

## ARTICLE

https://doi.org/10.1038/s41467-021-25708-y

OPEN



# Radiation-induced gliomas represent H3-/IDH-wild type pediatric gliomas with recurrent *PDGFRA* amplification and loss of *CDKN2A/B*

Maximilian Y. Deng <sup>1,2,18</sup>, Dominik Sturm <sup>1,2,3,18</sup>, Elke Pfaff<sup>1,2,3,18</sup>, Martin Sill<sup>1,4</sup>, Damian Stichel<sup>5,6</sup>, Gnana Prakash Balasubramanian<sup>1,4</sup>, Stephan Tippelt<sup>7</sup>, Christof Kramm<sup>8</sup>, Andrew M. Donson<sup>9</sup>, Adam L. Green <sup>9</sup>, Chris Jones<sup>10</sup>, Jens Schittenhelm <sup>11</sup>, Martin Ebinger <sup>12</sup>, Martin U. Schuhmann<sup>13</sup>, Barbara C. Jones<sup>1,2,3</sup>, Cornelis M. van Tilburg<sup>1,3,14</sup>, Andrea Wittmann<sup>1,2</sup>, Andrey Golanov<sup>15</sup>, Marina Ryzhova<sup>15</sup>, Jonas Ecker<sup>1,3,14</sup>, Till Milde <sup>1,3,14</sup>, Olaf Witt<sup>1,3,14</sup>, Felix Sahm <sup>1,5,6</sup>, David Reuss<sup>5,6</sup>, David Sumerauer<sup>16</sup>, Josef Zamecnik<sup>17</sup>, Andrey Korshunov<sup>5,6</sup>, Andreas von Deimling <sup>5,6</sup>, Stefan M. Pfister<sup>1,3,4,19</sup> & David T. W. Jones <sup>1,2,19</sup>

Long-term complications such as radiation-induced second malignancies occur in a subset of patients following radiation-therapy, particularly relevant in pediatric patients due to the long follow-up period in case of survival. Radiation-induced gliomas (RIGs) have been reported in patients after treatment with cranial irradiation for various primary malignancies such as acute lymphoblastic leukemia (ALL) and medulloblastoma (MB). We perform comprehensive (epi-) genetic and expression profiling of RIGs arising after cranial irradiation for MB (n = 23) and ALL (n = 9). Our study reveals a unifying molecular signature for the majority of RIGs, with recurrent *PDGFRA* amplification and loss of *CDKN2A/B* and an absence of somatic hotspot mutations in genes encoding histone 3 variants or *IDH1/2*, uncovering diagnostic markers and potentially actionable targets.

<sup>&</sup>lt;sup>1</sup>Hopp Children's Cancer Center Heidelberg (KiTZ), Heidelberg University Hospital and German Cancer Resarch Center (DKFZ), Heidelberg, Germany. <sup>2</sup> Division of Pediatric Glioma Research, German Cancer Research Center (DKFZ), Heidelberg, Germany, <sup>3</sup> Department of Pediatric Oncology, Hematology and Immunology, Heidelberg University Hospital, Heidelberg, Germany.<sup>4</sup> Division of Pediatric Neurooncology, German Cancer Research Center (DKFZ) and German Consortium for Translational Cancer Research (DKTK), Heidelberg, Germany. <sup>5</sup> Department of Neuropathology, Institute of Pathology, Heidelberg University Hospital, Heidelberg, Germany. <sup>6</sup> Clinical Cooperation Unit Neuropathology, German Cancer Research Center (DKFZ) and German Consortium for Translational Cancer Research (DKTK), Heidelberg, Germany. <sup>7</sup> Department of Pediatric Oncology and Hematology, Essen University Hospital, Essen, Germany. <sup>8</sup> Division of Pediatric Hematology and Oncology, University Medical Center Goettingen, Goettingen, Germany. <sup>9</sup> Morgan Adams Foundation Pediatric Brain Tumor Research Program, Department of Pediatrics, University of Colorado School of Medicine, Aurora, CO, USA. <sup>10</sup> Division of Molecular Pathology and Division of Cancer Therapeutics, The Institute of Cancer Research, London, United Kingdom. <sup>11</sup> Department of Neuropathology, Institute of Pathology and Neuropathology and Comprehensive Cancer Center Tübingen-Stuttgart, Tübingen University Hospital, Tübingen, Germany. <sup>12</sup> Department of Pediatric Hematology/Oncology, Children's University Hospital, Tübingen, Germany. <sup>13</sup> Department of Neurosurgery, Division of Pediatric Neurosurgery, Tübingen University Hospital, Tübingen, Germany.<sup>14</sup> Clinical Cooperation Unit Pediatric Oncology, German Cancer Research Center (DKFZ) and German Consortium for Translational Cancer Research (DKTK), Heidelberg, Germany.<sup>15</sup> Department of Neuropathology, NN Burdenko Neurosurgical Institute, Moscow, Russia. <sup>16</sup> Department of Pediatric Hematology and Oncology, Motol University Hospital, Charles University, Prague, Czech Republic. <sup>17</sup> Department of Pathology, Motol University Hospital, Charles University, Prague, Czech Republic. <sup>18</sup>These authors contributed equally: Maximilian Y. Deng, Dominik Sturm, Elke Pfaff. <sup>19</sup>These authors jointly supervised this work: Stefan M. Pfister, David T. W. Jones. <sup>22</sup>email: david.jones@kitz-heidelberg.de

R adiation therapy (RT) constitutes an essential element in the standard treatment of many cancers, improving clinical outcomes of patients with childhood malignancies including leukemia and central nervous system (CNS) tumors<sup>1</sup>. However, radiation-induced malignancies are observed post-RT in a subset of patients<sup>2,3</sup>, especially in pediatric patients due to their prolonged follow-up in case of long-term survival<sup>4,5</sup>.

Radiation-induced gliomas (RIGs) are known to arise in a subset of patients receiving cranial RT for the treatment of different primary malignancies including acute lymphoblastic leukemia (ALL) and medulloblastoma (MB), occurring after a variable latency ranging from 2.5 to 35 years after irradiation (Table 1)<sup>6–18</sup>.

RIGs commonly display an aggressive clinical course associated with poor prognosis, making early and precise diagnosis (including distinguishing RIGs from recurrences of the primary tumor) crucial for optimal treatment planning<sup>7,8</sup>. This is particularly challenging in cases where putative recurrent CNS tumors are not biopsied or resected, and where material for histological or molecular confirmation is therefore lacking.

