1 **Title page**

2

3 **The spectrum of *IDH*- and H3-wildtype high-grade glioma subgroups occurring across teenage and**
4 **young adult patient populations**

5

6 **Authors:**

7

8 Rita Pereira[1]*, Alan Mackay[1]*, Yura Grabovska[1]*, Amelia Bradley[2], Tabitha Bloom[2], James Nicoll[2],
9 Delphine Boche[2], John Procter[3], Mellissa Maybury[4,5], Johanna Schagen[5], Liam Walker[6,7], Federico Roncaroli[6,7],
10 Konstantina Karabatsou[8], Olumide Ogunbiyi[9], Thijs van Dalen[10], Jai Sidpra[17], Sabrina Rossi[11], Evelina Miele[12],
11 David S. Ziegler[13,14,15], Zhifeng Shi[16], Thomas S. Jacques[7], Darren Hargrave[17], Bassel Zebian[18], Cristina Bleil[18],
12 Joseph Yates[19], Emma Norton[19], Henry Mandeville[20], Antonia Creak[21], Liam Welsh[21], Lynley Marshall [22,20,23],
13 Fernando Carceller[20], Sucheta J. Vaidya[20], Zita Reisz[24], Safa Al-Sarraj[24], Angela Mastronuzzi[25], Andrea Carai[26],
14 Maria Vinci[27], Kathreena M. Kurian[28], Ho-keung Ng[29], Sebastian Brandner[30], Chris Jones[1]*, Matthew
15 Clarke[1,30]*

16

17 *Indicates that these authors contributed equally to the work.

18

19 **Affiliations:**
20   1. Division of Molecular Pathology, Institute of Cancer Research, United Kingdom
21   2. Clinical Neurosciences, Clinical and Experimental Sciences, Faculty of Medicine, University of
22   Southampton, Southampton, United Kingdom
23   3. Neuropathology and Pathology Research, Royal Preston Hospital
24   4. Oncology Services Group, Queensland Children's Hospital, Brisbane, Queensland, Australia.
25   5. Child Health Research Centre, The University of Queensland, Brisbane, Queensland, Australia
26   6. Department of Cellular Pathology, Northern Care Alliance, Salford, United Kingdom.
27   7. Geoffrey Jefferson Brain Research Centre, Division of Neuroscience, School of Biology, University of
28   Manchester, United Kingdom.
29   8. Department of Neurosurgery, Manchester Centre of Clinical Neuroscience, Northern Care Alliance
30   9. Neuropathology Department, Great Ormond Street Hospital for Children NHS Foundation Trust,
31   United Kingdom
32   10. Neurosurgery Department, National Hospital for Neurology and Neurosurgery, UCLH, United
33   Kingdom
34   11. Sabrina Rossi. Pathology Unit, Bambino Gesù Children's Hospital IRCCS, Rome, Italy
35   12. Evelina Miele. Pediatric Onco-Haematology, Cell Therapy, Gene Therapies and Hematopoietic
36   Transplant, Bambino Gesù Children's Hospital-IRCCS, Rome, Italy
37   13. Children's Cancer Institute, Lowy Cancer Centre, UNSW Sydney, Kensington, NSW, Australia
38   14. School of Clinical Medicine, UNSW Medicine & Health, UNSW Sydney, Kensington, NSW, Australia
39   15. Kids Cancer Centre, Sydney Children's Hospital, Randwick, NSW, Australia
40   16. Neurosurgery, Fudan University

41   17. Paediatric Neuro Oncology, UCL Great Ormond Street Institute of Child Health, London, UK

42   18. Neurosurgery Department, King's College Hospital, London, United Kingdom

43   19. Department of Cellular Pathology, University Hospital Southampton NHS Trust, Southampton, United
44   Kingdom

45   20. Department of Paediatric and Teenager & Young Adult Neuro-Oncology, The Royal Marsden Hospital
46   NHS Foundation Trust, United Kingdom; Division of Clinical Studies, The Institute of Cancer
47   Research, London, United Kingdom.

48   21. Department of Adult Neuro-Oncology, The Royal Marsden Hospital NHS Foundation Trust, United
49   Kingdom; Division of Clinical Studies, The Institute of Cancer Research, London, United Kingdom.

50   22. Paediatric and Adolescent Oncology Drug Development Unit, The Royal Marsden NHS Foundation
51   Trust, UK

52   23. Division of Clinical Studies, The Institute of Cancer Research, UK

53   24. Neuropathology Department, King's College Hospital, United Kingdom

54   25. Pediatric Onco-Hematology, Cell Therapy, Gene Therapies and Hematopoietic Transplant, Bambino
55   Gesù Children's Hospital - IRCCS, Rome, Italy

56   26. Neurosurgery Unit, Bambino Gesù Children's Hospital – IRCCS, Rome, Italy

57   27. Mara. Paediatric Cancer Genetics and Epigenetics Research Unit, Bambino Gesù Children's Hospital-
58   IRCCS, Rome, Italy

59   28. University of Bristol Brain Tumour Research Centre, Neuropathology, Southmead Hospital, Bristol,
60   United Kingdom

61   29. Chinese University of Hong Kong, Hong Kong

62   30. Division of Neuropathology, The National Hospital for Neurology and Neurosurgery, University
63   College London Hospitals NHS Foundation Trust, and Queen Square Institute of Neurology, United
64   Kingdom

65

66   **Corresponding author:**
67   Dr Matthew Clarke
68   Division of Molecular Pathology
69   Institute of Cancer Research
70   15 Cotswold Road
71   Sutton
72   London
73   SM2 5NG
74   Email: matthew.clarke01@icr.ac.uk

75

76   **Conflicts of interest**
77   No conflicts of interest to declare

78

79   **Other notes**
80   Word count abstract: 374

81   Word count text: 8225

82   Word count references: 1520

83   Word count figure and table legends: 1866

84   Number of figures: 6

85   Supplementary material: 6 supplementary figures, 4 supplementary tables

86

87   **Statement of translational relevance**

88
89   High-grade gliomas (HGG) in teenagers and young adults (TYA) are an understudied group of tumours.

90   Through national and international collaboration, we have gathered a large series of 207 cases, for the purpose

91   of molecular characterisation using DNA methylation profiling, whole exome sequencing and fusion panel

92   sequencing, alongside neuropathological review. We integrate histopathology, genetic and epigenetic profiling

93   to present the spectrum of different subtypes of HGG which occur in the TYA population. We identify that

94   HGGs occurring in TYAs comprise of methylation subgroups which occur in paediatric and adult age groups,

95   but also novel, poorly defined methylation classes which this study helps to characterise. The study highlights

96   mutational landscapes which may be targetable with immune checkpoint, MAPK-pathway and *PDGFRA*

97   inhibitors. Incidences of tumour predisposition syndromes (including constitutional mismatch repair disorder

98   (CMMRD)) are identified in the cohort, and we also identify tumours that have developed on a background of

99   previous treatment for childhood cancers. These data also highlight the value of undertaking DNA methylation

100   profiling and whole exome/panel sequencing on each HGG occurring in the TYA age group to accurately

101   diagnose and characterise these complex tumours.

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122
123
124
125   **Abstract**
126
127   <u>**Background**</u>

128   High-grade gliomas (HGG) occur in any central nervous system (CNS) location and at any age. HGGs in
129   teenagers/young adults (TYA) are understudied. This project aimed to characterise these tumours to support
130   accurate stratification of patients.
131

132   <u>**Methods**</u>
133   207 histone/IDH-wildtype tumours from patients aged 13–30 years were collected. DNA methylation profiling
134   (Illumina EPIC BeadArrays, brain tumour classifier (MNPv12.8 R package)) classified cases against reference
135   cohorts of HGG. Calibrated scores guided characterisation workflows (RNA-based ArcherDx fusion panel
136   (n=92), whole exome sequencing (WES) (n=107), histological review).
137

138   <u>**Results**</u>
139   53.4% (n=86) matched as paediatric-type subgroups (pedHGG_RTK1A/B/C (31.7%, n=51 *PDGFRA,*
140   *CDKN2A/B, SETD2, NF1* alterations), pedHGG_*MYCN* (8.1%, n=13, *MYCN/ID2* amplifications), and
141   pedHGG_RTK2A/B (7.5%, n=12, *TP53, BCOR, ATRX, EGFR* alterations)). 18.0% (n=29) classified as adult-
142   type subgroups (GBM_MES (15.5%, n=25, enriched for *RB1, PTEN*, *NF1* alterations) and GBM_RTK1/2
143   (2.5%, n=4, *CDK4* amplifications)). 23 cases (14.7%) classified as novel, poorly-characterised subgroups with
144   distinct methylation profiles and molecular features (pedHGG_A/B (n=10 6.2%), HGG_E (n=6 3.7%), HGG_B
145   (n=2 1.0%), GBM_CBM (n=5 3.1%)) with variable histological morphology. 8 cases (5.1%) showed
146   hypermutator phenotypes, enriched in HGG_E, one of which was associated with constitutional mismatch repair
147   deficiency (CMMRD), and their sibling who was diagnosed with the same syndrome, was diagnosed with a
148   tumour which classified as a pedHGG_RTK1B. HGGs which have developed on a background of previous
149   treatment for a childhood cancer are detected in the TYA population, classifying most frequently as
150   pedHGG_RTK1B, and contributing to the poor prognosis of this subgroup. Age-distribution/molecular profile
151   comparisons using publicly available methylation/sequencing data (and from local diagnostic cohorts) for
152   HGG_B (n=19), GBM_CBM (n=35) and GBM_MES_ATYP (n=102), irrespective of age, show that HGG_B is
153   a TYA-specific subgroup (median age 29 years) and that GBM_CBM and GBM_MES_ATYP show a peak of
154   distribution in the TYA population, but also have a wider age distribution (median age 35.7 and 50.5 years
155   respectively) with the latter showing distinct differences in copy number profiles compared to older adults in the
156   same subgroup, and containing fewer chr7 gains, chr10 losses, more *CDKN2A/B* deletions and *MET*
157   amplifications, and a worse survival compared with adult-specific GBM_MES_TYP.
158

159   <u>**Conclusion**</u>

4

160 TYA HGGs comprise novel methylation subgroups with distinct methylation and molecular profiles. Accurate
161 stratification of these patients will open opportunities to more effective treatments including immune check
162 point, MAPK-pathway and *PDGFRA* inhibitors.
163
164
165 **Introduction**
166
167 High-grade gliomas (HGG) comprise different subgroups of aggressive central nervous system (CNS) tumours
168 in both adults and children including infants (1,2). They can occur in any CNS location (3,4) within the cranio-
169 spinal axis and have a poor prognosis (20.8% 5-year survival for those diagnosed between 0-19 years and 21.9%
170 for patients aged 20-44 years) (3,5). Outcomes are considerably better than older adults with patients aged 45-54
171 and 55-64 years having a 5-year survival of 9.3% and 5.9% respectively, suggesting that underlying biology
172 may differ (3). It remains a significant clinical and therapeutic challenge to understand more about the
173 development, characteristics and evolution of these tumours to advance therapeutic interventions and improve
174 survival (6,7).
175
176 DNA methylation profiling and the development of the Molecular Neuropathology (MNP) brain tumour
177 classifier have transformed diagnostic practice (8,9), providing a molecular tool supporting delivery of accurate
178 diagnoses for challenging cases. Copy number profiles and prognostic information can be gained (9), and it
179 provides opportunities to discover and characterise new tumour entities (8,9) for more accurate patient
180 stratification (10,11).
181
182 The 5th edition of the World Health Organisation (WHO) classification of tumours of the CNS now recognises
183 that high- and low-grade diffuse glial tumours occurring in adults and children comprise distinct entities with
184 distinctive molecular profiles (12); histone-mutant CNS tumours (diffuse midline glioma, H3 K27-altered,
185 diffuse hemispheric glioma, H3 G34-mutant) are considered paediatric-type HGGs (12,13). The WHO
186 classification also recognises the newly defined paediatric entities including the infant-type hemispheric glioma
187 (IHG), seen predominantly in the infant population, and characterised by distinct methylation profiles and RTK
188 fusions (1,2).
189
190 Approximately 60% of CNS tumours diagnosed in teenagers and young adults (TYA) each year in the UK will
191 be a glial tumour (14), representing the fourth most frequent cancer type occurring in patients aged 15–24 years
192 (12). They account for 60% of cancer-related deaths in this age range (14). This group denotes an understudied
193 set of tumours with very limited published clinical, pathological and molecular data. No distinct HGG
194 subgroups are recognised as TYA-specific according to the WHO classification (12). Studies have explored a
195 variety of molecular-based techniques (immunohistochemistry, targeted/next generation sequencing, whole
196 exome sequencing and much less frequently epigenetic profiling) and show that a proportion of TYA HGGs can
197 be classified into well-recognised HGG entities including diffuse midline gliomas, H3 K27-altered, diffuse
198 hemispheric glioma, H3 G34-mutant, and IDH-mutant HGGs (15,16). However, other tumours remained
199 unclassifiable (15). A meta- analysis of >1000 paediatric HGG and diffuse midline gliomas (DMGs), including

5

200 patients aged up to 35 years, identified diverse clinical and biological subgroups characterised by different
201 somatic mutations (16), and showed that survival of TYA HGGs was significantly better compared to older
202 adults and H3-mutant groups (16), suggesting that novel entities of HGG with differing underlying biology
203 exist.

204

205 Gross total resection (GTR) is the mainstem HGG treatment, including for TYA patients. Given their infiltrative
206 nature, residual and radiologically undetectable tumour frequently remains, leading to recurrence and
207 progression. Post-operative chemotherapy and radiotherapy have considerable cumulative side effects (17–19),
208 including impaired cognitive function (14). Studies have also shown very limited benefit for HGG patients
209 undergoing re-resection of residual tumour (20), and highlight the psychosocial impact of a brain tumour
210 diagnosis for TYA patients (14). As efforts to improve survival continue, it remains imperative to minimise
211 short- and long-term side-effect profiles or find more targeted alternatives better suited to these complex
212 tumours.

