ORIGINAL PAPER

Genomic analysis reveals secondary glioblastoma after radiotherapy in a subset of recurrent medulloblastomas

Ji Hoon Phi1,2 · Ae Kyung Park3 · Semin Lee4,5 · Seung Ah Choi1,2 · In-Pyo Baek6 · Pora Kim7 · Eun-Hye Kim8 · Hee Chul Park9 · Byung Chul Kim9 · Jong Bhak4,5 · Sung-Hye Park10 · Ji Yeoun Lee1,2,11 · Kyu-Chang Wang1,2 · Dong-Seok Kim12 · Kyu Won Shim12 · Se Hoon Kim10 · Ji Yeoun Lee1,2 · Chae-Yong Kim14 · Seung-Ki Kim1,2

Received: 15 December 2017 / Revised: 2 April 2018 / Accepted: 2 April 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

Despite great advances in understanding of molecular pathogenesis and achievement of a high cure rate in medulloblastoma, recurrent medulloblastomas are still dismal. Additionally, misidentification of secondary malignancies due to histological ambiguity leads to misdiagnosis and eventually to inappropriate treatment. Nevertheless, the genomic characteristics of recurrent medulloblastomas are poorly understood, largely due to a lack of matched primary and recurrent tumor tissues. We performed a genomic analysis of recurrent tumors from 17 pediatric medulloblastoma patients. Whole transcriptome sequencing revealed that a subset of recurrent tumors initially diagnosed as locally recurrent medulloblastomas are secondary glioblastomas after radiotherapy, showing high similarity to the non-G-CIMP proneural subtype of glioblastoma. Further analysis, including whole exome sequencing, revealed missense mutations or complex gene fusion events in \( \text{PDGFRA} \) with augmented expression in the secondary glioblastomas after radiotherapy, implicating \( \text{PDGFRA} \) as a putative driver in the development of secondary glioblastomas after treatment exposure. This result provides insight into the possible application of PDGFRA-targeted therapy in these second malignancies. Furthermore, genomic alterations of \( \text{TP53} \) including 17p loss or germline/somatic mutations were also found in most of the secondary glioblastomas after radiotherapy, indicating a crucial role of \( \text{TP53} \) alteration in the process. On the other hand, analysis of recurrent medulloblastomas revealed that the most prevalent alterations are the loss of 17p region including \( \text{TP53} \) and gain of 7q region containing \( \text{EZH2} \) which already exist in primary tumors. The 7q gain events are frequently accompanied by high expression levels of \( \text{EZH2} \) in both primary and recurrent medulloblastomas, which provides a clue to a new therapeutic target to prevent recurrence. Considering the fact that it is often challenging to differentiate between recurrent medulloblastomas and secondary glioblastomas after radiotherapy, our findings have major clinical implications both for correct diagnosis and for potential therapeutic interventions in these devastating diseases.

Keywords Medulloblastoma · Recurrence · Secondary glioblastoma after radiotherapy · Genomic analysis

Introduction

Medulloblastoma (MB) is the most common malignant brain tumor in children. Decades of advances in surgery, adjuvant therapies, and patient care have led to a high cure rate of MB which was invariably fatal in Harvey Cushing’s time. Clinical trials are ongoing for reducing radiation dose to decrease late toxic effects in survivors. Molecular subgrouping of MB has been well-established by a series of genomic studies [27]. There are four main subgroups of MB, namely, WNT, SHH, Group 3, and Group 4, which exhibit distinct gene expression patterns, ages of onset, clinical behavior, and prognoses. More recently, genome-wide DNA methylation
profiling has been recognized as a more robust and accurate method for not only molecular subgrouping of MB but also CNS tumor classification [17]. Patients with WNT MB have a > 90% chance of survival [8, 37]. In contrast, patients with Group 3 tumors and MYC amplification face a far worse prognosis (< 50% survival) [39]. Molecular targeted therapy is also attempted for the treatment of refractory MBs with altered SHH signaling pathways [42, 43].

Notwithstanding the advancements in MB treatment and research, the situation regarding recurrent MB looks different. Currently, at least a quarter of all MB patients suffer from tumor recurrence, and the prognosis is dismal. There is no effective therapy for recurrent MB if standard treatments have already been delivered [21]. The genomic and molecular characteristics of recurrent MB are poorly understood, largely because of the lack of tumor tissues. It is known that recurrent MBs frequently acquire increased histological anaplasia and chromosomal imbalance [16]. A recent study has shown that the emergence of combined MYC and p53 pathway defects during recurrence leads to extremely poor outcomes [13]. Another study has shown that recurrent MBs are derived through clonal selection from existing clones in primary tumors and harbor a remarkable number of new genetic events [24]. Interestingly, despite the accumulation of genetic variations, the molecular subgroup of a MB is highly stable during recurrence [40]. Therefore, progressing genomic changes coexist with stable molecular subgroups during MB recurrence. To overcome the treatment resistance of recurrent MBs, it would be desirable to find and target emergent genomic aberrations in recurrent tumors.

