ* hybridization, which demonstrated a double minute
212 pattern, a finding typical of episomal amplification (Fig. 4C)[16]. There was a trend toward an increased
213 frequency of chromothripsis in cases with a coincident pathogenic *TP53* mutation (5/8 cases, 62.5%)
214 relative to those with wild-type *TP53* (0/4 cases, 0%, p=0.08).

215 In summary, we observed an elevated frequency of chromothripsis in TIHGG relative to *de novo*
216 pHGG, which facilitated extrachromosomal amplification of known oncogenes (*PDGFRA, CDK4).*

217

218 ***TIHGGs show recurrent abnormalities in genes that regulate cell proliferation, growth, and cell cycle***

219 In addition to copy-number variation, we also identified focal alterations in TIHGG using
220 WES/WGS data. Eighteen TIHGG cases underwent whole-exome (WES) or whole-genome sequencing
221 (WGS) (Fig. 5). Tier 1 mutations in the 5 non-hypermutator WES cases without matched germline data
222 are shown in Sup. Table 9; Tier 1 mutations for the 9 non-hypermutator WGS cases with matched
223 germline data are shown in Sup. Table 10. The most frequent focal alterations included *PDGFRA*,
224 *CDKN2A*, *BCOR*, *NF1*, *TP53*, and *CDK4* (Fig. 5, Sup. Tables 7, 9, 10). Missense mutations were noted in

225 *TP53* and *PDGFRA* in 7 of 8 and 4 of 15 cases with an observed alteration in *TP53* and *PDGFRA,*
226 respectively. Nonsense and frameshift mutations of *NF1* occurred in 8 of 9 observed cases with *NF1*
227 alterations. *MET* fusion gene products were identified from RNA-seq samples in three cases (Sup. Table
228 11). Structural variations and copy-number alterations from 10 WGS cases are shown in Sup. Tables 12
229 and 13.
230
231 ***Compared to pHGG, TIHGG are enriched for specific somatic and germline abnormalities***
232 We compared select recurrent alterations in the TIHGG cohort to *de novo* pHGG from the HERBY
233 cohort across each cohort as a whole and among pedRTK I tumors (Sup. Fig. 8, Sup. Table 14)[14].
234 Statistically significant increases in *BCOR* loss (p=0.0004) and *CDKN2A* loss (p=0.05) were observed in
235 TIHGG relative to pHGG (Sup. Table 7). No statistically significant differences were observed between
236 the proportions of molecular alterations in pedRTK I TIHGG relative to pHGG – pedRTK I tumors,
237 although power was limited by the sparse number of pedRTK I cases in the HERBY cohort (N=8). Of note,
238 the frequency of *TP53* mutations in pedRTK I TIHGG cases was 43.8% (7/16) and 62% (5/8) in the *de*
239 *novo* pHGG pedRTK I cases (p=0.68, Sup. Table 7). We observed a single *H3F3A* mutation (Fig. 5, Case 7)
240 (1/19, 5.6%) in the TIHGG cohort, compared to 37.8% (28/74) with mutant *H3F3A* (p=0.0053) in the
241 HERBY cohort (composed of non-brainstem pHGG).
242 We also compared variant frequency between the TIHGG cohort and pHGG cases in the HERBY
243 cohort. Germline mutations in the TIHGG cohort were analyzed in 10 cases for which WGS data were
244 available (Sup. Table 15). The total number of germline and somatic variants per megabase (Mb) were
245 calculated and stratified by pHGG versus TIHGG and are compared in Sup. Fig. 9A. The number of
246 variants per Mb was significantly higher in somatic coding regions of TIHGG DNA than in those of pHGG
247 (median 0.000784 ± 0.000286 (range) variants/Mb vs. median 0.0007 ± 0.000245 (range) variants/Mb,
248 p=0.031). The relative frequencies of base transitions were calculated per sample in pHGG and TIHGG

249 (Sup. Fig. 9B). Germline and somatic coding base changes were not significantly different across base
250 change type. Somatic noncoding region variants showed decreased A to C (median 14.4 ± 4.91 vs. 8.8 ±
251 4.1, p=0.007) and A to G (34.3 ± 9.67 vs. 20.2 ± 11.3, p=0.03), and increased C to A (6.79 ± 2.2 vs. 16.7 ±
252 5.47, p=0.004) base changes in TIHGG cases relative to *de novo* pHGG (Sup. Fig. 9B-C).
253 In summary, TIHGG have comparable burden of somatic and germline alterations to *de novo*
254 pHGG. Frequent pathogenic alterations were noted in *PDGFRA, BCOR, CDK4, TP53,* and *NF1.*
255
256 ***Transcriptomic analysis demonstrates that TIHGGs cluster into two subgroups that are separate from***
257 ***pHGG***
258 Transcriptomic data (array and RNA-seq) were available on 14 TIHGG cases and 42 *de novo*
259 glioblastoma (GBM) cases. TIHGG tumors clustered separately from *de novo* pediatric, infant, and adult
260 GBM based on transcriptomic data and formed two distinct subgroups (A and B) containing 6 and 8
261 tumor samples respectively (Fig. 6A). A comparison of the transcriptomic and methylation-based
262 clustering results showed that the two gene expression-based TIHGG subgroups were, with few
263 exceptions, recapitulated in the methylation analysis (Fig. 6B). Our data thus suggest two distinct
264 patterns of gene expression within the pedRTK I methylation-based subgroup into which most of the
265 TIHGGs cluster.
266 To further investigate gene expression in TIHGG, we performed Metascape/Cytoscape analysis,
267 which revealed several differences in gene ontology between TIHGG and the transcriptional reference
268 group of *de novo* GBM. The differences include fundamental cellular processes, such as RNA processing
269 and transport, protein translation and catabolism, and cellular signaling, as well as pathways controlling
270 neurogenesis and gliogenesis (Fig. 6C, Sup. Table 16). To determine the directionality of gene expression
271 patterns identified in the Metascape/Cytoscape analysis, we performed gene set enrichment analysis
272 (GSEA) using the gene ontology (GO) gene set collection[17,18]. The results of this analysis showed

273 similarities in differential gene expression to that identified in the Metascape/Cytoscape analysis, with
274 TIHGG samples showing relative enrichment of genes involved in DNA metabolism, cell cycle
275 progression, DNA repair, nervous system development, and protein catabolism compared to *de novo*
276 GBM, and relative depletion of genes related to immune response, signaling and cellular response to
277 external stimulus, receptor activity, and neurogenesis (Fig. 6D, Sup. Table 17).
278
279 ***Expression-based characterization of TIHGG subgroups***
280 To characterize the TIHGG gene expression subgroups, we analyzed the differences in their
281 relative gene expression, mutational status, and copy-number alterations. GSEA revealed that TIHGG
282 gene expression Group A shows stem cell (proneural) and neuronal characteristics and has enriched
283 expression of *MYC* pathway genes (Fig. 7A). Expression Group B has mesenchymal and astroglial
284 characteristics, enriched expression of inflammatory (particularly NF-κB pathway) and immune genes,
285 and depletion of DNA-repair pathway genes (Fig. 7A, Sup. Table 18). We did not observe relative
286 expression differences between the TIHGG subgroups in cell cycle or proliferative genes, consistent with
287 clinical experience showing that TIHGGs, in general, are highly proliferative.
288 Using the DNA methylation data, we analyzed whether there were differences in copy-number
289 alterations between the two gene expression subgroups. We found Group A to be enriched in 1p loss/1q
290 gain (5/6 tumor samples) compared to Group B (0/7 tumor samples, p=0.005). GSEA showed particular
291 differences at 1p34 (NES = −5.21, FDR=0) and 1p36 (NES = −7.77, FDR=0). The remaining gene- or
292 chromosome-level amplifications and deletions identified in the TIHGG cohort, including Ch.13 or 14
293 loss, *PDGFRA* amplification, and *CDKN2A* loss, were relatively evenly distributed between the two TIHGG
294 subgroups (Sup. Table 5). We did not identify patterns of significant point mutation or similar small-scale
295 genetic differences between Group A and B tumors, nor did we find any relationship between initial
296 malignancy and TIHGG subgroup (Fig. 5).

297 The relative depletion of DNA repair pathway genes that we identified in Group B relative to
298 Group A led us to investigate whether there were differences in mutation load between the two groups.
299 Excluding one hypermutator case, WGS data were available on matched blood and tumor samples for 3
300 cases from Group A and 6 from Group B. We found nearly identical levels of germline mutation load in
301 Groups A and B (Fig. 7B). We found that Group B tumors, however, had a nine-fold greater somatic
302 mutation load than Group A tumors (1.08 vs. 0.12 mut/Mb, $p<0.002$) (Fig. 7B).
303 As noted previously, relative to Group A, Group B is depleted in DNA repair pathway gene
304 expression (mean NES = −3.96, p=0.001 vs. expected value of 1), which led us to compare DNA repair
305 pathway gene expression in group A and B germline samples. We found that germline samples from
306 Group A patients were likewise enriched in DNA repair pathway gene expression (mean NES = -4.30,
307 p=0.001) (Fig. 7C). Individual DNA repair pathway genes with lower expression in Group B relative to
308 Group A include *GTF2H2,* a subunit of RNA polymerase II transcription initiation factor IIH that is
309 involved in nucleotide excision repair; *PAPD7,* a DNA polymerase involved in DNA repair; *ERCC4,* which is
310 essential for nucleotide excision repair; and *TP53* (Fig. 7D, Sup. Table 19.). In a pairwise comparison of
311 the individual DNA repair pathway genes shown in Figure 7D, the mean expression ratio (Group B/Group
312 A) was 0.76 ($p=9.79\times10^{-11}$ vs. expected value of 1) (Sup. Table 19).
313 In summary, TIHGGs split into two transcriptional subgroups with distinct gene expression
314 profiles relative to *de novo* GBM. Subgroup A resembles proneural GBM and is enriched in expression of
315 *MYC*-pathway genes, while subgroup B resembles mesenchymal GBM and has greater mutational
316 burden and decreased DNA repair gene expression.
317

318 **In silico *and* in vitro *drug screening identify potentially effective therapeutic agents for the treatment***
319 ***of TIHGG***

320 In order to investigate therapeutic susceptibilities in TIHGG tumors, we used *in silico* modeling
321 to identify potential TIHGG sensitivity to drug classes compared to a set of 23 *de novo* pediatric and 14
322 *de novo* adult GBMs. We also performed an *in vitro* drug screen of two primary tumor-derived cell lines,
323 one from gene expression Group A (MAF-496) and one from Group B (MAF-145), for sensitivity to a
324 panel of all FDA-approved anti-cancer agents. The *in silico* modeling predicted several classes of
325 therapeutic agents to have greater activity in TIHGG when compared with *de novo* GBM. These included
326 DNA-damaging agents, microtubule polymerization/de-polymerization inhibitors, proteasome inhibitors,
327 and histone deacetylase (HDAC) inhibitors (Fig. 8A, Sup. Table 20).

328 Our *in vitro* screen encompassed all FDA-approved anti-cancer agents, including multiple agents
329 from the classes identified in the *in silico* screen, all dosed at 1 μM concentration (Sup. Figure 10A, Sup.
330 Table 21). We found, combining results from the two cell lines, that 11/12 DNA intercalator results, 6/12
331 microtubule agent results, 4/4 proteasome inhibitor results, and 3/6 HDAC inhibitor results showed at
332 least 50% cell death compared to vehicle (Fig. 8B, Sup. Table 21). The screen also identified
333 topoisomerase I/II inhibitors (6/8 results) and PI3K-MTOR, BCR-ABL, and RAF-MEK pathway inhibitors
334 (12/24 results total) as potentially effective using the 50% cell death criterion (Fig. 8B, Sup. Table 21).
335 Folate analogs (0/6 results with at least 50% cell death), platinum drugs (0/6 results), and alkylators
336 (3/32 results) performed poorly in the *in vitro* screen (Fig. 8B, Sup. Table 21). We found that, with the
337 exception of proteasome inhibitors, which are further discussed below, the Group A cell line (MAF-496)
338 was more susceptible to the effective drugs than the Group B cell line (MAF-145) (Fig. 8C).

339 To validate the screening results, we performed single-drug assays of selected agents at a range
340 of concentrations in the two TIHGG cell lines. Drugs were selected for validation based upon the
341 screening results and included two DNA-damaging agents (aldoxorubicin and etoposide), two

342 microtubule agents (paclitaxel and vinblastine), three proteasome inhibitors (bortezomib, carfilzomib,
343 and marizomib), and two kinase inhibitors (sunitinib, a multi-TKI inhibitor, and trametinib, a MEK
344 inhibitor) (Fig. 8D, Sup. Fig. 10B-G). In the validation testing, aldoxorubicin (an anthracycline drug
345 capable of blood-brain barrier penetration), etoposide, paclitaxel, and vinblastine showed sub-
346 micromolar IC_{50} values in both cell lines (Fig. 8D, Sup. Fig. 10B-C), as did the MEK inhibitor trametinib
347 (Fig. 8D, Sup. Fig. 10D). Sunitinib had an IC_{50} of 5,256 nM in the MAF-145 (Group B) cell line but did not
348 reach an IC_{50} level in MAF-496 cells (Group A) (Sup. Fig. 10E). Validation testing of bortezomib and
349 carfilzomib showed *in vitro* effectiveness in Group B cells, consistent with the enrichment of the NF-κB
350 pathway in Group B tumors (Fig. 8D, Sup. Fig. 10F). In the Group A cell line, however, 25-45% of cells
351 survived the highest concentration of each drug, suggesting the presence of a substantial drug-resistant
352 population in Group A (Fig. 8D, Sup. Fig. 10F). We found that the proteasome inhibitor marizomib, which
353 is capable of penetrating the blood-brain barrier, had IC_{50} values of 366 and 379 nM in the Group A and
354 Group B lines, respectively, but also had a drug resistant population of approximately 30% of cells in
355 both cell lines (Sup. Fig. 10G), which was not seen with bortezomib or carfilzomib. Further investigation
356 showed that proteasome inhibition reduced nuclear levels of NF-κB, a known target of proteasome
357 inhibition, in the Group B cell line (MAF-145) but had a lesser effect on already low nuclear levels of NF-
358 κB in the Group A line (MAF-496) (Sup. Fig. 10H-I).
359 Taken together, our drug screening results in TIHGG tumors showed that drugs that interfere
360 with S-phase (aldoxorubicin and etoposide) or M-phase (vinblastine and paclitaxel), and the MEK
361 inhibitor trametinib, were effective *in vitro* in both TIHGG subtypes. In addition, proteasome inhibition
362 was effective in the Group B cell line.
363

364 **Discussion**

365 The present TIHGG cohort represents the largest and most comprehensive studied to date and
366 extends prior knowledge from past studies that largely included targeted analyses of sequencing,
367 expression, and staining of smaller tumor collections. The use of methylation profiling, WES/WGS of
368 tumor and matched germline samples, RNA-seq, and drug screening help advance understanding of
369 long-held questions regarding TIHGG, including similarities and differences from *de novo* pHGG,
370 distinguishing TIHGG from recurrent tumors, TIHGG oncogenesis, subgroups within TIHGG, and potential
371 therapeutic susceptibilities.

372 All but one of our cases had a clinical history of cranial radiation exposure. The exception was a
373 high-risk neuroblastoma patient who received multiple rounds of multi-agent chemotherapy, including
374 two cycles of high-dose chemotherapy, but only received radiation to her primary tumor site in the
375 mediastinum, far removed from her frontal lobe TIHGG. Her latency and tumor findings from our assays
376 cluster closely with the remainder of the cohort, suggesting chemotherapy was most likely causative.
377 While it was exceptional that chemotherapy exposure alone should give rise to a TIHGG, the term
378 treatment-induced is preferred to radiation-induced to reflect this possibility and the known interaction
379 between combined modality therapy and increased incidence of subsequent malignancies[1]. In addition,
380 since these tumors were a mix of WHO grade III, IV, and HGG NOS, we have replaced the term
381 glioblastoma with high-grade glioma.

382 The results of our methylation and transcriptomic analyses delineate TIHGG as an entity
383 separate from pHGG and identify two distinct transcriptional subgroups within the TIHGG cohort.
384 Additional genomic and transcriptomic analyses identify characteristic genetic alterations and gene
385 expression profiles that define the two subgroups. TIHGG were most frequently assigned to the pedRTK I
386 methylation subgroup of pHGG, despite their varied clinical origin. Of note, resolving pedRTK I from
387 IDHwt, GBM-mid subtype pHGG, and TIHGG was not possible in our analysis, thus highlighting the

388 molecular and epigenetic similarities between these groups. One TIHGG case in our series clustered with
389 diffuse midline glioma (DMG) but lacked the defining *H3K27M* mutation (Case 13). One recent case
390 series described three patients who had received prior radiotherapy for medulloblastoma and later
391 developed tumors clinically consistent with DMG[19]. While one case in our series had the H3K27M
392 mutation (Case 7), this case lacked methylation data and was not be able to be clustered against
393 reference cohorts. Because of the close association between the pedRTK I subgroup with TIHGG, specific
394 consideration may be useful in future brain tumor taxonomy. The relative contribution of *de novo* HGG
395 and TIHGG remains to be established. Additionally, whether exposure to ionizing radiation outside the
396 therapeutic setting also predisposes to the tumor subclass is currently unclear.

397 We compared recurrent genetic alterations in TIHGG and *de novo* pHGG cases in the pedRTK I
398 group to define potential differentiating characteristics. We observed an increased frequency of
399 *CDKN2A* and *BCOR* losses in TIHGG. *BCOR* alterations are infrequently observed in pHGG without co-
400 segregating *HIST1H3B* K27M mutations and are not described in pedRTK I pHGG[20]. While *BCOR*
401 alterations noted in the cohort were predominately non-focal, the region showed a high frequency of
402 overlap across TIHGG cases, suggesting potential biologic relevance. Of note, *BCOR* alterations noted in
403 the Pediatric Cancer Genome Project were primarily due to frameshift or nonsense mutations and were
404 often coupled with loss of heterozygosity[21]. Despite the increased frequency of *CDKN2A* deletions, we
405 did not observe accompanying *BRAF* V600E alterations, a finding common in pediatric low-grade glioma
406 and some pHGG[22,22]. Because we observed recurrent activation of MAP kinase pathway in our cohort,
407 concomitant CDKN2A loss may be a necessary mechanism to prevent oncogene induced senescence, as
408 has been described in other tumor types[23].

409 The differentiation of TIHGG from recurrent pHGG or transformed low-grade glioma based on
410 DNA methylation profiling remains challenging. Two cases in our cohort arose following focal
411 radiotherapy for an optic pathway glioma but were localized to a region within the radiotherapy low

412 dose region but distant from the original cancer and not coincident with a metastasis. Prior to review,
413 these were expected to be transformation events, but after reviewing location of the TIHGG in each case
414 relative to the original tumor and the radiotherapy field, it is clear that the TIHGG arose in previously
415 uninvolved but irradiated normal brain. It remains unclear whether modern methods such as IMRT,
416 which expose larger regions of normal tissue to non-therapeutic, albeit potentially mutagenic doses of
417 radiation, will yield an increased incidence of TIHGG despite the increased conformality of the
418 therapeutic dose region.

419 RNA-seq analysis differentiated the overall TIHGG cohort most distinctly from pHGG, which has
420 been previously documented in array expression analysis[24]. Our results furthermore show distinct
421 subgroups within the TIHGG cohort: a low mutation load group A that is enriched for Ch.1p loss/1q gain,
422 has neuronal characteristics, and is enriched in expression of stemness genes and transcriptional targets
423 of *MYC*, and a higher mutation load group B that is enriched in expression of inflammatory and immune
424 pathway genes, has astroglial and mesenchymal characteristics, and is impaired in DNA repair processes.
425 The subgroups do not correlate with original tumor type or histology. Instead, the depleted germline
426 DNA repair pathway gene expression in the Group B patients suggests their potential susceptibility to
427 develop TIHGG following mutagenic cancer treatment.