Histopathologically, most RIGs present as high-grade gliomas (HGG) reminiscent of their sporadic counterparts<sup>7,8,13,14,19</sup>, and histopathological features to distinguishing them from de-novo HGG are missing<sup>6,7</sup>. A report of morphological features of both MB and HGG within secondary tumors led to the hypothesis of transformation from MB (stem-)cells as a possible route for RIG development<sup>11</sup>, while the higher genetic homogeneity of RIGs compared to de-novo pediatric HGG is suggestive of a common origin for RIGs<sup>7</sup>. Genetic alterations including gain of chromosome arm 1q, loss of CDKN2A/B, and PDGFRA amplification or overexpression have been frequently observed in RIGs13,20-22, and-in contrast to sporadic pediatric HGG-recurrent somatic hotspot mutations in genes encoding histone 3 variants or IDH1/2 were found to be uncommon<sup>13,20–22</sup>. To date, however, most series have been relatively small, and there remains some lack of clarity about the origin of RIGs and the underlying molecular and biological mechanisms leading to their formation.

Recent studies have shown that DNA methylation profiling represents a robust and reproducible approach to classify CNS tumors into clinically meaningful entities<sup>23–26</sup>. Thereby, pediatric HGG represent a heterogeneous group of CNS tumors, clearly distinguishable from their adult counterparts<sup>21,27–29</sup>, that can be classified into distinct subgroups with characteristic genetic and epigenetic alterations and clinical associations<sup>30–33</sup>. Next-generation sequencing and transcriptomic profiling approaches represent complementary tools for molecular characterization of tumorigenic processes.

Here, we performed comprehensive molecular characterization of RIGs to detect distinctive, (epi-)genetic features which might predict or explain their formation after RT, and/or which may act as diagnostic or therapeutic markers.

## Results

**Clinical characteristics.** Patients in our series developing a radiation-induced glioma (RIG, n = 32) were previously treated with cranial irradiation for MB (MB-RIG, n = 23) or ALL (ALL-RIG, n = 9). A variety of different RT regimens was administered, including craniospinal radiotherapy in MB patients or prophylactic cranial irradiation in ALL patients. In line with a higher probability of RIG occurrence in areas receiving a higher effective radiation dose, most MB-RIG tumors were encountered in the cerebellum (20/23, 87% vs. 3/23 in the cerebral hemisphere), the location with the highest dose in RT. ALL-RIG tumors predominantly appeared in the cerebral hemisphere (7/9, 78%, vs 3/23 in MB-RIGs, p < 0.01; Fig.1a), with only two cases arising in

the posterior fossa. The gender ratio was slightly shifted towards males (male:female ratio: 1.28:1) (Fig. 1b). The latency for RIG occurrence ranged from 2 to 30 years in MB patients (n = 23, median: 5 years) and from 3–17 years in ALL patients (n = 9, median: 8 years; Fig. 1c).

For 16/23 MB-RIG and 5/9 ALL-RIG patients with available follow-up data, we observed an aggressive clinical course that was as bad as or even slightly worse than histone 3 K27M-mutant tumors – with 13 out of 16 MB-RIG and four out of five ALL-RIG patients showing fatal outcome during the follow-up period (OS: 6 months, range: 3–10 months vs 8.5 months, range: 2–18 months in MB-RIG patients; Suppl. Table 1). One ALL-RIG (RIG\_09, follow-up period: 3 months) and 3 MB-RIG patients diagnosed at the age of 14 year (RIG\_22, follow-up period: 9 months), 15 years (RIG\_30, follow-up period: 3 months), and 39 years (RIG\_14, follow-up period: 18 months) were known to be alive at last follow-up, with the latter showing signs of progression after 12 months (Fig. 2).

**DNA methylation profiling of radiation-induced gliomas and primary medulloblastomas.** DNA methylation patterns of both MB- and ALL-RIGs were analyzed by unsupervised clustering and t-SNE alongside 120 reference samples from pediatric HGG subgroups with known molecular features. The majority of RIGs after both primary indications closely resembled those of pedGBM\_RTK1 tumors (29/32, 91%; Fig. 3a, b)<sup>32</sup>. Two MB-RIG and one ALL-RIG showed a methylation pattern more similar to pleomorphic xanthoastrocytomas (PXA) (Fig. 3b). Thus, RIG DNA methylation patterns suggest a high degree of similarity (and likely common origins) between tumors, and were clearly distinct from any other HGG DNA methylation class.

The 11 pre-RIG primary MBs for which DNA was available were also profiled, and found to represent all four major MB subgroups (MB-WNT: 1/11; MB-SHH: 3/11; MB-Group 3: 2/11, MB-Group 4: 5/11), with no obvious enrichment for a particular primary subtype relative to their overall incidence rate. DNA methylation and CNV profiling confirmed the distinct biological nature between the primary MB and secondary RIG using t-SNE and copy-number analyses (Fig. 4a, b). Histological staining for hematoxylin and eosin (Fig. 4c) of two paired primary medulloblastoma and radiation-induced glioma couples (MB/RIG\_28, MB/RIG\_29) display distinct morphological characteristics.

RIGs display recurrent PDGFRA amplification and loss of CDKN2A/B. Amplification of PDGFRA (6/9 ALL-RIG; 11/23 MB-RIG) and loss of CDKN2A/B (4/9 ALL-RIG; 17/23 MB-RIG) represented the most common copy-number alterations (CNAs) in RIGs, with 2/8 ALL-RIG and 8/23 MB-RIG patients exhibiting cooccurrence of both (Fig. 5). This is in keeping with the methylation profile of these tumors matching to the pedGBM\_RTK1 group, where both changes are also well-known characteristic features. Indeed, the overall frequency of PDGFRA amplification in the RIG cohort (17/32, 55%) is higher than in an unselected pedHGG cohort, but is not significantly higher than that reported in the pedGBM\_RTK1 subgroup (33%, p = 0.14)<sup>32</sup>. Amplification of MET (9/32, 28%; previously also linked to MET fusion in pediatric GBMs<sup>34</sup>) and CDK4 (5/32, 16%) were frequently encountered in our cohort. Additional broader chromosomal CNAs included gain of chromosome arm 1q (16/32, 50%), 1p deletion (19/32, 59%), 6q deletion (18/32, 56%), 13q deletion (23/32, 72%), and 14q deletion (16/32, 45%) (Fig. 5). A more focused analysis was performed to illustrate the potential discrepancies between RIGs (n = 32) and the non-RIG pedGBM\_RTK1 tumors (n = 20) regarding their respective characteristic copy-number changes, with both groups