A confounding problem in the management of MB recurrence is the presence of secondary malignancies [15]. Whole neuro-axis radiation is the mainstay of treatment for MBs. Therefore, secondary malignancy, especially malignant gliomas, can develop in a small percentage of survivors [5]. Clinically, it is often challenging to differentiate between the two entities based solely on imaging results. Furthermore, matched primary and recurrent tumor tissues are uncommon because surgery is usually not indicated for recurrent MBs. Even pathological examinations often provide misleading information due to histological ambiguity, such as high degrees of anaplasia during recurrence. However, misdiagnosis of secondary glioblastoma (GBM) after radiotherapy (RT) as recurrent MB can result in grim outcomes because the treatment protocols and regimens are different.

Genomic and epigenomic analyses have shown great potential not only in stratifying diseases but also in correcting misdiagnoses [49]. In the genomic analyses of primary and recurrent MBs, we found that a subset of recurrent tumors contains mutations or gene fusion events in PDGFRA, which is not characteristic of MB. The gene expression pattern of these recurrent tumors co-segregated with that of GBM rather than MB. Re-evaluation of histological images revealed a misdiagnosis of secondary GBM after RT. PDGFRA is a frequently altered gene in GBM [4, 36].

This study demonstrates the power of genomic analyses for diagnosis of brain tumors, especially during recurrence. Integrated diagnosis, combining histology and molecular characteristics, is recommended for recurrent MBs. Furthermore, altered PDGFRA signaling pathways can be targeted by receptor tyrosine kinase inhibitors, and outcomes of patients with secondary GBM after RT with PDGFRA mutations can be improved with appropriate diagnosis and therapies.

**Materials and methods**

**Patients and samples**

Specimens were collected from 17 patients from three centers, namely, Seoul National Children’s Hospital (SNUCH), Severance Hospital, and Seoul National University Bundang Hospital (SNUBH), with appropriate consent from institutional review boards. The primary tumor in all the patients was MB. Details of patient characteristics and samples are described in Table S1. The biospecimens were composed of 14 tumors paired with patient-matched primary MB samples (11 pairs of primary–recurrent MB and 3 pairs of primary MB–secondary GBM after RT); one pair of locally recurrent-distantly metastasized MB without a primary tumor sample; one pair of secondary GBM after RT, from two independent operations, without a primary MB sample; and one unpaired secondary GBM after RT without a primary MB sample. Normal DNA from peripheral whole blood cells (WBCs) was available for 12 patients. For DNA and RNA extraction from frozen tissues, DNeasy Blood and Tissue Kit (Qiagen) and mirVana miRNA Isolation Kit (Invitrogen) were used, respectively. Extraction of DNA and RNA from FFPE tissues was carried out using RecoverAll Total Nucleic Acid Isolation Kit (Thermo Fisher Scientific).

**RNA sequencing and analysis**

mRNA sequencing was performed using a TruSeq RNA Library Preparation Kit and TruSeq RNA Access Kit (Illumina) and sequenced on an Illumina HiSeq 2000 or 2500 platform. We used an in-house script to trim low-quality reads [proportion of ambiguous nucleotide (N ratio) > 0.1 or proportion of nucleotide with low Phred quality score (< 20) > 0.4]. The filtered sequence data were analyzed using STAR [7] for alignment, Cufflinks v2.2.1 1 [47] for assembly, and a known set of reference transcripts from Ensembl v72 for expression estimation. The details of RNA sequencing and alignments are summarized in Table S2. Gene fusions were identified by searching for the spanning
reads and split reads by ChimeraScan [14], deFuse [22], and FusionMap [11] using default settings. To detect robust candidates of gene fusions accompanied by change in expression, we used several selection criteria: detection with more than two software programs, high changes in expression in recurrent tumors compared with primary tumors (log₂FC ≥ 1) in either of two genes involved in fusion event, and high expression levels in either of the two genes in at least one tumor of the paired samples (maximum expression of FPKM in either of the two genes across tumor pairs ≥ 50).