213 In this study, we have investigated a series of TYA HGGs integrating DNA methylation, mutational analysis,
214 neuropathological and clinical annotations, and identified novel tumour subgroups with distinctive methylation
215 profiles.

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

Materials and Methods

Cases

A total of 398 samples were collected from national and international collaborators (tumour samples: University College London Hospitals, London, n = 63; internal biobank, n = 58; King's College Hospital, London, n = 39; Southampton General Hospital, n = 38; The Chinese University of Hong Kong, n = 37; Salford NHS Foundation Trust, n = 33; Royal Preston Hospital n = 29; Bambino Gesù Children's Hospital, Rome, n = 26; Queensland Children's Tumor Bank, Brisbane, n = 23; University Hospitals Bristol, n = 17; Great Ormond Street Hospital, London, n = 15; Poland, n = 11; Children's Cancer Institute and Sydney Children's Hospital, n = 9). Where possible, both formalin-fixed paraffin-embedded (FFPE) and/or frozen tissue was provided for each case. Data were also retrieved and integrated from previously published studies from the Jones Lab (2,16,21). Patients were selected according to age (≥13 - ≤30 years old) and diagnosis (IDH/H3-wildtype glioma, WHO grade 2, 3 or 4). Definitions of TYA age are variable, with the UK stating that the age range is between 16 – 24 years (22); for this study, we decided to broaden the age to ≥13 - ≤30 years old to ensure the age distribution for each tumour subtype could be captured. Cases were excluded if they were diagnosed at the centre of origin/classified as a well-characterised subtype of high-grade glioma according to the WHO 2021 (astrocytoma, IDH-mutant; diffuse midline glioma, H3 K27-altered; diffuse hemispheric glioma H3 G34-mutant). All CNS WHO grade 1 gliomas and glioneuronal tumour diagnoses were excluded, alongside non-gliomas (ependymal, embryonal, mesenchymal, and germ cell tumours). Tissue samples obtained from the University Hospital Southampton NHS Foundation Trust, Royal Preston Hospital, Great Ormond Street Institute for Child Health, University College London Hospitals, Southmead Hospital, and the Northern Care Alliance as part of BRAIN UK, were supported by Brain Tumour Research, the British Neuropathological Society and the Medical Research Council.

A final cohort of 207 samples was investigated.

Consent and ethics statement

Written informed consent was obtained for all samples included in this study, under Research Ethics Committee approval at each participating centre. Project specific Research Ethical Approval was also received via application to BRAIN UK (23).

7

Nucleic Acid Extraction

DNA and RNA were extracted from frozen tissue using the *Quick*-DNA/RNA Miniprep Plus Kit (Zymo) according to manufacturer's instructions. For FFPE tissue, DNA was extracted where possible after manual macro-dissection using the QIAamp DNA FFPE Tissue Kit protocol (QIAGEN) and RNA was extracted using the AllPrep DNA/RNA kit (Qiagen) according to manufacturer's instructions.

Methylation Profiling

Analysis was performed using Illumina EPIC BeadArrays at University College London (UCL) Great Ormond Street Institute of Child Health when more than 140 ng of DNA was extracted. Data from EPIC arrays was pre-processed using the minifi package in R (v1.48.0). 15 Italian cases were analysed at Bambino Gesù Children's Hospital and raw data were shared and included for the bioinformatic analyses. Data from EPIC arrays was pre-processed using the minifi package in R (v1.48.0). A cohort of 207 cases met the entry criteria and could progress into the study. The Heidelberg brain tumour classifier (molecularneuropathology.org) was used to assign a calibrated score to each case, associating it with one of the >100 tumour entities which feature within the classifier (v12.8). Tumours were subclassified into one of 17 methylation subgroups; diffuse paediatric-type high grade glioma, H3 wildtype and IDH wild type, Subtype A/B (pedHGG_A/B), diffuse paediatric-type high grade glioma, MYCN subtype (pedHGG_MYCN), diffuse paediatric-type high grade glioma, RTK1 subtype, subclass A/B/C (pedHGG_RTK1A/B/C), diffuse paediatric-type high grade glioma, RTK2 subtype, subclass A/B (pedHGG_RTK2A/B), glioblastoma, IDH-wildtype, subtype posterior fossa (GBM_CBM), adult-type diffuse high grade glioma, IDH-wildtype, subtype E (HGG_E), adult-type diffuse high grade glioma, IDH-wildtype, subtype B (HGG_B), glioblastoma, IDH-wildtype, RTK1 subtype (GBM_RTK1), glioblastoma, IDH-wildtype, RTK2 subtype (GBM_RTK2), glioblastoma, IDH-wildtype, mesenchymal subtype, typical/atypical (GBM_MES_TYP/ATYP), pleomorphic xanthoastrocytoma (PXA), and high-grade astrocytoma with piloid features (HGAP). Clustering of beta values from methylation arrays was performed based upon correlation distance using a ward algorithm. Quality control was undertaken; the methylation data for samples where the probe failure rate was >20% or bisulphite conversion failed were excluded from the methylation analysis. However, they were not excluded from the study and were analysed with other workflow streams. Mean p-value detection is a common QC filtering criteria and is advised by Illumina. However, the probe failure rate is not stringent because of the inclusion of historical FFPE cases. To date, Illumina do not provide specific guidelines on the cut-off and advise discretion. The MNP classifier uses 10,000 probes; cases with a probe failure rate ≤20% showed concordance with MNP assignments, t-SNE location, and clinical/genomic case annotations.

Samples were assigned a methylation subgroup when the calibrated score was >0.5, and after correlating with the histological diagnosis and available molecular data. Samples with a calibrated score of <0.5, or if the assignment was not appropriate, were assigned as NEC (not elsewhere classified), as per the WHO classification definition, and underwent further characterisation to determine their subtype.

Reference cohorts

8

305

306 Classifier results for the TYA cohort were plotted against a t-SNE reference cohort of CNS tumours, composed
307 of 8488 pairs of publicly available methylation iDats for different brain tumour types that are part of the
308 classifier (derived from accession numbers E-MTAB-3476, E-MTAB-4969, E-MTAB-5797, E-MTAB-7490, E-
309 MTAB-8390, E-MTAB-8888, GSE103659, GSE104210, GSE104293, GSE109330, GSE109381, GSE116298,
310 GSE117130, GSE119774, GSE122038, GSE122920, GSE122994, GSE123678, GSE124617, GSE125450,
311 GSE128654, GSE131482, GSE133801, GSE135017, GSE136361, GSE137845, GSE138221, GSE140124,
312 GSE143843, GSE147391, GSE152653, GSE156012, GSE156090, GSE157397, GSE164994, GSE166569,
313 GSE183656, GSE183972, GSE184900, GSE188547, GSE190362, GSE193196, GSE196490, GSE197378,
314 GSE198855, GSE200647, GSE215240, GSE36278, GSE52556, GSE55712, GSE60274, GSE61160,
315 GSE65362, GSE70460, GSE73801, GSE73895, GSE85218, GSE92577, GSE92579), TCGA_GBM and
316 TCGA_LGG. Entry into the reference cohort was determined by the same quality control metrics as for the
317 TYA cohort. Data was derived from 450K and 850K array data and was combined using combineArrays()
318 functionality from minfi, reducing both arrays to their common set of probes. The reference cohort therefore
319 comprised methylation profiles for the spectrum of different brain tumour types, irrespective of age, that the
320 TYA methylation cohort could be compared with in terms of their t-SNE cluster locations, and allowed the
321 exclusion of any cases that classified or aligned with non-glial tumours. Gender information was available for
322 4653 cases, 2109 cases were from female patients, and 2544 from male patients. Age was available for 5078
323 cases (median 19 years, range 0.1-93 years). The glioma reference cohort was derived by filtering the CNS
324 reference cohort for the different glial subtypes (specifically high-grade glial tumours). Cohorts of all the
325 existing HGG entities that are currently part of the v12.8 classifier were represented. Gender information was
326 available for 1446 cases (604 female, 842, male). Age was available for 1434 cases (median 25 years, range 0.1-
327 86 years). These were plotted in a separate t-SNE, and the TYA cohort was then plotted with them, to observe
328 more clearly the clustering of the different HGG subgroups in the TYA population.

329

330 **DNA copy number**

331

332 DNA copy number was recovered from combined intensities using the conumee package (v1.36.0). This was
333 derived from combined $\log_2$ intensity data based upon an internal median processed using the R packages minfi
334 and conumee to call copy number in 15,431 bins across the genome. Samples were arranged in columns
335 clustered by contiguous categorical copy number states based upon log ratio thresholds of ±0.1 for gain/loss and
336 ±0.5 for amplification and deletion, and organised by their DNA methylation subgroups. Clustering used
337 Euclidean distance and a ward algorithm.

338

339 **Fusion Panel**

340

341 RNA was successfully extracted from 128 samples. Library preparation was completed using FUSIONPlex™
342 Pan Solid Tumor v2 panel (Archer) following manufacturer's instructions. Libraries were pooled and sequenced
343 aiming to target 2.5M reads per sample. Analysis was performed using Archer® Analysis Unlimited (software
344 v7.1).

9

DNA Sequencing

Whole exome sequencing was performed for 107 cases using Agilent SureSelect whole exome v8 at the Institute of Cancer Research core genomics facility. Libraries were pooled and sequenced appropriately aiming for 300x depth. Matched germline DNA was available for 27 samples, which were prepared as above and sequenced at 100x depth. Exome capture reads were aligned to the hg19 build of the human genome using bwa v0.7.12 (bio-bwa.sourceforge.net), and PCR duplicates removed with PicardTools 1.94 (pcard.sourceforge.net). Single-nucleotide variants were called using the Genome Analysis Tool Kit v3.4–46 based upon current best practices using local realignment around indels, down sampling, and base recalibration with variants called by the Unified Genotyper (broadinstitute.org/gatk/). Variants were annotated using the Ensembl Variant Effect Predictor v74 (ensembl.org/info/docs/variation/vep) incorporating SIFT (sift.jcvi.org) and PolyPhen (genetics.bwh.harvard.edu/pph2) predictions, COSMIC v64 (sanger.ac.uk/genetics/CGP/cosmic/), dbSNP build 137 (ncbi.nlm.nih.gov/sites/SNP), ExAc, and ANNOVAR annotations.

Neuropathology assessment

The original slides were centrally reviewed, blinded to the molecular profile. Tumours were reviewed with reference to microscopic criteria reported in the WHO Classification of Tumours of the Central Nervous System, 5th Edition 2021 including the degree of cellularity, atypia, presence of mitoses, microvascular proliferation and necrosis (10). Differences in cytological or architectural appearances were noted and subsequently reviewed in the context of molecular profile.

Statistical Analysis

Statistical analysis was carried out using R 4.3.2 (www.r-project.org) and GraphPad Prism 9. Categorical comparisons of counts were carried out using Fisher exact test; comparisons between groups of continuous variables employed Student t test or ANOVA. Univariate and multivariate differences in survival were analysed by cox regression to determine the hazard ratios and significance. All tests were two-sided and a P value of less than 0.05 was considered significant after multiple testing correction (FDR).

Data Availability

All newly generated data have been deposited in the European Genome–phenome Archive (www.ebi.ac.uk/ega) with accession number EGAS50000000641 (sequencing) or ArrayExpress (www.ebi.ac.uk/arrayexpress/) with accession numbers E-MTAB-13974 and E-MTAB-13975 (methylation arrays). Curated gene-level copy number and mutation data are provided as part of the paediatric-specific implementation of the cBioPortal genomic data visualisation portal (pedcbioportal.org). The summary data can also be accessed via Synapse (accession number 64620613). Raw data from this study are available upon request to the corresponding author.

385

386

387

388

389

390

391

392

393

394 **Results**

395

396 **The TYA HGG cohort**

397

398 Two-hundred and seven of the collected samples, from patients aged 13-30 years at diagnosis and from any

399 location within the CNS, underwent DNA methylation profiling, of which 107 cases also underwent whole

400 exome sequencing, and were confirmed to be *IDH-* and histone-wildtype high-grade glioma (**figure 1A**).

401 Tumour location was available for 189 cases; 84.7% (n=160) were hemispheric, and 15.3% (n=29) were

402 identified in a midline anatomical location (**figure 1B**). 44.4% (n=92) of the collected cohort were female, and

403 55.6% (n=115) were male (**supplementary figure 1A**). The median age was 17.0 years (range 13–30 years).

404 Where the histological diagnosis was available, 83.7% (n=154/184) were diagnosed histologically as a

405 glioblastoma or high-grade glioma; rare cases formerly known as PNET and anaplastic pleomorphic

406 xanthoastrocytoma (PXA) were included due to no available molecular confirmation of these diagnoses. Review

407 of the histology (where available) showed high cellularity, microvascular proliferation, necrosis, and increased

408 mitotic activity. 7.6% (n=14) were diagnosed as a form of anaplastic glioma (a term no-longer recognised by the

409 WHO classification in the context of glioma subtypes) (**supplementary figure 1B**).

410

411 Outputs from DNA methylation profiling was used to filter cases with poor quality control (QC), and those

412 classified as 'NOS' due to calibrated scores lower than 0.5. A pan-CNS reference cohort, consisting of 8488

413 cases from existing publicly available datasets (2,8,16,21,24,25), was constructed and plotted using t-SNE

414 projections (**supplementary figure 2A**).