428 Given the potential for mutational profile to identify malignancy type and the difficulty of
429 distinguishing recurrent primary from secondary tumors, we evaluated the frequency and type of
430 somatic variants in TIHGG and pHGG. The relative frequency of pathogenic variants in all TIHGG in our
431 cohort was not substantially increased relative to pHGG, as other groups have found[7], but Group B
432 TIHGG had increased mutational burden compared to Group A, suggesting heterogeneity in the pedRTK
433 I/TIHGG subgroup. We also observed differential rates of select base changes in somatic non-coding
434 regions. Increased frequency of variations in non-coding regions have also been observed in syndromes

435 associated with genomic instability[25], as mutations in coding regions are likely to result in negative
436 selective pressure.

437 The therapeutic potential of photon radiotherapy in tumor tissue has always come at the cost of
438 sub-lethal damage in adjacent normal tissue[26]. Irradiated normal tissue damaged directly by secondary
439 electrons or indirectly by reactive oxygen species is known to produce non-lethal double-stranded DNA
440 breaks, which lead to impaired function and survival in long-term survivors of childhood cancer[26].
441 Incomplete error-prone repair of genomic DNA via non-homologous end joining and non-error prone
442 repair via homologous recombination result in genomic instability, which may be further modified by
443 host normal tissue radiosensitivity, dictated by germline DNA damage response functionality.

444 TIHGG exhibits signs of a mutagenic origin, with oncogenic amplifications of Ch.1q, *PDGFRA* and
445 *CDK4*, copy number losses in tumor suppressors including *CDKN2A* and *BCOR*, and pathogenic mutations
446 in *TP53, NF1* and *MET* (fusion events). TIHGGs have more frequent *PDGFRA* gain and *CDKN2A* loss
447 relative to *de novo* pHGG; TIHGG also displays increased expression of oncogenic pathways indicative of
448 increased proliferation, including DNA metabolism and cell cycle genes. The characteristics of the two
449 TIHGG gene expression subgroups suggest differences in tumorigenesis related to differing types of DNA
450 damage and their respective DNA repair capabilities. Group A TIHGGs all have large-scale, chromosome-
451 level abnormalities that are associated with poor outcome in pediatric brain tumors, including Ch. 1q
452 gain and Ch. 13 loss. Gain of Ch. 1q distinguishes pediatric from adult HGG, has been recognized as a
453 feature of radiation-induced tumors, and is an indicator of progression and poor outcome in
454 ependymoma, neuroblastoma, and Wilms tumors[5,27-30]. Specific associations between chromosome-level
455 abnormalities and the other characteristics of group A tumors have not been reported but are an area
456 for further inquiry. The higher mutation load of group B tumors could arise through the combination of
457 mutagenic treatment-induced DNA damage and the downregulation of DNA repair pathways in these
458 cases. Further insight into potential germline susceptibilities in group B patients related to impaired DNA

459 repair may even lead to identification of patients at risk for TIHGG prior to primary tumor therapy and
460 allow treatment modification to prevent TIHGG. Based on our germline data, we hypothesize these
461 germline susceptibilities are different from known mutations causing tumor predisposition syndromes in
462 that they may only manifest in the setting of highly mutagenic treatment.

463 Large-scale, localized DNA damage in the form of chromothripsis has also been recognized as a
464 means of carcinogenesis following radiotherapy, as it can perpetuate a string of subsequent random
465 molecular alterations. While pHGG exhibit genomic instability secondary to alterations in histone- and
466 chromatin-modifying machinery, we observed an increased rate of chromothripsis in TIHGG relative to
467 that of *de novo* pHGG. Chromothripsis was more common in the setting of pathogenic *TP53* mutations.
468 This has also been observed in pediatric and adult malignancies not associated with therapeutic ionizing
469 radiation[31]. Although chromothripsis is known to result in complex chromosomal alterations that can
470 perpetuate a high frequency of copy-number alterations, we also observed characteristic broad and
471 focal copy-number changes in non-chromothripsis cases (+1q, −6q, −13, −14, *+PDGFRA*, *+CDK4,*
472 *−CDKN2A, −BCOR)* that have known oncogenic potential[27 22,32-34]. We observed several instances of
473 oncogene amplification, likely facilitated by inclusion in eccDNA, which has been identified as a
474 mechanism of resistance to targeted therapies. In tumor cells with eccDNA containing the oncogene
475 *EGFR,* eccDNA levels increase in response to EGFR inhibition[35]. The eccDNA increase is frequently
476 independent of and more efficient, in terms of onset and magnitude of response, than chromosomal
477 copy-number alterations[35-37]. Similarly, serial samples from patients exposed to standard-of-care
478 therapy show increases in *EGFR* levels during treatment with *EGFR* inhibitors, also likely in part
479 facilitated by the presence of eccDNA[16]. These findings may explain the lack of observed efficacy of well-
480 characterized, CNS-penetrant, targeted drugs such as PDGFRA inhibitors in our preclinical drug screens.

481 *In silico* and *in vitro* drug screening identified several FDA-approved drugs that merit further
482 study as potential therapeutic agents to treat TIHGG. Drug classes identified as most effective against

483 TIHGG (DNA-damaging agents, anti-mitotic drugs that target microtubules, and histone deacetylase
484 (HDAC) inhibitors) have not been clinically effective in *de novo* pHGG, though clinical studies of HDAC
485 inhibitors are ongoing.[38,39] Mechanistically, Group B TIHGG may show vulnerability to these DNA-
486 damaging and anti-mitotic agents because of its DNA repair pathway deficiencies, as cells with
487 substantial drug-induced DNA damage prevented from completing mitosis would be expected to
488 undergo cell death associated with mitotic catastrophe.[40] Interestingly, in the *in vitro* drug screen, the
489 PARP2 inhibitor olaparib enhanced proliferation compared to vehicle in the Group B cell line (Sup. Table
490 21). Where DNA damage surveillance is intact, PARP2 inhibition is ineffective in inducing apoptosis and
491 instead induces a choice of repair pathway, favoring error-prone, non-homologous end resection over
492 homology-directed repair[41,42]. Thus, in Group B TIHGG, in which homology-directed repair pathways are
493 already deficient, PARP2 inhibition may enhance, rather than retard, proliferation by favoring non-
494 homologous resection.

495 Proteasome inhibitors, which have not been studied beyond preclinical phases in pHGG, were
496 effective *in vitro* in the Group B cell line, possibly through a mechanism involving inhibition of NF-κB-
497 mediated inflammatory pathways. Clinically, proteasome inhibition as a therapeutic target is established
498 in multiple myeloma and mantle cell lymphoma; however, its effectiveness is limited by acquired
499 resistance[43]. This experience suggests that the efficacy of proteasome inhibition for TIHGG may depend
500 upon identification of effective combination therapies. An interesting possibility for inhibiting the NF-κB
501 pathway in conjunction with proteasome inhibition is through inhibition of MDA-9/Syntenin, which has
502 shown promise in preclinical studies in reducing the invasiveness of radiation-induced GBM in adults.[44]
503 Targeting the MAPK pathway may also be an effective strategy given the frequent *NF1* mutations in
504 TIHGG; MEK inhibition is a therapeutic strategy of interest in HGG and was effective (trametinib) in both
505 TIHGG derived cell lines[45,46].

506 Strengths of the current study include the size and multi-institutional nature of the cohort, as
507 well as the comprehensive, orthogonal, unbiased assays performed. Recent characterization of large
508 cohorts of *de novo* HGG provide basis for comparing TIHGG to the HGG landscape in a way that has not
509 previously been possible. We also characterized two primary human TIHGG cell lines, allowing for
510 interrogation of therapeutic vulnerabilities. Weaknesses include the heterogeneity of the molecular
511 assays performed across the cohort samples; this has been partially ameliorated through the
512 performance of separate assays and analyses at only one institution per modality in order to standardize
513 results. In addition, *in vivo* models of TIHGG are still in development, and should be the focus of future
514 work. With cell lines available and common molecular alterations better understood, both patient-
515 derived xenograft and genetically engineered mouse models are possible. Our study demonstrates
516 similarities and differences between TIHGG and pHGG, as well as both homogeneity and heterogeneity
517 within TIHGG. Therefore, more efficacious treatment regimens may include a common backbone
518 incorporating some aspects of pHGG treatment, in combination with targeted treatments directed at
519 specific molecular alterations and susceptibilities characterizing TIHGG subgroups and individual patient
520 tumors.

521

522 **Methods**

523 ***TIHGG case review***

524 Fifty-four cases were reviewed across multiple organizations (Children's Hospital Colorado
525 (CHCO), St. Jude Children's Research Hospital, Childhood Cancer Survivors' Study (CCSS), the University
526 of Hamburg, and the University of Florida) from 1981 to 2015 to determine cohort eligibility. Initial
527 query of the CCSS institutional tumor tissue bank was based on the history of prior radiation with
528 subsequent development of HGG. Patients were enrolled on Institutional Review Board-approved

529 protocols for the harvest and study of tissue for research, including Colorado Multi-Institutional Review
530 Board (COMIRB 95-500) and SJCRH IRB Number: Pro00007403; Mnemonic: XPD17-029; Reference
531 Number: 001628. After clinical review, 7 cases were judged to be recurrent primary pHGG, two cases
532 were recurrent primary ependymoma, one case was found to be a recurrent vs. malignant
533 transformation of a juvenile pilocytic astrocytoma, and one case was found to be a recurrent
534 glioneuronal tumor. Tissue was not available for 6 cases, and tissue was of insufficient quality or
535 quantity in 2 cases (Sup. Fig. 11).

536 We used a modified version of Cahan's criteria to determine eligibility for tumors induced by
537 radiation[47]. All radiation-induced tumors arose within the initial irradiated field. Although Cahan's
538 criteria specify that TIHGG must have a histologically proven difference between the initial and
539 subsequent tumors, 7 cases had an initial diagnosis of glial origin (1 anaplastic astrocytoma, 2
540 astrocytomas, 3 ependymomas, and 1 ganglioglioma). The two cases with astrocytoma both presented
541 as low-grade pilocytic astrocytomas arising in the optic tract. Both cases were treated with radiotherapy
542 after progressive visual loss following a failed trial of chemotherapy. The resultant TIHGGs originated in
543 the posterior fossa and occipital lobe, respectively, both within a region of normal but previously
544 irradiated brain.

545 We also included tumors arising in two patients with germline predisposition syndromes in
546 which clinical evidence supported the diagnosis of induction by radiation. One patient with known
547 mismatch repair deficiency and germline *PMS2* mutation developed a primary anaplastic astrocytoma
548 (AA) localized to the right frontal lobe at the age of 14 years and was managed with a gross total
549 resection followed by erlotinib and adjuvant radiation as a part of SJHG04 (NCT00124657). Following 9.5
550 years of controlled disease, a new tumor was discovered centered in the left occipital lobe. The area was
551 judged to be of sufficient latency to be designated a TIHGG. After review of the clinical history,
552 radiographic imaging, initial radiotherapy plan, and pathology, it was felt that the subsequent GBM was

553 more likely to be treatment induced, as it arose in the prior radiotherapy field, differed from the initial
554 AA, arose from a region of normal brain parenchyma (except for previous irradiation), and occurred with
555 significant delay, making a late recurrence of AA exceedingly unlikely. The second patient had a
556 mismatch repair deficiency arising from a heterozygous loss of *MSH*. Upon presentation with
557 Philadelphia chromosome-positive ALL at age 9 years, the patient received a bone marrow transplant
558 and total body irradiation (12 Gy). At age 20 years, the patient developed a right frontal anaplastic
559 oligodendroglioma (WHO grade III). The clinical diagnosis of a radiation-induced tumor was based upon
560 the occurrence of the oligodendroglioma in the radiation field and the likelihood that radiation-induced
561 DNA damage played a significant role in the tumor formation.
562 One tumor clinically diagnosed as being induced by prior exposure to chemotherapeutic agents
563 was also included in the TIHGG cohort (Sup. Table 1)[47]. Details regarding the clinical history of all
564 reviewed and included cases are in Sup. Fig. 1.
565
566 ***TIHGG material available for analyses***
567 Deidentified tumor tissue specimens from frozen or formalin-fixed paraffin embedded (FFPE)
568 sections and frozen patient blood samples were processed following appropriate institutional review
569 board approval (Fig. 1E). Whole-genome methylation analysis was completed in 31 cases. RNA was of
570 sufficient quality for RNA-seq analysis in 14 cases. We conducted whole-genome sequencing on twelve
571 frozen tumor and matched blood samples (to obtain germline genomic information) and whole-exome
572 sequencing on 5 additional FFPE tumor samples[48].
573
574 ***Voxel-wise lesion mapping of TIHGG***

575 T1 images were registered to the Montreal Neurological Institute template[49] by using ANTsR[50]
576 and analyzed with voxel-based lesion symptom mapping (VLSM)[51] to assess similarity between statistical
577 maps by calculating the correlation between t-scores, treating lesion voxels as subjects.
578
579 ***Patient-level statistical analysis***
580 Patient- and sample-level statistical analyses were performed using Rstudio Version 1.1.463.
581 Packages used for the presented analyses include "survminer", "ggplot2", "survival", and "networkD3".
582 Continuous data are described using non-parametric measures of central tendency and are tested
583 across strata using the Wilcoxon-Mann-Whitney test. Frequency data across groups were evaluated
584 using either the Fisher's exact test or Chi-square test. Time to event endpoints are summarized using the
585 Kaplan Meier estimator. Differences in time to event strata are compared using the log-rank test.
586
587 ***Methylation array processing***
588 Tumor DNA was extracted from FFPE material by using the Maxwell16 FFPE Plus LEV purification
589 kit and the Maxwell 16 instrument (Promega, Madison, WI) according to the manufacturer's
590 instructions. Extracted DNA from FFPE tissue underwent quality control assessment using the Illumina
591 Infinium FFPE QC Assay kit for qPCR. The Delta Cq values for all samples were less than four. DNA
592 concentration was assessed using PicoGreen. At least 300 ng of DNA was used per sample for the
593 subsequent bisulfite conversion with the Zymo EZ-96 DNA Methylation kit. After the bisulfite
594 conversion, the Infinium HD FFPE Restoration kit was used to restore degraded FFPE DNA to a state that
595 is amplifiable by the Infinium HD FFPE methylation whole-genome amplification kit. Restored DNA was
596 then plate-purified (with the Zymo ZR-96 DNA Clean & Concentrator-5), amplified, fragmented,
597 precipitated, re-suspended, and hybridized to an Illumina Infinium Methylation EPIC 850K BeadChip
598 array for 22 hours and 30 minutes (by using the Illumina Infinium Methylation EPIC assay kit). Following

599 hybridization, the arrays were manually disassembled and washed. The subsequent X-Staining of the
600 array features was processed on a Tecan Freedom EVO robotics system. The arrays were then manually
601 coated and imaged using an Illumina iScan system with autoloader.

602 DNA methylation data analysis was performed using the open source statistical programming
603 language R (R Core Team, 2016). Raw data files generated by the iScan array scanner were read and
604 preprocessed using minfi Bioconductor package[52]. With the minfi package, the same preprocessing steps
605 as in Illumina's Genomestudio software were performed. In addition, the following filtering criteria were
606 applied: removal of probes targeting the X and Y chromosomes, removal of probes containing -
607 nucleotide polymorphism (dbSNP132 Common) within five base pairs of and including the targeted CpG-
608 site, and removal of probes not mapping uniquely to the human reference genome (hg19), allowing for
609 one mismatch. In total, 394,848 common probes of Illumina 450K and EPIC arrays were kept for
610 clustering analysis.

611

612 ***Statistical analysis of DNA methylation***

613 To determine the subgroup affiliation of our TIHGG samples, we used the reference DNA
614 methylation cohort published by Capper *et al*. and available from the gene expression omnibus
615 (GSE90496)[12]. TIHGG samples were combined with the 2,801 reference IDATs containing CNS tumors
616 and control brain tissues for unsupervised hierarchical clustering, as previously described[12]. In brief, the
617 32,000 most variable methylated CpG probes measured by standard deviation across combined samples
618 were selected. Pearson correlation was calculated as the distance measured between samples, and the
619 unsupervised hierarchical clustering was performed by the average linkage agglomeration method. The
620 probe-level beta values were also analyzed using t-stochastic neighbor embedding (t-SNE)[53]. Hierarchical
621 clustering and t-SNE analyses were repeated by using the top 12,000 differentially methylated probes
622 against a reduced reference set of 666 tumors representing 16 different previously described

623 methylation classes (Sup. Fig. 3, Sup. Table 7)[12]. The reduced set contained normal control methylation
624 classes, high-grade diffuse astrocytic tumors, high-grade neuroepithelial tumor with MN1 alteration,
625 pleomorphic xanthoastrocytoma, and anaplastic pilocytic astrocytoma. Supervised analysis was
626 performed using the random forest DNA methylation class prediction algorithm (V11b2) by uploading
627 raw IDAT files to www.molecularneuropathology.org[12].
628

629 ***Detection of copy-number alterations with methylation array data***

630 Copy-number variation was analyzed with Illumina methylation arrays using the conumee
631 Bioconductor package in R using default settings. The combined intensities of all available CpG probes
632 were normalized against control samples from normal brain tissue using a linear regression approach.
633 Platform-matched normal brain tissue controls were used to reference against EPIC 850k array or 450k
634 array tumor samples. Mean segment value of −0.2 and 0.2 were used as thresholds for losses and gains,
635 respectively. Copy-number plots were manually examined for selected copy-number alterations. When
636 copy-number information was also available from sequencing data, the two results were compared and
637 adjudicated. The adjudicated results are shown in Fig. 5 for selected recurrent somatic alterations.

638

639 ***RNA-seq analysis***

640 Library preparation was performed using the TruSeq Library Preparation Kit v2 (Agilent);
641 directional mRNA sequencing was performed at the University of Colorado Anschutz Medical Campus
642 Genomics and Microarray Core on a HiSeq 2500 sequencing system (Illumina) using single pass 125 bp
643 reads (1x125) and approximately 50 million reads per sample. The resulting data were mapped to the
644 human genome (hg19) by gSNAP, expression (FPKM) was derived by Cufflinks, and differential
645 expression was analyzed with ANOVA in R. Output files contained read-depth data and FPKM expression
646 levels for each sample and, where gene expression levels were compared between groups of samples,
647 the ratio of expression in $\log_2$ format and a p-value for each gene. CICERO was used to detect fusion
648 genes on the RNA-seq data[54].