Reference	Number of cases	Primary malignancy	Latency period (years)	Age at RIG diagnosis (years)	Gender ratio (m:f)	Death (y/ number of cases)	OS (months)	PDGFRA status	CDKN2A/ B status	Further genetic alterations
Brat at al 6	9	ALL, Pineal tumors, Lymphoblastic Lymphoma, Pituary	5-23	9-60	2:1	N.D.	N.D.	N.D.	N.D.	1
Donson	ъ	actioner, viracooniyosa kone, cramphiary iscona Burkitt's Lymphoma, MB, low-grade astrocytoma, energytoma All	3-15	11-23	3:2	5/5	1-10	Overexpressed	N.D.	ı
	12	Craniopharyngioma, Germinoma, Medulloblastoma, Pisconiopharyngioma, Germinoma, Medulloblastoma,	4-41	7-48	8:4	5/7	<24	Amplified in 42%,	Loss in 33%	Amplified CDK4 in 33%,
et al. Nakao	4	rineocytorina, ALL, HOOGKIII'S LYINPIONA Pituary adenoma, Meduloblastoma, Cominano Mitt	22-29	32-55	3:1	4/4	11-30	M.D.	N.D.	IDH1 and H3F3A wild type
Paugh	10	Crainopriaryngionia, rivel ALL, Germinoma, MB, Ependymoma	N.A.	8-19	N.A.	6/2	8-91	Amplified in 50%	Loss in 50%	1q gain (50%), 1p loss (50%),
et al. <sup></sup> Phi	IJ	Medulloblastoma (SHH, Group3)	4.3-10	9.2-17	2:3	N.D.	N.D	Missense mutation,	N.D.	TP53 (somatic mutations or
et al.	7	ALL, MB	7-14	9-19	4:3	L/T	9-27	gene iusion N.D.	N.D	17P 10SS), 7Q gain (EZMZ) -
et al. <sup>15</sup> Walter et al. <sup>18</sup>	7	ALL	7-14	10-24	4:3	5/7	0.1-93	N.D.	N.D.	1
OS overall st	rvival, N.D. nc	ot determined.								
US OVERAIL S.	Irvival, N.U. no	ot determined.								

exhibiting a similar pattern of chromosomal gains and losses (Suppl. Fig. 1a, c). The number of chromosomal breakpoints in RIGs and non-RIG pedGBM\_RTK1 tumors were further visually determined to estimate the extent of genomic rearrangement caused by double-strand DNA breaks. Overall, RIGs display a significantly wider range and a higher number of chromosomal breakpoints (median: 38.5, range: 4–134) compared to non-RIG pedGBM\_RTK1 tumors (median: 23, range: 13–51; two-sided *t*-test: p < 0.04) (Suppl. Fig. 1 d).

Genomic profile of radiation-induced gliomas. RNA sequencing (n = 9), gene panel (n = 5), and whole-exome (n = 13) sequencing with matched blood samples were performed for a subset of RIG samples. None of the RIGs harbored somatic hotspot mutations in genes encoding histone 3 variants, *IDH1/2*, *BRAF*, or in the *TERT* promoter region. Four out of 18 tested samples (22%) exhibited somatic mutations in the *TP53* gene. Additional somatic mutations included *CBL*, *PDGFRA*, *NTRK2*, *EGFR*, *RAF1*, *ATRX*, and *BCOR* (Suppl. Table 1). One patient (RIG\_25) harbored a *TP53* germline splice site mutation. One patient (RIG\_14) was clinically diagnosed with Li-Fraumeni syndrome, but tumor material was insufficient for molecular testing. Further germline alterations associated with cancer predisposition syndromes were absent in the remainder of cases with germline material available (n = 13).

Analysis of RNA sequencing data (n = 9) revealed relevant gene fusions including PTPRZ1:MET (RIG\_11, RIG\_25), CAP-ZA2:MET (RIG\_09), FYCO1:RAF1 (RIG\_08), and GFAP1:NTRK2 (RIG\_13) (Fig. 6). Interestingly, the FYCO1:RAF1 and GKAP1:NTRK2 fusions were detected in two of the three tumors from the PXA-like subgroup based on DNA methylation profiling (Fig. 3). Both rearrangements are predicted to lead to a constitutive activation of the kinase domain, via loss of the N-terminal regulatory domain (RAF1) or constitutive dimerization (NTRK2). Deriving from a complex structural rearrangement, PRPRZ1:MET fusion proteins harbor an almost full length MET protein driven by the highly active PTPRZ1 promoter. All MET-fused cases (RIG\_09, RIG\_11, RIG\_25) demonstrated amplification of MET on copy-number profiles derived from DNA methylation arrays. While all four genetic rearrangements, which result in aberrant MAPK/ERK pathway activation, were reported in high-grade gliomas, none of them were previously discovered in RIGs34-36.

Transcriptional profiling of radiation-induced gliomas. Arraybased gene expression analysis was performed on seven RIG samples as well as 24 additional tumors representing the H3/ IDH-wildtype pediatric HGG pedGBM\_RTK1 (n = 5), pedGBM\_RTK2 (n = 3), pediatric HGG MYCN (n = 8)<sup>32</sup>, and PXA (n = 8) subgroups. Unsupervised hierarchical clustering analysis based on the 100 most differentially expressed genes recapitulated the methylation-defined subgrouping, with the largest fraction of the RIGs forming a homogenous expression pattern, resembling pedGBM\_RTK1 tumors. Two cases (RIG\_08, RIG\_13), previously categorized into the PXA methylation group, also displayed an expression profile similar to PXAs (Fig. 7). The level of expression glioma-characteristic genes (e.g. CD34, GFAP, OLIG2, MAP2, MKI67, RBFOX3, SOX2, SYN; Suppl. Fig. 2) between RIGs and sporadic H3/IDH-wildtype pediatric HGGs did not reveal any notable differences, supporting the conclusion of a general similarity in their pathway activation.

