ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

ESMO OPEN SCIENCE FOR OPTIMAL CANCER CARE

ORIGINAL RESEARCH

Methylation subgroup and molecular heterogeneity is a hallmark of glioblastoma: implications for biopsy targeting, classification and therapy

J. Gempt[1], F. Withake[2], A. K. Aftahy[1], H. S. Meyer[1], M. Barz[1], C. Delbridge[3], F. Liesche-Starnecker[3], G. Prokop[3], N. Pfarr[4], J. Schlegel[3], B. Meyer[1], C. Zimmer[2], B. H. Menze[5†] & B. Wiestler[2,6†*]

Departments of [1]Neurosurgery; [2]Neuroradiology, Klinikum rechts der Isar, School of Medicine, Technical University Munich, Munich; Departments of [3]Neuropathology; [4]Pathology, Institute of Pathology, School of Medicine, Technical University Munich, Munich, Germany; [5]Department of Quantitative Biomedicine, University of Zurich, Zurich, Switzerland; [6]TranslaTUM, Technical University Munich, Munich, Germany

Available online 30 August 2022

Background: Intratumoral heterogeneity at the cellular and molecular level is a hallmark of glioblastoma (GB) that contributes to treatment resistance and poor clinical outcome. Little is known regarding epigenetic heterogeneity and intratumoral phylogeny and their implication for molecular classification and targeted therapies.

Patients and methods: Multiple tissue biopsies (238 in total) were sampled from 56 newly-diagnosed, treatment-naive GB patients from a prospective in-house cohort and publicly available data and profiled for DNA methylation using the Illumina MethylationEPIC array. Methylation-based classification using the glioma classifier developed by Ceccarelli et al. and estimation of the *MGMT* promoter methylation status via the MGMT-STP27 model were carried out. In addition, copy number variations (CNVs) and phylogeny were analyzed.

Results: Almost half of the patients (22/56, 39%) harbored tumors composed of heterogeneous methylation subtypes. We found two predominant subtype combinations: classic-/mesenchymal-like, and mesenchymal-/pilocytic astrocytoma-like. Nine patients (16%) had tumors composed of subvolumes with and without *MGMT* promoter methylation, whereas 20 patients (36%) were homogeneously methylated, and 27 patients (48%) were homogeneously unmethylated. CNV analysis revealed high variations in many genes, including *CDKN2A/B*, *EGFR*, and *PTEN*. Phylogenetic analysis correspondingly showed a general pattern of *CDKN2A/B* loss and gain of *EGFR*, *PDGFRA*, and *CDK4* during early stages of tumor development.

Conclusions: (Epi)genetic intratumoral heterogeneity is a hallmark of GB, both at DNA methylation and CNV level. This intratumoral heterogeneity is of utmost importance for molecular classification as well as for defining therapeutic targets in this disease, as single biopsies might underestimate the true molecular diversity in a tumor.

Key words: glioblastoma, heterogeneity, methylation, phylogeny

INTRODUCTION

Glioblastoma [GB, World Health Organization (WHO) grade 4] is a rapidly progressing malignant, primary brain tumor and is the most common such tumor in adults.[1,2] GB is characterized by extensive heterogeneity at the cellular level, which is also exemplified in its old name 'glioblastoma multiforme'. This heterogeneity is reflected in genomic aberrations and transcriptomic expression, which adds to the complexity of treating GB, as therapeutic response to radiation and chemotherapy varies between tumor clones.[3,4] Despite extensive research efforts motivated in part by high-throughput discoveries of promising molecular targets in GB, to date almost all phase II or phase III trials of molecularly targeted therapies in GB failed to improve survival.[5]

Molecular markers were integrated for the first time in the WHO 2016 classification and were refined in the recently published 2021 WHO classification of central nervous system tumors. Genetic alterations most frequently observed in *IDH* wild-type GB are *TERT* promoter mutations, *MGMT* promoter methylation, and mutations in the *PTEN* gene.[6,7] Copy number variation (CNV) is another key mechanism through which tumors undergo functional adaptation, frequently through loss of tumor suppressor genes or amplification of oncogenes. Besides classifications based on individual molecular markers, computational methods to leverage the rich information contained in (epi)

*Correspondence to: PD Dr med Benedikt Wiestler, Department of Neuroradiology, Klinikum rechts der Isar, School of Medicine, Technical University Munich, Ismaninger Str. 22, 81675 Munich, Germany. Tel: +49-89-4140-4651; Fax: +49-89-4140-4653

E-mail: b.wiestler@tum.de (B. Wiestler).

†These authors contributed equally as senior authors.

genome-wide data have increasingly been used for molecularly-driven stratification of tumors. As one such example, Ceccarelli et al.[7] report the discovery of six methylation subgroups in adult gliomas (WHO grades 2-4).