649

650 ***Analysis of transcriptomic data***

651 Using micro-array data (Affymetrix Human Genome U133 Plus 2.0 Array) previously acquired
652 from tumor samples of patients treated at Children's Hospital Colorado, we compared patterns of gene

653 expression in 14 TIHGG samples (13 radiation-induced and one chemotherapy-induced) to those of a
654 cohort of non-treatment-induced tumors consisting of 24 primary pHGG, 4 infant HGG, and 14 adult
655 HGG. We performed clustering analysis using the t-SNE method available in the RTSNE package and
656 confirmed the TIHGG subgrouping obtained through t-SNE by using non-negative matrix factorization
657 (NMF)[53,55]. Principal component analysis with 30 initial dimensions preceded the t-SNE analysis, in which
658 we empirically selected perplexity of 3 and 50,000 iterations as providing optimal results. For the NMF
659 analysis, we utilized the identical micro-array dataset used in the t-SNE analysis, using k (number of
660 clusters) of 2–5. We performed Metascape analysis on the micro-array data using as input a list of 2,162
661 genes differentially expressed ($p<0.01$) between the TIHGG and HGG samples, followed by Cytoscape to
662 identify differentially enriched pathways[56,57]. We performed geneset enrichment analysis (GSEA, Broad
663 Institute) to identify the direction of enrichment between the HGG and TIHGG groups and GSEA and
664 Ingenuity Pathway Analysis (IPA, Qiagen) to identify gene expression patterns within the TIHGG
665 cohort[58,59]. GSEA results were evaluated using the normalized enrichment score (NES), in which
666 increased expression results in a positive score, and reduced expression in a negative score. Scores were
667 considered potentially informative from a statistical perspective if the false-discovery rate (FDR) was less
668 than 0.25.
669
670 ***Whole-genome sequencing***
671 WGS library preparation and sequencing were performed by BGI Americas; 100-fold mean
672 coverage data were acquired. BGI performed initial quality testing of the sample, including
673 concentration and sample integrity/purity. Concentration was detected by fluorometer or microplate
674 reader (*e.g.* Qubit Fluorometer, Invitrogen). Sample integrity and purity were detected by agarose gel
675 electrophoresis (agarose gel concentration: 1%, voltage: 150 V, electrophoresis time: 40 min). Following
676 quality confirmation, 1 µg genomic DNA was randomly fragmented by Covaris. The fragmented genomic

677 DNA was selected by Agencourt AMPure XP-Medium kit to an average size of 200-400 bp. Fragments
678 were end repaired and then 3' adenylated. Adaptors were ligated to the ends of these 3' adenylated
679 fragments to facilitate amplification by PCR. PCR products were purified using the Agencourt AMPure
680 XP-Medium kit. The double-stranded PCR products were heat denatured and circularized by the splint
681 oligonucleotide sequence. The single-strand circular DNA (ssCir DNA) was formed as the final library
682 qualified via a quality control procedure. The qualified libraries were sequenced by BGISEQ-500. Briefly,
683 each ssCir DNA molecule was formed into a DNA nanoball (DNB) containing more than 300 copies
684 through rolling-cycle replication. The DNBs were loaded into the patterned nanoarray using high density
685 DNA nanochip technology. Finally, pair-end 100-bp reads were obtained by combinatorial Probe-Anchor
686 Synthesis (cPAS). WGS mapping, quality control, and somatic mutation (SNV and INDEL) calling and
687 classification have been described previously[60,61].[60,61]. The tumor and germline reads were mapped to
688 GRCh37-lite and Tier1 mutations (i.e. coding somatic mutations). CNA regions were identified using
689 CONSERTING on the paired tumor-germline WGS samples with Log Ratio > 0.25 or < -0.25 reported[62].[62].
690 Structural variants (SVs) of tumor WGS samples were identified using CREST based on soft-clipped reads
691 evidence[63].[63]. The SVs that have discordant reads support in tumor sample but not in paired germline
692 sample were reported.
693
694 ***Comparison of mutation frequency in de novo pHGG and TIHGG by expression group***
695 We analyzed mutation load, small scale variants, and structural variants in the TIHGG tumor
696 samples and matched blood samples. We noticed, based upon our initial review of the genome
697 sequencing data, that several tumor samples appeared to have anomalously large numbers of mutations
698 and decided to perform a mutation load analysis based upon that observation. We prepared vcf files
699 containing QUAL-filtered (threshold of 50) and unfiltered putatively damaging variants. Germline
700 mutation load was defined as the total number of mutations per sample in the QUAL filtered dataset.

701 Somatic mutation load was determined from tumor samples without QUAL-filtering. The mutation calls
702 were filtered to require at least 4 reads per called variant in order to constitute a mutation. This
703 approach was deemed reasonable because of the likely heterogeneity present in the tumor samples
704 such that a clone constituting a fraction of the total sample could have an allelic-level mutation that
705 would be detected in only a small fraction of the overall reads at a particular locus, resulting in a low
706 QUAL score. Data were sensitivity tested to determine the thresholds for fold requirement and also to
707 determine whether a QUAL threshold should be imposed. The tumor sample from the patient with
708 Lynch syndrome was not included in this analysis because of the existence of a known DNA repair defect
709 unrelated to the therapeutic radiation treatment.
710
711 ***Whole-exome sequencing***
712 Human genomic libraries were generated using the SureSelectXT kit specific for the Illumina
713 HiSeq instrument (Catalog No. G9611B; Agilent Technologies, Santa Clara, CA), followed by exome
714 enrichment using the SureSelectXT Human All Exon V6+COSMIC bait set (Catalog No. 5190-9307). The
715 resulting exome-enriched libraries were then sequenced by the Genome Sequencing Facility on a HiSeq
716 4000 (Illumina). WES mapping and quality assessment have been described previously[61].[60,61]. The tumor
717 reads were mapped to GRCh37-lite, and variants were called by Bambino[18] and annotated by Medal
718 Ceremony as “Gold”, “Silver”, “Bronze” or “unknown”[64]. “Gold” mutations and variants matching with
719 COSMIC database were retained[65]. For other coding variants, those that were low-frequency (<0.001) or
720 absent in ExAC/1000Genome/NHLBI databases were reported if the variant was supported with at least
721 5 mutant alleles and at least 30% VAF[66,67]. The significance of mutated genes was assessed using the
722 Significantly Mutated Gene test[68]. Mutation frequency and composition were analyzed by comparing
723 the number and type of mutations across primary and TIHGG samples. The absolute number of
724 mutations, and frequencies of base pair substitutions in SNVs were compared across TIHGG and primary

725 HGG using a *t*-test and Chi-square test, respectively (Fig. 7C, Sup. Table 7, and Sup. Table 8). The
726 frequencies of commonly altered genes in TIHGG and primary HGG were compared by using Fisher's
727 exact test and are listed in (Sup. Table 7).
728
729 ***Evaluation of chromothripsis events and structure prediction of eccDNA***
730 The presence or absence of chromothripsis was evaluated in twelve samples' WGS data (Sup.
731 Table 8). Four key criteria were used to infer chromothripsis as described by Korbel *et al*; (oscillating
732 CNA regions, clustering of breakpoints, randomness of DNA fragment joins, and randomness of DNA
733 fragment order)[69]. Chromothripsis was called when at least two criteria were satisfied and further
734 evaluated by manual review. The eccDNA structures were constructed following the procedures
735 described in Xu *et al.* by identifying the cyclic graphs composed of highly amplified CNA segments and
736 their associated SVs[16].
737
738 ***FISH analysis***
739 Dual-color FISH was performed on 4-µm thick paraffin-embedded tissue sections. Probes were
740 derived from BAC clones (BACPAC Resources, Oakland, CA) and labeled with either AlexaFluor-488 or
741 AlexaFluor-555 fluorochromes. BAC clones were used to construct probes for the following genes:
742 *PDGFRA (*laboratory-developed probe (RP11-231C18 & 601I15); 4p control (CTD-2057N12 & CTD-
743 2588A19), and *CDK4* (Empire Genomics, Williamsville, New York, Cat# CDK4-CHR12-20-ORGR). Probes
744 were co-denatured with the target cells on a slide moat at 90°C for 12 minutes. The slides were
745 incubated overnight at 37°C on a slide moat and then washed in 4M Urea/2xSSC at 25°C for 1
746 minute. Nuclei were counterstained with DAPI (200ng/mL) (Vector Labs) for viewing on an Olympus
747 BX51 fluorescence microscope equipped with a 100-watt mercury lamp; FITC, rhodamine, and DAPI

748 filters; 100X PlanApo (1.40) oil objective; and a Jai CV digital camera. Images were captured and
749 processed using the Cytovision software from Leica Biosystems (Richmond, IL)[70].
750
751 **In silico *and* in vitro *drug screening and validations***
752 In silico *screen.* Micro-array data for 12 TIHGG samples and 37 *de novo* GBM samples (23
753 pediatric, 14 adult) were compared to target genes of several drug classes. A ratio of mean expression
754 and p-value for each target gene in the TIHGG and *de novo* GBM samples were computed, and an overall
755 ratio and p-value of all target genes used for each therapeutic classification for each target gene were
756 computed and reported (Supp. Table 20).
757 In vitro *drug screen and validation:* For the drug screen performed in TIHGG cell lines MAF-145
758 and MAF-496, we utilized the Approved Oncology Drugs Set VI (National Cancer Institute), comprising
759 129 drugs, supplemented by selinexor (Karyopharm Therapeutics) and AZD2014 (Astra Zeneca). The
760 complete list of drugs included in the screen is included in Sup. Table 21. Cells were plated at a density
761 of 5,000 cells per well in in 90 μL medium in a 96-well treated cell culture plate (Corning #3595) and
762 allowed to gain adhesion overnight. Drugs were applied in 10 μL of medium/1% DMSO at a
763 concentration of 10 μM, resulting in a final concentration of 1 μM and 0.1% DMSO. Cells were incubated
764 in drug for 5 days. DMSO (0.1%) was utilized as a control. Cell viability was assayed following 5 days of
765 treatment using incubation with tritiated thymidine and quantification using a scintillation counter.
766 Results were collected as counts/min and converted to survival using the formula (sample –
767 medum)/(DMSO – medium) where "sample" is the scintillation count for each drug-treated sample,
768 "medium" is the scintillation count for a well containing medium only, and "DMSO" is the scintillation
769 count for a three-well average of cells treated with 0.1% DMSO only (Sup. Table 21). The drug screen
770 was conducted twice in MAF-145 cells and once in MAF-496 due to limitations on cell availability.

771 Validation tests of single drugs were conducted using drug concentrations ranging from 0.316 nM to 10
772 μM in half-$\log_{10}$ increments. Cells were plated as described above and incubated in drug for 120 hours.
773 Three biological replicates were used for each drug concentration. Results were assessed using CellTiter
774 96 Aqueous One Solution Cell Proliferation Assay (Promega), according to manufacturer instructions.
775 Survival was computed as above. IC_{50} values were calculated using a variable slope four parameter non-
776 linear model with maximum survival constrained at 100% (Prism 7, Graphpad).
777 *Immunofluorescence staining of in vitro samples:* Cells were plated at a density of 20,000 cells
778 per well in BioCoat chamber slides coated with poly-D-lysine or poly-D-lysine and laminin (Corning) and
779 allowed approximately 24-48 hours in which to develop adhesion before subjecting them to
780 experimental conditions. Cells to be stained were fixed for 20 minutes in formaldehyde diluted to 3.7%
781 in PBS (Sigma), permeabilized in 0.1% Triton-X in PBS for 10 minutes, and blocked for 45 minutes in 4%
782 BSA in PBS supplemented with 0.05% Triton-X. Cells were incubated in primary antibody to the p65
783 subunit of NF-κB (Cell Signaling, #6956, 1:400) diluted with 4% BSA (in PBS and 0.05% Triton-X) for 1h at
784 room temperature or overnight at 4°C. Following multiple rinses with PBS, cells were incubated in a
785 secondary fluorophore (Alexa Fluor 488) for 1 hour, rinsed, and coverslips were then adhered using
786 ProLong Gold antifade reagent with DAPI (Invitrogen). Confocal imaging was performed at 400x using
787 405 nM (DAPI) and 488 nM (Alexa Fluor 488) lasers on a 3I Marianas imaging system (Intelligent Imaging
788 Innovations). Images were captured using an Evolve 16-bit EMCCD camera (Photometrics).

789

790 **Figures (8 total)**

791 Figure 1. TIHGG cohort characteristics.

792 A) Venn diagram illustrating the number of cases with each type of historical data/molecular analysis. B)
793 Radiotherapy field type used for initial cancer. C) Time to diagnosis of TIHGG from initial cancer
794 diagnosis. D) Time to death following TIHGG diagnosis. *Two patients survived beyond 35 months but
795 were deceased at study closure. E) Anatomic location of TIHGG. Color scale indicates the number of
796 cases anatomically overlapping at each point in space.

797

798 Figure 2. Methylation-based classification of TIHGG.

799 A) Circular dendrogram indicating location of TIHGG (black bars) relative to diagnoses. B) Localization of
800 TIHGG relative to other CNS cancers in t-distributed stochastic neighboring (t-SNE) space. C)
801 Dendrogram and heatmap of TIHGG cases clustered against other pediatric brain tumor entities. A-IDH=
802 Astrocytoma, subclass IDH-mutant; A-IDH-HG= High-Grade Astrocytoma, subclass IDH-mutant; ANA PA =
803 Anaplastic Pilocytic Astrocytoma, CONTR-CEBM= Control Cerebellum; CONTR-HEMI= Control Cerebral
804 Cortex; DMG-K27 = Diffuse Midline Glioma, H3 K27M-mutant; GBM-G34= Glioblastoma, subclass H3.3
805 p.G34R-mutant; GBM-MES= Glioblastoma, subclass mesenchymal; GBM-MID = Glioblastoma, IDH-Wild
806 type, subclass midline; GBM-MYCN= Glioblastoma, subclass MYCN-amplified; GBM-RTK-I = Adult
807 Glioblastoma, subclass RTK I; GBM-RTK-II= Adult Glioblastoma, subclass RTK II; GBM-RTK-II = Adult
808 Glioblastoma, subclass RTK III; HGNET-MN1 = High Grade Neuroepithelial Tumor with MN1 alteration;
809 IHG= Infantile High Grade Glioma; pedRTK I= Pediatric Glioblastoma, subclass RTK I; pedRTK II = Pediatric
810 Glioblastoma, subclass RTK II; PXA = Pleomorphic Xanthoastrocytoma; TIHGG = Treatment-Induced
811 High-Grade Glioma.

812

813

814 Figure 3. Frequent copy number alterations in TIHGG.

815 A) Copy-number frequency plot. B) Expression-based confirmation of high frequency copy number

816 alterations in *PDGFRα, CDKN2A,* and *BCOR;* mean and standard error of the mean (SEM) shown. C)

817 Frequency of copy number gain and loss across all TIHGG; boxes show median and $1^{st}/3^{rd}$ quartiles, with

818 whiskers representing range limited to 1.5x the interquartile range from the box edge.

819

820 Figure 4. Frequency of chromothripsis and resultant eccDNA structures, oncogene incorporation and

821 FISH analysis in Case_29.

822 A) Frequency of chromothripsis in primary non-brainstem high-grade glioma (NBS-HGG), primary DIPG,

823 TIHGG, and combined NBS-HGG and DIPG. B) Circos plot illustrating the intra- and inter-chromosomal

824 rearrangements of chromosomes 4 and 14, resulting in dm1, and intra-chromosomal rearrangements in

825 chromosome 12 resulting in dm2. Byproduct of chromothripsis events between chromosomes 4 and 12

826 (dm1) and intra-chromosomal rearrangements in chromosome 14 (dm2). The innermost circle shows

827 the highly amplified CNA segments, as evidenced by their extremely high coverage in the tumor sample

828 (blue circle) compared to that in the paired germline sample (green circle). The outermost circle shows

829 the involved chromosomes. C) Two-color FISH experiment from Case_29 shows *PDGFRα* (upper panel)

830 and *CDK4* amplification (lower panel) with diffuse punctate staining relative to the control centromeric

831 DNA in each cell, suggesting the presence of multiple double-minutes. D) The predicted structures of the

832 eccDNA dm1 and dm2 showing the directions of each of the joined CNA segments and their break point

833 positions. "Bridge reads support" means SVs identified through common discordant reads support

834 between two segment boundaries as described in Xu *et al*[16].

835

836 Figure 5. Recurrent molecular alterations in TIHGG.

837 Oncoprint describing the clinical characteristics, histopathological features, methylation profile, tier1
838 mutations, genes affected by copy-number gain/loss, and fusion genes. ANA PA = Anaplastic Pilocytic
839 Astrocytoma; CONTR-CEBM = Control Cerebellum; DMG-K27 = Diffuse Midline Glioma H3.3 K27M; GBM-
840 MID = Glioblastoma IDH-wild type, subclass midline; HGNET-MN1 = High Grade Neuroepithelial Tumor
841 with MN1 alteration; PXA = Pleomorphic Xanthoastrocytoma. Sequencing/array platforms for each case
842 are also shown on the bottom row. The second sample for Case 2 is shown as Case 2B while the second
843 sample for Case 14 and Case 12 are not shown but are annotated in Sup. Table 5.

844

845 Figure 6. Gene expression profiling of TIHGG versus *de novo* GBM.

846 A) Clustering of TIHGGs with microarray-based transcriptomic data (N=14) versus *de novo* pediatric GBM
847 (N=24), infant GBM (N=4) and adult GBM (N=14) using t-SNE analysis. B) Methylation clustering showing
848 locations of TIHGGs included in panel A; all TIHGGs included in panel A cluster in pedRTK I; methylation
849 and transcriptomic clusters overlap with the exception of two samples. C) Cytoscape analysis based on
850 gene ontology (GO) genesets show cellular pathways and processes that differ based upon gene
851 expression between TIHGG and *de novo* GBM (from panel A); color scale represents p-value, and the
852 comparisons are non-directional between sample sets. D) geneset enrichment analysis (GSEA) using GO
853 genesets identifies differences in gene expression between TIHGG and *de novo* GBM (from panel A);
854 number scale represents log of the false discovery rate (with FDR = 0 set to a value of 10 for purposes of
855 plotting); colors represent the ratio of GSEA normalized enrichment score (NES) of TIHGG/*de novo* GBM.

856

857 Figure 7. Gene expression and genetic profiling of Group A versus Group B TIHGG.

858 A) GSEA results showing GO differences organized by category between Group A (blue) and Group B
859 (red) TIHGG tumors; horizontal axis is NES for the comparison of TIHGG Group B vs. TIHGG Group A (NES

860 > 0 means enrichment in Group B and NES < 0 means enrichment in Group A) (FDR values for
861 comparisons are listed in Sup. Table 18); mean and SEM shown. B) Frequency of 1p loss and 1q gain
862 based upon methylation profiling in TIHGG gene expression groups A and B; note that one TIHGG tumor
863 for which there was transcriptomic data lacked a DNA sample and was not included in the methylation
864 profiling; mean and SEM shown. C) For TIHGG samples with germline somatic genome sequencing data
865 and transcriptomic data, germline variant load was identical between subgroups, but somatic load was
866 approximately 9-fold greater in Group B ($p<0.002$); TIHGG sample with *MSH2* mismatch repair defect
867 was excluded from this analysis. D) GSEA results for DNA repair pathways for Group B versus Group A
868 germline (based on blood) and tumor samples; negative numbers represent depletion in Group B;
869 Individual genes from the GO_DNA_REPAIR geneset with the most significant differences in expression
870 between Group B and Group A are shown by decreasing p-value (top to bottom); colors represent the
871 ratio of mean expression (ER) of Group B/Group A.
872
873 Figure 8. TIHGG preclinical drug screen.
874 A) *In silico*-predicted response of TIHGG vs. *de novo* GBM tumors to drug classes based upon gene
875 expression data. B) *In vitro* drug screening results by drug class for TIHGG cell lines MAF-145 (gene
876 expression Group B) and MAF-496 (gene expression Group A) showing reduced survival by drug class
877 relative to mean survival; screen performed using FDA-approved anti-cancer agents at 1 μM
878 concentration. C) *In vitro* drug screen results in TIHGG cell lines using 1 μM concentration; Group A
879 (MAF-496) cell line response is plotted on the x-axis and Group B cell line (MAF-145) on the y-axis. Dot
880 colors correspond to drug class as shown in the legend. D) *In vitro* validation results in TIHGG cell lines
881 MAF-145 and MAF-496 for candidate drugs identified through *in vitro* drug screen (panels B, D)
882

883 **Supplementary Figures** (11 total)

884 Supplementary Figure 1. Case Treatment Histories.

885 Supplementary Figure 2. Anatomical distribution of *de novo* pHGG and TIHGG at initial presentation. A)
886 (Top Row) Distribution of 88 "*de novo*" pHGGs; (2nd Row) 18 TIHGG; (3rd Row) Combined distribution of
887 both *de novo* pHGG and TIHGGs; (Bottom Row) Spatial differences in distribution between *de novo*
888 pHGGs and TIHGG using voxel-based lesion mapping (VLSM). B) Imaging analysis schema.

889 Supplementary Figure 3. Position of TIHGG among *de novo* CNS tumors. A) Localization of TIHGG relative
890 to other CNS cancers in t-SNE space. B) Magnified region indicated by A1 illustrating the close
891 relationship between GBM-MID and TIHGG. C) Diagnosis key.

892 Supplementary Figure 4. Results of clustering of TIHGG cases. A) Sankey diagram illustrating the inter-
893 relationship between initial diagnosis and methylation-defined CNS tumor subgroup. B) Latency of
894 TIHGG by consensus methylation cluster. C) Overall survival of TIHGG cases stratified by pedRTK I vs.
895 other vs. unknown.