415

416 Twelve cases did not meet the methylation data QC requirements and so were not included in the methylation

417 data analysis, but remained in the cohort and were processed as per other elements of the study. Thirty-seven

418 cases were assigned 'NOS', requiring further characterisation using alternative integrated diagnostic approaches

419 (**figure 1C**). The NOS group did not form distinct methylation clusters using t-SNE, suggesting they did not

420 represent a novel subgroup(s). The remaining 158 cases (excluding NOS cases) were projected onto a reference

421 set of gliomas comprising several HGG entities (n=3357) (**figure 2A**, (**supplementary table 1**). Twenty-nine

422 cases (18.0%) were classified as subgroups more frequently seen in older adults (by current classifier

423 definitions) including GBM_MES (n=25, 15.5%), GBM_RTK1/2 (n=4, 2.5%). 15.5% (n=25) cases classified as

11

424   a pleomorphic xanthoastrocytoma (PXA) and 5.0% (n=8) as high-grade astrocytoma with piloid features
425   (HGAP), a new entity in the WHO classification.

426

427   Clinical treatment information was available for nine cases classifying as PXA. 7/9 PXA cases achieved a GTR,
428   a single tumour was debulked and one received a biopsy only. All 9 cases were treated with chemotherapy
429   (reported to be temozolomide in 5 cases) and radiotherapy, and two cases were treated with radiotherapy alone.
430   Two cases were treated with trametinib and dabrafenib. A combination of dabrafenib and vemurafenib was used
431   for two cases diagnosed as 'anaplastic PXA'. Survival data comparing PXA with HGAP showed a similar
432   overall survival profile (median OS 28 months vs 34 months respectively, p=0.5393 ns) (**supplementary figure**
433   **2B**).

434

435   Overall, this demonstrated that a relatively small proportion of TYA *IDH-* and H3-wildtype HGGs represent
436   tumour types associated with the adult population.

437

438   **Paediatric-specific and novel methylation subgroups**

439

440   When reviewing the classifier assignments and tSNE distribution of the 158 cases (excluding NOS cases),
441   eighty-six (54.4%) classified as paediatric-type subgroups (by current classifier definitions) including pedHGG-
442   RTK1A/B/C (n=51, 32.3% (47.1% n=24 RTK1A, 31.4% n=16 RTK1B, 21.6% n=11 RTK1C), pedHGG-*MYCN*
443   (n=13, 8.2%), and pedHGG-RTK2A/B (n=12, 7.6% (75.0% n=9 pedHGG_RTK2A, 25.0% n=3
444   pedHGG_RTK2B). Twenty-three cases were assigned to novel, recently identified, poorly-characterised
445   subgroups with distinct methylation profiles including adult-type diffuse HGG, IDH-wildtype, subtype B
446   (HGG_B, n=2, 1.3%), adult-type diffuse HGG, IDH-wildtype, subtype E (HGG_E, n=6, 3.8%), glioblastoma,
447   IDH-wildtype, subtype posterior fossa (GBM-CBM, n=5, 3.2%) and pedHGG-A/B (n=10, 6.3%) (**figure 2A**).

448

449   The 158 cases were then plotted in a t-SNE alone (**figure 2B**). They maintained the expected pattern of
450   clustering according to their methylation profiles, including clustering of those subgroups associated with
451   paediatric populations (e.g., pedHGG_A/B/RTK/*MYCN* etc.) and adult population (GBM_MES_TYP/ATYP).
452   Similar patterns of clustering to that seen in the glioma reference t-SNE was observed, helping to validate the
453   assignments. This was also confirmed using unsupervised clustering; the pedHGG_RTK tumours clustered
454   together as did the MAPK-pathway driven tumours (PXA and HGAP cases). Interestingly, the HGG_E cohort
455   formed a distinct cluster with a hypo-methylated profile (**supplementary figure 2C**).

456

457   The majority of TYA HGG *IDH-* and H3-wildtype HGGs therefore classify as either a paediatric-type or a
458   poorly understood novel methylation class.

459

460   **Incidence of tumour subtypes in the TYA age group**

461

462 To explore the relationship between age and frequency across the different HGG groups and subgroups, we

463 investigated publicly available methylation datasets in combination with our data (n=8488) to create age-density

464 plots for each methylation-defined subgroup of HGG (**figure 2C**).

465

466 PedHGG_RTK1B (n=37, median age 16.6 years, age-density peak 13.1 years), and pedHGG_B (n=23, median

467 age 31.4 years, age-density peak 16.2 years) show a large proportion of their frequency (n=27/38 and n=11/23

468 respectively) within the TYA age spectrum of 13-30 years. pedHGG_RTK1A (n=124, median 16.6 years, age-

469 density peak 13.1 years) and pedHGG_RTK1C (n=53, median 22.4 years, age-density peak 13.9 years) also

470 frequently occurred within the TYA age group (n=54/124 and n=24/53 respectively). Of those cases classifying

471 as pedHGG_RTK1, seven cases underwent GTR, three debulking procedures, two subtotal resections and a

472 single case of biopsy only. Fifteen cases were treated with adjuvant chemotherapy (most frequently

473 temozolomide) and fourteen treated with combined radiotherapy. Two patients were treated with chemotherapy

474 or radiotherapy alone. Single cases were treated with a peptide vaccine, bevacizumab, and palliative

475 immunotherapy (nivolumab).

476

477 Glioblastoma, RTK1 (n=213, median age 59.4 years, age-density peak 61.4 years), RTK2 (n=350, median age

478 58.4 years, age density peak 54.2 years), and MES_TYP (n=311, median age 54.5 years, age-density peak 61.6

479 years) subtypes show an age distribution of patients aged >50 years. GBM_MES_ATYP (n=40, median age

480 32.7 years, age-density peak 15.0 years), distinct from the GBM_MES_TYP group, show that 38% (n=15/40) of

481 cases fall within the TYA age group. The pedHGG_RTK1 subgroups also showed a similar distribution, with a

482 higher incidence in the TYA age group compared to the paediatric age group (n=105/215 vs n=86/215

483 pedHGG_RTK1), but they are not exclusive to this age. No subgroups show all of their distribution within the

484 TYA age group. With consideration to the number of cases in each subgroup cohort, many of the curves showed

485 a tail characterised by smaller peaks of density occurring outside their peak distribution, and in some subgroups

486 their expected range (e.g., paediatric-type tumours occurring in older adults and adult-specific tumours

487 occurring in the paediatric setting). pedHGG_RTK2A tumours (n=58, median age 13.3 years, age-density peak

488 12.7 years) showed a greater proportion of cases in the TYA age group than the pedHGG_RTK2B tumours

489 (n=41, median age 19.7 years, age-density peak 9.1 years) with n=25/58 and n=9/41 respectively. Clinical

490 annotations for cases classifying as pedHGG_RTK2, indicated one case had a subtotal resection and two

491 underwent biopsy only. Three were treated with both chemotherapy and radiotherapy, one of which was

492 craniospinal due to the identification of metastatic spinal lesions. Chemotherapy consisted of adjuvant

493 temozolomide and second line procarbazine and lomustine. Re-irradiation was reported in a single case.

494 Survival data for the paediatric-type HGG subgroups showed similar survival profiles (median OS for

495 pedHGG_MYCN 13.0 months, pedHGG_RTK1C 23.0 months, pedHGG_RTK2A 14.7 months, pedHGG_A

496 14.0 months, pedHGG_B 20.1 months) (**supplementary figure 2D**).

497

498 For the poorly characterised methylation subgroups, HGG_E cases show a broad range of incidence with a peak

499 at approximately 9.4 years but a median age of 15.8 years (n=23, range 2 - 64 years), with cases identified

500 across both the paediatric, adolescent and young adult ages but declining to very low incidences after age 30

501 (**figure 2C**). HGG_B cases (n=8, median age 35.1 years, age density peak 29.1 years) showed a higher

13

502   frequency of cases in the TYA age group, though numbers are small (n=6/7). According to the current classifier

503   definition of the HGG_B and HGG_E groups, they were both 'adult-type' diffuse HGG but such a definition

504   was not reflected by either of the density profiles. The GBM_CBM group (n=26, median age 26.8 years)

505   showed n=13/26 cases occurring in the TYA age group, with a peak at 15.6 years and a range of 2 – 74 years.

506   The pedHGG_B subgroup mimics the trend seen with the GBM_CBM group. The pedHGG_A group (n=25)

507   has a median age of 13.8 years, and the incidence then declines after age 30 years, with a range of 1.3 to 57

508   years. Clinical annotations for novel entities showed that three cases achieved GTR, and single incidences of

509   STR, tumour debulking and biopsy alone. One case of GBM_CBM was treated with adjuvant radiotherapy in

510   addition to bevacizumab and is alive at the time of writing. 4/5 HGG_E cases were treated with adjuvant

511   chemotherapy (temozolomide) and radiotherapy, with the remaining case treated with chemotherapy alone.

512

513   When comparing to other molecularly-defined HGG subtypes, the infant-type hemispheric gliomas (IHG)

514   (n=66, median age 2.3, age-density peak 0.5 years) characteristically occur in the infant population, the diffuse

515   hemispheric glioma, G34-mutant (DHG_G34) (n=181, median age 19.1 years, age-peak 14.3 years) and IDH-

516   mutant high-grade astrocytomas (A_IDH_HG) (n=221, median age 35.6 years, age-density peak 32.9 years),

517   and diffuse midline glioma, H3 K27-altered (DMG_K27) (n=357, median age 11.4 years, age-density peak 6.5

518   years) show a proportion of their frequency within the TYA age spectrum of 13-30 years (**supplementary**

519   **figure 2E**).

520

521   Therefore, no methylation subclasses of HGG are exclusive to the TYA age group, although some show a peak

522   in their distribution in the 13-30 year range (including HGG_B, GBM_CBM and GBM_MES_ATYP), and

523   more generally, paediatric- and adult-type HGG groups show a wider age distribution than perhaps would be

524   expected.

525

526   **Copy number profiling**

527

528   Copy number data derived from the Illumina array, was available for all 207 cases. Review of the clinical

529   annotations associated with the glioma reference cohort identified that 8310/11933 (69.6%) cases were

530   annotated for age, and of these 276 cases correlated with the study inclusion criteria and were within the TYA

531   age range. After excluding cases classified as NOS or failing to meet the methylation QC parameters, 158 cases

532   were combined with those derived from the glioma reference cohort (giving a total of n=434) and were clustered

533   within their respective methylation subgroups (**figure 3A**). Subtype-enriched profiles were observed, including

534   the classical gain of chr7 and loss of chr10 seen in the GBM_MES_TYP and GBM_RTK1 subgroups,

535   consistent with the pattern of adult HGG. The methylation-defined paediatric-specific subgroups showed more

536   variable profiles with changes frequently involving chr1. The pedHGG_*MYCN* tumours were enriched for gains

537   of chr1, 2 and 7.

538

539   When the same cohort was clustered by copy number change across the methylation subgroups, frequent

540   changes across the cohort included chromosomal gains (chr1q (54%), chr2 (22%), chr7 (41%)) and losses

541   (chr10 (40%), chr13 (64%)) (**supplementary figure 3**).

14

542

543   Focal amplifications seen in the cohort included *PDGFRA* (5.8%), *CDK4* (3.9%), *MYCN/ID2* (4.3%), *MYC*

544   (1.4%), *CDK6* (1.4%) and *EGFR* (1.0%) and the most common focal deletion was *CDKN2A/p16* (19.2%).

545   Across the combined cohort of 434 cases, similar patterns were seen, but also amplifications in *KIT* (n=14, most

546   frequent in pedHGG_RTK1B subgroup), *MET* amplifications (n=4, seen in pedHGG_RTK1C (n=2), HGG_B

547   (n=1) and PXA (n=1) subgroups) and a higher frequency of *EGFR* amplifications (n=15, seen across

548   GBM_RTK2 (n=4), pedHGG_*MYCN* (n=4), pedHGG_RTK2A (n=4) and pedHGG_A (n=3) subgroups) were

549   identified (**figure 3B, 3C, supplementary tables 2 & 3**).

550

551   TYA HGG therefore show the characteristic copy number changes, and gene amplifications and deletions,

552   which are known to occur across the HGG spectrum, but are enriched in specific methylation subclasses.

553

554   **Somatic mutational landscape of TYA HGG subgroups**

555

556   Whole exome sequencing data was available in 107/207 cases (**figure 4, supplementary table 4**). Twenty-

557   seven cases assigned as NOS (calibrated score <0.5) and 12 cases whose methylation data were excluded due to

558   inadequate QC parameters also underwent WES. Cases were grouped according to their HGG subgroup based

559   on their DNA methylation profiles; some of the variants detected (and the pattern of incidence) from the NOS

560   and methylation QC-failed group paralleled the known methylation profiles. The essential and desirable

561   diagnostic criteria from the current WHO classification (12) were applied when grouping these cases

562   accordingly. If the methylation or variant profiles could not support a reliable assignment, they were grouped

563   according to the variant type seen e.g., MAPK pathway-altered (n=8), *MYCN*-amplified (n=2), and

564   *CDKN2A/B/p16*-deleted (n=6).

565

566   Pathognomonic mutations of tumour subgroups were detected, including the presence of *CDKN2A/B* alterations

567   and MAPK pathway alterations (e.g., *BRAF* V600E mutations) in PXA (n=9), and *CDKN2A/B* alterations and

568   *ATRX* mutations in HGAP (n=3). The most frequent glioma-associated variants included *TP53* (n=65), *NF1*

569   (n=35), *PDGFRA* (n=28), and *EGFR* (n=18) (**supplementary figure 4A**).

570

571   Eight cases from the cohort showed a hypermutator phenotype (mutation frequency ranging from 2817-5273)

572   with frequent mutations in *TP53, ATRX, PDGFRA, BCOR, SEDT2, NF1,* and *RB1* (**figure 4**). Three of these

573   cases classified as the novel HGG_E methylation subgroup, suggesting a somatic mutational signature of this

574   group. Others included pedHGG_RTK1 (n=1) and GBM_CBM (n=1) cases, and also tumours which did not

575   classify as any of the currently recognised entities n=3, NOS).

576

577   Current adult defined subtypes including GBM_MES (n=15) and GBM_RTK1 (n=3) were enriched for the

578   presence of *PTEN* variants (n=8/18) and *CDK4* amplifications (n=2/3 in GBM_RTK1), in addition to frequent

579   variants in *TP53* (n=12/18).