### Discussion

RIGs represent a fatal long-term side effect of cranial irradiation $^{2,3,19,37-43}$ . A better understanding of the occurrence of







MB-RIG (23)

Cerebral hemisphere

c. Latency period between irradiation and RIG occurrence



**Fig. 1 Clinical patient characteristics.** Tumor location for 23 MB-RIG and 9 ALL-RIG illustrated with circled numbers indicating the number of patients with RIG occurrence in the respective brain region (**a**). Significant differences in tumor location were observed between MB-RIG and ALL-RIG patients (p < 0.01). Gender distributions, age at RIG diagnosis (**b**) and latency period (**c**) between cranial irradiation and RIG diagnosis are shown with numbers in brackets indicating group size. MB-RIG patients displayed a wider distribution regarding the age at RIG diagnosis and latency period. RIG radiation-induced glioma, MB-RIG post-medulloblastoma RIG, ALL-RIG post-ALL RIG.

these tumors and their molecular background is therefore of important clinical relevance. Our study supports the findings that RIGs harbor recurrent genetic alterations converging on an aberrant activation of the MAPK/ERK pathway (particularly via PDGFRA) together with loss of cell cycle control, facilitating tumorigenesis<sup>7,20,35,36</sup>. Importantly, we show that RIGs harbor a largely homogenous genetic and epigenetic profile, closely resembling sporadic pediatric GBM RTK1 tumors<sup>32</sup>. This suggests that both radiation-induced and sporadic pedGBM RTK1 tumors might share a common cell of origin, which could be particularly vulnerable to ionizing radiation. However, the outcome of RIG patients is particularly poor, and worse than what has previously been described for sporadically occurring PDGFRA-amplified, 9p21-deleted pediatric GBM RTK1 tumors<sup>32</sup>. The discrepancy could suggest that additional oncogenic mechanisms might play a role during RIG formation, e.g. the trend towards greater genomic instability in the RIG cohort, as indicated by the elevated frequency of chromosomal breakpoints compared to their sporadic counterparts, potentially caused by radiation-induced DNA double-strand breaks. On the other hand, the RIG patient cohort also typically represents children with a long history of often aggressive prior treatment for their primary malignancy, which means that certain treatment options typically available at first diagnosis of GBM (e.g. radiotherapy) may no longer be an option for RIG patients<sup>44,45</sup>. It is also anticipated that the general health status of RIG patients is likely to be worse than that of primary-diagnosis counterparts.

Further, some RIGs arising after MB were treated with an MB relapse protocol on the assumption that the new lesion was a recurrence of the primary tumor in the absence of biopsy confirmation, posing a possible risk that these protocols may be less efficacious against high-grade glioma than upfront GBM protocols. Thus, there are a number of key clinical reasons why an inferior survival might be expected, in addition to molecular mechanistic differences.

A small subset of RIGs with a more PXA-like profile harbored potentially druggable alterations in *RAF1* and *NTRK2*, and it will be of interest to investigate the precise characteristics of this group in larger cohorts. Our findings also do not at present suggest a substantial contribution of germline alterations in known tumor-associated genes in the occurrence of RIGs. However, further comprehensive germline analyses with matched blood samples are required to elucidate in full the potential role of hereditary predisposition in RIG formation.

RIG following MB of various subtypes as well as after ALL displayed converging fingerprints, despite the diversity of applied radiation protocols (MB-RIGs arose mostly at the site of local boost, while ALL-RIGs occurred throughout the cranial radiation field). All of the four main molecular MB subgroups were represented in the primary lesions, with no recurrent alterations observed in the primary MBs. Our findings unfortunately suggest that no conclusions can currently be drawn regarding predictive features in the patient or primary tumor for assessing the probability of subsequent RIG formation. An important question to



Fig. 2 Kaplan-Meier curves illustrating overall survival for RIGs with representative CNS tumor entities. PXA pleiomorphic xanthoastrocytoma, pedGBM\_RTK1 pediatric glioblastoma subclass RTK1, pedGBM\_RTK2 pediatric glioblastoma subclass RTK2, pedGBM\_MYCN pediatric glioblastoma subclass MYCN, DMG K27 diffuse midline glioma H3 K27 mutant, GBM G34 glioblastoma, H3.3 G34 mutant.

address in future will be the precise role of radiotherapy protocols (craniospinal vs local boost, fractionation schema, photons/protons/heavy ions) in determining the risk of secondary malignancy<sup>46-50</sup>.

The ALL-RIGs included in our series were all diagnosed with their second tumor between 9 and 18 years of age, suggesting a particularly vulnerable but relatively narrow time frame for RIG occurrence after ALL. If confirmed in larger series, this could be an important variable for the planning of surveillance in this patient population. In contrast, both the age at diagnosis and latency after the primary tumor were much more varied for medulloblastoma. The earliest RIG occurred after a latency period of only 2 years following treatment for MB - within a timeframe which could easily allow misidentification as a recurrence of the primary tumor. Molecular characterization of 31 clinicallypresumed relapse MBs within the prospective INFORM study<sup>51</sup> revealed that 6/31 'relapsed' MBs (19%) were in fact secondary HGG (unpublished data). In light of their aggressive clinical course and fundamentally different biology, it is essential to distinguish newly-arising RIGs from true recurrences in order to adjust treatment planning as early as possible to increase treatment efficacy. Thus, biopsy of recurrent lesions after MB is clearly warranted and should be considered whenever surgically feasible. Furthermore, the observation that a subset of prognostically favorable MBs (e.g. the MB-WNT subgroup) are also at risk of developing RIGs, highlights the urgency of current efforts to

further examine the application of alternative treatment modalities (e.g. proton/heavy ion beam therapy) or the general deintensification of radiation therapy in these patients<sup>47,52–58</sup>.

In summary, our findings demonstrate that RIGs are an aggressive, relatively homogenous group of CNS tumors with recurrent amplification of *PDGFRA* and loss of *CDKN2A/B* in the absence of somatic H3/IDH hotspot mutations. Our study uncovers possible similarities in origins with the pedGBM\_RTK1 group of sporadic HGG, which will be important for further understanding the mechanisms by which these secondary tumors arise.

#### Methods

**Patient population and tumor samples**. Criteria for patient selection adapted from Cahan et al.<sup>59,60</sup> included (1) a glioma emerging from the previously irradiated field, and (2) a non-glial primary malignancy, in this series MB (n = 23) and ALL (n = 9), to exclude potential tumor progression/recurrence not directly related to radiotherapy<sup>61–63</sup>.

This study is covered by the ethical approval of the Heidelberg University's medical faculty. Additional clinical data and tumor material were collected by international collaborating centers according to local ethical and institutional review board approval and collected at the German Cancer Research Center (DKFZ, Heidelberg, Germany). Informed consent for molecular profiling was obtained from all patients and their legal representatives.