896 Supplementary Figure 5. Comparison of TIHGG to *de novo* pHGG. A) Proportion of consensus cluster
897 subtypes of TIHGG and primary pHGG cases from the HERBY dataset. B) Survival of TIHGG and pHGG
898 stratified by consensus methylation group. C) Total genomic length (Mb) lost and gained by methylation
899 subgroup; boxes show median and $1^{st}/3^{rd}$ quartiles, with whiskers representing range limited to 1.5x the
900 interquartile range from the box edge.

901 Supplementary Figure 6. Spectral Plot illustrating chromosome-level copy-number alterations across
902 TIHGG cases.

903 Supplementary Figure 7. Comparison of copy-number alterations in TIHGG and *de novo* pHGG stratified
904 by consensus methylation class.

905 Supplementary Figure 8. Recurrent molecular alterations in TIHGG relative to *de novo* pHGG (HERBY
906 dataset). Oncoprint describing the clinical characteristics, histopathological features, methylation
907 profile, tier1- mutations, genes affected by copy-number gain/loss, and fusion genes in TIHGG and *de*
908 *novo* pHGG.

909 Supplementary Figure 9. Germline and somatic variants in TIHGG relative to *de novo* pHGG (Wu G
910 reference cohort) stratified by base transition type. A) Total number of germline, somatic coding, and
911 somatic non-coding variants per Mb stratified by *de novo* pHGG vs. TIHGG. B) Proportion of base
912 transitions in somatic non-coding variants per Mb stratified by primary pHGG vs. TIHGG. The relative
913 frequency of base transitions in germline coding SNVs. C) Somatic coding SNVs and somatic non-
914 noncoding SNVs. For violin plots, box shows median and interquartile range, whiskers show 95%
915 confidence interval, total length of violin represents range, and width of violin shows frequency.

916 Supplementary Figure 10. TIHGG *in vitro* sensitivities by expression profile. A) *In vitro* drug screen
917 consort diagram; primary cell lines derived from TIHGG tumors were MAF-145 (gene expression TIHGG
918 Group B) and MAF-496 (gene expression TIHGG Group A). B) MTS assay results for indicated drugs in
919 MAF-496 (IC_{50} values listed in Fig. 8D). C) MTS results for indicated drugs in MAF-145 (IC_{50} values listed in
920 Fig. 8D). D) MTS results for trametinib. E) MTS results for sunitinib. F) MTS assay results for proteasome
921 inhibitors bortezomib and carfilzomib in MAF-145 and MAF-496 cell lines. G) MTS results for marizomib.
922 H) NF-κB levels in Group A (MAF-496) and B (MAF-145) TIHGG cell lines treated with vehicle and
923 bortezomib. I) Immunofluorescence of NF-κB nuclear localization in vehicle bortezomib-treated Group A
924 (MAF-496) and Group B (MAF-145) TIHGG cell lines. Mean and SEM shown.

925 Supplementary Figure 11. Study consort diagram. CCSS = Childhood Cancer Survivor Study, SJCRH = St.
926 Jude Children's Research Hospital, U. Colorado = University of Colorado, WGS = whole-genome
927 sequencing, WES = whole-exome sequencing, RNA-seq = RNA sequencing.

928 **Supplementary Tables** (21 total)

929 Supplementary Table 1. Patient, treatment, initial, and TIHGG disease characteristics.

930 Supplementary Table 2. TIHGG treatment and outcomes.

931 Supplementary Table 3. Methylation-based classification probabilities for the reference dataset for
932 methylation-based classification (Korshunov A and Capper D).

933 Supplementary Table 4. Methylation-based classification probabilities.

934 Supplementary Table 5. Methylation-based classification summary.

935 Supplementary Table 6. Comparison of TIHGG and *de novo* pHGG methylation class.

936 Supplementary Table 7. Comparison of TIHGG and *de novo* pHGG from the HERBY cohort.

937 Supplementary Table 8. Chromothripsis evaluation criteria.

938 Supplementary Table 9. Tier 1 sequence mutations identified from 5 non-hypermutator WES cases
939 without matched germline data.

940 Supplementary Table 10. Tier 1 sequence mutations identified from 9 non-hypermutator WGS cases
941 with matched germline data.

942 Supplementary Table 11. TIHGG fusion genes in RNA-seq samples.

943 Supplementary Table 12. TIHGG structural variations in 10 WGS cases.

944 Supplementary Table 13. TIHGG copy-number alteration segments in 10 WGS cases.

945 Supplementary Table 14. Methylation-based classification probabilities for the HERBY dataset

946 Supplementary Table 15. Germline variants.

947 Supplementary Table 16. Cytoscape/Metascape results

948 Supplementary Table 17. GSEA results in TIHGG vs. *de novo* GBM for GO, with all genesets

949 Supplementary Table 18. GSEA results by category as shown in Fig. 7A.

950 Supplementary Table 19. Expression Ratio (ER) and p-value (Group B TIHGGs/Group A TIHGGs) for

951 selected genes from GO_DNA_REPAIR geneset.

952 Supplementary Table 20. *In silico* drug screening results.

953 Supplementary Table 21. *In vitro* drug screening results.

954 **References**

955 1. Armstrong, G.T. *et al.* Long-term outcomes among adult survivors of childhood central nervous
956 system malignancies in the Childhood Cancer Survivor Study. *J Natl Cancer Inst* **101**, 946-58
957 (2009).
958 2. Armstrong, G.T. *et al.* Late mortality among 5-year survivors of childhood cancer: a summary
959 from the Childhood Cancer Survivor Study. *J Clin Oncol* **27**, 2328-38 (2009).
960 3. Maluf, F.C., DeAngelis, L.M., Raizer, J.J. & Abrey, L.E. High-grade gliomas in patients with prior
961 systemic malignancies. *Cancer* **94**, 3219-24 (2002).
962 4. Carret, A.S. *et al.* Outcome of secondary high-grade glioma in children previously treated for a
963 malignant condition: a study of the Canadian Pediatric Brain Tumour Consortium. *Radiother*
964 *Oncol* **81**, 33-8 (2006).
965 5. Paugh, B.S. *et al.* Integrated molecular genetic profiling of pediatric high-grade gliomas reveals
966 key differences with the adult disease. *J Clin Oncol* **28**, 3061-8 (2010).
967 6. Behjati, S. *et al.* Mutational signatures of ionizing radiation in second malignancies. *Nat Commun*
968 **7**, 12605 (2016).
969 7. Lopez, G.Y. *et al.* The genetic landscape of gliomas arising after therapeutic radiation. *Acta*
970 *Neuropathol* **137**, 139-150 (2019).
971 8. Sholl, L.M., Barletta, J.A. & Hornick, J.L. Radiation-associated neoplasia: clinical, pathological and
972 genomic correlates. *Histopathology* **70**, 70-80 (2017).
973 9. Leeman-Neill, R.J. *et al.* ETV6-NTRK3 is a common chromosomal rearrangement in radiation-
974 associated thyroid cancer. *Cancer* **120**, 799-807 (2014).
975 10. Leeman-Neill, R.J. *et al.* RET/PTC and PAX8/PPARgamma chromosomal rearrangements in post-
976 Chernobyl thyroid cancer and their association with iodine-131 radiation dose and other
977 characteristics. *Cancer* **119**, 1792-9 (2013).
978 11. Korshunov, A. *et al.* H3-/IDH-wild type pediatric glioblastoma is comprised of molecularly and
979 prognostically distinct subtypes with associated oncogenic drivers. *Acta Neuropathol* **134**, 507-
980 516 (2017).
981 12. Capper, D. *et al.* DNA methylation-based classification of central nervous system tumours.
982 *Nature* **555**, 469-474 (2018).
983 13. Sturm, D. *et al.* Paediatric and adult glioblastoma: multiform (epi)genomic culprits emerge. *Nat*
984 *Rev Cancer* **14**, 92-107 (2014).
985 14. Mackay, A. *et al.* Molecular, Pathological, Radiological, and Immune Profiling of Non-brainstem
986 Pediatric High-Grade Glioma from the HERBY Phase II Randomized Trial. *Cancer Cell* **33**, 829-842
987 e5 (2018).
988 15. Wu, G. *et al.* The genomic landscape of diffuse intrinsic pontine glioma and pediatric non-
989 brainstem high-grade glioma. *Nat Genet* **46**, 444-450 (2014).
990 16. Xu, K. *et al.* Structure and evolution of double minutes in diagnosis and relapse brain tumors.
991 *Acta Neuropathol* **137**, 123-137 (2019).
992 17. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology
993 Consortium.
994 18. Edmonson, M.N. *et al.* Bambino: a variant detector and alignment viewer for next-generation
995 sequencing data in the SAM/BAM format. *Bioinformatics* **27**, 865-6 (2011).
996 19. Gits, H.C. *et al.* Medulloblastoma therapy generates risk of a poorly-prognostic H3 wild-type
997 subgroup of diffuse intrinsic pontine glioma: a report from the International DIPG Registry. *Acta*
998 *Neuropathol Commun* **6**, 67 (2018).

20. Mackay, A. *et al.* Integrated Molecular Meta-Analysis of 1,000 Pediatric High-Grade and Diffuse Intrinsic Pontine Glioma. *Cancer Cell* **32**, 520-537 e5 (2017).
21. Downing, J.R. *et al.* The Pediatric Cancer Genome Project. *Nat Genet* **44**, 619-22 (2012).
22. Mistry, M. *et al.* BRAF mutation and CDKN2A deletion define a clinically distinct subgroup of childhood secondary high-grade glioma. *J Clin Oncol* **33**, 1015-22 (2015).
23. Collado, M. & Serrano, M. Senescence in tumours: evidence from mice and humans. *Nat Rev Cancer* **10**, 51-7 (2010).
24. Lassaletta, A. *et al.* Therapeutic and Prognostic Implications of BRAF V600E in Pediatric Low-Grade Gliomas. *J Clin Oncol* **35**, 2934-2941 (2017).
25. Hudler, P. *et al.* Mutations in the hMLH1 gene in Slovenian patients with gastric carcinoma. *Clin Genet* **65**, 405-11 (2004).
26. De Ruysscher, D. *et al.* Radiotherapy toxicity. *Nat Rev Dis Primers* **5**, 13 (2019).
27. Kilday, J.P. *et al.* Copy number gain of 1q25 predicts poor progression-free survival for pediatric intracranial ependymomas and enables patient risk stratification: a prospective European clinical trial cohort analysis on behalf of the Children's Cancer Leukaemia Group (CCLG), Societe Francaise d'Oncologie Pediatrique (SFOP), and International Society for Pediatric Oncology (SIOP). *Clin Cancer Res* **18**, 2001-11 (2012).
28. Hing, S. *et al.* Gain of 1q is associated with adverse outcome in favorable histology Wilms' tumors. *The American journal of pathology* **158**, 393-398 (2001).
29. Gratias, E.J. *et al.* Association of Chromosome 1q Gain With Inferior Survival in Favorable-Histology Wilms Tumor: A Report From the Children’s Oncology Group. *Journal of Clinical Oncology* **34**, 3189-3194 (2016).
30. Hirai, M. *et al.* 1q23 gain is associated with progressive neuroblastoma resistant to aggressive treatment.
31. Molenaar, J.J. *et al.* Sequencing of neuroblastoma identifies chromothripsis and defects in neuritogenesis genes. *Nature* **483**, 589-93 (2012).
32. Roy, A. *et al.* Recurrent internal tandem duplications of BCOR in clear cell sarcoma of the kidney. *Nat Commun* **6**, 8891 (2015).
33. Olanich, M.E. *et al.* CDK4 Amplification Reduces Sensitivity to CDK4/6 Inhibition in Fusion-Positive Rhabdomyosarcoma. *Clin Cancer Res* **21**, 4947-59 (2015).
34. Szerlip, N.J. *et al.* Intratumoral heterogeneity of receptor tyrosine kinases EGFR and PDGFRA amplification in glioblastoma defines subpopulations with distinct growth factor response. *Proc Natl Acad Sci U S A* **109**, 3041-6 (2012).
35. Nathanson, D.A. *et al.* Targeted therapy resistance mediated by dynamic regulation of extrachromosomal mutant EGFR DNA. *Science* **343**, 72-6 (2014).
36. Turner, K.M. *et al.* Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. *Nature* **543**, 122-125 (2017).
37. L'Abbate, A. *et al.* Genomic organization and evolution of double minutes/homogeneously staining regions with MYC amplification in human cancer. *Nucleic Acids Res* **42**, 9131-45 (2014).
38. Zhang, D., Liu, X., Fan, C. & Chen, J. Novel drugs in pediatric gliomas. *Oncol Lett* **13**, 2881-2885 (2017).
39. Williams, M.J., Singleton, W.G.B., Lowis, S.P., Malik, K. & Kurian, K.M. Therapeutic Targeting of Histone Modifications in Adult and Pediatric High-Grade Glioma. *Frontiers in oncology* **7**, 45-45 (2017).
40. Huang, X., Tran, T., Zhang, L., Hatcher, R. & Zhang, P. DNA damage-induced mitotic catastrophe is mediated by the Chk1-dependent mitotic exit DNA damage checkpoint. *Proc Natl Acad Sci U S A* **102**, 1065-70 (2005).

1046 41. Takagi, M. *et al.* Loss of DNA Damage Response in Neuroblastoma and Utility of a PARP
1047 Inhibitor. *JNCI: Journal of the National Cancer Institute* **109**(2017).
1048 42. Fouquin, A. *et al.* PARP2 controls double-strand break repair pathway choice by limiting 53BP1
1049 accumulation at DNA damage sites and promoting end-resection. *Nucleic Acids Research* **45**,
1050 12325-12339 (2017).
1051 43. Acosta-Alvear, D. *et al.* Paradoxical resistance of multiple myeloma to proteasome inhibitors by
1052 decreased levels of 19S proteasomal subunits.
1053 44. Kegelman, T.P. *et al.* Inhibition of radiation-induced glioblastoma invasion by genetic and
1054 pharmacological targeting of MDA-9/Syntenin. *Proc Natl Acad Sci U S A* **114**, 370-375 (2017).
1055 45. See, W.L., Tan, I.L., Mukherjee, J., Nicolaides, T. & Pieper, R.O. Sensitivity of glioblastomas to
1056 clinically available MEK inhibitors is defined by neurofibromin 1 deficiency. *Cancer Res* **72**, 3350-
1057 9 (2012).
1058 46. Sunayama, J. *et al.* Crosstalk between the PI3K/mTOR and MEK/ERK pathways involved in the
1059 maintenance of self-renewal and tumorigenicity of glioblastoma stem-like cells. *Stem Cells* **28**,
1060 1930-9 (2010).
1061 47. Cahan, W.G., Woodard, H.Q. & et al. Sarcoma arising in irradiated bone; report of 11 cases.
1062 *Cancer* **1**, 3-29 (1948).
1063 48. Donson, A.M. *et al.* Unique molecular characteristics of radiation-induced glioblastoma. *J*
1064 *Neuropathol Exp Neurol* **66**, 740-9 (2007).
1065 49. Fonov, V. *et al.* Unbiased average age-appropriate atlases for pediatric studies. *Neuroimage* **54**,
1066 313-27 (2011).
1067 50. Avants, B.B. *et al.* A reproducible evaluation of ANTs similarity metric performance in brain
1068 image registration. *Neuroimage* **54**, 2033-44 (2011).
1069 51. Bates, E. *et al.* Voxel-based lesion-symptom mapping. *Nat Neurosci* **6**, 448-50 (2003).
1070 52. Aryee, M.J. *et al.* Minfi: a flexible and comprehensive Bioconductor package for the analysis of
1071 Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363-9 (2014).
1072 53. van der Maaten, L. & Hinton, G. Visualizing Data using t-SNE. *Journal of Machine Learning*
1073 *Research* **9**, 2579--2605 (2008).
1074 54. Parker, M. *et al.* C11orf95-RELA fusions drive oncogenic NF-kappaB signalling in ependymoma.
1075 *Nature* **506**, 451-5 (2014).
1076 55. Brunet, J.-P., Tamayo, P., Golub, T.R. & Mesirov, J.P. Metagenes and molecular pattern discovery
1077 using matrix factorization. *Proceedings of the National Academy of Sciences* **101**, 4164-4169
1078 (2004).
1079 56. Tripathi, S. *et al.* Meta- and Orthogonal Integration of Influenza "OMICs" Data Defines a Role for
1080 UBR4 in Virus Budding. *Cell Host Microbe* **18**, 723-35 (2015).
1081 57. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular
1082 interaction networks. *Genome Res* **13**, 2498-504 (2003).
1083 58. Mootha, V.K. *et al.* PGC-1α-responsive genes involved in oxidative phosphorylation are
1084 coordinately downregulated in human diabetes. *Nature Genetics* **34**, 267 (2003).
1085 59. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for
1086 interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*
1087 **102**, 15545-15550 (2005).
1088 60. Zhang, J. *et al.* The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature*
1089 **481**, 157-63 (2012).
1090 61. Zhang, J. *et al.* A novel retinoblastoma therapy from genomic and epigenetic analyses. *Nature*
1091 **481**, 329-34 (2012).
1092 62. Chen, X. *et al.* CONSERTING: integrating copy-number analysis with structural-variation
1093 detection. *Nat Methods* **12**, 527-30 (2015).

1094 63. Wang, J. *et al.* CREST maps somatic structural variation in cancer genomes with base-pair
1095 resolution. *Nat Methods* **8**, 652-4 (2011).
1096 64. Zhang, J. *et al.* Germline Mutations in Predisposition Genes in Pediatric Cancer. *N Engl J Med*
1097 **373**, 2336-2346 (2015).
1098 65. Tate, J.G. *et al.* COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res* **47**,
1099 D941-D947 (2019).
1100 66. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285-91
1101 (2016).
1102 67. Genomes Project, C. *et al.* A global reference for human genetic variation. *Nature* **526**, 68-74
1103 (2015).
1104 68. Dees, N.D. *et al.* MuSiC: identifying mutational significance in cancer genomes. *Genome Res* **22**,
1105 1589-98 (2012).
1106 69. Korbel, J.O. & Campbell, P.J. Criteria for inference of chromothripsis in cancer genomes. *Cell* **152**,
1107 1226-36 (2013).
1108 70. Ellison, D.W. *et al.* Definition of disease-risk stratification groups in childhood medulloblastoma
1109 using combined clinical, pathologic, and molecular variables. *J Clin Oncol* **29**, 1400-7 (2011).