Whole exome sequencing and analysis

Whole exome sequencing (WES) was performed using the 71 Mb SureSelect V4 plus UTR Kit or the 51 Mb SureSelect V5 Kit (Agilent Technologies) according to the manufacturer's instructions. DNA libraries were constructed according to the protocol provided by the manufacturer and WES was performed using an Illumina HiSeq 2000 or 2500 platform to generate 101 bp paired-end reads. We used cutadapt [20] and sickle (https://github.com/najoshi/sickle) to remove adapter sequences and low-quality sequence reads. Burrows–Wheeler aligner [18] was used to align the sequencing reads onto the human reference genome (hg19). We used a Genome Analysis ToolKit (GATK) (http://www.broadinstitute.org/gatk) for local realignment and score recalibration and to filter the sequence data. Picard (http://broadinstitute.github.io/picard) and Samtools (http://samtools.sourceforge.net/) were also used for basic processing and management of the sequencing data. The details of WES and alignments are summarized in Table S3. Somatic single nucleotide variants (SNVs) and short insertions and deletions (indels) were analyzed using MuTect [6] and SomaticIndelDetector in GATK, respectively, with matched normal samples as controls. For comparisons between normal and tumor samples without any available patient-matched normal DNA, we generated variants from 12 normal WBC samples using GATK with UnifiedGenotyper, VariantRecalibrator (hapmap, 1000g, dbsnp) and ApplyRecalibration options, and we identified intersections between the normal variants and each tumor sample using BEDtools (http://sourceforge.net/projects/bedtools/) with the intersect option. SnpEff (http://snpeff.sourceforge.net/) was used to select somatic variants located in coding sequences and predict their functional consequences, such as silent or non-silent variants. Amplicon sequencing was also performed using an Ion AmpliSeq Comprehensive Cancer Panel (Life Technologies) according to the manufacturer’s protocols to validate the missense mutations in PDGFRA detected by WES. Sequencing was carried out on an Ion Proton sequencer (Life Technologies) with the Ion PI chip. Somatic copy number alterations (SCNAs) were detected based on WES data using ControlFREEC (v10.4) with default options [2] in 12 tumor samples for which matched normal data were available. The annotation of known cancer genes in the segments of SCNAs was performed using the Catalogue of Somatic Mutations in Cancer (COSMIC) [10].

Results

Samples

To investigate genomic changes in recurrent MB, tumor tissues of 17 patients were collected from three centers (Table S1), including 11 pairs of patient-matched primary–recurrent MB (D2–D9, L1–L3), three pairs of primary MB–secondary GBM after RT (L4–L6), one pair of locally recurrent and distantly metastasized MB without its primary MB sample (D1), and one pair of secondary GBM after RT obtained from separate operations without its primary MB sample (G1). The final sample was a single secondary GBM after RT without its primary MB sample (G2). All initial diagnoses were based on pathological examinations of the original specimens. Three pairs of primary MB–secondary GBM after RT (L4–L6) were initially diagnosed as primary–recurrent MBs. However, genomic analysis in this study revealed that the recurrent tumors were not locally recurrent MBs but newly developed secondary GBMs after RT. Other GBM tumors (G1 and G2), which occurred in two MB patients and were originally confirmed as secondary GBMs after RT based on their pathologies, were used as positive controls.

Subgroup affiliation of recurrent medulloblastoma and secondary glioblastoma after RT

We first examined subgroup affiliation of tumor samples using unsupervised hierarchical clustering and partitioning around medoid (PAM) methods based on RNA-seq expression profiles of 17 marker genes from a CodeSet [30] excluding WNT group-specific genes (Fig. 1a). The subgroups of primary MBs were well differentiated by both methods, however, the clustering of one primary MB sample, L6.1, was not consistent. In addition, a negative silhouette width was observed in the PAM analysis (Fig. 1a, right) even though the expression level of Group 3-specific genes was considerably high in L6.1. This ambiguous result might be ascribed to the low RNA quality of the sample (the RNA Integrity Number = 2.8). Therefore, the L6.1 sample was assigned to Group 3 according to the high expression level of Group 3-specific genes. All 11 primary–recurrent MB pairs showed the same subgroup affiliation regardless of recurrent location, distant (D2–D9) or local (L1–L3), which is consistent with a previous report [40]. Even the subgroup affiliation of a pair of locally recurrent (3rd operation) (D1.1)
and distantly metastasized (4th operation) (D1.2) MBs remained unchanged. However, the expression patterns of three secondary GBM tumors (L4–L6) initially misdiagnosed as recurrent MBs clearly showed strong similarity to those of pathologically confirmed secondary GBMs after RT (G1 and G2). In the principal component analysis (PCA) plot using the 3000 most variably expressed genes, all the secondary GBMs after RT were segregated into one cluster (Fig. 1b). We further confirmed the entity of the secondary GBMs after RT based on microarray expression profiles that were available for two patients, L4 and L5, combined with two public datasets containing 30 pediatric primary MBs (GSE30074) [33] and 34 adult GBMs (GSE58399) [32]. Hierarchical clustering and PAM analysis with 22 genes of a CodeSet (Fig. 1c) and PCA plot with top 3000 genes with the high expression variability (Fig. 1d) clearly showed that the secondary GBMs were grouped together with adult GBMs, not with pediatric MBs. Furthermore, we found that most of the secondary GBM tumors were most similar to the proneural subtype GBM when considering the expression correlations calculated between each secondary GBM sample and 163 TCGA GBM tumor samples based on previously reported GBM subtype-specific genes [48] (Fig. 1e). We compared the time interval between the first diagnosis...
coefficients calculated between log2 (FPKM + 1) values of each sample (IQR) of 68 samples of a method.