**DNA methylation profiling**. The Illumina Infinium HumanMethylation450 (450k) array and Illumina Infinium MethylationEPIC (EPIC) array were used to obtain genome-wide DNA methylation profiles, according to the manufacturer's



**Fig. 3 Molecular classification of RIGs by DNA methylation profiling.** Unsupervised analysis of 32 RIG tumors was conducted based on the 10,000 most variably methylated probes. Through a t-SNE analysis, the majority of RIG tumors form a homogenous group of tumors resembling the pediatric GBM RTK1 group. ALL-RIG and MB-RIG samples were compared with 120 well-characterized reference samples representing CNS tumors of known histological and/ or molecular subtype. PXA pleiomorphic xanthoastrocytoma, pedGBM\_RTK1 pediatric glioblastoma subclass RTK1, pedGBM\_RTK2 pediatric glioblastoma subclass RTK2, pedGBM\_MYCN pediatric glioblastoma subclass MYCN, DMG\_K27 diffuse midline glioma H3 K27 mutant, GBM\_G34 glioblastoma, H3.3 G34 mutant.



**Fig. 4 Comparison of matched primary MB and post-radiation glioma samples.** T-SNE analysis of primary MB and post-MB RIG compared with existing MB subgroups (**a**) and representative copy-number profile (**b**) reveal distinct genome-wide methylation patterns and recurrent high-level *PDGFRA* amplification and *CDKN2A/B* deletion. Available primary MB and post-MB RIG samples (*n* = 11 pairs) were investigated with 100 MB reference samples representing established MB subtypes<sup>57,70</sup>. MB and RIG pairs from our series were encircled and linked with dotted lines. Histological staining for hematoxylin and eosin (**c**) of two paired primary medulloblastoma and radiation-induced gliomas (MB/RIG\_28, MB/RIG\_29) display distinct morphological characteristics. Scale bar represents 200 μm. T-SNE t-stochastic neighboring embedding, MB medulloblastoma, MB-SHH-INF medulloblastoma sonic-hedgehog adult subgroup, CNP copy-number profile.

instructions (Illumina, San Diego, USA). Data were generated at the Genomics and Proteomics Core Facility of the DKFZ (Heidelberg, Germany) and St. Jude Children's Research Hospital (Memphis, USA). DNA methylation data was generated from both fresh-frozen and formalin-fixed paraffin-embedded (FFPE) tissue samples. On-chip quality metrics of all samples were carefully controlled. Copy-number variation (CNV) analysis from 450k and EPIC methylation array data was performed using the conumee Bioconductor package version 1.12.0. All computational analyses were performed in R version 3.4.4 (R Development Core Team, 2019). Raw signal intensities were obtained from IDAT-files using the minfi Bioconductor package version 1.24.0. Illumina EPIC and 450k samples were



**Fig. 5 Oncoplot illustrating the genomic landscape in RIGs.** Summary of recurrent, characteristic genetic alterations (focal copy-number alterations, mutations) identified through DNA methylation arrays and whole-exome/gene panel sequencing. *MGMT* O6-methylguanine-DNA methyltransferase. ALL acute lymphoblastic leukemia, MB medulloblastoma.

merged to a combined data set by selecting the intersection of probes present on both arrays (combineArrays function, minfi). Each sample was individually normalized by performing a background correction (shifting of the 5th percentile of negative control probe intensities to 0) and a dye-bias correction (scaling of the mean of normalization control probe intensities to 10,000) for both color channels. Subsequently, a correction for the type of material tissue (FFPE/frozen) and array (450k/EPIC) was performed by fitting univariate, linear models to the log2transformed intensity values (removeBatchEffect function, limma package version 3.34.5). The methylated and unmethylated signals were corrected individually. Beta-values were calculated from the retransformed intensities using an offset of 100 (as recommended by Illumina).

Before further analysis, the following filtering criteria were applied: Removal of probes targeting the X and Y chromosomes (n = 11,551), removal of probes containing a single-nucleotide polymorphism (dbSNP132 Common) within five base pairs of and including the targeted CpG-site (n = 7998), probes not mapping uniquely to the human reference genome (hg19) allowing for one mismatch (n = 3965), and 450k array probes not included on the EPIC array. In total, 428,230 probes were kept for downstream analysis<sup>64–66</sup>.

To perform unsupervised dimension reduction, the remaining probes were used to calculate the 1-variance weighted Pearson correlation between the samples. The resulting distance matrix was used as input for t-SNE analysis (t-Distributed Stochastic Neighbor Embedding; Rtsne package version 0.13). The following non-default parameters were applied: theta = 0, pca = F, max\_iter = 2500 perplexity = 20.

To perform unsupervised hierarchical clustering, the 10.000 probes with highest standard deviation were selected to calculate the Euclidean distance between samples, followed by applying Wards linkage method for sample clustering. In the heatmap, representation probes were reordered by complete linkage hierarchical clustering of the Euclidean distance between probes.

To evaluate focal amplifications and deletions and chromosomal gains and losses, we visually inspected copy-number profiles of each case. Candidate genes and their 3' and 5' intergenic neighborhood were further investigated using the Integrative Genomic Viewer (IGV) for the presence of breakpoints, as an indication for potential gene fusions.

**Gene expression profiling.** Tumor samples with sufficient high-quality RNA were analyzed on the Affymetrix GeneChip Human Genome U133 Plus (v.2.0) Array (Affymetrix, Santa Clara, USA) at the Microarray Department of the University of Amsterdam, the Netherlands or the Genomics ad Proteomics Core Facility of the German Cancer Research Center (DKFZ). Subsequent library preparation, hybridization, and quality control was conducted following the manufacturer's guideline. Expression data were normalized using the MAS5.0 algorithm. Identification of significantly differentially expressed genes between subgroups (test: anova; correction for multiple testing: false discovery rate, transform: *z*-score) and unsupervised hierarchical clustering were performed within the R2: Genomics Analysis and Visualization Platform (http://r2.amc.nl).

**Next-generation gene panel, whole-exome, low-coverage whole genome, and RNA sequencing.** For the detection of single-nucleotide variations (SNVs), small insertions and deletions (indels), and gene fusions, a subset of tumor samples (n = 5) were analyzed via a customized enrichment/hybrid-capture-based next-generation sequencing gene panel<sup>67</sup>. DNA from FFPE tissue was extracted using the Promega Maxwell device (Promega), sheared on Covaris M220 (Covaris), according to the manufacturer's guidelines. Following a successful quality control using a Bioanalyzer 2100 (Agilent), sequencing was applied to increase the detection sensitivity of duplicates and possible gene fusions.