1110

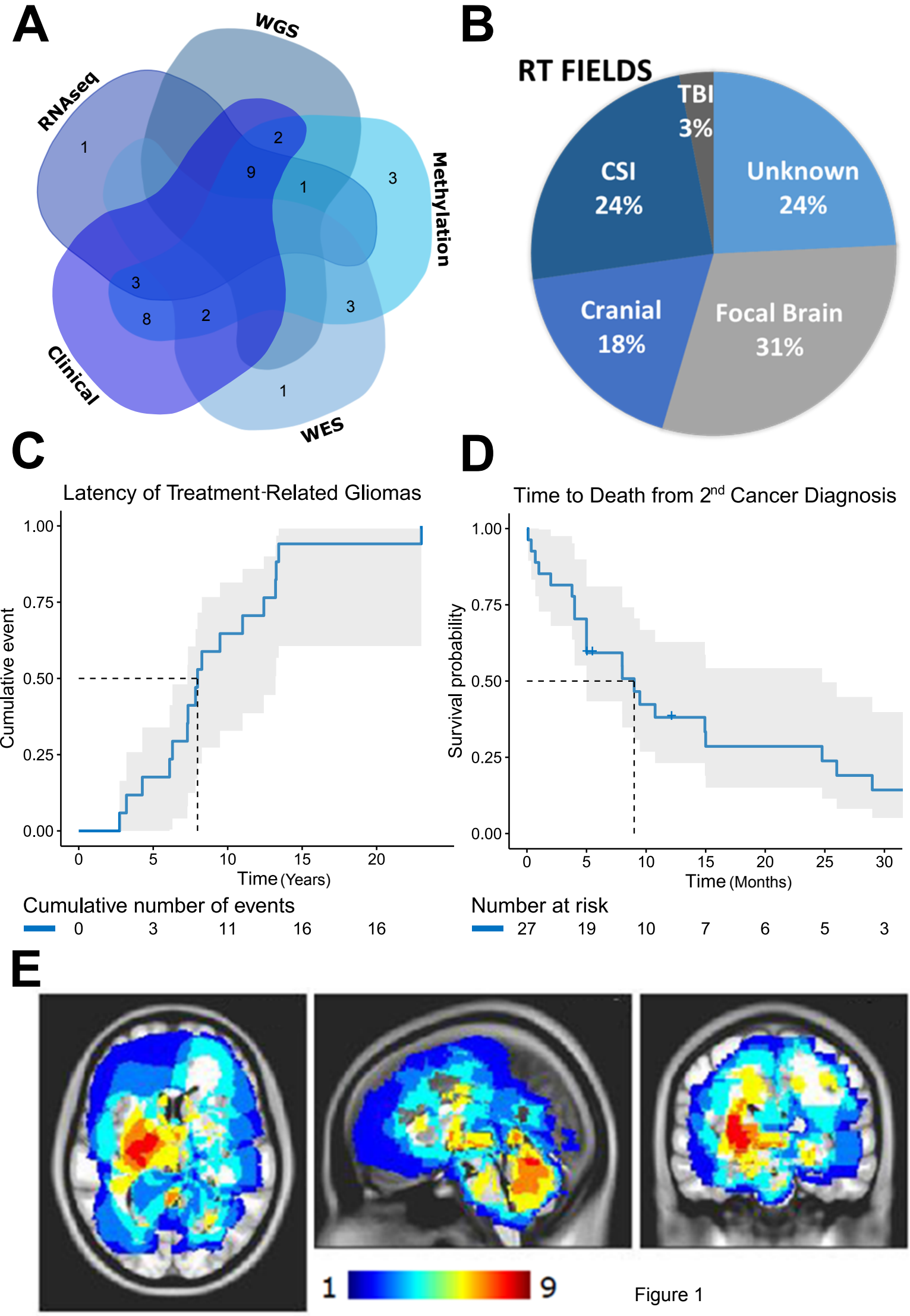

A
WGS
RNAseq
Methylation
Clinical
WES
1
2
9
1
3
3
3
8
2
1
B
RT FIELDS
TBI
3%
CSI
24%
Unknown
24%
Cranial
18%
Focal Brain
31%
C
Latency of Treatment-Related Gliomas
Cumulative event
1.00
0.75
0.50
0.25
0.00
0
5
10
15
20
Time (Years)
Cumulative number of events
0
3
11
16
16
D
Time to Death from 2nd Cancer Diagnosis
Survival probability
1.00
0.75
0.50
0.25
0.00
0
5
10
15
20
25
30
Time (Months)
Number at risk
27
19
10
7
6
5
3
E
1
9

Figure 1

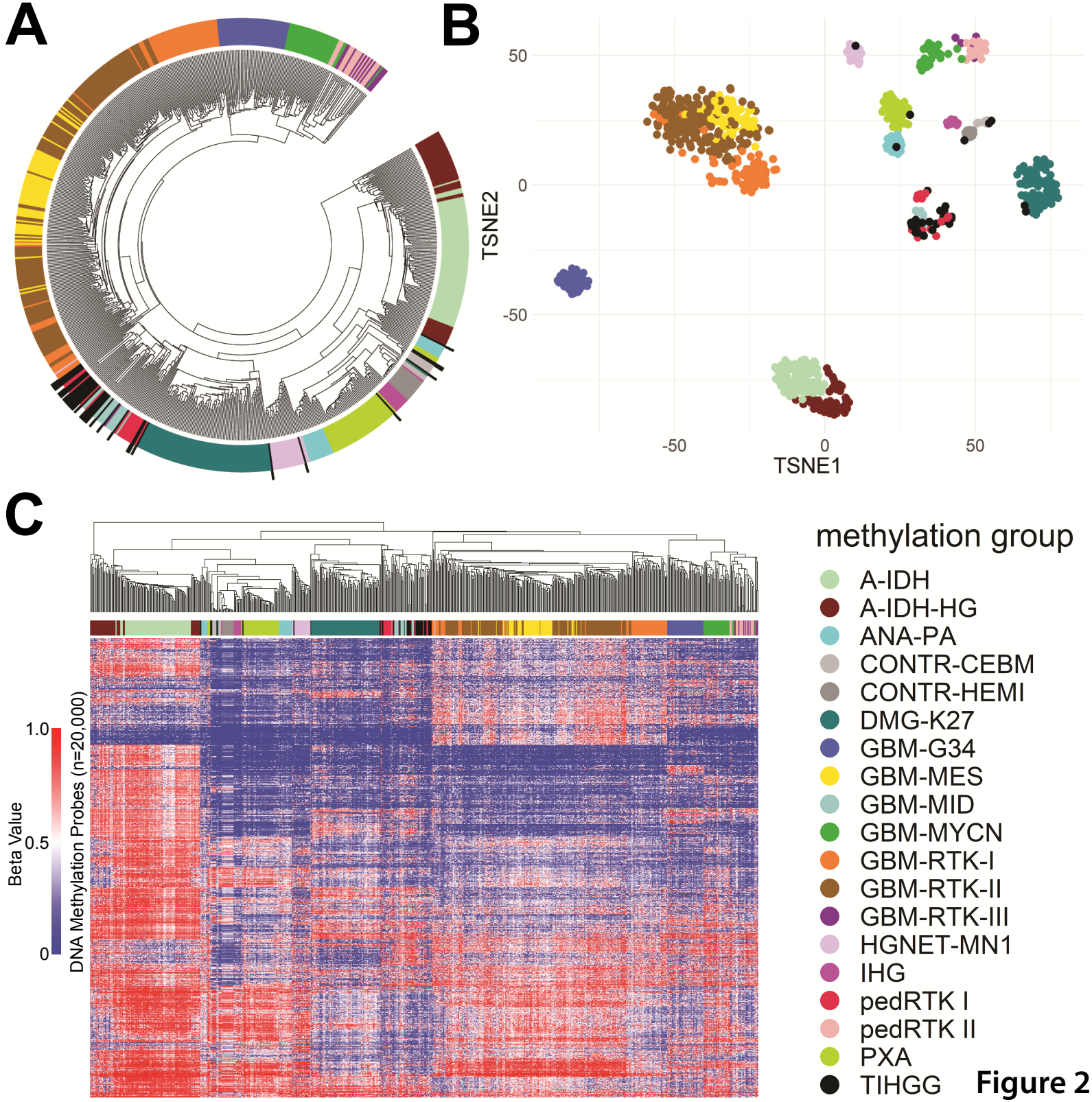
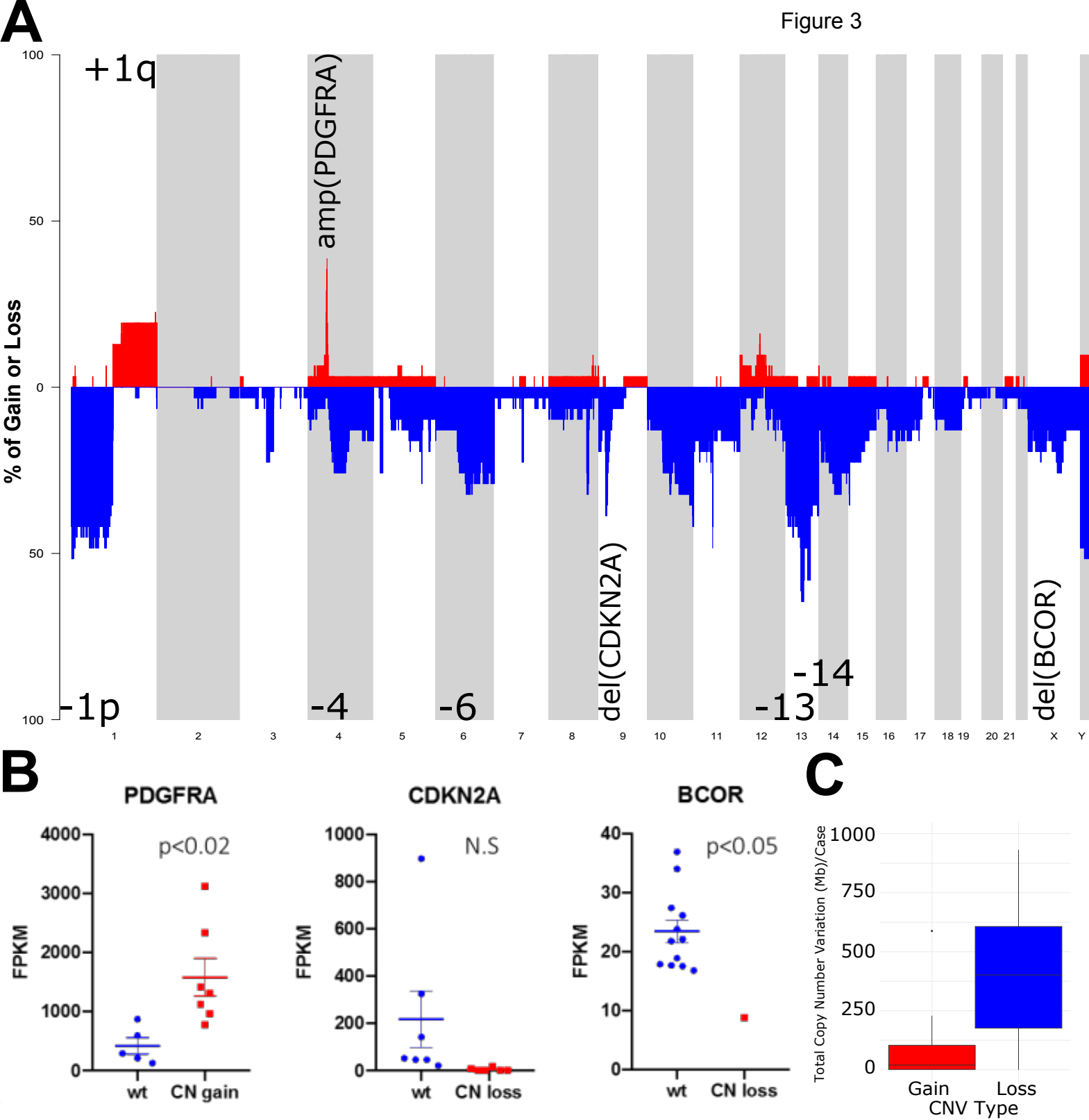

A
B
50
0
-50
TSNE2
-50
0
50
TSNE1
C
1.0
0.5
0
Beta Value
DNA Methylation Probes (n=20,000)
methylation group
A-IDH
A-IDH-HG
ANA-PA
CONTR-CEBM
CONTR-HEMI
DMG-K27
GBM-G34
GBM-MES
GBM-MID
GBM-MYCN
GBM-RTK-I
GBM-RTK-II
GBM-RTK-III
HGNET-MN1
IHG
pedRTK I
pedRTK II
PXA
TIHGG

Figure 2

Figure 3


A
+1q
amp(PDGFRA)
-1p
-4
-6
del(CDKN2A)
-13
-14
del(BCOR)
% of Gain or Loss
100
50
0
50
100
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 X Y
B
PDGFRA
p<0.02
FPKM
4000
3000
2000
1000
0
wt
CN gain
CDKN2A
N.S
1000
800
600
400
200
0
wt
CN loss
BCOR
p<0.05
40
30
20
10
0
wt
CN loss
C
Total Copy Number Variation (Mb)/Case
1000
750
500
250
0
Gain
Loss
CNV Type

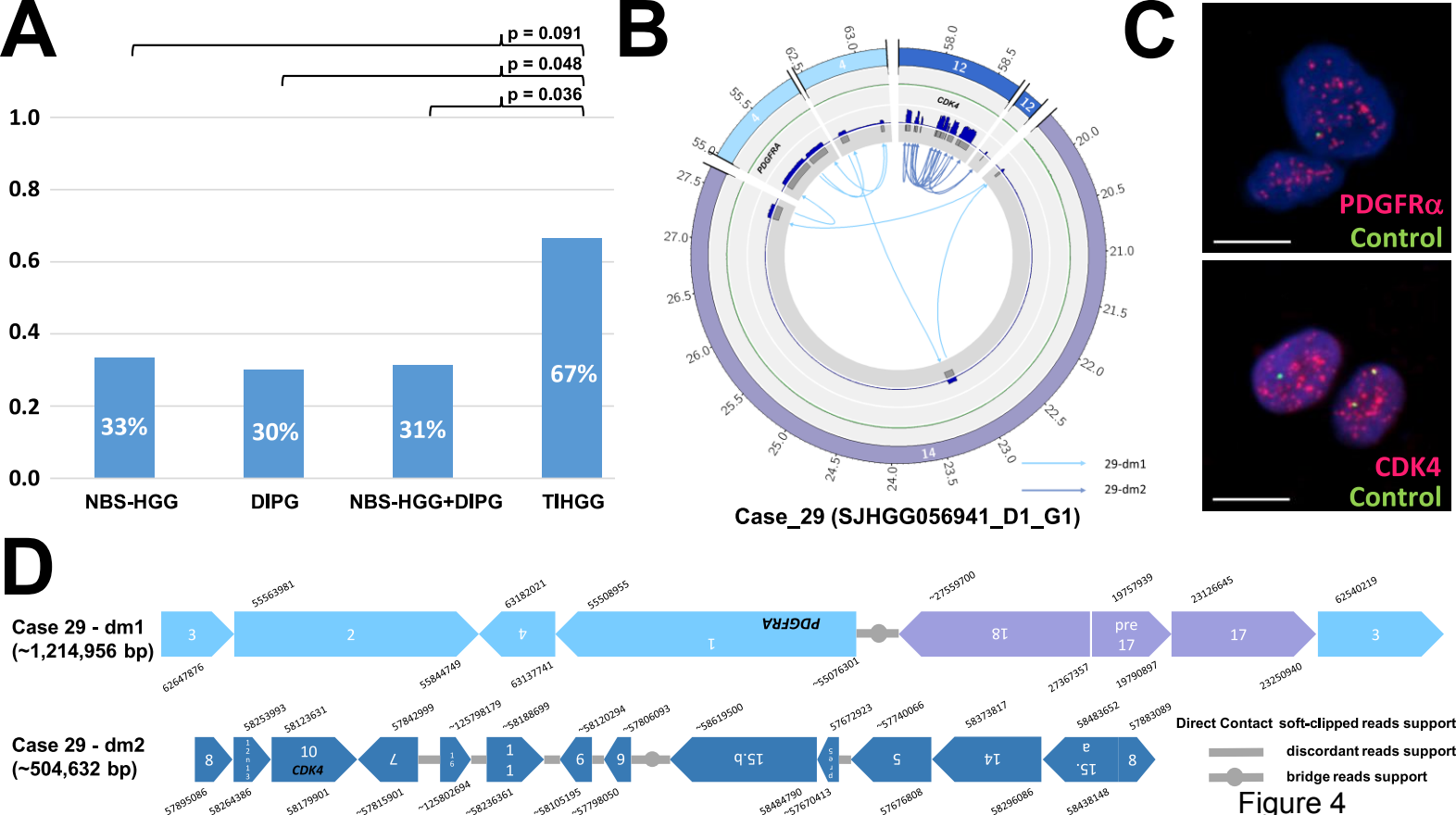

A
p = 0.091
p = 0.048
p = 0.036
1.0
0.8
0.6
0.4
0.2
0.0
33%
30%
31%
67%
NBS-HGG
DIPG
NBS-HGG+DIPG
TIHGG
B
CDK4
PDGFRA
29-dm1
29-dm2
Case_29 (SJHGG056941_D1_G1)
C
PDGFRα
Control
CDK4
Control
D
Case 29 - dm1
(~1,214,956 bp)
Case 29 - dm2
(~504,632 bp)
Direct Contact
soft-clipped reads support
discordant reads support
bridge reads support

Figure 4

Prior Diagnosis
Prior RT Fields
Latency
TIHGG Grade
Relation to Prior RT
Consensus Cluster
Chromothripsis
Expression Group
PDGFRA (15)
NF1 (9)
TP53 (8)
BCOR (10)
CDKN2A (11)
KMT2D (3)
MET (5)
SMARCA4 (4)
SETD2 (4)
CDK4 (6)
HMCN1 (4)
FLNC (4)
PTPN11 (4)
ATRX (2)
BRAF (3)
GHR (2)
TRPM2 (3)
PIK3CA (2)
RB1 (2)
CTCF (2)
H3F3A (1)
IDH2 (1)
Case ID
25 26 30 09 28 27 29 16 14 12 15 31 08 33 21 2B 02 07 01 23 34 05 24 43 20 06 13 22 44 42 32 11 18 19
Platform
Methylation Array
RNAseq
WGS
WES
Mutation Types
Missense
Frameshift
Nonsense
Splice
Fusion transcript
Copy-number gain
Copy-number loss
Prior Diagnosis
Craniopharyngioma
Embryonal tumors
Germinoma
Glial/Ependymal tumors
Leukemia/Lymphoma
Neuroblastoma
* Hypermutator
Prior RT Fields
CSI
Focal
TBI
Unknown
WBRT
None
Latency
10+ years
2-5 years
5-10 years
Unknown
TIHGG Grade
III
III/IV
IV
Relation to Prior RT
In field, Gradient
In field, Rx Dose
Unknown
None
Consensus Cluster
pedRTK I
Control-Cerebellum
PXA
HGNET-MN1
DMG-K27
ANA-PA
Chromothripsis
No
Yes
Undefined
Expression Group
A
B
Undefined