PCA plot of top 3000 genes with high RNA-seq expression variability (IQR) of 33 samples of a. Heatmap with unsupervised hierarchical clustering and silhouette plot of partitioning around medoid (PAM) of 33 samples of a. PCA plot of top 3000 genes with high expression variability (IQR) of 68 samples of c. e Box plots of Pearson’s correlation coefficients calculated between log2 (FPKM + 1) values of each sample of secondary GBM after RT and log2 (normalized counts + 1) values from RNA-seq data of 163 TCGA GBM samples (n=39, 28, 42, and 54 for proneural, neural, classical, and mesenchymal subtypes, respectively) based on 763 subtype-specific genes [48]. Using a similar procedure applied in Sandmann et al. [44], the similarity between each sample of the secondary GBMs after RT and the TCGA samples was investigated by Pearson’s correlation coefficients based on the expression profiles of 763 subtype-specific genes that are commonly available in both datasets. f Recurrence free survival analysis with Kaplan–Meier plot and generalized Wilcoxon test. Comparisons of median time to recurrence were made between secondary GBMs after RT and recurrent MBs (left) or across subgroups of MB (right) of MB and the advent of recurrence. Median time to recurrence was significantly longer for secondary GBM after RT (median 9.0 years) compared to recurrent MBs (median 3.0 years) (generalized Wilcoxon p = 0.0034) (Fig. 1f, left). When the comparison was made across subgroups of MB, no statistical significance was obtained (Fig. 1f, right), which is not congruent to the previous report that Group 4 MBs recurred significantly later than Group 3 and SHH tumors [40].

Somatic mutations in recurrent medulloblastoma and secondary glioblastoma after RT

To identify the genomic alterations that drive recurrent MBs and secondary GBMs after RT, somatic SNVs and indels were called on the exomes of tumor samples. Various comparisons were made based on the availability of the samples (Table S4). For the tumor samples without matched normal DNA, comparisons were made between the tumor and multiple variants calls generated from 12 normal WBC DNA samples (designated combWBC calls). Coding mutation rates and the number of mutations are presented in Table S5. The median non-silent mutation rate in primary MBs was 0.27 per megabase compared with patient-matched genomic DNA (Table S5), which is lower than the previously reported rate of 0.35 non-silent mutations per megabase [38]. To further select the damaging mutations, we excluded known polymorphic variants (allele frequency ≥ 0.1%) from the non-silent mutations based on the 1000 Genomes Project, Exome Sequencing Project (ESP6500), and Exome Aggregation Consortium (ExAC). The details of the damaging mutations are listed in Table S6. Consistent with a previous report [24], the number of damaging mutations increased in recurrent MBs, and only a part of these mutations was shared in primary and recurrent MBs (Fig. 2a). In contrast, only one damaging mutation overlapped between the primary MBs and the secondary GBMs after RT, indicating that the secondary GBMs after RT did not originate from primary MBs. In the secondary GBMs after RT, non-silent mutations in PDGFRA were observed in all five patients (Fig. 2b), including eight missense mutations and one splice region mutation (Table S6). This implies that genomic alterations in PDGFRA might be one of the key drivers of secondary GBM after RT. Most of the missense mutations in PDGFRA occurred in extracellular immunoglobulin-like domains (L5, G1, and G2), and one other missense mutation was observed in the intracellular tyrosine kinase domain, i.e., the D842V substitution (L4) (Fig. 2c), which is the most frequently found mutation in PDGFRA and is related to imatinib resistance. Two missense mutations in PDGFRA found in tumors from two patients (L4.2 and L5.2) were confirmed by deep sequencing (×1000) using Ion AmpliSeq Comprehensive Cancer Panel (Table S7). The mutations in PDGFRA were accompanied by significantly increased mRNA expression in secondary GBMs after RT compared with primary MBs (Fig. 2d). Additionally, TP53 and several genes involved in response to ionizing radiation or DNA repair such as ATM, DNMT3A, and UIMC1 were detected as mutated genes in the secondary GBMs after RT, indicating failure of protective mechanisms against postoperative therapeutic radiation (Fig. 2b).

In the analysis of ten patient-matched primary–recurrent MBs, no specific common mutated gene was recognized. However, in recurrent MBs, new damaging mutations were identified in the genes involved in cancer pathways (PIK3CA, PIK3CG, SOS2, HDAC1, E2F2, TGFBR2, PTEN, and MTOR) or stem cell maintenance (LIF, NOTCH2, and DDX6) (Fig. 2b). In addition, the mutated genes included the genes, CESR3, ABCA13, and, WDFY3, which were previously detected in recurrent human MBs [24]. Other mutated genes involved in neurogenesis (ASPM and FAT4) were found in two recurrent MBs, respectively. In comparisons between WBC DNAs and primary MBs, we detected several mutated genes such as TP53, PTC1, SMO, LDB1,
PRKACA, and KDM6A that were previously reported by exome sequencing of a large cohort of primary MBs [38]. The damaging mutations in TP53 and PTCH1 were observed in two patients, respectively, and some mutations including TP53 were retained in recurrent MBs, while others were lost.