For whole exome (WES; n = 13) and RNA sequencing (n = 9), library preparation, sequencing and data processing were conducted following the pipeline established in the INFORM study<sup>51</sup>. In brief, library preparation for WES was performed using the Agilent SureSelectXT Human V5 kit. Prepared libraries were sequenced together with a tumor cDNA library (poly(A) + RNA, Illumina TruSeq RNA Kit v2) on an Illumina HiSeq. The 1000 Genomes phase 2 human reference assembly (NCBI build 37.1) was selected for mapping the sequencing reads using BWA (version 0.6.2). For detection of SNVs and indels, custom pipelines were used<sup>68</sup>. In brief, the computational analysis was performed using SAMtools mpileup and bctools version 0.1.19 to detect somatic variants. To annotate variants, ANNOVAR was applied with Gencode version 17. All high confidence coding or splice site germline variants in a selected panel of cancer predisposition genes were extracted using a custom Python script. To call short insertions or deletions (InDels) in tumor and control blood BAM files, the extracted data was annotated using Platypus (version 0.5.2/0.7.4 with parameters genInDels = 1, genSNPs = 0, ploidy = 2, nIndividuals = 2). Poor genotype quality and low variant counts in the tumor were excluded for the subsequent analysis. RNA sequencing data were analyzed with deFuse to detect gene fusions<sup>69</sup>

Potentially relevant somatic and germline alterations were manually assessed and cross-examined through various databases (http://www.ncbi.nlm.nih.gov/SNP build 135, http://www.1000genomes.org, http://exac.broadinstitute.org, https:// cancer.sanger.ac.uk/cosmic, https://www.ncbi.nlm.nih.gov/clinvar/).

**Statistics**. Overall survival (OS) and progression-free survival (PFS) were analyzed by Kaplan–Meier analysis and tested for significant differences using a log-rank test. Binary and categorical patient characteristics between different subgroups were compared by a two-sided Fisher's exact test. *P*-values < 0.05 were considered significant.



**Fig. 6 Alternative genetic rearrangements leading to aberrant MAPK/ERK pathway signaling.** Partial deletions of chromosomal arm 3p induce formation of a *FYCO1:RAF1* fusion (**a**), with occurrence of a *GKAP1:NTRK2* fusion (**d**) emerging through inversion of a chromosomal arm 9q segment. Both rearrangements are predicted to lead to constitutive activation of the kinase domain (KD). Induced by complex rearrangements, *CAPZA2:MET* (**b**) and *PTPRZ1:MET* fusion (**c**, **e**) contained almost the full length MET protein, with overexpression driven by the highly active *CAPZA2* or *PTPRZ1* promoter.



Fig. 7 Heatmap representing the expression levels of the 100 most significant, differentially expressed genes comparing RIG samples with reference groups (PXA, pedGBM\_RTK1, pedGBM\_RTK2, and pedGBM\_MYCN). Unsupervised hierarchical clustering reveals that gene expression profiling reflects our aforementioned findings based on DNA methylation profiling, with most RIGs resembling the pedGBM\_RTK1 subgroup and a small subset more similar to PXAs. Each column represents one sample and each lane represents one gene. Expression levels are indicated by a color scale as indicated. PXA pleiomorphic xanthoastrocytoma, pedGBM\_RTK1 pediatric glioblastoma subclass RTK1, pedGBM\_RTK2 pediatric glioblastoma subclass RTK2, pedGBM\_MYCN pediatric glioblastoma subclass MYCN.

# ARTICLE

**Reporting summary**. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

#### Data availability

Methylation array data from this study have been deposited in GEO: GSE131482. Gene expression array were deposited in GEO: GSE168457. DNA sequencing data are available from the European Genome-Phenome Archive (EGA) under accession EGAS00001005243. Source data are provided with this paper.

Received: 22 May 2020; Accepted: 12 August 2021; Published online: 20 September 2021

#### References

- Taylor, R. E. et al. Results of a randomized study of preradiation chemotherapy versus radiotherapy alone for nonmetastatic medulloblastoma: The International Society of Paediatric Oncology/United Kingdom Children's Cancer Study Group PNET-3 Study. J. Clin. Oncol. 21, 1581–91 (2003).
- Armstrong, G. T. et al. Reduction in late mortality among 5-year survivors of childhood cancer. New Engl. J. Med. 374, 833-42 (2016).
- 3. Travis, L. B. et al. Second malignant neoplasms and cardiovascular disease following radiotherapy. *J. Natl Cancer Inst.* **104**, 357–70 (2012).
- Christopherson, K. M. et al. Late toxicity following craniospinal radiation for early-stage medulloblastoma. *Acta Oncol.* 53, 471–80 (2014).
- Relling, M. V. et al. High incidence of secondary brain tumours after radiotherapy and antimetabolites. *Lancet* 354, 34–9 (1999).
- Brat, D. J. et al. Molecular genetic alterations in radiation-induced astrocytomas. Am. J. Pathol. 154, 1431–8 (1999).
- Donson, A. M. et al. Unique molecular characteristics of radiation-induced glioblastoma. J. Neuropathol. Exp. Neurol. 66, 740–9 (2007).
- Elsamadicy, A. A., Babu, R., Kirkpatrick, J. P. & Adamson, D. C. Radiationinduced malignant gliomas: a current review. *World Neurosurg.* 83, 530–42 (2015).
- Gessi, M. et al. Radiation-induced glioblastoma in a medulloblastoma patient: a case report with molecular features. *Neuropathology* 28, 633–9 (2008).
- Hamasaki, K., Nakamura, H., Ueda, Y., Makino, K. & Kuratsu, J. Radiationinduced glioblastoma occurring 35 years after radiation therapy for medulloblastoma: case report. *Brain Tumor Pathol.* 27, 39–43 (2010).
- Madden, J. R. et al. Radiation-induced glioblastoma multiforme in children treated for medulloblastoma with characteristics of both medulloblastoma and glioblastoma multiforme. J. Pediatr. Hematol. Oncol. 32, e272–8 (2010).
- Menon, G., Nair, S., Rajesh, B. J., Rao, B. R. & Radhakrishnan, V. V. Malignant astrocytoma following radiotherapy for craniopharyngioma. *J. Cancer Res. Ther.* 3, 50–2 (2007).
- Nakao, T. et al. Radiation-induced gliomas: a report of four cases and analysis of molecular biomarkers. *Brain Tumor Pathol.* 34, 149–154 (2017).
- Ng, I., Tan, C. L., Yeo, T. T. & Vellayappan, B. Rapidly fatal radiation-induced glioblastoma. *Cureus* 9, e1336 (2017).
- Prasad, G. & Haas-Kogan, D. A. Radiation-induced gliomas. *Expert Rev.* Neurother. 9, 1511–7 (2009).
- Romeike, B. F., Kim, Y. J., Steudel, W. I. & Graf, N. Diffuse high-grade gliomas as second malignant neoplasms after radio-chemotherapy for pediatric malignancies. *Childs Nerv. Syst.* 23, 185–93 (2007).
- Salvati, M. et al. Radio-induced gliomas: 20-year experience and critical review of the pathology. J. Neurooncol 89, 169–77 (2008).
- Walter, A. W. et al. Secondary brain tumors in children treated for acute lymphoblastic leukemia at St Jude Children's Research Hospital. *J. Clin. Oncol.* 16, 3761–7 (1998).
- Yamanaka, R., Hayano, A. & Kanayama, T. Radiation-induced gliomas: a comprehensive review and meta-analysis. *Neurosurg. Rev.* 41, 719–731 (2016).
- Lopez, G. Y. et al. The genetic landscape of gliomas arising after therapeutic radiation. Acta Neuropathol. 137, 139–150 (2019).
- Paugh, B. S. et al. Integrated molecular genetic profiling of pediatric highgrade gliomas reveals key differences with the adult disease. *J. Clin. Oncol.* 28, 3061–8 (2010).
- Phi, J. H. et al. Genomic analysis reveals secondary glioblastoma after radiotherapy in a subset of recurrent medulloblastomas. *Acta Neuropathol* 135, 939–53 (2018).
- Capper, D. et al. DNA methylation-based classification of central nervous system tumours. *Nature* 555, 469–474 (2018).
- 24. Hovestadt, V. et al. Decoding the regulatory landscape of medulloblastoma using DNA methylation sequencing. *Nature* **510**, 537–41 (2014).
- 25. Pajtler, K. W. et al. Molecular classification of ependymal tumors across all CNS compartments, histopathological grades, and age groups. *Cancer Cell* **27**, 728–43 (2015).