580

15

581 The current paediatric subtypes also showed distinct signatures compared to other subgroups of HGG; as
582 expected, the pedHGG_RTK1A/B/C subgroups were all enriched for *PDGFRA* alterations (mostly
583 amplifications) (n=11/16) as well as *TP53* (n=11/16) and *ATRX* (n=6/16) variants, and variable incidences of
584 *CDKN2A/B, BCOR, SETD2* and *NF1* alterations (**figure 4**). Nine cases were assigned to a 'pedRTK1-like'
585 group based on *PDGFRA* (n=8) alteration. The pedHGG_MYCN was associated with *MYCN* amplifications,
586 although this was not a universal feature of high-scoring tumours in the *MYCN* subgroup (n=3/7). 2 further
587 NOS/methylation QC-failed cases contained a *MYCN* amplification and in the absence of other subgroup-
588 identifiable features, were assigned to a *MYCN*-like group.

590 Three cases classifying as the novel GBM_CBM group showed a variable incidence of variants including
591 hypermutator phenotype in one case, *ATRX* and *PDGFRA* variants, and non-specific range of variants in genes
592 frequently mutated in HGG including *BCOR, SETD2, NF1, RB1* and *CDK4* amplifications.

594 Eight cases contained variants frequently seen in HGG, but the profile of these and the available
595 neuropathological or molecular data did not support a reliable assignment and they were therefore grouped as
596 'HGG_Other' (**figure 4**).

598 ArcherDX fusion panel sequencing was performed on 92/207 (44.4%) cases (**supplementary figure 4B**), and
599 identified isolated examples of RTK fusions (*PRPTZ1::MET* and *GOPC::ROS1* fusions). These alterations were
600 validated by indicative small copy number changes, showing that RTK fusions are not a frequent feature in
601 TYA HGG. A single *KIAA1549::BRAF* fusion was identified in an HGAP case, which is a recognised molecular
602 feature in a subset of these tumours.

604 TYA HGG therefore show mutational landscapes which may be targetable, and an integrated diagnostic
605 approach can be used to assign cases to methylation- or variant-based cohorts to help guide management.

607 **Characteristics of poorly-described methylation-defined subgroups and clinical correlations**

609 This study has identified three HGG subgroups occurring more frequently in the TYAs including GBM_CBM,
610 HGG_B and GBM_MES_ATYP, all of which were considered poorly-described methylation subgroups. We
611 gathered all publicly available data, and from a local adult-based diagnostic centre, for these rare groups,
612 irrespective of age, to review their characteristic features (HGG_B (n=19), GBM_CBM (n=34) and
613 GBM_MES_ATYP (n=102)). The GBM_MES_ATYP subgroup occurred more frequently in males (n=62/102)
614 (**figure 5A**). The median age for GBM_CBM was 36.8 (range 9–77 years), GBM_MES_ATYP was 50.4 (range
615 1–82 years) and HGG_B was 29 (range 23–78 years) (**figure 5B**). There is a peak of distribution in the TYA age
616 group for GBM_CBM and GBM_MES_ATYP. However, for the GBM_CBM subgroup, there is a long tail of
617 distribution extending into older adults, and for the GBM_MES_ATYP, a large, broad peak in older patients.
618 The HGG_B subgroup however, is a TYA-specific subgroup. Although the current naming of the GBM_CBM
619 group implies they occur in the cerebellum, this is not always the case with 4/27 cases (14.8%) with available

16

620   location information reported in a hemispheric location (either frontal or temporal lobes), and single cases in the

621   thalamus and spinal cord (**figure 5C**).

622

623   Histology showed features consistent with a HGG including the presence of multinucleated giant cells for all

624   subtypes (**Figure 5D, supplementary figure 5A**). The TYA-specific HGG_B subgroup showed heterogenous

625   appearances. Some tumours showed perinuclear clearing (**figure 5D1-4**) with round to oval atypical nuclei.

626   However, they also showed variable architectural features; rare cases showed a nested architecture surrounded

627   by a more fibrous-like stroma (**figure 5D4**) and others showed a low to moderate cellularity with a more

628   infiltrative pattern (**figure 5D5**). A single case showed focal areas with an embryonal-like appearance

629   comprising hyperchromatic and pleomorphic irregular shaped nuclei which showed a distinct palisading

630   architectural pattern (**figure 5D6**). 3/9 cases showed the presence of microcalcifications. GBM_CBM histology

631   showed a highly cellular and infiltrative glial tumour, with a subset of cases showing perinuclear clearing, set

632   within a fibrillary stroma (**supplementary figure 5A1-5**). The nuclei were round to oval, with variable degrees

633   of nuclear pleomorphism (**supplementary figure 5A5**). Some cases showed oval-shaped or spindled nuclei, and

634   others contained cells with a 'pennies on a plate' or multinucleated morphology (**supplementary figure 5A1-5**).

635   Small, thin-walled branching vessels were interspersed within the tumours. GBM_MES_ATYP cases comprised

636   a highly cellular glial tumour set within a fibrillary stroma. The tumour cells formed a dense sheet-like

637   architecture. The nuclei showed variable degrees of atypia (including very markedly atypical and large nuclei),

638   and frequently displayed pale chromatin with prominent nucleoli, surrounded by pink eosinophilic pale, granular

639   or glassy cytoplasm (**supplementary figure 5A6-10**). Multinucleated cells were also common (**supplementary**

640   **figure 5A6-8**). Rare cases showed focal collections of more primitive appearing cells (**supplementary figure**

641   **5A9**), and frequently cases showed distinct areas of inflammatory cell infiltrates (12/39 cases) (**supplementary**

642   **figure 5A10**). Limited available survival data shows that the HGG_B, GBM_CBM and GBM_MES_ATYP

643   groups have a poor prognosis (median OS 19.1, 22.6 and 7 months respectively) (**figure 5E**).

644

645   Copy number profiling data for HGG_B (n=19) showed considerable variability with no distinctive patterns

646   (**figure 5F**). GBM_CBM (n=34) also showed considerable variability with no distinctive differences between

647   patients aged <30 or >30 years (**supplementary figure 5B**). The GBM_MES_ATYP cases (n=102) showed

648   frequent gains of chr7 associated with chr10 loss (18.6%, n=19/102) compared to GBM_MES_TYP (15.2%,

649   n=5/33, p=0.7960), but also a proportion of cases which contained just chr7 gains (19.6%, n=20/102), lower

650   than the proportion with this change seen in GBM_MES_TYP (27.3%, n=9/33, p=0.3421). When looking at

651   cases of GBM_MES_ATYP aged <30, these cases were more frequently associated with a gain of chr7 alone

652   (36.8% n=7/19) compared to those aged >30 (15.6%, n=13/83, p=0.0528), and a subset of these were also

653   associated with a gain of chr1 (21.1%, n=4/19), and a proportion of cases containing gains of chr6 (21.1%,

654   4/19). There was a higher frequency of cases which contained neither change in the GBM_MES_ATYP group

655   (40.2%, n=41/102) compared to GBM_MES_TYP (30.3%, n=10/33, p=0.4092) (**supplementary figure 5B**).

656   Sequencing data shows *TP53* variants (n=5/6) and amplifications including *MYCN* (n=1/6), *EGFR* (2/6), *MET*

657   (1/6), *CDK4* (1/6) and *PDGFRA* (1/6) in the HGG_B cases (**figure 5G**). There is enrichment of *CDKN2A/B*

658   deletions in the GBM_CBM group (n=7/10), and *RB1* (n=3/13), *PTEN* (n=6/13), *NF1* (n=5/13) and *TP53*

659   (10/13) alterations occurring more frequently in the GBM_MES_ATYP group (**supplementary figure 5C**).

17

660 When plotting methylation data of GBM_MES_ATYP cases alone into a tSNE and annotating them by age, no

661 age-defined clustering was seen (**supplementary figure 5D**).

662

663 Clinically, 3/5 GBM_MES cases achieved a GTR with two cases of subtotal resection. Three cases reported

664 treatment with adjuvant chemotherapy and radiotherapy. In addition, one patient received nivolumab,

665 ipilimumab, and regorafenib during the course of their treatment. Re-irradiation was used in a single case. Cases

666 classifying as GBM_MES_TYP showed an improved survival when compared to the GBM_MES_ATYP group

667 (median OS 18 months vs 7 months, p=0.05463) (**supplementary figure 5E**).

668

669 Several cases associated with tumour predisposition syndromes were identified; interestingly, two patients were

670 siblings (Glio_0021, Glio_0029). Glio_0021 was reported to have Turcot's/mismatch repair syndrome and

671 Glio_0029 was reported to have constitutional mismatch repair deficiency syndrome (CMMRD). Both tumours

672 showed histological features of a high-grade glioma, with the presence of multinucleated cells (**figure 6A**), and

673 were hypermutant, but classified differently by methylation profiling (pedHGG_RTK1A and HGG_E

674 respectively) (**figure 6B,C,D**). Glio_0004 was identified to have Li Fraumeni syndrome, and Glio_0272 (HGG

675 classifying as a pedHGG_RTK1A) was found to have bi-allelic somatic inactivation of *MSH2*, but germline

676 samples were not available.

677

678 6/70 cases (8.6%) were treated for a previous childhood malignancy before the diagnosis of an HGG (**figure

679 6E**); Glio_0228 was treated for a tectal plate pilocytic astrocytoma (including radiotherapy), but was

680 subsequently diagnosed with an HGG eleven years later, in a similar location, classifying as an HGG_E tumour.

681 Glio_0219 had a history of acute lymphoblastic leukaemia (ALL) as a child, treated with chemotherapy and

682 whole body radiotherapy, and was subsequently diagnosed with an HGG aged 22 which classified as

683 pedHGG_RTK1B. Similarly, Glio_0226 was also treated for ALL as a child with chemotherapy, and then

684 diagnosed with a pedHGG_RTK1B tumour aged 20. Glio_0272 was diagnosed with ALL aged 7 months and

685 was initially treated with chemotherapy, but underwent radiotherapy treatment aged 5 years when the disease

686 recurred; they were later diagnosed with a pedHGG_RTK1C tumour. Glio_0224 was treated with radiotherapy

687 for ALL aged 2 years, and then diagnosed with a tumour classifying as a PXA. An NOS case (0.22 scoring

688 GBM_CBM, Glio_0218) was treated for a desmoplastic medulloblastoma as a child with chemotherapy and

689 radiotherapy, and subsequently diagnosed with an HGG aged 20. HGG_E cases showed a worse overall survival

690 when compared to pedHGG_RTK1A and pedHGG_RTK1B cases in the cohort (median OS survival of 7.9,

691 16.0 and 16.0 months respectively) (**figure 6F**).

692

693 Multivariate analyses show that the pedHGG_RT1B and HGG_E methylation classes were associated with poor

694 survival (hazard ratio 1.69 (0.827-5) p=0.15 and 3.53 (2.26-13.4) p=0.041 respectively) (**supplementary figure

695 6A**). The GBM_MES_ATYP methylation subclass also was significantly associated with poor outcomes (3.86

696 (2.47-14.6) p=0.03). Mutations in *SETD2* and *MET* were also associated with worse survival compared to

697 wildtype cases (hazard ratios 2.02 (1.03-6.12) p=0.07 and 2.19 (1.41-8.33) p=0.18 respectively)

698 (**supplementary figure 6B**). Amplifications in *MET, CDK4, MYCN* and *PDGFRA* were associated with worse

699 survival, whereas *CDKN2A/B* deletions did not impact survival (**supplementary figure 6C**). GTR with

18

700   adjuvant chemotherapy alone was significantly associated with worse outcomes (hazard ratio 4.26 (2.92-17.8)
701   p=0.03), as was the decision to undertake a biopsy rather than a resection (hazard ratio 2.63 (1.42-8.33)
702   p=0.026) (**supplementary figure 6D**). *PDGFRA* and *TP53* mutations were significantly associated with a worse
703   overall survival within the TYA cohort compared to wildtype cases (hazard ratios 3.36 (1.96-11.4) p=0.014, and
704   2.07 (0.922-5.8) p=0.013 respectively) (**supplementary figure 6B,E,F**). However, in many cases throughout
705   these analyses, the numbers are small and therefore these results should be treated with caution.

706
707
708
709
710
711
712
713
714
715
716
717
718
719
720

721   **Discussion**

722

723   TYA HGGs are a diverse group of tumours, sharing many clinical challenges of HGGs occurring in other age
724   groups. The limited research in this age group, with few studies and limited molecular data currently available
725   has further impacted on management (15). The current study investigated a large international multi-institutional
726   cohort of TYA HGG to provide greater insights into their spectrum. An integrated approach to tumour
727   classification and characterisation, using DNA methylation profiling, fusion panel, WES, and histological
728   review highlighted the importance of this approach when adopted in diagnosis and tumour stratification (26).

729

730   In the UK and Europe, the age definition of TYA is (15-24 years) (27). The WHO classification recognises that
731   there are adult-type and paediatric-type HGGs, but that these tumours can occur outside of these predicted
732   profiles. The rationale to use the age range of 13-30 years was to explore whether there were specific tumour
733   types with a median incidence occurring within the TYA window, but also to explore the possible tails of
734   distribution beyond the classic age limit; pedHGG_A (median age 13.8), HGG_E (median age 15.8),
735   pedHGG_RTK1A/1B/1C (median ages 16.6, 16.6, 22.4 respectively), pedHGG_RTK2B (median age 19.7), and
736   PXA (median age 17.3) all had a median ages within the 13-30 window. Although the median age for the
737   HGG_B group was 35.1 years, the age density peak was 29.5 years; the MNP classifier currently defines this as
738   an 'adult-type diffuse high-grade glioma, IDH-wildtype tumour' but our study highlights that this may not be
739   the case, and that it may correlate more closely with other paediatric-type subgroups or be a true TYA subgroup.