**Fig. 2** Somatic SNV/indels in secondary glioblastomas after radiotherapy, recurrent medulloblastomas, and primary medulloblastomas. 

a Venn diagrams of damaging mutations detected in primary, recurrent medulloblastomas (MBs), and secondary glioblastomas (GBMs) after radiotherapy (RT). b Representative mutated genes with damaging mutations in secondary GBMs after RT, recurrent MBs, and primary MBs. Asterisks indicate the genes identified in previous studies in recurrent MBs (asterisk) [24] or primary MBs (double asterisks) [38]. c Lollipop of missense mutations detected in secondary GBM tumors. d Expression levels of PDGFRA

**Somatic copy number alterations in recurrent medulloblastoma and secondary glioblastoma**

We investigated SCNAs occurring in primary/recurrent MBs and secondary GBMs after RT using WES data from 12 patients (D1–D4, D6–D8, L1, L4–L6, and G1) with available matched normal data (Table S8). The SCNAs
observed in recurrent MBs were largely concordant with those in primary MBs (Fig. 3a). However, SCNAs observed in primary MBs were rarely shared by their respective patient-matched secondary GBMs after RT, confirming the distinct origin of the tumors. On the other hand, in the locally recurrent-distant metastasized MB pair (D1) and the secondary GBM after RT tumor pair (G1), new SCNAs occurred in latter samples while many of initial SCNAs remained.

Fig. 3 Somatic copy number alterations in primary medulloblastomas, recurrent medulloblastomas, and secondary glioblastomas after radiotherapy. a Copy number change in 12 patients whose normal DNA samples were available. Red and blue indicate copy number gain and loss, respectively. b Overview of the two most frequently observed cytogenetic gains and losses with suspected driver genes. Focal aberrations were defined as 5 Mb or less in size. c Expression level of TP53, PTEN, and EZH2 in FPKM of RNA-seq data. Asterisk denotes congruency between copy number change and expression level.
The most frequent SCNAs in tumors were deletions in the 17p telomeric region that includes TP53, which was observed in ten out of 12 patients (Fig. 3b). In primary-recurrent MB pairs, the 17p deletion was observed in both primary and recurrent tumors. However, the appearance of the 17p deletion in secondary GBMs after RT (L4 and L5) was a new observation. These findings imply that TP53 loss plays a central role not only in the recurrence of MBs but also in the novel occurrence of secondary GBMs after RT. The expression level of TP53 was largely consistent with the SCNA result, showing considerably low expression levels of TP53 (FPKM value ranges from 0.92 to 8.92) in five patients (D1, D3, D4, D8, and L1) (Fig. 3c). The next most frequent copy number loss was detected in the wide 10q region that includes tumor suppressor PTEN as a suspected driver gene, and the loss was also associated with low expression of PTEN in the tumors containing a copy number loss of 10q (D6, L1, L4, and G1.1) (Fig. 3c). The PTEN locus has been previously reported as the most frequently homozygously deleted region in MBs [29]. The two most frequent copy number gain events were detected in the 7q and 17q regions. The gain of the 7q region includes EZH2 as a putative driver gene was observed in four patients with recurrent MB, and the expression of EZH2 was considerably elevated in three patients with distant metastasis (D1, D2, and D4) (Fig. 3c). High-level expression of EZH2 was also observed in one patient with distant metastasis without SCNA information (D5). The gain of the 17q region was not consistently observed in primary and recurrent MBs; this region was newly detected (D7) or disappeared in recurrent MB (D3).

![Fig. 4](https://example.com/fig4.png)

**Fig. 4** Complex gene fusion events in the PDGFRA gene region in chromosome 4 in a secondary glioblastoma tumor after radiotherapy. **a** Schematic diagram of gene fusion events with aligned multiple RNA-seq reads mapped on fusion junctions. The fusion gene KIT–LNX1 (NM_001126328) was spliced using a normal splice site of exon 1 of KIT and a cryptic splice site of intron 1 of LNX1 (NM_001126328). **b** Read depth coverage plot of RNA-seq data of LNX1. **c** Expression of PDGFRA–LNX1 fusion transcript, PDGFRA and LNX1 in a primary medulloblastoma (L6.1) and a secondary glioblastoma after radiotherapy (L6.2), confirmed by RT-qPCR.
Intrachromosomal gene fusion in the PDGFRA region in a secondary glioblastoma

A total of 20 most putative fusion genes with high expression changes were detected in five tumor samples (Table S9). Among them, complex intrachromosomal gene fusion events were observed in chromosome 4 in one secondary glioblastoma after RT (L6.2) involving four neighboring genes, namely, FIP1L1, LNX1, and two receptor tyrosine kinases, PDGFRA and KIT (Fig. 4a). We identified three different fusion transcripts and multiple RNA-seq reads that spanned the fusion junctions in which LNX1 was involved in all three events: LNX1 (NM_032622)–FIP1L1, PDGFRA–LNX1 (NM_032622), and KIT–LNX1 (NM_001126328). The expression of all four genes was considerably elevated (Table S9). Furthermore, in the fusion with KIT, a cryptic splice site in intron 1 of LNX1 (NM_001126328) was used for splicing with exon 1 of KIT, and the depth coverage plot of the RNA-seq data showed high transcription levels of the intronic sequence of intron 1 of LNX1 (Fig. 4b). In addition, the abrupt change of coverage depth of exon 7 of LNX1 (NM_032622) after the break point is also consistent with the presence of the fusion transcript of LNX1 (NM_032622)–FIP1L1. The PDGFRA–LNX1 fusion cDNA was successfully amplified in the secondary GBM after RT (L6.2) but not in the primary MB (L6.1) (Fig. 4c).