Figure 5

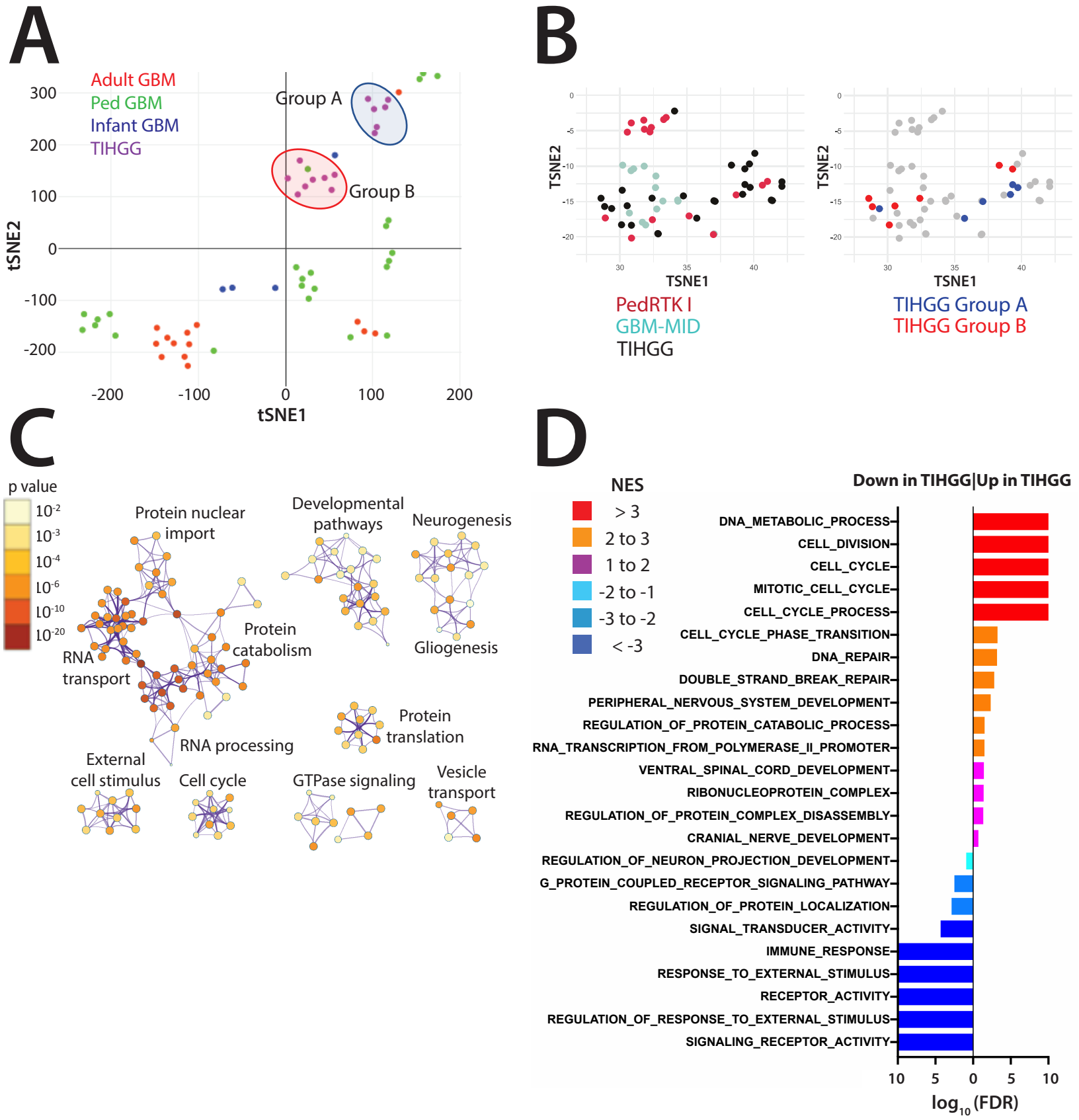
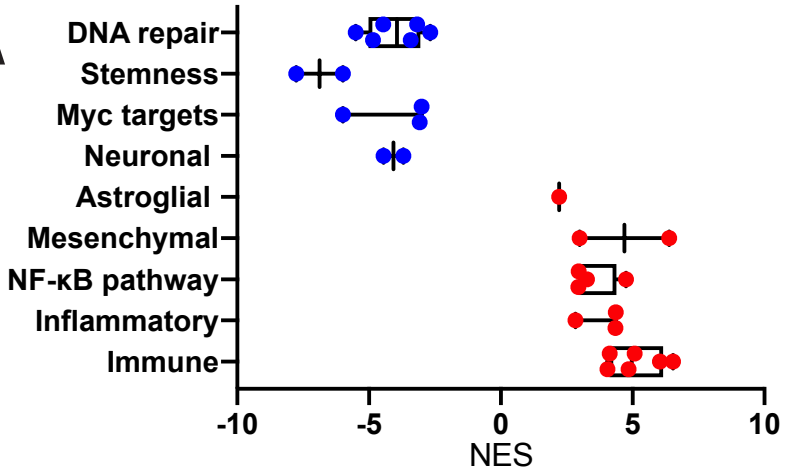
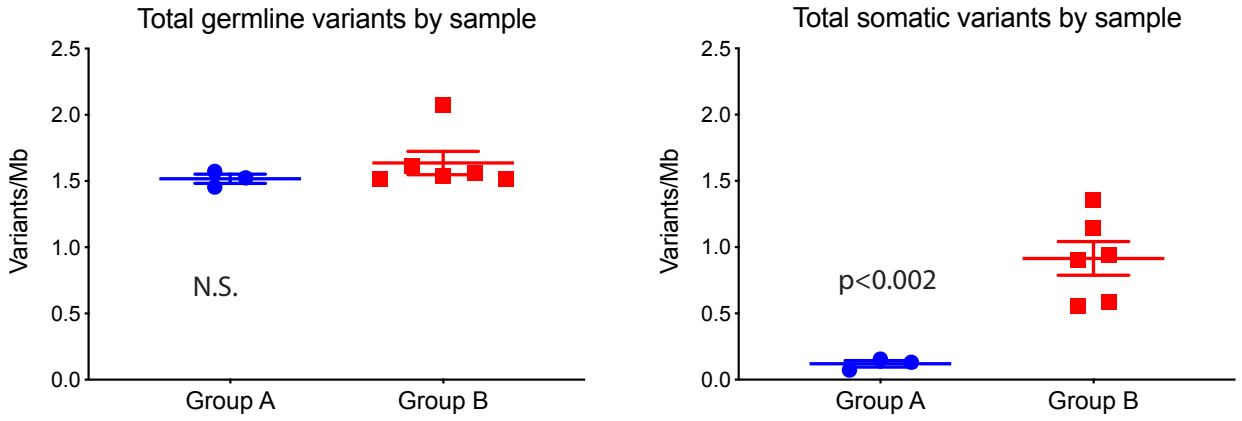

A
Adult GBM
Ped GBM
Infant GBM
TIHGG
Group A
Group B
tSNE2
tSNE1
B
TSNE2
TSNE1
PedRTK I
GBM-MID
TIHGG
TIHGG Group A
TIHGG Group B
C
p value
Protein nuclear import
Developmental pathways
Neurogenesis
Protein catabolism
Gliogenesis
RNA transport
Protein translation
RNA processing
External cell stimulus
Cell cycle
GTPase signaling
Vesicle transport
D
NES
> 3
2 to 3
1 to 2
-2 to -1
-3 to -2
< -3
Down in TIHGG|Up in TIHGG
DNA_METABOLIC_PROCESS
CELL_DIVISION
CELL_CYCLE
MITOTIC_CELL_CYCLE
CELL_CYCLE_PROCESS
CELL_CYCLE_PHASE_TRANSITION
DNA_REPAIR
DOUBLE_STRAND_BREAK_REPAIR
PERIPHERAL_NERVOUS_SYSTEM_DEVELOPMENT
REGULATION_OF_PROTEIN_CATABOLIC_PROCESS
RNA_TRANSCRIPTION_FROM_POLYMERASE_II_PROMOTER
VENTRAL_SPINAL_CORD_DEVELOPMENT
RIBONUCLEOPROTEIN_COMPLEX
REGULATION_OF_PROTEIN_COMPLEX_DISASSEMBLY
CRANIAL_NERVE_DEVELOPMENT
REGULATION_OF_NEURON_PROJECTION_DEVELOPMENT
G_PROTEIN_COUPLED_RECEPTOR_SIGNALING_PATHWAY
REGULATION_OF_PROTEIN_LOCALIZATION
SIGNAL_TRANSDUCER_ACTIVITY
IMMUNE_RESPONSE
RESPONSE_TO_EXTERNAL_STIMULUS
RECEPTOR_ACTIVITY
REGULATION_OF_RESPONSE_TO_EXTERNAL_STIMULUS
SIGNALING_RECEPTOR_ACTIVITY
log10 (FDR)

Figure 6

A



B



C

Geneset	Germline		Somatic	
	NES	FDR	NES	FDR
GO_DNA_REPAIR	-6.97	0.0	-2.68	5.88e-4
HALLMARK_DNA_REPAIR	-4.37	0.0	-4.84	0.0
KAUFFMANN_DNA_REPAIR_GENES	-4.70	0.0	-5.50	0.0
REACTOME_DNA_REPAIR	-3.78	0.0	-4.46	0.0
REACTOME_NUCLEOTIDE_EXCISION_REPAIR	-2.72	1.25e-4	-3.42	0.0
REACTOME_P53_DEPENDENT_G1_DNA_DAMAGE_RESPONSE	-3.80	0.0	-3.17	0.0

D

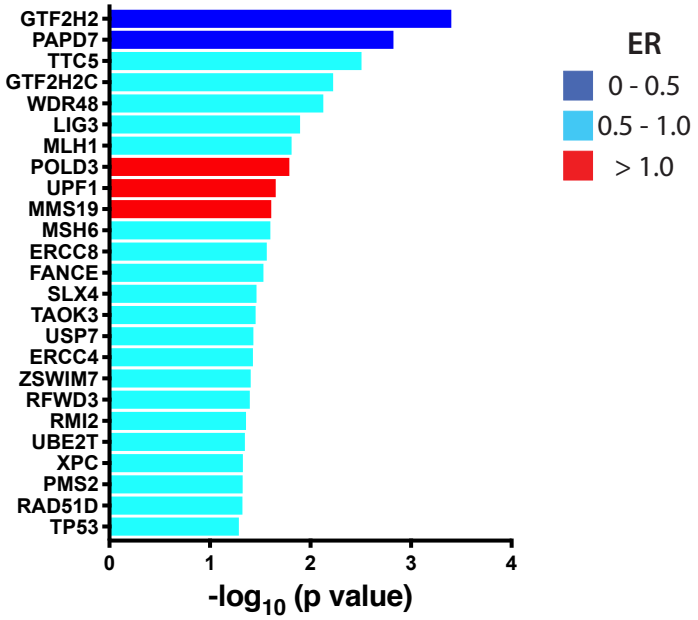


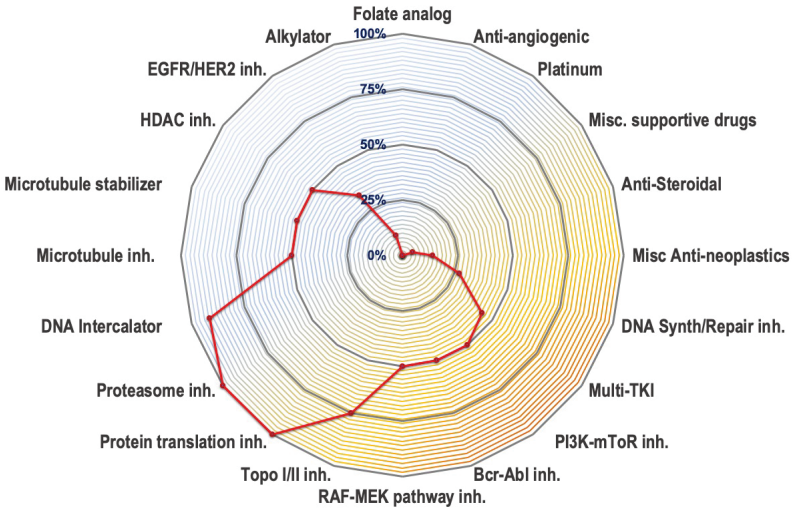
Figure 7

A

Genes	TIHG vs GBM	p-value	Drug Class
DNA Pol & Primase	1.45	1.7e-15	DNA Damage
Mitosis	1.39	1.7e-5	Microtubule
Proteasome	1.33	3.7e-4	Proteasome Inh
Epigenetic Regulation	1.24	2.0 e-5	HDAC Inhibitors

B

Sensitivity by drug class



D

Drug	IC_{50}	
	Group A (95% CI)	Group B (95% CI)
Aldoxorubicin	197 (154-251)	182 (112-298)
Bortezomib	5.38 (4.31-6.92)	10.3 (10.0-10.7)
Carfilzomib	0.893 (0.693-1.28)	4.06 (3.79-4.42)
Etoposide	286 (236-345)	797 (367-1780)
Marizomib	366 (182-1136)	379 (260-586)
Paclitaxel	82.9 (51.1-132)	26.9(14.3-50.6)
Trametinib	57.5 (32.4-106)	93.9 (71.2-121)
Vinblastine	3.28 (2.60-4.16)	1.05 (0.73-1.61)