740   With the acquisition of additional methylation data for GBM_CBM and GBM_MES_ATYP cases, irrespective

741   of age, their median ages were 36.8 and 50.4 years respectively. However, despite this not being a population-

742   based study, GBM_CBM shows a long tail of distribution extending into older adults, and the

743   GBM_MES_ATYP shows a bimodal distribution, including the TYA window. The age-density plots for

744   different HGG subgroups in this study highlight that age is important, but not an arbitrary singular cut-off; it

745   should be more appropriately defined by distribution medians and confidence intervals. Broad age inclusion

746   criteria (or no age exclusions) will be helpful to consider in future studies; HGG subtype-specific studies could

747   focus on identifying cases via their molecular characteristics, rather than an age definition. Comparative studies

748   with other published cohorts will therefore require the prioritisation of molecular profiling (including the use of

749   DNA methylation profiling), opposed to comparative age inclusion criteria.

750

751   Within this study, we accepted an MNP12 calibrated score of >0.5 as acceptable for classification. The use of

752   scores >0.9 in Capper *et al* 2018 as well as the MNP website is a conservative threshold (8). For example, if a

753   calibrated score of >0.9 is not achievable, the classifier can still assign the case to a glioma family group. Also,

754   tumour classification was undertaken by integrating MNP classifier outputs, sequencing data, and clinical

755   information, utilising the integrated approach used in neuropathology diagnostic setting. It is also important to

756   consider the context of this study which is an exploratory research project, helping to refine future diagnostic

757   practice. Therefore, the need for diagnostic precision, where the output will influence treatment in individual

758   cases, is not the goal. A calibrated score of <0.9 should not automatically mean that the class assignment is not

759   of value (9); many factors can influence the score including the cellularity, tumour percentage and the presence

760   of necrosis. However, if the assignment is still comparable with the histological features or other sequencing

761   results (which provide a further level of scrutiny), then this can still support the diagnosis. The copy number

762   data derived from the Illumina array data is also very valuable, illustrating characteristic copy number changes

763   which can either support or refute the suggestion by the classifier, or the neuropathologist's suspicions. If

764   genomics data, copy number changes, histological features, and the t-SNE position align, the inclusion of cases

765   with a score <0.9 was justified.

766

767   Rare subgroups showed a peak age distribution in the TYA age group (GBM_CBM, GBM_MES_ATYP,

768   HGG_B) compared to paediatric (<13) and adult (>30) ages. Distinct patterns of copy number are seen between

769   all three of these subgroups when comparing data from tumours occurring in those aged <30 vs >30 years, most

770   prominently with the GBM_MES_ATYP group. At this stage, methylation does not seem to suggest that those

771   occurring in patients aged <30 are a different subgroup, but further sequencing data and clinical annotations are

772   needed to explore if the distinctions in copy number represent a solitary molecular difference.

773

774   The GBM_CBM group are currently defined by the methylation classifier as occurring in the cerebellum. Early

775   studies on cerebellar glioblastoma showed that they were seen more frequently in younger patients with a

776   median age of 52 years (range 5-88 years) and classified predominantly as anaplastic astrocytoma with piloid

777   features (now known as HGAP) and the former GBM_MID methylation class, with the tumours demonstrating

778   similar methylation profiles to DMG_K27 tumours, but with an absence of the histone mutation (28). However,

779   there are other methylation subgroups which are found in this location but of lower frequency, including

780 GBM_MES, GBM_RTK1 and GBM_RTK2, leading authors to suggest that they don't represent a molecularly

781 uniform tumour (28). However, our study suggests otherwise, that there is in fact a distinct methylation

782 subgroup, with a peak of distribution in the TYA window, and some potentially occurring post-radiotherapy

783 after treatment for medulloblastoma in childhood. Interestingly, despite implications of current terminology in

784 the classifier, they do not occur exclusively in the cerebellum.

785

786 The HGG_B group appear to be the only true TYA subgroup identified within this study, and there are currently

787 no published studies focussed on this subgroup. They can occur in hemispheric, midline and posterior fossa

788 locations, and display a variable copy number profile, with some changes more specific to those patients aged

789 <30. Alongside frequent *TP53* alterations and variable incidences of *CDKN2A/B* deletions, *BCOR* alterations,

790 amplifications across different genes including *MYCN, MET* and *CDK4* appear to be a characteristic feature.

791 Histologically, they showed variable appearances, but a proportion were noted to be associated with

792 microcalcifications. Therefore, DNA methylation profiling remains the crucial molecular test for diagnosis.

793 Larger cohort studies including *in vitro* and *in vivo* work are needed to further explore targetable associations.

794

795 Although clinical annotations were incomplete, there is a suggestion of better survival in some paediatric-type

796 subgroups including pedHGG_RTK1A, but more data is needed to validate this trend. There are several tumour

797 predisposition syndromes associated with HGGs including CMMRD and Li Fraumeni (29,30), and this study

798 identified several incidences of these. Tumours which occur as a result of DNA replication-repair deficiency

799 contain pathogenic variants in polymerase-proofreading genes and/or mismatch repair genes (31). Existing

800 studies report a median age of 50 years (range 27-78), an association with the histological giant cell variant of

801 GBM, and most frequently classifying as the pedHGG_RTK1A methylation group (32). They contain high

802 mutation and microsatellite burdens, but their association with high T-cell infiltration leads to the optional use of

803 immune checkpoint inhibition; synergistic combinations (the addition of ipilimumab and the use of MEK-

804 inhibition) have shown promise in those where immune checkpoint inhibition does not work (33,34). In one

805 study, 4/5 cases classifying as pedHGG_RTK1A and treated with immune check point blockage survived

806 greater than 3 years, and there was an overall survival of 36.8 months compared to 15.5 months for other HGGs

807 in the study (32). Hypermutation can also be induced by temozolomide treatment in both IDH-mutant and IDH-

808 wildtype LGG and HGG (31,35), which is an important consideration for tumours that may have undergone

809 transformation to a higher-grade glioma, and the potential value of longitudinal sampling to monitor molecular

810 changes as treatment progresses. One such case in our cohort of a tectal plate pilocytic astrocytoma which was

811 treated with radiotherapy and years later was diagnosed with an HGG is a case in point. Similarly, we report

812 siblings with CMMRD; novel homozygous *MSH6* mutations have been previously reported in siblings, and

813 resulted in an HGG and T-ALL in an 11-year-old female, and a GBM in her 10-year-old brother, both of which

814 rapidly progressed (36).

815

816 Ionising radiation is the only known risk factor for the development of a secondary HGG, but the incidence is

817 only at 1% (37,38), with variable latency periods seen across all age spectrums (39). A striking feature of our

818 cohort was that 8.6% of cases had been treated for a childhood malignancy (most frequently ALL or a

819 medulloblastoma of the cerebellum), with a subsequent HGG diagnosis in their adolescent years. The risk that

21

820 previous radiation poses for the development of HGGs has been reported in the literature (39–42), and it is
821 important to recognise that these secondary tumours occur in the TYA population, classifying as primarily
822 pedHGG_RTK1B or HGG_E subgroups in our cohort. Existing studies also show that these tumours cluster
823 with the pedHGG_RTK1 subgroups, and are associated with *PDGFRA* and *CDK4* gains and losses of
824 *CDKN2A/B* and *BCOR* (42). Transcriptome analysis has also identified two subgroups (stem-like and pro-
825 inflammatory), with drug assays suggesting that protease inhibitors may be effective in the pro-inflammatory
826 subgroup (42). It highlights the need for further studies of these biological and molecular alterations, occurring
827 at the time of, and post-treatment. Better risk stratification is also needed for those most vulnerable and earlier
828 identification for patients who have been treated for a childhood cancer (including non-CNS tumours),
829 potentially with the use of screening during their TYA years. Avoiding the harmful use of radiotherapy to treat
830 childhood cancers should be the goal (43).

831

832 A cohort of 207 cases represents one of the largest cohorts of TYA cases published in the literature. However, a
833 limitation of this study is that some of the rarer subgroups identified within this set comprise very small
834 numbers, and with limited clinical annotations. Where material is available (and after the provision of molecular
835 testing has been considered), additional immunohistochemical studies may be helpful, particularly for novel
836 subgroups, and may assist the neuropathologist in identifying these cases at an earlier stage in the diagnostic
837 workflow, and for those neuropathologists working in lower-middle-income countries. This study also
838 showcases the value of detailed molecular profiling of all HGG cases, irrespective of age, so that we can
839 identify examples of these different tumour types which may be occurring outside of their typical or expected
840 age ranges; this is, and will become, more valuable when further targeted treatments are developed, which are
841 specific to the alterations we see in these cases. For example, hypermutant cases may be eligible for treatment
842 with check point inhibitors (44), and the use of *PDGFRA* (45–47) and MAPK-pathway inhibitors (48,49) are
843 now more established. Other frequently occurring alterations which this study highlights are *CDKN2A/B*
844 deletions, *PTEN* and *MYCN* alterations; although not specific to HGG, therapies targeting these frequent
845 changes may help to improve outcomes as part of combination therapies. One of the most significant challenges
846 with cohort studies is acquiring clinical and survival data. It is important that this data is available in a more
847 accessible way to researchers to enable direct comparisons between the treatment strategies employed, and to
848 explore the differences in response. Preclinical studies using patient-derived cell cultures and patient-derived
849 and cell-line derived xenografts (PDX, CDX) are needed to support this work. Establishing robust
850 collaborations with local neurosurgical teams will be important for all teams across the neuro-oncology research
851 network to help meet this urgent need.

852

853 The development of new molecular tests and the refinement of existing versions is important to consider. For
854 this study, we have used the most recently updated v12.8 MNP classifier. However, future iterations will be
855 produced, and alternative classifiers made available, which can lead to differences in assignments between
856 different versions. It is hoped that the use of allied molecular tests in addition to DNA methylation profiling,
857 already well established in most centres, will not impact on patient safety or outcomes and maintained the
858 desired accuracy of diagnosis. It should also be considered that the current version of the classifier is
859 unpublished, and developed using data which is not entirely publicly available; within this study we have

860 attempted to mitigate these challenges through the direct use of the MNP classifier website (the current classifier
861 methodology in diagnostic practice), the development of our own reference cohort (using publicly available
862 data), but also the use of the t-SNE projections and unsupervised clustering as supplementary tools, to visually
863 see the distribution of the subgroups. Until the eventual publication of the classifier, this remains the most
864 robust method we have available. The continual development of such evolving platforms must be encouraged
865 and supported. But it is important for all clinical teams and researchers to be mindful of such changes to see how
866 existing cohorts are impacted, and whether these changes may be beneficial for TYA HGG patients.

867

868 **Conclusion**

869 High-grade gliomas occurring in TYAs, perhaps unsurprisingly, overlap with methylation subgroups in the
870 paediatric and adult age groups. Some entities cluster within the TYA age group, whilst others display peaks
871 close to or involving the TYA age range. Novel methylation-defined classes are a feature of TYA HGGs, some
872 of which show mutational landscapes which may be targetable with immune checkpoint, *PDGFRA* and MAPK-
873 pathway inhibitors. Neuro-oncologists should be cautious in assigning tumour subgroups as 'adult-type' or
874 'paediatric-type' as this may influence the level of molecular profiling that will be undertaken in certain clinical
875 centres in relation to the age of the patient. Patients who have previously been treated for a childhood
876 malignancy but then develop a secondary HGG, and also backgrounds of tumour predisposition syndromes, are
877 enriched in the TYA population. DNA methylation profiling, WES/panel sequencing should be performed on
878 each TYA HGG to accurately diagnose and characterise these complex tumours.

879
880
881
882

23

907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923

924 **References**

925 1.   Guerreiro Stucklin AS, Ryall S, Fukuoka K, Zapotocky M, Lassaletta A, Li C, et al. Alterations in
926   ALK/ROS1/NTRK/MET drive a group of infantile hemispheric gliomas. Nat Commun. 2019;10(1):1–
927   13.

928 2.   Clarke M, Mackay A, Ismer B, Pickles JC, Tatevossian RG, Newman S, et al. Infant High-Grade
929   Gliomas Comprise Multiple Subgroups Characterized by Novel Targetable Gene Fusions and Favorable
930   Outcomes. Cancer Discov. 2020;10(7):942–63.

931 3.   Ostrom QT, Cioffi G, Gittleman H, Patil N, Waite K, Kruchko C, et al. CBTRUS Statistical Report:
932   Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2012-2016.
933   Neuro Oncol. 2019;21:V1–100.

934 4.   Gianno F, Antonelli M, Ferretti E, Massimino M, Arcella A, Giangaspero F. Pediatric high-grade
935   glioma: A heterogeneous group of neoplasms with different molecular drivers. Glioma. 2018;1(4):117.

936 5.   Ostrom QT, Cote DJ, Ascha M, Kruchko C, Barnholtz-Sloan JS. Adult glioma incidence and survival
937   by race or ethnicity in the United States from 2000 to 2014. JAMA Oncol. 2018;4(9):1254–62.

938 6.   Rong L, Li N, Zhang Z. Emerging therapies for glioblastoma: current state and future directions. J Exp
939   Clin Cancer Res. 2022;41(1):1–18.

940   7.   Pollack IF, Agnihotri S, Broniscer A. Childhood brain tumors: Current management, biological insights,
941   and future directions. J Neurosurg Pediatr. 2019;23(3):261–73.

942   8.   Capper D, Jones DTW, Sill M, Hovestadt V, Schrimpf D, Sturm D, et al. DNA methylation-based
943   classification of central nervous system tumours. Nature. 2018;555(7697):469–74.