Pathological confirmation of secondary glioblastomas after radiotherapy

The original pathology slides of the presumed secondary GBMs after RT (L4–L6) were re-evaluated by a pathologist (S.P.). Morphological features of secondary GBMs after RT (L4.2, L5.2 and L6.2) showed polymorphic nuclei, long cytoplasmic processes, microvascular proliferation and necrosis which were different from the features of MBs (L4.1) (Fig. 5a). However, the anaplastic/large cell features observed in the secondary GBMs after RT mimicked anaplastic/large cell variant of MBs with large atypical nuclei. Therefore, the presence of edematous stroma

Fig. 5 Pathological confirmation of previously misdiagnosed secondary glioblastomas after radiotherapy. a Hematoxylin and eosin (H&E) (×200) slides of a primary medulloblastoma (MB) (L4.1) as a negative control, three secondary glioblastomas (GBMs) after radiotherapy (RT) initially misdiagnosed as recurrent MBs (L4.2, L5.2, and L6.2), and two pathologically confirmed secondary GBMs after RT (G1.1 and G2.1) as positive controls. b Olig2 immunohistochemistry of MB (L4.1) and secondary GBMs after RT (L4.2, L5.2, G1.1, and G2.1) and ultrastructural imaging (uranyl acetate and lead citrate staining; ×12,300) of secondary GBM after RT (L6.2)
| Tumor label | Gender | Age | M stage | Subgroup | Time interval (year) primary to recurrence | Tumor label | Characteristics | PDGFRA Mutation/gene fusion | Expression (FPKM) | TP53 Mutation | Expression (FPKM) | EZH2 Mutation | Expression (FPKM) |
|-------------|--------|-----|---------|----------|-------------------------------------------|-------------|----------------|--------------------------|----------------|-------------|----------------|--------------|----------------|----------------|
| Recurrent MB |         |     |         |          |                                           |             |                |                          |                |             |                |              |                |                |
| D1          | M      | 9   | –       | Group 4  | 3.4                                       | D1.1        | Local          | 0.7                      | Loss           | 8.2         | Gain           | 75.7         |                |                |
|             |        |     |         |          |                                           | D1.2        | Meta           | 0.6                      | Loss           | 6.0         | Gain           | 79.9         |                |                |
| D2          | M      | 9   | 3       | Group 4  | 2.5                                       | D2.1        | Primary        | 12.4                     | Neutral        | 17.4        | Gain           | 98.9         |                |                |
|             |        |     |         |          |                                           | D2.2        | Meta           | 4.2                      | Neutral        | 15.0        | Gain           | 88.9         |                |                |
| D3          | M      | 9   | 0       | Group 4  | 6.8                                       | D3.1        | Primary        | 0.1                      | Loss           | 7.7         | Neutral        | 26.1         |                |                |
|             |        |     |         |          |                                           | D3.2        | Meta           | 0.5                      | Loss           | 8.9         | Neutral        | 21.8         |                |                |
| D4          | F      | 14  | 0       | Group 4  | 3.0                                       | D4.1        | Primary        | 0.5                      | Loss           | 8.7         | Gain           | 112.2        |                |                |
|             |        |     |         |          |                                           | D4.2        | Meta           | 0.6                      | Loss           | 7.5         | Gain           | 128.1        |                |                |
| D5          | F      | 16  | 0       | Group 4  | 3.0                                       | D5.1        | Primary        | 3.0                      | –              | 162.9       | –              | 179.7        |                |                |
|             |        |     |         |          |                                           | D5.2        | Meta           | 11.0                     | –              | 19.5        | –              | 203.7        |                |                |
| D6          | F      | 15  | 0       | SHH      | 1.3                                       | D6.1        | Primary        | 5.3                      | Loss           | 33.7        | Neutral        | 46.9         |                |                |
|             |        |     |         |          |                                           | D6.2        | Meta           | 13.5                     | Loss           | 39.8        | Neutral        | 54.9         |                |                |
| D7          | M      | 12  | 1       | SHH      | 3.6                                       | D7.1        | Primary        | 8.2                      | Loss           | 32.6        | Neutral        | 36.9         |                |                |
|             |        |     |         |          |                                           | D7.2        | Meta           | 33.7                     | Loss           | 36.3        | Neutral        | 27.8         |                |                |
| D8          | M      | 8   | 0       | SHH      | 1.0                                       | D8.1        | Primary        | 4.9                      | Loss           | 0.9         | Neutral        | 32.8         |                |                |
|             |        |     |         |          |                                           | D8.2        | Meta           | 13.3                     | Loss           | 1.8         | Neutral        | 41.1         |                |                |
| D9          | F      | 9   | 0       | SHH      | 1.6                                       | D9.1        | Primary        | 14.0                     | –              | 38.8        | –              | 29.0         |                |                |
|             |        |     |         |          |                                           | D9.2        | Meta           | 3.9                      | –              | 52.9        | –              | 17.8         |                |                |
| Local recurrence | | | | | | | | | | | | | |
| L1          | F      | 9   | 3       | SHH      | 3.9                                       | L1.1        | Primary        | 5.0                      | Somatic (missense) | Loss | 5.9 | Gain | 40.4 | |
|             |        |     |         |          |                                           | L1.2        | Local          | 2.3                      | –              | 6.0 | Gain | 26.0 | |
| L2          | M      | 15  | 0       | Group 3  | 4.0                                       | L2.1        | Primary        | 7.8                      | –              | 31.3 | – | 39.0 | |
|             |        |     |         |          |                                           | L2.2        | Local          | 5.4                      | –              | 44.7 | – | 41.6 | |
| L3          | M      | 9   | 3       | Group 3  | 2.0                                       | L3.1        | Primary        | 18.3                     | –              | 147.8       | – | 13.3 | |
was also observed. Among various immunomarkers, Olig2 was the most valuable marker to distinguish GBMs from MBs because only gliomas are always positive for this antibody (Fig. 5b). In the electron microscopy study, secondary GBMs after RT had cytoplasm which was abundant with glial intermediate filaments (L6.2) (Fig. 5b). In addition, all secondary GBMs after RT were negative for the IDH1 mutation-specific antibody. The MIB1 (Ki67) index ranged from 12 to 48%.