C

***In vitro* drug screen**

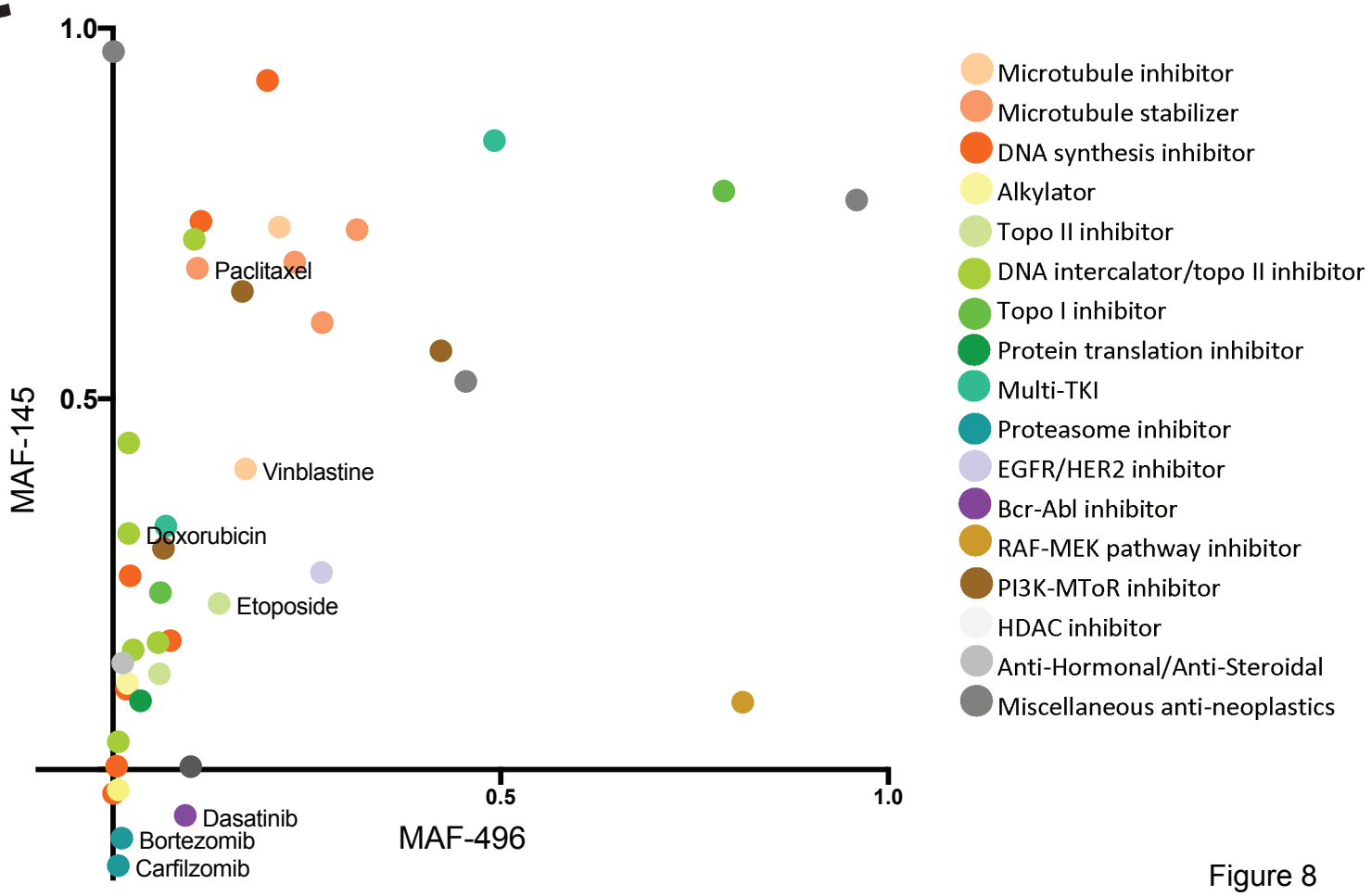


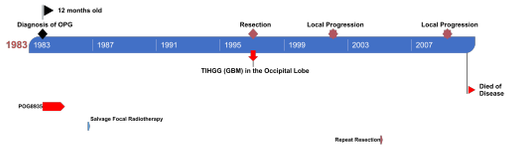
Figure 8

Case_2

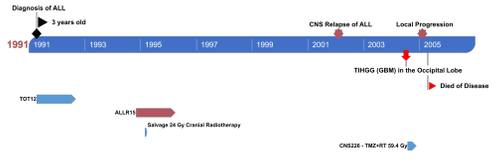
Case_5

Case_6

Case_11



Case_12



Case_13

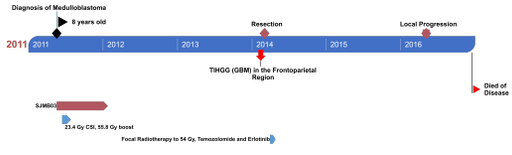
Case_14

Case_15

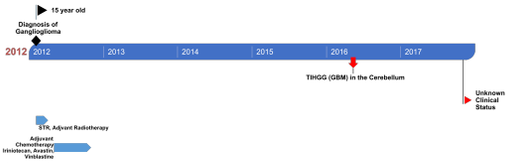
Case_16

Case_18

Case_19



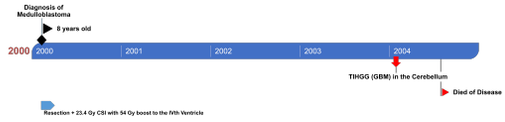
Case_20



Case_21

Case_22

Case_23



Case_24

Case_25

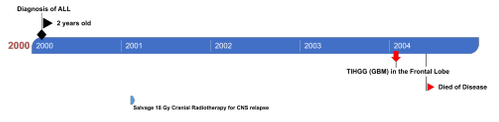
Case_26

Case_27

Case_28

Case_29

Case_30

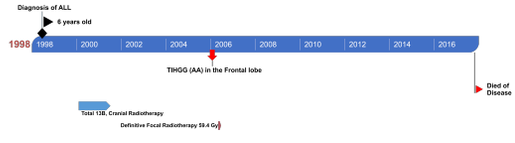


Case_31

Case_32

Case_33

Case_34

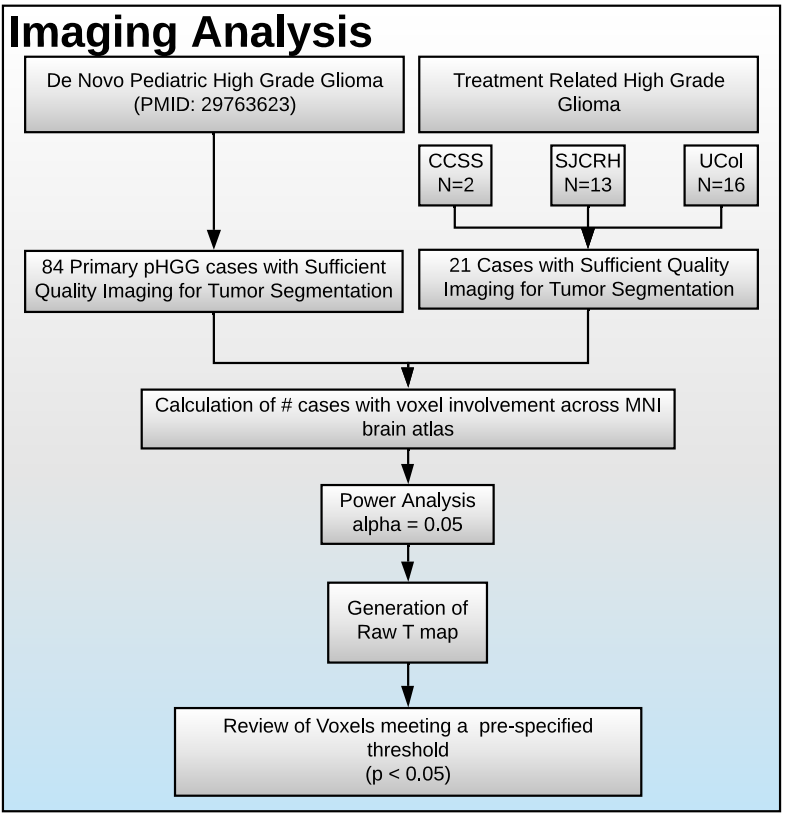


Case_42

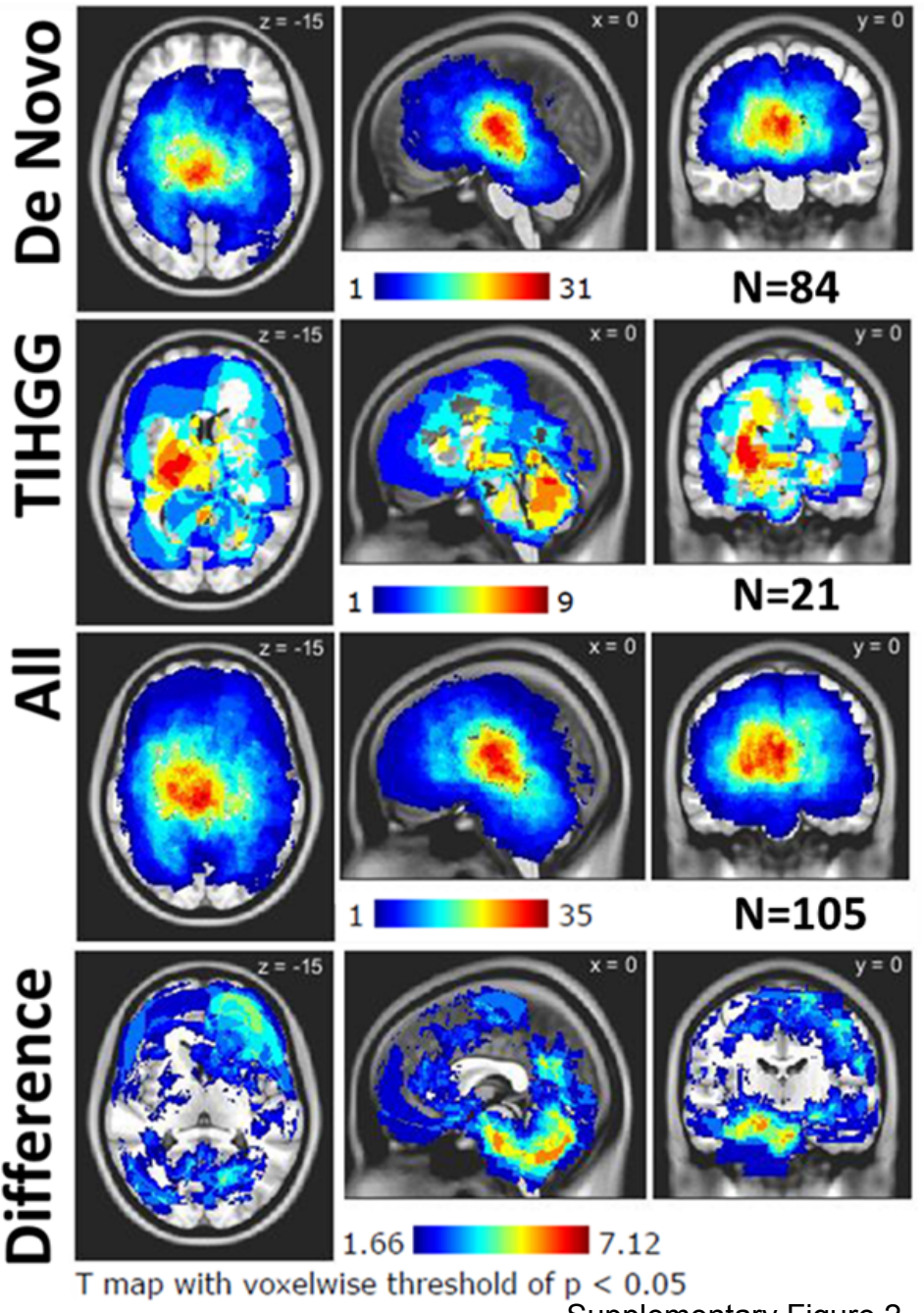
Case_43

Supplementary Figure 1