- Sturm, D. et al. New brain tumor entities emerge from molecular classification of CNS-PNETs. *Cell* 164, 1060–1072 (2016).
- Jones, C., Perryman, L. & Hargrave, D. Paediatric and adult malignant glioma: close relatives or distant cousins? *Nat. Rev. Clin. Oncol.* 9, 400–13 (2012).
- Northcott, P. A., Pfister, S. M. & Jones, D. T. Next-generation (epi)genetic drivers of childhood brain tumours and the outlook for targeted therapies. *Lancet Oncol.* 16, e293–302 (2015).
- 29. Sturm, D. et al. Paediatric and adult glioblastoma: multiform (epi)genomic culprits emerge. *Nat. Rev. Cancer* 14, 92–107 (2014).
- Buczkowicz, P. et al. Genomic analysis of diffuse intrinsic pontine gliomas identifies three molecular subgroups and recurrent activating ACVR1 mutations. *Nat. Genet.* 46, 451-6 (2014).
- Jones, C. & Baker, S. J. Unique genetic and epigenetic mechanisms driving paediatric diffuse high-grade glioma. *Nat Rev Cancer* 14, 10.1038/nrc3811 (2014).
- Korshunov, A. et al. H3-/IDH-wild type pediatric glioblastoma is comprised of molecularly and prognostically distinct subtypes with associated oncogenic drivers. Acta Neuropathol. 134, 507–516 (2017).
- Sturm, D. et al. Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. *Cancer Cell* 22, 425–37 (2012).
- International Cancer Genome Consortium PedBrain Tumor Project. Recurrent MET fusion genes represent a drug target in pediatric glioblastoma. *Nat. Med.* 22, 1314–1320 (2016).
- Jones, D. T. et al. Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. *Nat. Genet.* 45, 927–32 (2013).
- Jones, D. T. et al. Oncogenic RAF1 rearrangement and a novel BRAF mutation as alternatives to KIAA1549:BRAF fusion in activating the MAPK pathway in pilocytic astrocytoma. *Oncogene* 28, 2119–23 (2009).
- 37. Agnihotri, S. et al. Therapeutic radiation for childhood cancer drives structural aberrations of NF2 in meningiomas. *Nat. Commun.* **8**, 186 (2017).
- Bhuller, K. S. et al. Late mortality, secondary malignancy and hospitalisation in teenage and young adult survivors of Hodgkin lymphoma: report of the Childhood/Adolescent/Young Adult Cancer Survivors Research Program and the BC Cancer Agency Centre for Lymphoid Cancer. *Br. J. Haematol.* 172, 757–68 (2016).
- Inskip, P. D. et al. Radiation-related new primary solid cancers in the childhood cancer survivor study: comparative radiation dose response and modification of treatment effects. *Int. J. Radiat. Oncol. Biol. Phys.* 94, 800–7 (2016).
- Morton, L. M. et al. Genome-Wide Association Study to identify susceptibility loci that modify radiation-related risk for breast cancer after childhood cancer. *J. Natl Cancer Inst.* **109**, djx058 (2017).
- Schwartz, B. et al. Risk of second bone sarcoma following childhood cancer: role of radiation therapy treatment. *Radiat. Environ. Biophys.* 53, 381–90 (2014).
- van Dijk, I. W. et al. Risk of symptomatic stroke after radiation therapy for childhood cancer: a long-term follow-up cohort analysis. *Int. J. Radiat. Oncol. Biol. Phys.* 96, 597–605 (2016).
- Winther, J. F. et al. Childhood cancer survivor cohorts in Europe. Acta Oncol. 54, 655–68 (2015).
- Bakst, R. L. et al. Reirradiation for recurrent medulloblastoma. *Cancer* 117, 4977–82 (2011).
- Wetmore, C. et al. Reirradiation of recurrent medulloblastoma: does clinical benefit outweigh risk for toxicity? *Cancer* 120, 3731–7 (2014).
- Konig, L. et al. Secondary malignancy risk following proton vs. X-ray treatment of mediastinal malignant lymphoma: a comparative modeling study of thoracic organ-specific cancer risk. *Front. Oncol.* 10, 989 (2020).
- Newhauser, W. D. et al. The risk of developing a second cancer after receiving craniospinal proton irradiation. *Phys. Med. Biol.* 54, 2277–91 (2009).
- Dasu, A., Toma-Dasu, I., Olofsson, J. & Karlsson, M. The use of risk estimation models for the induction of secondary cancers following radiotherapy. *Acta Oncol.* 44, 339–47 (2005).
- Schneider, U., Zwahlen, D., Ross, D. & Kaser-Hotz, B. Estimation of radiationinduced cancer from three-dimensional dose distributions: Concept of organ equivalent dose. *Int. J. Radiat. Oncol. Biol. Phys.* 61, 1510–5 (2005).
- 50. Zhang, Q. et al. Secondary cancer risk after radiation therapy for breast cancer with different radiotherapy techniques. *Sci. Rep.* **10**, 1220 (2020).
- 51. Worst, B. C. et al. Next-generation personalised medicine for high-risk paediatric cancer patients The INFORM pilot study. *Eur. J. Cancer* **65**, 91–101 (2016).
- Harrabi, S. B. et al. Dosimetric advantages of proton therapy over conventional radiotherapy with photons in young patients and adults with low-grade glioma. *Strahlenther. Onkol.* **192**, 759–769 (2016).
- Jones, D. T. et al. Dissecting the genomic complexity underlying medulloblastoma. *Nature* 488, 100–5 (2012).
- 54. Louis, D. N. et al. WHO Classification of Tumours of the Central Nervous System, (International Agency for Research on Cancer, 2016).