Discussion

In this study, we investigated the genomic changes in recurrent MBs. A relatively low mutation rate is characteristic of pediatric brain tumors such as atypical teratoid/rhabdoid tumor and MBs [3]. However, in recurrent MBs, the mutation rate increased several fold with the accumulation of non-silent mutations (Table S5). Therapy-induced DNA damage and clonal selection of mutated tumor cells are considered to be the underlying mechanisms for the increased mutation rate in recurrent MBs [24]. Recurrent MB is refractory to most chemotherapeutic regimens currently available [21]. The accelerated genomic changes may contribute to the failure of treatments. Nearly all patients with recurrent MBs exhibited genomic alterations of TP53 including somatic mutations and/or loss of the 17p region containing TP53 (Table 1), indicating the crucial role of TP53 in MB recurrence. A recent study reported that deleterious alterations in TP53 or TP53 pathway genes were primarily responsible for the recurrence of SHH MBs in both humans and mice [24]. Gain of 7q with high expression of EZH2, was also prevalent in recurrent MBs, particularly enriched in Group 4 (Fig. 3b). It has been noticed that the high expression of EZH2, encoding a component of the polycomb repressive complex and histone H3K27 methyltransferase, is associated with gain of chromosome 7, which is observed specifically in Group 3 and Group 4 and implicated in deregulated chromatin modification in these groups of MB [28, 41].

Secondary malignancies account for 11.8% of the late death of childhood MB patients [26]. Brain tumors, mostly gliomas, constitute one-third of all secondary malignancies [26]. Ionizing radiation can cause DNA cleavage and can induce transformation and oncogenesis [46]. All clinical protocols for childhood MB include treatment with ionizing radiation along the entire craniospinal axis to eliminate microscopically disseminated tumor cells. The median latency from initial radiation therapy to the diagnosis of secondary glioma is 9–11 years [9, 50], which is consistent with our data (median 9 years) (Table 1). In fact, a longer time to recurrence is an important clue to suspect secondary malignancy in childhood brain tumor
survivors. However, one patient (D3) developed very late recurrence of MB (6.8 years) and another (L6) developed secondary GBM earlier than usual (4.3 years) in our study. Therefore, molecular and genomic profiling can be a great aid in correct diagnosis.

A recent study reported an excess of deletions relative to insertions as one of the mutational signatures of ionizing radiation in secondary malignancies [1]. Accordingly, we examined whether the deletion/insertion ratio increased in newly occurred mutations in recurrent MBs and secondary GBMs after RT (Fig. 6). We observed that the deletion/insertion ratio was significantly increased in recurrent MBs compared with primary MBs (Wilcoxon rank sum test $p = 0.015$) but not in secondary GBMs after RT (Fig. 6a). Consistently, in pairwise comparisons, the deletion/insertion ratio was elevated in most of the recurrent MBs (Fig. 6b). However, the increase of deletion/insertion ratio was not obvious in two secondary GBMs after RT, L4 and L5. In conjunction with treatment exposure, other etiological factors such as genetic predisposition contribute to the development of secondary malignancies in pediatric cancers [25]. Indeed, a germline TP53 mutation was found in one patient, L5.