A

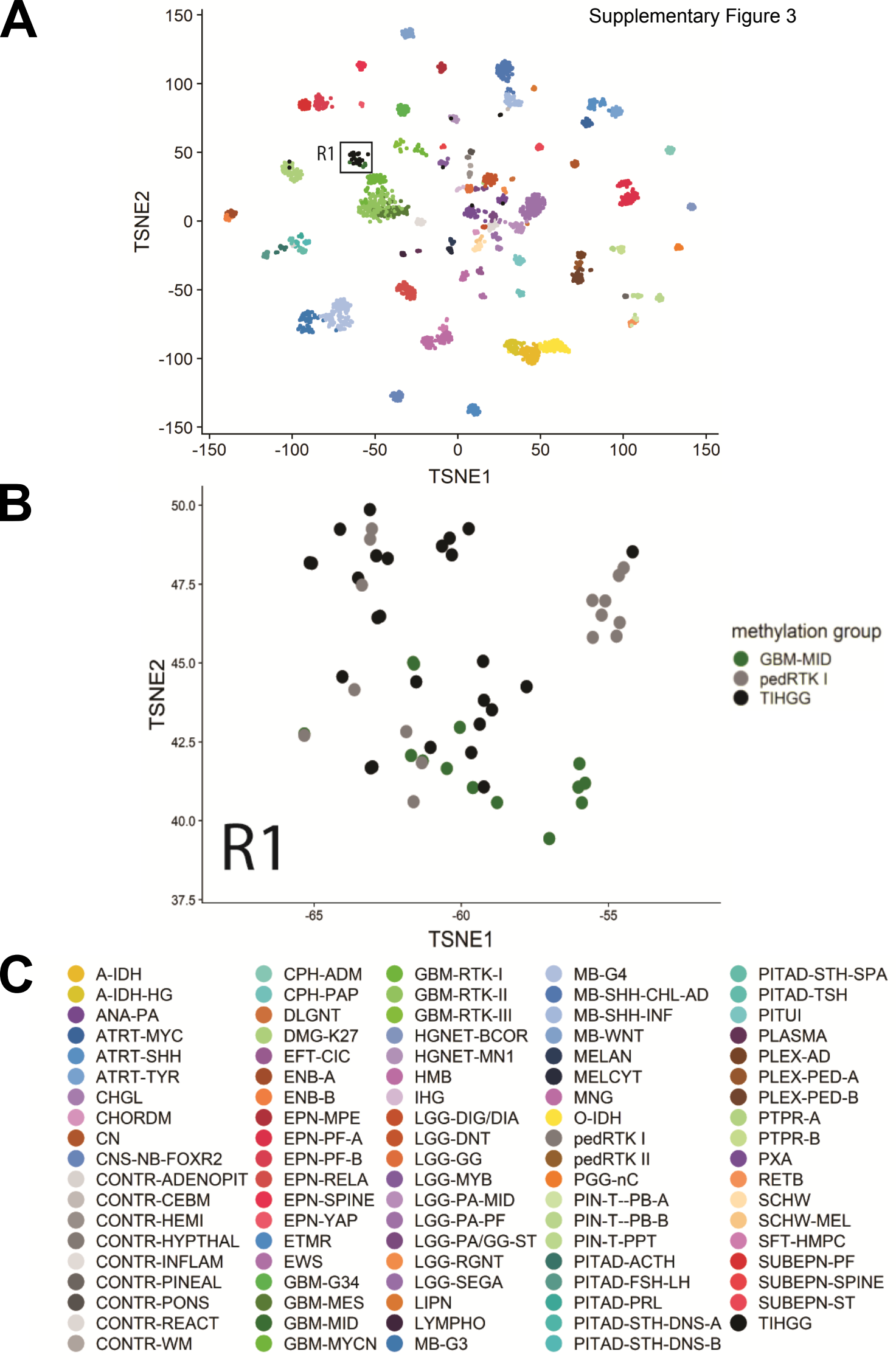


B

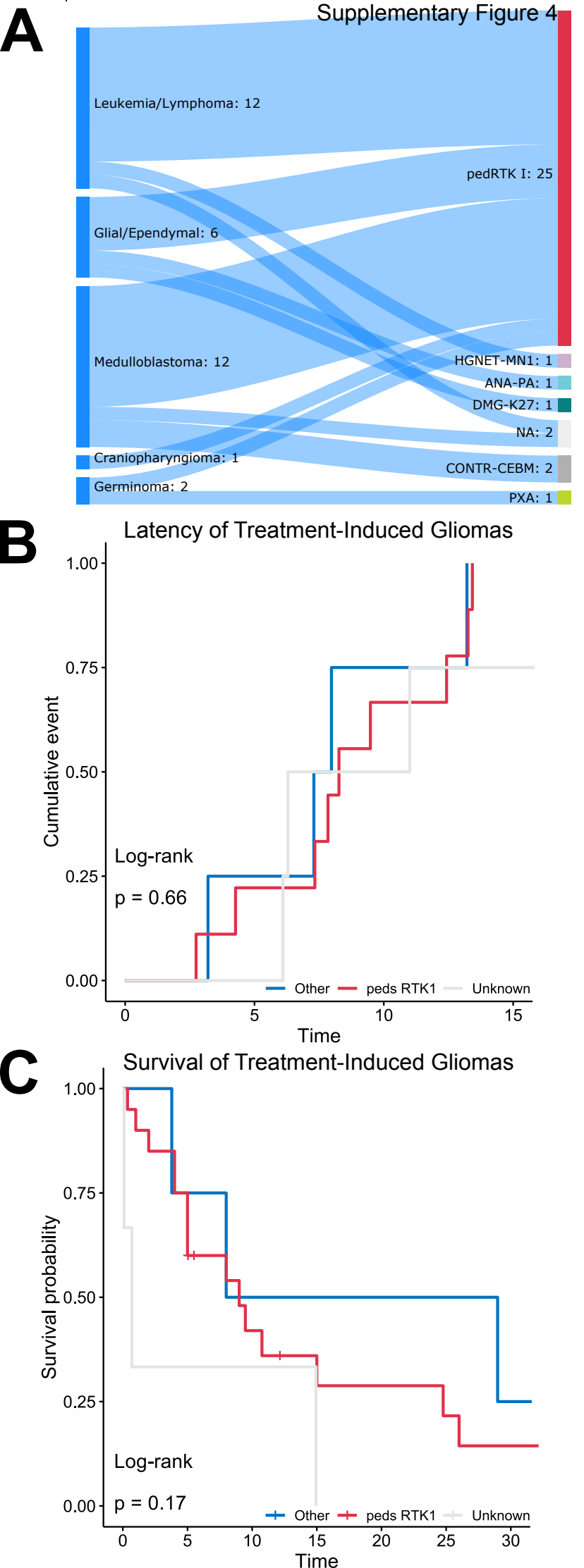


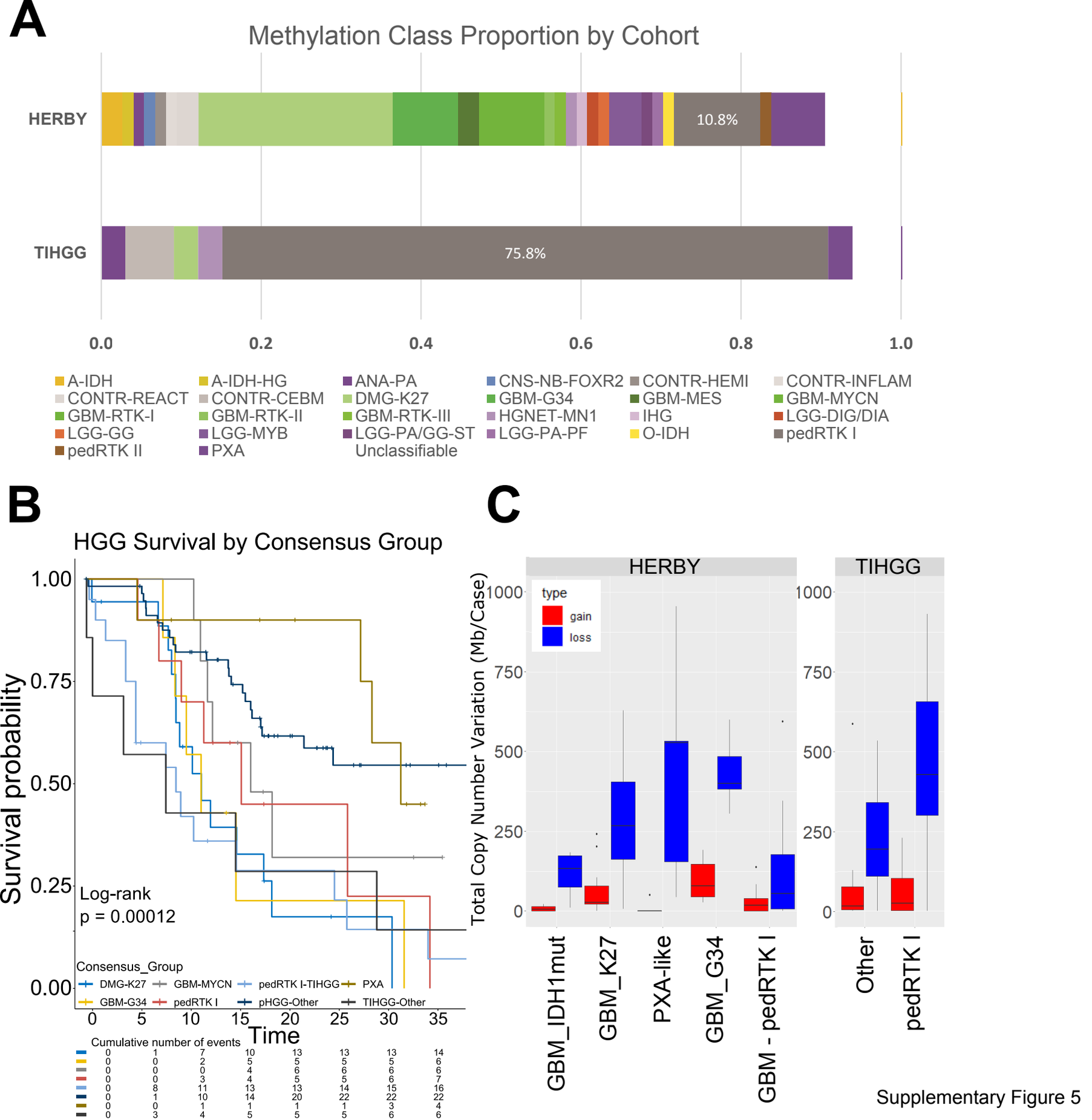
Supplementary Figure 2

Supplementary Figure 3

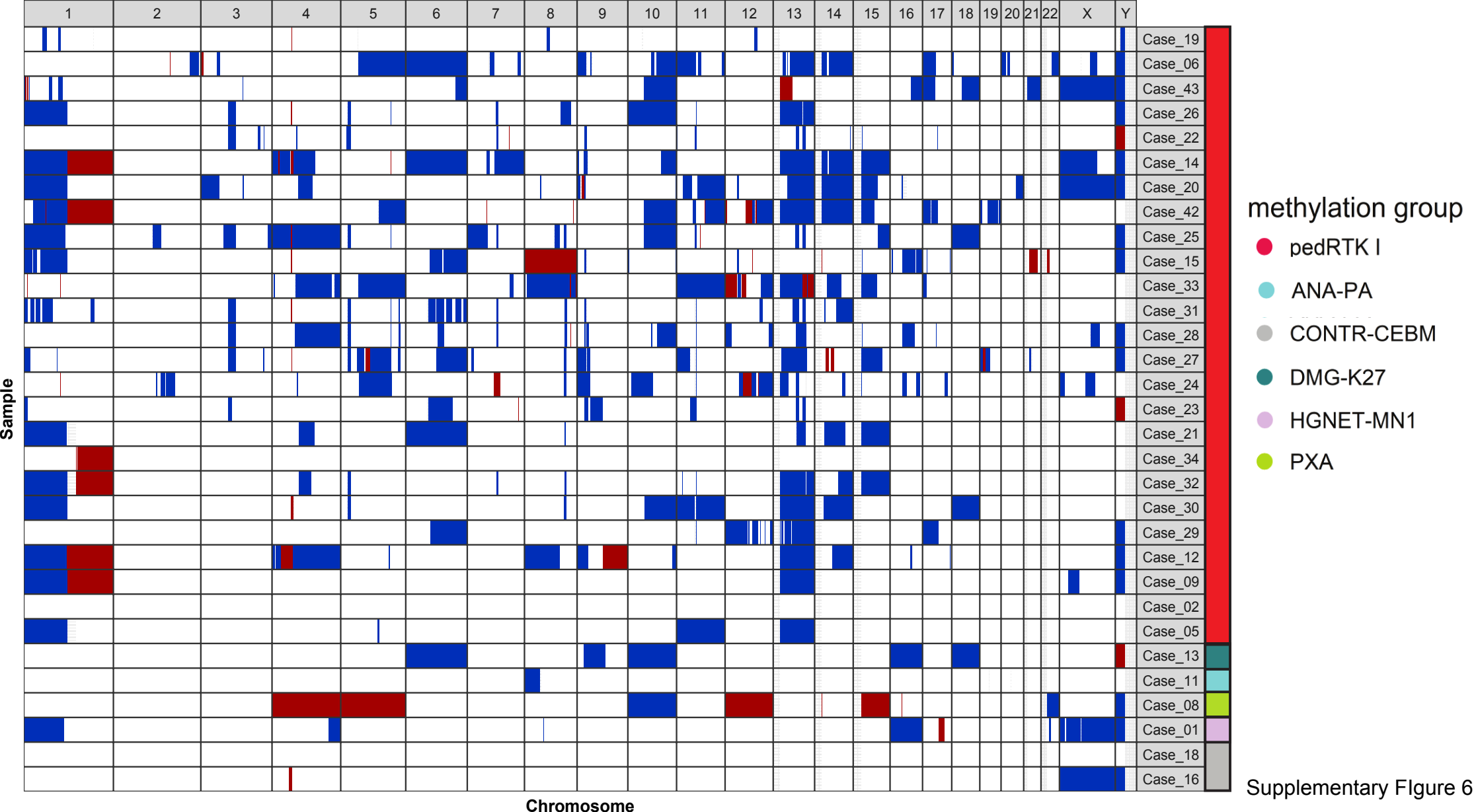


Supplementary Figure 4



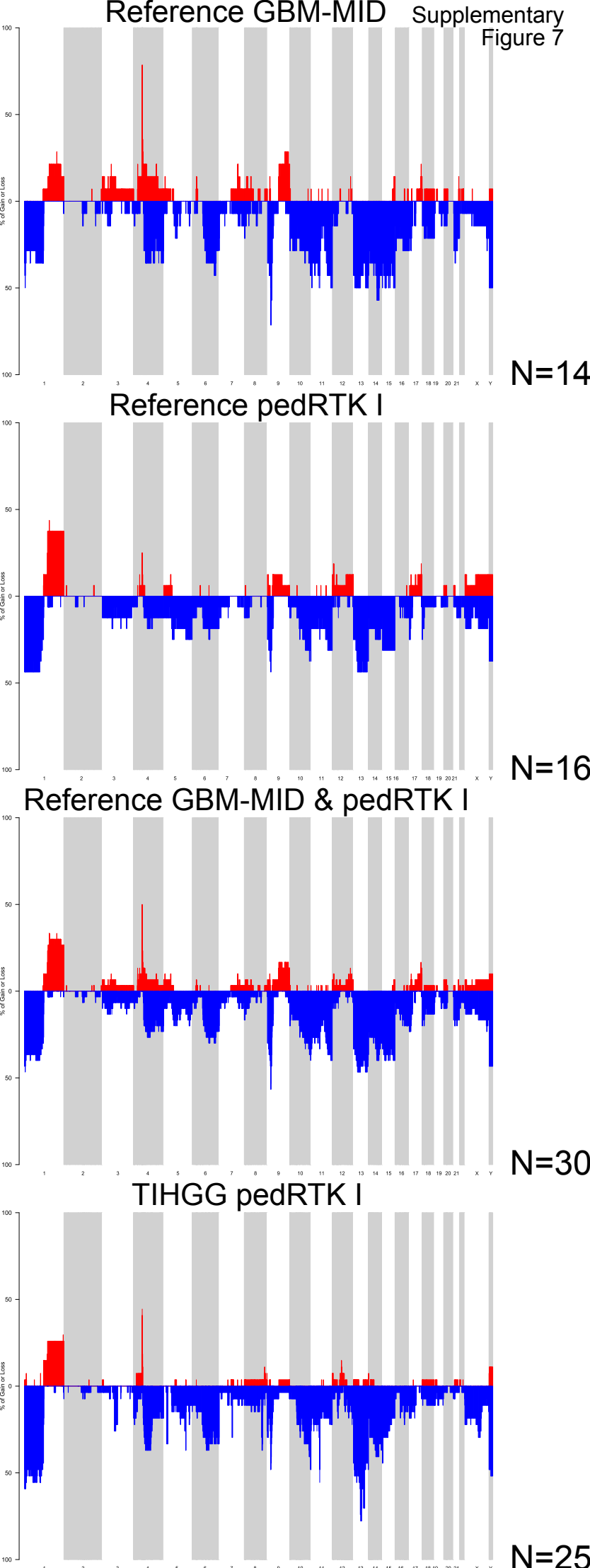


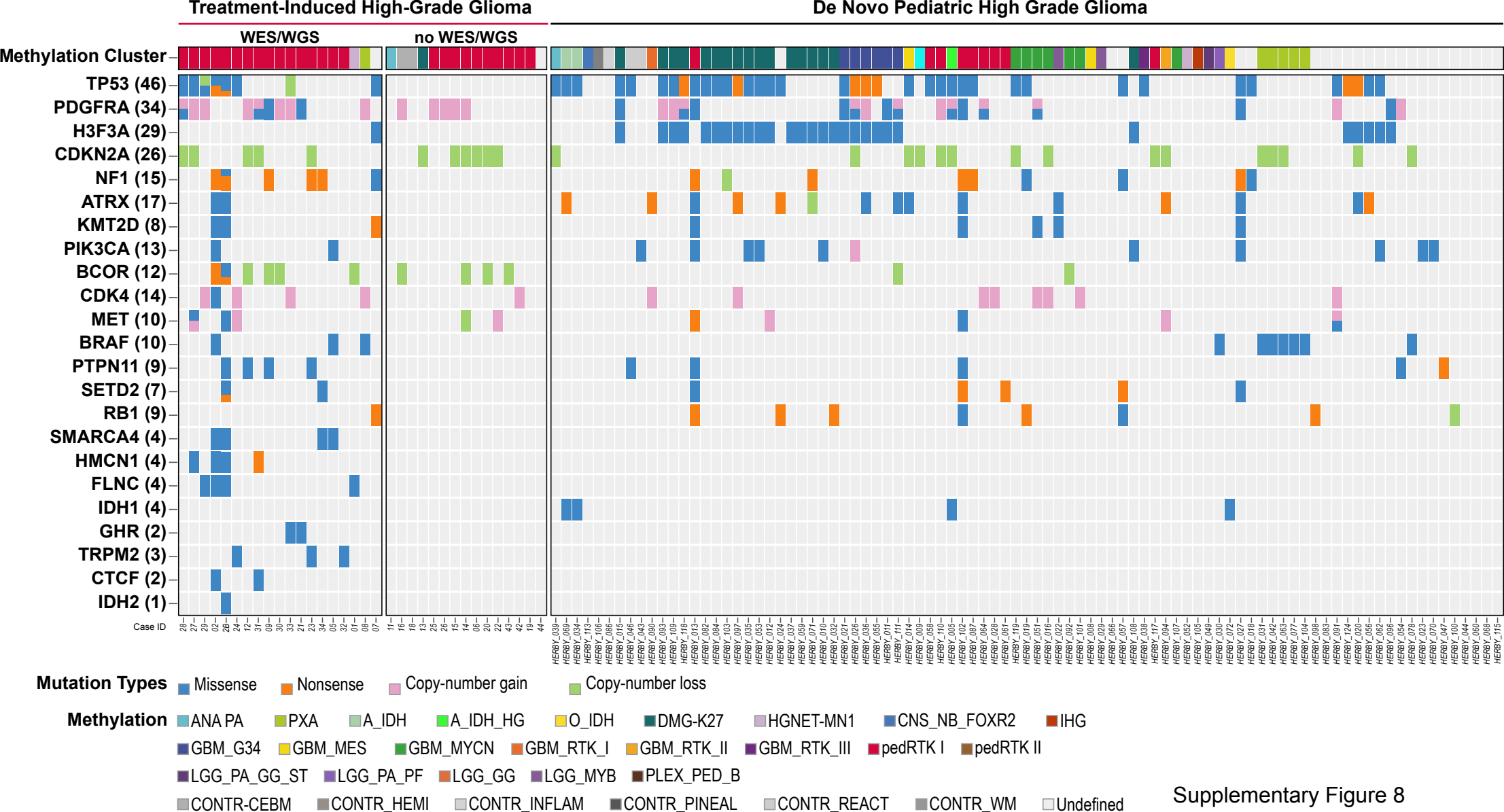
Supplementary Figure 5



Supplementary FIgure 6

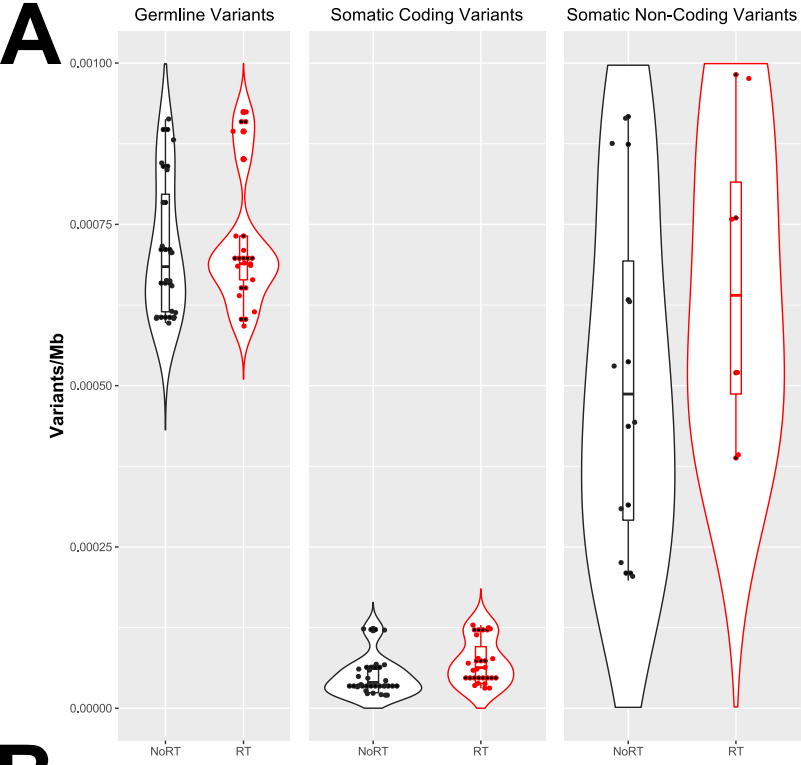
Supplementary Figure 7

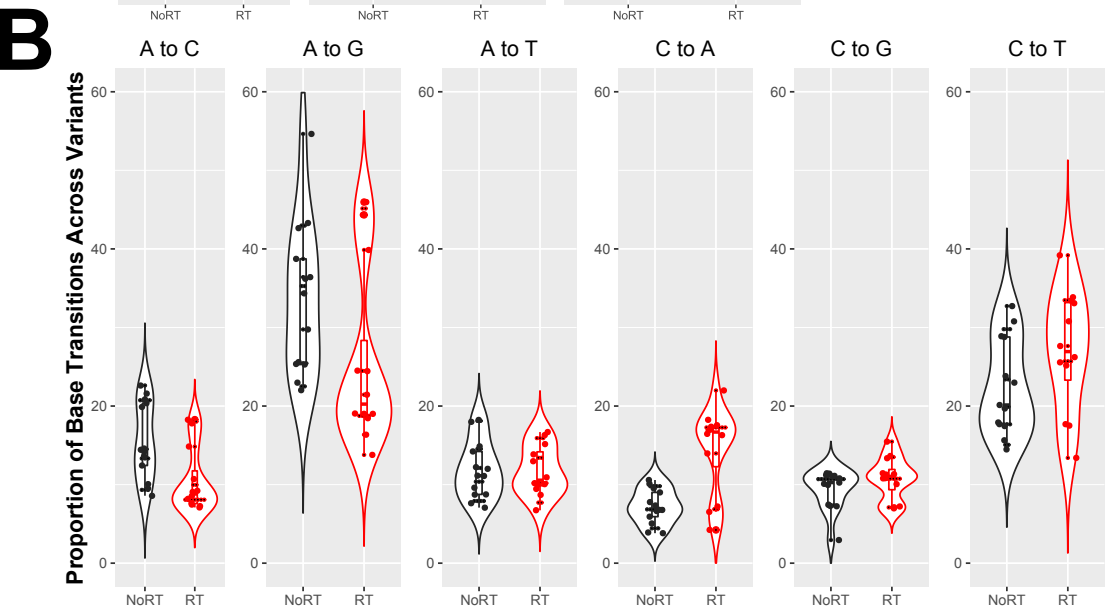


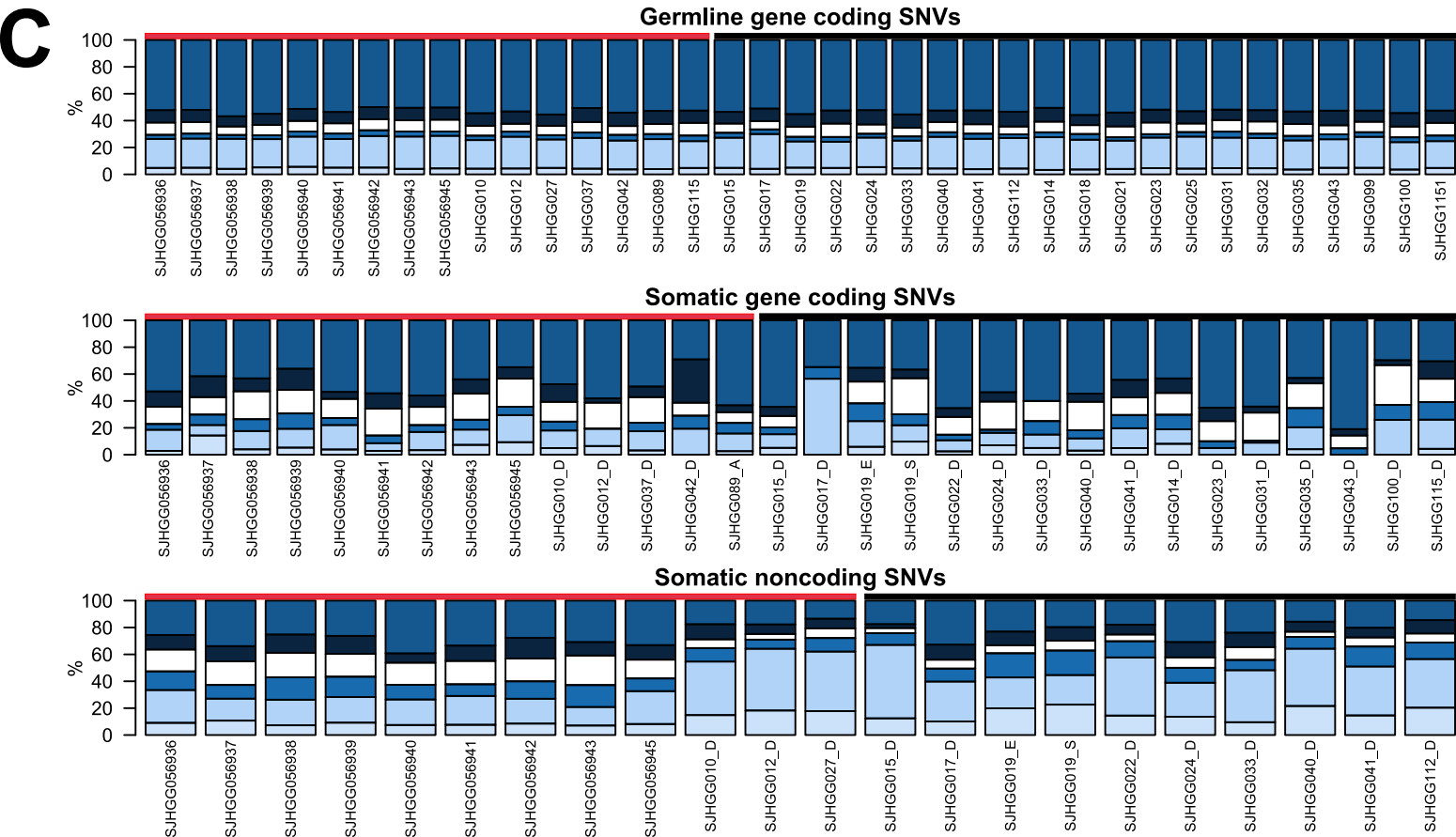


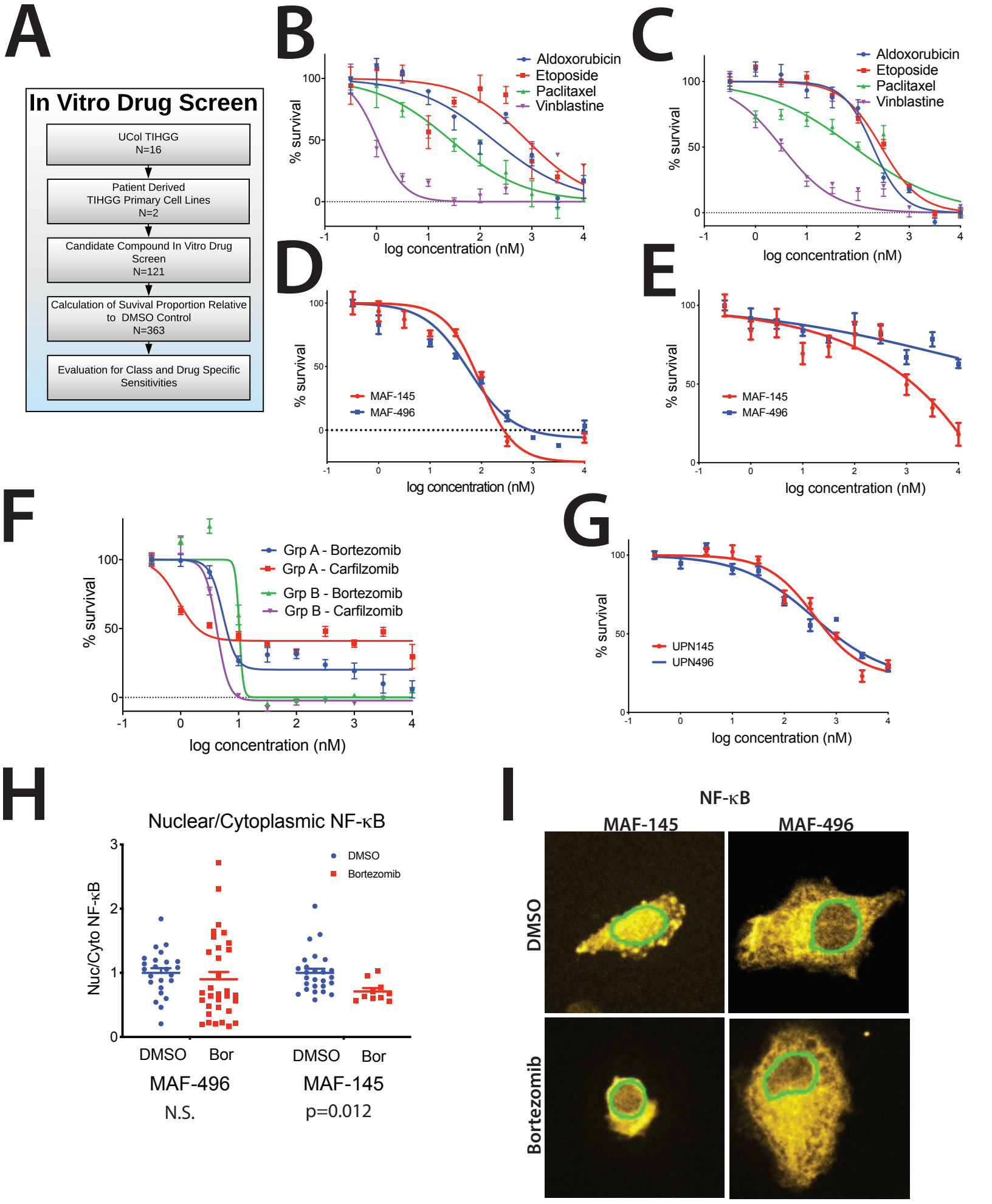
Supplementary Figure 8

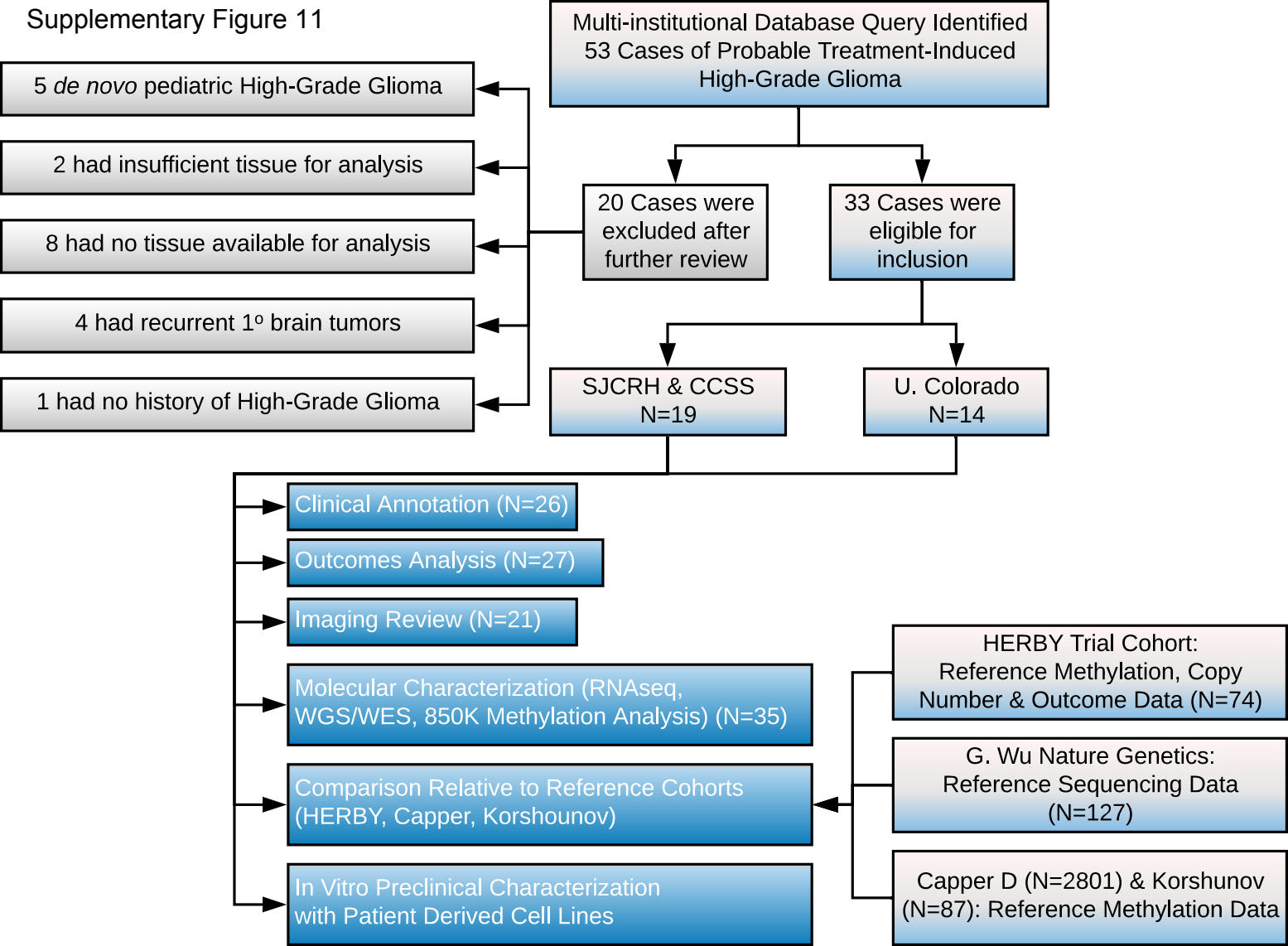
Supplementary Figure 9









Supplementary Figure 10

Supplementary Figure 11