944   9.   Capper D, Stichel D, Sahm F, Jones DTW, Schrimpf D, Sill M, et al. Practical implementation of DNA
945   methylation and copy-number-based CNS tumor diagnostics: the Heidelberg experience. Acta
946   Neuropathol [Internet]. 2018;136(2):181–210. Available from: https://doi.org/10.1007/s00401-018-
947   1879-y

948   10.   Drexler R, Brembach F, Sauvigny J, Ricklefs FL, Eckhardt A, Bode H, et al. Unclassifiable CNS tumors
949   in DNA methylation-based classification: clinical challenges and prognostic impact. Acta Neuropathol
950   Commun [Internet]. 2024;12(1):4–11. Available from: https://doi.org/10.1186/s40478-024-01728-9

951   11.   Karimi S, Zuccato JA, Mamatjan Y, Mansouri S, Suppiah S, Nassiri F, et al. The central nervous system
952   tumor methylation classifier changes neuro-oncology practice for challenging brain tumor diagnoses and
953   directly impacts patient care. Clin Epigenetics. 2019;11(1):1–10.

954   12.   Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, et al. The 2021 WHO
955   classification of tumors of the central nervous system: A summary. Neuro Oncol. 2021;23(8):1231–51.

956   13.   Gianno F, Giovannoni I, Cafferata B, Diomedi-Camassei F, Minasi S, Barresi S, et al. Paediatric-type
957   diffuse high-grade gliomas in the 5th CNS WHO Classification. Pathologica. 2022;114(6):422–35.

958   14.   Alken SP, D'Urso P, Saran FH. Managing teenage/young adult (TYA) brain tumors: a UK perspective.
959   CNS Oncol. 2015;4(4):235–46.

960   15.   Roux A, Pallud J, Saffroy R, Edjlali-Goujon M, Debily MA, Boddaert N, et al. High-grade gliomas in
961   adolescents and young adults highlight histomolecular differences from their adult and pediatric
962   counterparts. Neuro Oncol. 2020;22(8):1190–202.

963   16.   Mackay A, Burford A, Carvalho D, Izquierdo E, Fazal-Salom J, Taylor KR, et al. Integrated Molecular
964   Meta-Analysis of 1,000 Pediatric High-Grade and Diffuse Intrinsic Pontine Glioma. Cancer Cell.
965   2017;32(4):520-537.e5.

966   17.   Gottardo NG, Gajjar A. Chemotherapy for malignant brain tumors of childhood. J Child Neurol.
967   2008;23(10):1149–59.

968   18.   Terziev R, Psimaras D, Marie Y, Feuvret L, Berzero G, Jacob J, et al. Cumulative incidence and risk
969   factors for radiation induced leukoencephalopathy in high grade glioma long term survivors. Sci Rep
970   [Internet]. 2021;11(1):1–9. Available from: https://doi.org/10.1038/s41598-021-89216-1

971   19.   Wu J, Heidelberg RE, Gajjar A. Adolescents and Young Adults With Cancer: CNS Tumors. J Clin
972   Oncol. 2024;42(6):686–95.

973   20.   da Costa MDS, Camargo NC, Dastoli PA, Nicácio JM, Benevides Silva FA, Sucharski Figueiredo ML,
974   et al. High-grade gliomas in children and adolescents: Is there a role for reoperation? J Neurosurg
975   Pediatr. 2021;27(2):160–9.

976   21.   Mackay A, Burford A, Molinari V, Jones DTW, Izquierdo E, Brouwer-Visser J, et al. Molecular,
977   Pathological, Radiological, and Immune Profiling of Non-brainstem Pediatric High-Grade Glioma from
978   the HERBY Phase II Randomized Trial. Cancer Cell. 2018;33(5):829-842.e5.

979   22.   Specialist cancer services for children and young people: Teenage and Young Adults Designated

980   Hospitals. 2023.

981   23.   Nicoll JAR, Bloom T, Clarke A, Boche D, Hilton D. BRAIN UK: Accessing NHS tissue archives for

982   neuroscience research. Neuropathol Appl Neurobiol. 2022;48(2):1–8.

983   24.   Sturm D, Orr BA, Toprak UH, Hovestadt V, Jones DTW, Capper D, et al. New Brain Tumor Entities

984   Emerge from Molecular Classification of CNS-PNETs. Cell. 2016;164(5):1060–72.

985   25.   Sturm D, Capper D, Andreiuolo F, Gessi M, Kölsche C, Reinhardt A, et al. Multiomic neuropathology

986   improves diagnostic accuracy in pediatric neuro-oncology. Nat Med. 2023;29(4):917–26.

987   26.   Gritsch S, Batchelor TT, Gonzalez Castro LN. Diagnostic, therapeutic, and prognostic implications of

988   the 2021 World Health Organization classification of tumors of the central nervous system. Cancer.

989   2022;128(1):47–58.

990   27.   Ferrari A, Stark D, Peccatori FA, Fern L, Laurence V, Gaspar N, et al. Adolescents and young adults

991   (AYA) with cancer: a position paper from the AYA Working Group of the European Society for

992   Medical Oncology (ESMO) and the European Society for Paediatric Oncology (SIOPE). ESMO Open

993   [Internet]. 2021;6(2):100096. Available from: https://doi.org/10.1016/j.esmoop.2021.100096

994   28.   Reinhardt A, Stichel D, Schrimpf D, Koelsche C, Wefers AK, Ebrahimi A, et al. Tumors diagnosed as

995   cerebellarglioblastoma comprise distinct molecularentitiesAnnekathrin Reinhar. 2019;8:1–12.

996   29.   Hansford JR, Das A, McGee RB, Nakano Y, Brzezinski J, Scollon SR, et al. Update on Cancer

997   Predisposition Syndromes and Surveillance Guidelines for Childhood Brain Tumors. Clin Cancer Res.

998   2024;30(11):2342–50.

999   30.   Michaeli O, Tabori U. Pediatric High Grade Gliomas in the Context of Cancer Predisposition

1000   Syndromes. J Korean Neurosurg Soc [Internet]. 2018 May 1;61(3):319–32. Available from:

1001   https://doi.org/10.3340/jkns.2018.0031

1002   31.   Choi S, Yu Y, Grimmer MR, Wahl M, Chang SM, Costello JF. Temozolomide-associated

1003   hypermutation in gliomas. Neuro Oncol. 2018;20(10):1300–9.

1004   32.   Hadad S, Gupta R, Oberheim Bush NA, Taylor JW, Villanueva-Meyer JE, Young JS, et al. "De novo

1005   replication repair deficient glioblastoma, IDH-wildtype" is a distinct glioblastoma subtype in adults that

1006   may benefit from immune checkpoint blockade. Acta Neuropathol [Internet]. 2024;147(1):1–24.

1007   Available from: https://doi.org/10.1007/s00401-023-02654-1

1008   33.   Das A, Ercan AB, Tabori U. Neuro-Oncology Advances germline replication-repair deficiency

1009   syndromes. 2024;6(June):1–6.

1010   34.   Das A, Fernandez NR, Levine A, Bianchi V, Stengs LK, Chung J, et al. Combined Immunotherapy

1011   Improves Outcome for Replication-Repair-Deficient (RRD) High-Grade Glioma Failing Anti-PD-1

1012   Monotherapy: A Report from the International RRD Consortium. Cancer Discov. 2024;14(2):259–73.

1013   35.   Touat M, Li YY, Boynton AN, Spurr LF, Iorgulescu JB, Bohrson CL, et al. Mechanisms and therapeutic

1014   implications of hypermutation in gliomas. Nature [Internet]. 2020;580(7804):517–23. Available from:

1015   http://dx.doi.org/10.1038/s41586-020-2209-9

1016   36.   Ilencikova D, Sejnova D, Jindrova J, Babal P. High-grade brain tumors in siblings with biallelic MSH6

1017   mutations. Pediatr Blood Cancer. 2011 Dec;57(6):1067–70.

1018   37.   Carret AS, Tabori U, Crooks B, Hukin J, Odame I, Johnston DL, et al. Outcome of secondary high-

1019   grade glioma in children previously treated for a malignant condition: A study of the Canadian Pediatric

1020   Brain Tumour Consortium. Radiother Oncol. 2006;81(1):33–8.

1021 38. Maluf FC, DeAngelis LM, Raizer JJ, Abrey LE. High-grade gliomas in patients with prior systemic
1022   malignancies. Cancer. 2002;94(12):3219–24.

1023 39. Deng MY, Sturm D, Pfaff E, Sill M, Stichel D, Balasubramanian GP, et al. Radiation-induced gliomas
1024   represent H3-/IDH-wild type pediatric gliomas with recurrent PDGFRA amplification and loss of
1025   CDKN2A/B. Nat Commun. 2021;12(1).

1026 40. Whitehouse JP, Howlett M, Federico A, Kool M, Endersby R, Gottardo NG. Defining the molecular
1027   features of radiation-induced glioma: A systematic review and meta-analysis. Neuro-Oncology Adv.
1028   2021;3(1):1–16.

1029 41. López GY, Van Ziffle J, Onodera C, Grenert JP, Yeh I, Bastian BC, et al. The genetic landscape of
1030   gliomas arising after therapeutic radiation. Acta Neuropathol [Internet]. 2019;137(1):139–50. Available
1031   from: https://doi.org/10.1007/s00401-018-1906-z

1032 42. DeSisto J, Lucas JT, Xu K, Donson A, Lin T, Sanford B, et al. Comprehensive molecular
1033   characterization of pediatric radiation-induced high-grade glioma. Nat Commun. 2021;12(1):1–16.

1034 43. Jeha S, Pei D, Choi J, Cheng C, Sandlund JT, Coustan-Smith E, et al. Improved CNS control of
1035   childhood acute lymphoblastic leukemia without cranial irradiation: St Jude Total Therapy Study 16. J
1036   Clin Oncol. 2019;37(35):3377–91.

1037 44. Bouffet E, Larouche V, Campbell BB, Merico D, De Borja R, Aronson M, et al. Immune checkpoint
1038   inhibition for hypermutant glioblastoma multiforme resulting from germline biallelic mismatch repair
1039   deficiency. J Clin Oncol. 2016;34(19):2206–11.

1040 45. Yang W, Wang S, Zhang X, Sun H, Zhang M, Chen H, et al. New natural compound inhibitors of
1041   PDGFRA (platelet-derived growth factor receptor α) based on computational study for high−grade
1042   glioma therapy. Front Neurosci. 2023;16(January):1–14.

1043 46. Lane R, Cilibrasi C, Chen J, Shah K, Messuti E, Mazarakis NK, et al. PDGF-R inhibition induces
1044   glioblastoma cell differentiation via DUSP1/p38MAPK signalling. Oncogene. 2022;41(19):2749–63.

1045 47. Paugh BS, Zhu X, Qu C, Endersby R, Diaz AK, Zhang J, et al. Novel oncogenic PDGFRA mutations in
1046   pediatric high-grade gliomas. Cancer Res. 2013;73(20):6219–29.

1047 48. Stitzlein LM, Adams JT, Stitzlein EN, Dudley RW, Chandra J. Current and future therapeutic strategies
1048   for high-grade gliomas leveraging the interplay between epigenetic regulators and kinase signaling
1049   networks. J Exp Clin Cancer Res. 2024;43(1):1–20.

1050 49. Izquierdo E, Carvalho DM, Mackay A, Temelso S, Boult JKR, Pericoli G, et al. DIPG Harbors
1051   Alterations Targetable by MEK Inhibitors, with Acquired Resistance Mechanisms Overcome by
1052   Combinatorial Inhibition. Cancer Discov. 2022;12(3):712–29.

1053

1054

1055

1056

1057

1058

1059

1061

1062

1063

1064

1065

1066

1067

1068

1069

1070

1071

1072

1073

1074

1075

1076

1077

1078

1079

1080

1081

1082

1083

Figure Legends

1085

1086 **Figure 1**

1087 Defining an intrinsic set of TYA high-grade gliomas. A, Flow diagram providing an overview of the inclusion

1088 and exclusion criteria for the assembled cohort of 207 samples from patients aged >13 - $\leq$30 years. B, Anatomic

1089 location of TYA HGG after exclusion of non-glioma entities by methylation profiling (n = 189). Left, sagittal

1090 section showing internal structures; right, external view highlighting cerebral lobes. Each circle represents the

1091 proportion of cases occurring in this location, and is coloured by the generic locations for hemispheric (brown),

1092 midline (red) and brainstem (pink). C, Pie chart showing the proportion of different methylation subclasses

1093 within the cohort, before exclusion of non-glioma entities (n = 195). Each subgroup is represented by a different

1094 colour, as indicated by the key. Cases which scored <0.5 using the MNP classifier v12.8 were classed as 'NOS'.

1095 12 cases were excluded based on poor scoring quality control parameters of the methylation data.

1096

1097 **Figure 2**

1098 DNA methylation array profiling of the TYA HGG cohort. A, Methylation array profiling and analysis by the

1099 Heidelberg classifier excluded 12 cases which failed QC, and identified 34 with a calibrated score of <0.5 which

28

1100   were assigned as NOS. t-statistic based stochastic neighbour embedding (t-SNE) projection of the remaining

1101   158 cases highlighted cohorts which clustered with both adult and paediatric-type HGG subgroups, including

1102   some novel methylation-defined subgroups. Cases then underwent further histopathological assessment and in-

1103   depth sequencing to either confirm the methylation assignment, or to further characterise the different

1104   subgroups. B, t-statistic based stochastic neighbour embedding (t-SNE) projection of the collected cohort of 158

1105   TYA cases alone, without the glioma reference set cases. Samples are represented by dots coloured by subtype.

1106   C, Age-density plot showing the age distribution and peak incidences of different high-grade glioma subtypes

1107   by DNA methylation profiling, comprising 1704 cases derived from the collected TYA cohort and published

1108   datasets. Tumours are grouped according to methylation subclasses and predicted age-distribution according to

1109   the WHO classification. Dotted lines define the age cut-off for the TYA group in this project (13 and 30 years).

1110

1111   **Figure 3**

1112   DNA copy number profiling of TYA HGG. A, Heatmap representation of segmented DNA copy number for

1113   434 cases of TYA glioma profiled on the Illumina 450kor EPIC BeadArray platform (dark red, amplification;

1114   red, gain; dark blue, deletion; blue, loss). Samples are arranged in columns clustered by contiguous categorical

1115   copy number states based upon log ratio thresholds of ± 0.1 for gain/loss and ± 0.5 for amplification and deletion,

1116   and organised by their DNA methylation subgroups. Clustering used Euclidean distance and a ward algorithm.