- Merchant, T. E. et al. Proton versus photon radiotherapy for common pediatric brain tumors: comparison of models of dose characteristics and their relationship to cognitive function. *Pediatr. Blood Cancer* 51, 110–7 (2008).
- Northcott, P. A. et al. The whole-genome landscape of medulloblastoma subtypes. *Nature* 547, 311–317 (2017).
- 57. Taylor, M. D. et al. Molecular subgroups of medulloblastoma: the current consensus. *Acta Neuropathol.* **123**, 465–72 (2012).
- Yock, T. I. et al. Long-term toxic effects of proton radiotherapy for paediatric medulloblastoma: a phase 2 single-arm study. *Lancet Oncol.* 17, 287–298 (2016).
- Cahan, W. G. & Woodard, H. Q. et al. Sarcoma arising in irradiated bone; report of 11 cases. *Cancer* 1, 3–29 (1948).
- Kleinschmidt-Demasters, B. K., Kang, J. S. & Lillehei, K. O. The burden of radiation-induced central nervous system tumors: a single institution s experience. J. Neuropathol. Exp. Neurol. 65, 204–16 (2006).
- Mistry, M. et al. BRAF mutation and CDKN2A deletion define a clinically distinct subgroup of childhood secondary high-grade glioma. J. Clin. Oncol. 33, 1015–22 (2015).
- Krieger, M. D., Gonzalez-Gomez, I., Levy, M. L. & McComb, J. G. Recurrence patterns and anaplastic change in a long-term study of pilocytic astrocytomas. *Pediatr. Neurosurg.* 27, 1–11 (1997).
- Broniscer, A. et al. Clinical and molecular characteristics of malignant transformation of low-grade glioma in children. J. Clin. Oncol. 25, 682–9 (2007).
- 64. Cavalli, F. M. G. et al. Heterogeneity within the PF-EPN-B ependymoma subgroup. Acta Neuropathol. 136, 227–237 (2018).
- Pajtler, K. W. et al. Molecular heterogeneity and CXorf67 alterations in posterior fossa group A (PFA) ependymomas. *Acta Neuropathol.* 136, 211–226 (2018).
- Deng, M. Y. et al. Molecularly defined diffuse leptomeningeal glioneuronal tumor (DLGNT) comprises two subgroups with distinct clinical and genetic features. *Acta Neuropathol.* 136, 239–253 (2018).
- Sahm, F. et al. Next-generation sequencing in routine brain tumor diagnostics enables an integrated diagnosis and identifies actionable targets. *Acta Neuropathol.* 131, 903–10 (2016).
- Alioto, T. S. et al. A comprehensive assessment of somatic mutation detection in cancer using whole-genome sequencing. *Nat. Commun.* 6, 10001 (2015).
- 69. McPherson, A. et al. deFuse: an algorithm for gene fusion discovery in tumor RNA-Seq data. *PLoS Comput. Biol.* 7, e1001138 (2011).
- Ramaswamy, V. et al. Risk stratification of childhood medulloblastoma in the molecular era: the current consensus. *Acta Neuropathol.* 131, 821–31 (2016).

#### Acknowledgements

This study was supported by the German Consortium for Translational Cancer Research (DKTK), the Shantella Foundation (Vaduz, Liechtenstein), 'Ein Herz für Kinder', German Cancer Aid (DKH), Deutsche Kinderkrebsstiftung (DKKS), the German Federal Ministry of Science and Research (BMBF), the Ministry of Research and the Arts of the State of Baden-Württemberg and a generous donation from the Scheu family. M.Y.D. was supported by fellowships from the Mildred Scheel doctoral program of the German Cancer Aid and the German National Scholarship Foundation. A. K. is supported by the Helmholtz Association Research Grant (Germany). M.R. is supported by RSF Research

Grant 18-45-06012. For excellent technical support and expertize we thank the Genomics and Proteomics Core Facility (GPCF) and Omics, IT and Data Management Core Facility (ODCF) of the German Cancer Research Center (DKFZ).

#### Author contributions

D.T.W.J. and S.M.P. conceived the original idea and jointly supervised this study. M.Y.D., M.S., D.Stu., E.P., S.M.P. and D.T.W.J. wrote the manuscript with support from all authors. S.T., C.K., A.M.D., A.I.G., C.J., J.S., M.E., M.U.S., B.C.J., C.M.v.T., A.W., A.G., M.R., J.E., T.M., O.W., F.S., D.R., D.Su., J.Z., A.K., and A.v.D. provided samples and clinical data. M.Y.D., D.Stu., E.P., and A.W. prepared and processed the samples. M.S., D.Sti., and G.P.B. performed the computational processing of raw data. M.Y.D., D.Stu., E.P., S.M.P., and D.T.W.J. performed the data analysis and interpretation. All authors discussed the results and contributed to the final version of the manuscript.

#### Funding

Open Access funding enabled and organized by Projekt DEAL.

#### **Competing interests**

The authors declare no competing interests.

#### Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-25708-y.

Correspondence and requests for materials should be addressed to David T. W. Jones.

**Peer review information** *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/ licenses/by/4.0/.

© The Author(s) 2021