Brain tumors are very heterogeneous diseases and many entities mimic each other in clinical and histopathological features. Gene expression patterns, chromosomal aberrations, DNA methylation profiling, and specific mutations can provide valuable information for diagnoses of brain tumors [12, 17, 23, 45]. All five secondary GBMs after RT exhibited alterations in the PDGFRA locus, missense mutations in four patients and a PDGFRA-LNX1 gene fusion in one patient. PDGFRA expression was augmented in all cases. PDGFRA is a receptor tyrosine kinase and one of the principal upstream regulators of the RAS/MEK/PI3K signaling pathway. PDGFRA alterations such as gene amplification and oncogenic mutations are prominent in pediatric high-grade gliomas [19, 34, 35]. Secondary GBMs after RT share these characteristics with pediatric malignant gliomas, i.e., overexpression and alterations of the PDGFRA gene. Frequent TP53 mutations and/or loss of the 17p telomeric region containing TP53 were observed in four out of five secondary GBMs after RT (Table 1), implying that TP53 alteration is critical in the development of secondary GBMs after RT. In one secondary tumor after RT, L6.2, neither TP53 alteration nor PDGFRA missense mutation was observed, but complex gene fusion events involving the PDGFRA region occurred. Furthermore, the expression profile of this tumor, L6.2, showed similarity to neural and mesenchymal subtype GBMs in addition to proneural subtype (Fig. 1e). Clinical data also revealed that the secondary malignancy after RT occurred considerably earlier (4.3 years) in this patient (L6) than in other patients (L4, L5, G1, and G2) (7.3–10.0 years) (Table 1). Taken together, in the patient L6, the secondary GBM after RT might have occurred by a different mechanism. Nevertheless, the occurrence of the GBM involved PDGFRA.
alteration. On the other hand, no secondary tumor after RT harbored an IDH1 mutation, which is congruent to a previous investigation of pediatric high-grade gliomas [34]. In adult GBMs, IDH1 mutations are mostly found in the glioma-CpG island methylator phenotype (G-CIMP), a subset of the proneural subtype [31]. Therefore, all the secondary GBMs after RT in this study are considered to be similar to non-G-CIMP proneural subtype of GBM.

In this study, we explored the genomic landscape of recurrent MB. Recurrent MBs have higher rates of genomic alterations. TP53 alteration and EZH2 overexpression are observed in recurrent MBs. Notably, we found unexpected secondary gliomas in a subset of presumed recurrent MBs. However, the high proportion of the secondary tumor observed in our study (29.4%, 5 out of 17 patients) does not reflect the actual proportion in the population of MB patients which should be determined in a controlled trial cohort. Conclusively, high-throughput sequencing can be a useful supplement to histological diagnoses for recurrent MBs.

Accession codes
RNA sequencing datasets have been deposited to Sequence Read Archive (SRA) with bioproject number PRJNA418677. Whole exome sequencing datasets have been deposited to SRA with bioproject number PRJNA418669.

Acknowledgements
This study was supported by a grant from the National R&D Program for Cancer Control, Ministry for Health and Welfare, Republic of Korea (1420020). This work was also supported by the 2016 Research Fund (1.160052.01) of Ulsan National Institute of Science and Technology (UNIST) and the Collaborative Genome Program for Fostering New Post-Genome Industry of the National Research Foundation (NRF) funded by the Ministry of Science and ICT (MSIT) (NRF-2017M3C9A5031004).

Compliance with ethical standards
Conflict of interest The authors declare no potential conflicts of interest.

References
Affiliations

Ji Hoon Phi1,2 · Ae Kyung Park3 · Semin Lee4,5 · Seung Ah Choi1,2 · In-Pyo Baek6 · Pora Kim7 · Eun-Hye Kim8 · Hee Chul Park9 · Byung Chul Kim9 · Jong Bhak4,5 · Sung-Hye Park10 · Ji Yeoun Lee1,2,11 · Kyu-Chang Wang1,2 · Dong-Seok Kim12 · Kyu Won Shim12 · Se Hoon Kim13 · Chae-Yong Kim14 · Seung-Ki Kim1,2

1 Division of Pediatric Neurosurgery, Pediatric Clinical Neuroscience Center, Seoul National University Children’s Hospital, Seoul 03080, Republic of Korea
2 Department of Neurosurgery, Seoul National University Hospital, Seoul National University College of Medicine, Seoul 03080, Republic of Korea
3 College of Pharmacy and Research Institute of Life and Pharmaceutical Sciences, Sunchon National University, Suncheon 57922, Republic of Korea
4 Department of Biomedical Engineering, School of Life Sciences, Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, Republic of Korea
5 Korean Genomics Industrialization and Commercialization Center (KOGIC), Ulsan 44919, Republic of Korea
6 TheragenEtex Bio Institute, Suwon 16229, Republic of Korea
7 School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA
8 Gerotech Inc., Ulsan 44919, Republic of Korea
9 Clinomics Inc., Ulsan 44919, Republic of Korea
10 Department of Pathology, Seoul National University College of Medicine, Seoul 03080, Republic of Korea
11 Department of Anatomy, Seoul National University College of Medicine, Seoul 03080, Republic of Korea
12 Department of Pediatric Neurosurgery, Severance Children’s Hospital, Yonsei University College of Medicine, Brain Korea 21 Project for Medical Science, Seoul 03722, Republic of Korea
13 Department of Pathology, Severance Hospital, Yonsei University College of Medicine, Seoul 03722, Republic of Korea
14 Department of Neurosurgery, Seoul National University Bundang Hospital, Seongnam 13620, Republic of Korea