1117   Methylation annotations are provided as a bar according to the included key. B, Frequency bar plot showing the

1118   most frequent amplifications identified from the copy number profiles across the cohort (n=434). C, Frequency

1119   bar plot showing the most frequent deletions identified from copy number profiles across the cohort (n=434).

1120

1121   **Figure 4**

1122   Mutations in TYA gliomas. OncoPrint representation of an integrated annotation of single-nucleotide variants,

1123   DNA copy-number changes, and structural variants for TYA gliomas (n = 107). Samples are arranged in

1124   columns with genes labelled along rows, and are grouped by methylation subclass and landscape of variants.

1125   Clinicopathologic and molecular annotations are provided as bars according to the included key.

1126

1127   **Figure 5**

1128   Characterisation of a true TYA HGG subgroup using the collected cohort and publicly available cases with

1129   DNA methylation profiling data. A, Gender bar plots showing the gender distribution of the available cases for

1130   GBM_CBM, HGG_B and GBM_MES_ATYP subgroups. B, Violin plots showing the age distribution and

1131   median age of the available cases for GBM_CBM, HGG_B and GBM_MES_ATYP subgroups. C, Anatomic

1132   location of poorly-characterised HGG subgroups irrespective of age. Left, sagittal section showing internal

1133   structures; right, external view highlighting cerebral lobes. Each circle represents proportion of cases occurring

1134   in this location, and is coloured by the tumour subgroups. D, Representative haematoxylin and eosin (H&E)

1135   images of the HGG_B subgroup. E, Kaplan-Meier showing overall survival available by methylation subclass

1136   (line colour as per key) for TYA-specific subgroups (n=12). F, Heatmap representation of segmented DNA copy

1137   number for the HGG_B subgroup (dark red, amplification; red, gain; dark blue, deletion; blue, loss). Samples

1138   are arranged in columns clustered by contiguous categorical copy number states based upon log ratio thresholds

1139   of ± 0.1 for gain/loss and ± 0.5 for amplification and deletion. Clustering used Euclidean distance and a ward

29

1140 algorithm. Methylation annotations are provided as a bar according to the included key. G, OncoPrint
1141 representation of an integrated annotation of single-nucleotide variants, DNA copy-number changes, and
1142 structural variants for the HGG_B subgroup. Samples are arranged in columns with genes labelled along rows,
1143 and are grouped by methylation subclass. Clinicopathologic and molecular annotations are provided as bars
1144 according to the included key.

1145

1146 **Figure 6**

1147 Tumour predisposition syndromes and treatment for childhood malignancies within the TYA cohort. A,
1148 Haematoxylin and eosin (H&E) images showing different cytological and architectural features of the two cases.
1149 B, t-statistic based stochastic neighbour embedding (t-SNE) projection of selected subgroups from the glioma
1150 reference cohort. Samples are represented by dots coloured by subtype. The sibling cases are highlighted by the
1151 black circles and labelled. C, OncoPrint representation of an integrated annotation of single-nucleotide variants,
1152 DNA copy-number changes, and structural variants for the sibling cases. Samples are arranged in columns with
1153 genes labelled along rows. Clinicopathologic and molecular annotations are provided as bars according to the
1154 included key. D, Circos plots demonstrating the hypermutator phenotypes of the sibling cases. Chromosomal
1155 locations are represented by ideograms arranged around the circle. E, Patient timelines for five patients
1156 identified within the cohort which were treated for a childhood malignancy. Gender is annotated using symbols
1157 and a sagittal brain cartoon demonstrates the location of the tumour. A timeline provides details of key events
1158 throughout the course of treatment. F, Kaplan-Meier showing overall survival available for HGG_E,
1159 pedHGG_RTK1B and pedHGG_RTK1A cases, (line colour as per key) for the collected cohort (n=34).
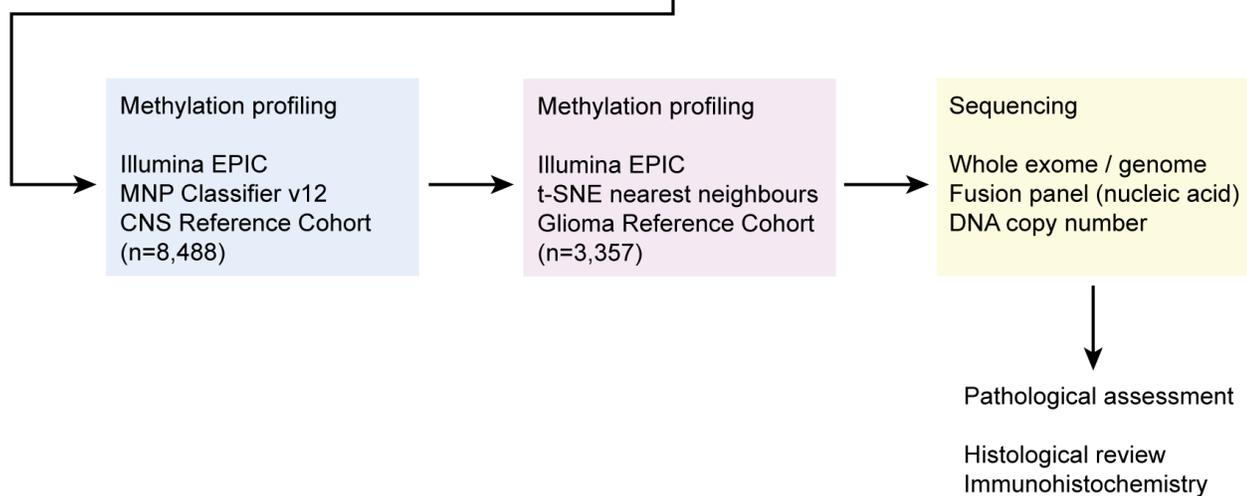
1160

1161

Figure 1

A

Inclusions

13-30 years
Any CNS locations
CNS WHO grade 2, 3, 4
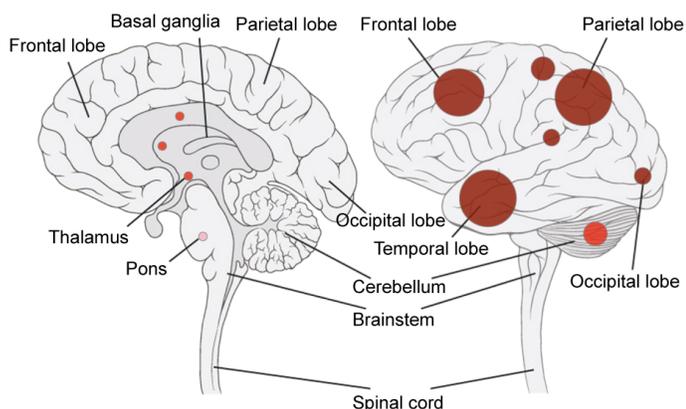High-grade glioma
IDH-wildtype, histone-wildtype

n=207

Exclusions

Established tumour entities;
LGG, histone-mutant glioma,
IDH-mutant glioma, ependymal,
embryonal, germ cell tumours,
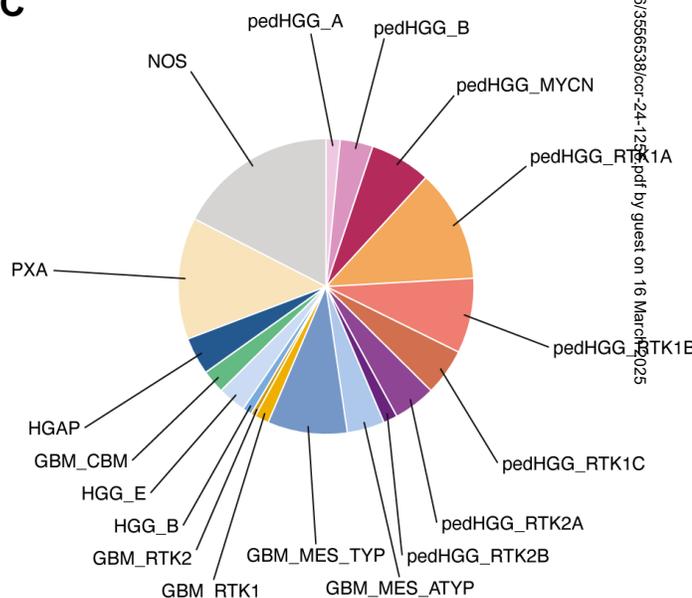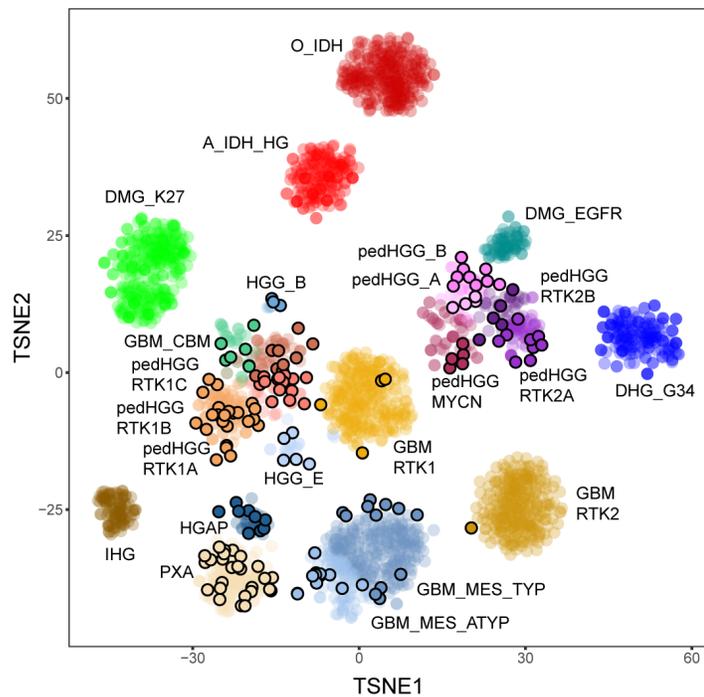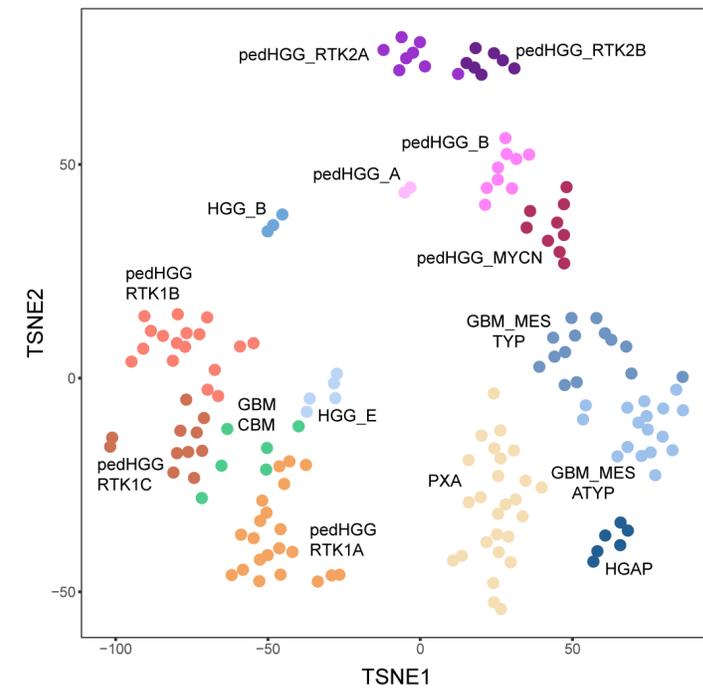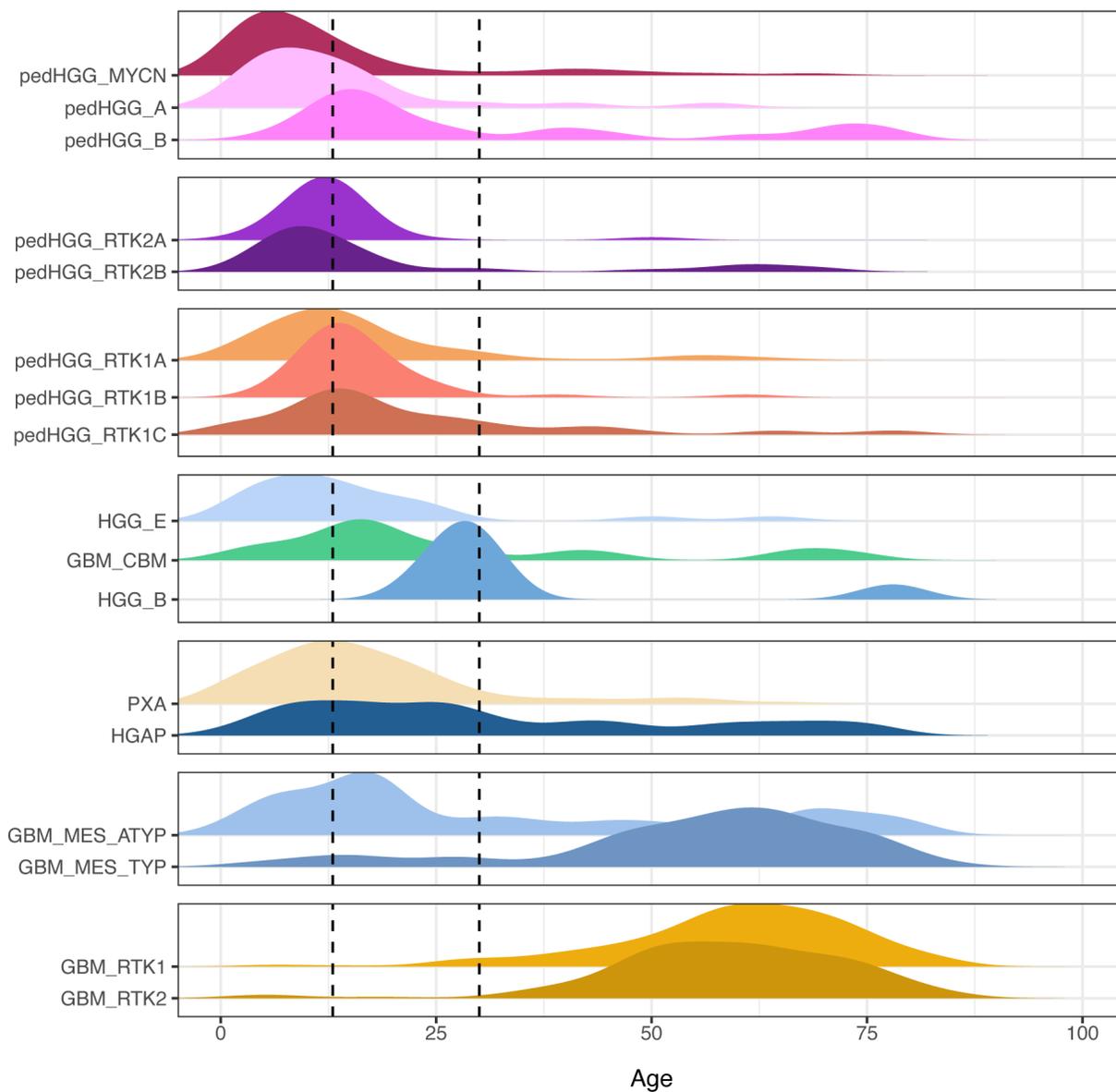lymphomas, glioneuronal,
neuroepithelial tumours

Methylation profiling

Illumina EPIC
MNP Classifier v12
CNS Reference Cohort
(n=8,488)

Methylation profiling

Illumina EPIC
t-SNE nearest neighbours
Glioma Reference Cohort
(n=3,357)

Sequencing

Whole exome / genome
Fusion panel (nucleic acid)
DNA copy number

Pathological assessment

Histological review
Immunohistochemistry

B

Frontal lobe
Basal ganglia
Parietal lobe
Frontal lobe
Parietal lobe
Thalamus
Pons
Occipital lobe
Temporal lobe
Cerebellum
Brainstem
Occipital lobe
Spinal cord

C

NOS
pedHGG_A
pedHGG_B
pedHGG_MYCN
pedHGG_RTK1A
pedHGG_RTK1B
PXA
pedHGG_RTK1C
HGAP
GBM_CBM
HGG_E
HGG_B
GBM_RTK2
GBM_RTK1
GBM_MES_TYP
GBM_MES_ATYP
pedHGG_RTK2B
pedHGG_RTK2A

Figure 2

Figure 3

A

Amp | Gain | NC | Loss | Del
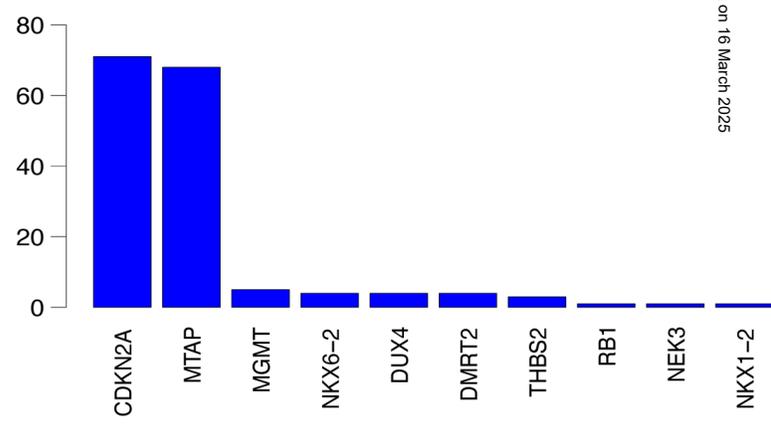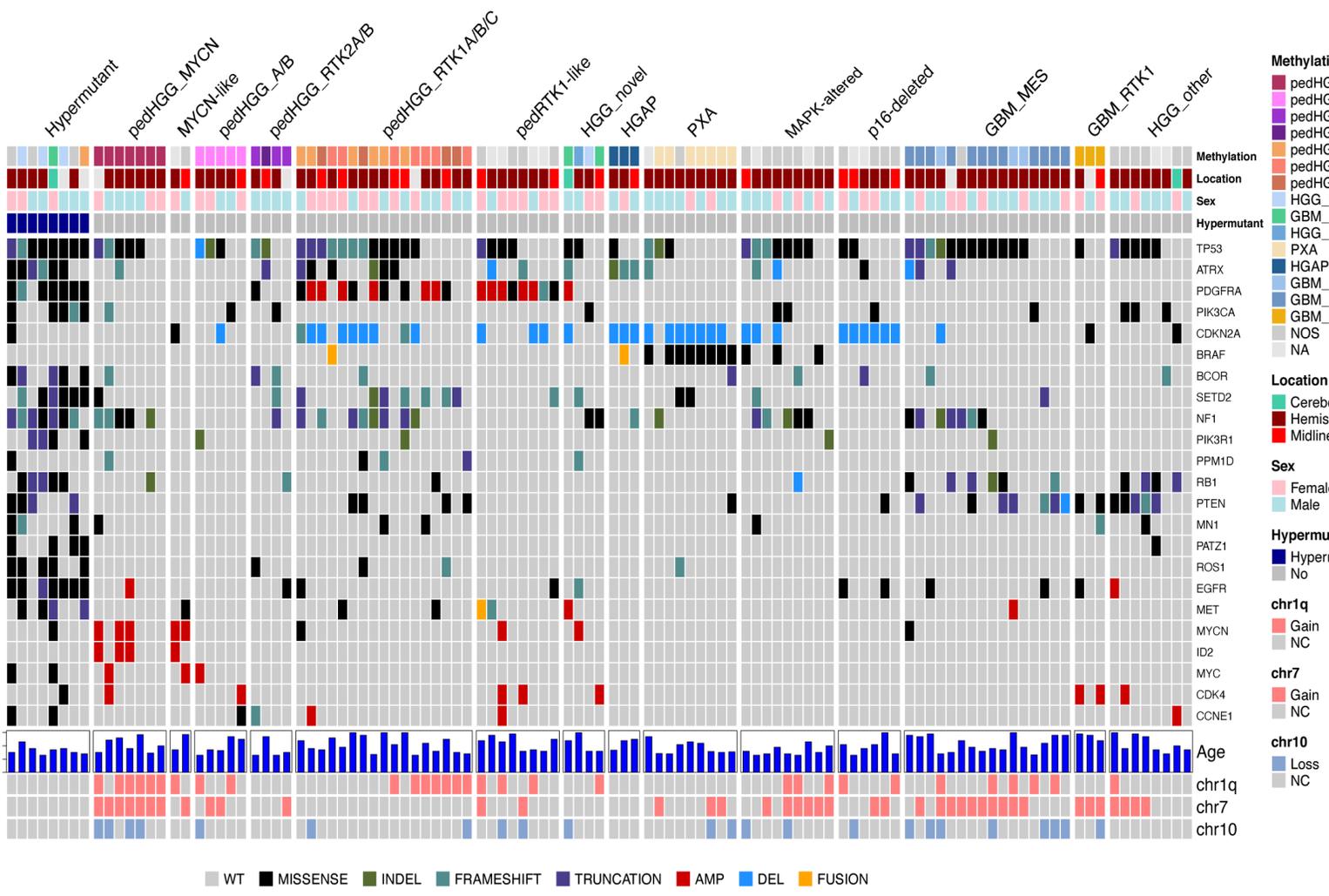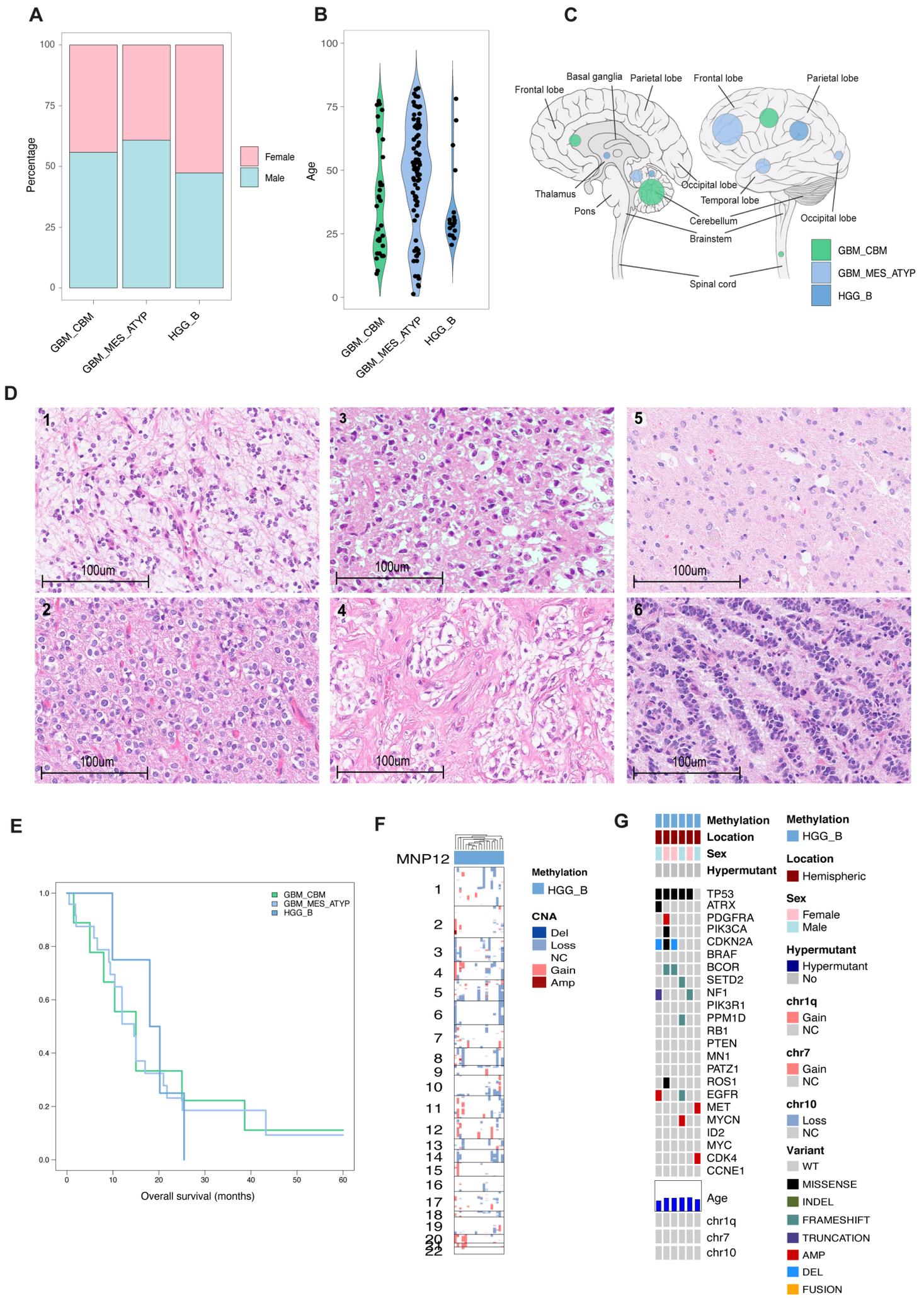
B

C

Figure 4

Figure 5

Figure 6

A

Glio_0021

Glio_0029

B

HGG_B

GBM_CBM

pedHGG_RTK1C

pedHGG_RTK1B

pedHGG_RTK1A

Glio_0021

Glio_0029

HGG_E

TSNE2

TSNE1

C

Methylation
Location
Sex
Hypermutant

MSH2
MSH6
MLH1
POLD1
POLE

TP53
ATRX
PDGFRA
PIK3CA
CDKN2A
BRAF
BCOR
SETD2
NF1
PIK3R1
PPM1D
RB1
PTEN
MN1
PATZ1
ROS1
EGFR
MET
MYCN
ID2
MYC
CDK4
CCNE1

Age

Glio_0029 Glio_0021

Methylation
pedHGG_RTK1A
HGG_E

Location
NA

Sex
Male

Hypermutant
Hypermutant

WT
MISSENSE
INDEL
FRAMESHIFT
TRUNCATION
AMP
DEL
FUSION

D

Glio_0021

Glio_0029

E

Glio_0226 ♂
pedHGG_RTK1B

ALL diagnosis
Chemotherapy

Relapse ALL

Relapse ALL
Bone marrow transplant

Diagnosis HGG - GTR
Radiotherapy
TMZ
Deceased

0   3yrs   7yrs   12yrs   20yrs 21yrs

Glio_0219 ♂
pedHGG_RTK1B

ALL diagnosis
Chemotherapy
Whole-body radiotherapy

Diagnosis HGG
Lobectomy (GTR)
TMZ & Radiotherapy
Deceased

0   12yrs   22yrs 24yrs

Glio_0231 ♂
pedHGG_RTK1B

Medulloblastoma
Surgery - GTR
Chemotherapy
Radiotherapy

Diagnosis HGG
Debulking
TMZ & Radiotherapy
Deceased

0   5yrs   19yrs 20yrs

Glio_0218 ♀
NOS

Medulloblastoma
Surgery (GTR)
Chemotherapy
Radiotherapy

Diagnosis HGG
Surgery (GTR)
TMZ
Deceased

0   9yrs   20yrs 22yrs

Glio_0228 ♀
HGG_E

Diagnosis tectal plate PA
Debulking
Radiotherapy

Progression
Diagnosis HGG (Bx)
TMZ & Radiotherapy
Progression (x2)
Deceased

0   10yrs   21yrs 22yrs

F

Survival probability

pedHGG_RTK1A
pedHGG_RTK1B
HGG_E

OS_months (months)