Many of these large studies which have greatly advanced our understanding of the molecular basis of GB rely, however, on a single biopsy per patient, thereby neglecting the possibly significant contribution of molecular tumor heterogeneity to GB biology. Importantly, the presence of multiple diverse subclones in a single tumor is very likely a major reason for treatment failure.[8] In order to successfully translate molecularly targeted therapies into clinical use, a better understanding of intratumoral heterogeneity and its implications for molecular classification and therapy of GB is necessary.

In GB, two studies have reported conflicting results on the presence of intratumoral methylation heterogeneity: whereas Wenger et al.[10] detected intratumoral heterogeneity in 5/12 tumors, Verburg et al.[9] reported a stable intratumoral methylation profile, in particular when adjusting for tumor purity.

To shed light on this matter, we investigated intratumoral DNA methylation heterogeneity in detail. Multiple tissue biopsies from 56 newly diagnosed, treatment-naive GB patients were sampled both from a prospective in-house cohort as well as publicly available data from Wenger et al.[10] and profiled for DNA methylation and CNV, forming the largest investigated cohort to date. We show that GB exhibits extensive methylation heterogeneity, ranging from global epigenome-wide level in terms of molecular classification down to single molecular targets. Based on our assessment of intratumoral phylogeny, we propose a theoretical evolutionary trajectory featuring distinct CNVs during early and late stages of disease progression.

MATERIAL AND METHODS

This study was approved by our local ethics committee (284/16S). All local patients were part of a prospective GB cohort from February 2018 to March 2021 and gave written informed consent.

Cohort and sampling

Multiple spatially separated biopsies (range 2-9, median = 4) were collected from 45 primary adult, newly diagnosed, treatment-naive GB patients. Biopsies from clearly separated targets as determined by the neurosurgeon were collected. All patients were classified as GB *IDH* wild type according to the 2016 WHO classification.[11] The tumor content of the samples was assessed by a neuropathologist and only samples with tumor content >70% were used for further analysis. Genome-wide methylation profiling using the Infinium MethylationEPIC BeadChip platform (Illumina, San Diego, CA) was carried out according to the manufacturer's specifications. Additionally, a similar dataset consisting of 35 GB samples obtained from 11 patients published by Wenger et al.[10] (GEO accession GSE116298) was included in the data analysis. In total, 238 samples from 56 patients were available for computational analysis. For all patients with available follow-up data, overall survival was noted.

Methylation array processing

All subsequent analyses were carried out using R (version 4.0.5) and Bioconductor (version 3.12).[12,13] Raw Illumina MethylationEPIC array data were preprocessed with the 'minfi' package.[14] Beta values were calculated in accordance with the reference implementation by the manufacturer.[15] Based on the total probe intensity, a detection *P* value was calculated for each CpG site that quantifies a confidence measure for the reported beta value. The total signal intensity for each CpG site was compared with the background signal intensity using negative control probes on the array. Samples with an average detection *P* value >0.01 were excluded from further analysis. In order to allow comparison of methylation levels obtained from probes of different types, subset-quantile within-array normalization (SWAN) was applied to the dataset.[16]

Methylation data analysis

For genomic annotations like CpG islands and genes, the UCSC genome browser and genome assembly GRCh37 (hg19) were used. EPIC array probes were mapped to the genome using the mapToGenome function in minfi.

Tumor purity of the tissue biopsies was assessed using the InfiniumPurify package.[17] To increase confidence in these estimates, additional purity predictions obtained from the RFpurify package[18] were compared, resulting in a strong correlation between the two (Pearson's $r = 0.923$; Supplementary Figure S1, available at https://doi.org/10.1016/j.esmoop.2022.100566).

CNV analysis was carried out via the conumee package.[19] Dr Martin Sill from the German Cancer Research Center (personal communication) provided a reference cohort consisting of 50 healthy male and 50 healthy female CNV profiles. To correct for different tumor purities, the resulting log2-ratios that quantify the copy number of a given genomic segment were scaled by the inverse of tumor purity, assuming a linear relationship between log2-ratio magnitude and purity.

For global tumor classification/subtyping of GB samples, the glioma classifier based on the work by Ceccarelli et al.[7] as implemented in the TCGAbiolinks package was used.[15] Possible classifications were mesenchymal-like, classic-like, pilocytic astrocytoma-like, G-CIMP-high, G-CIMP-low, and codel as explained in detail in the original publication.

Methylation status of the *MGMT* promoter was predicted using the MGMT-STP27 model with the default probability cut-off of 0.358. As recommended by the authors, non-normalized methylation values were used as input for this task.[20]

t-Distributed stochastic neighbor embedding (t-SNE) was carried out to transform high-dimensional data into a lower-dimensional space for clustering and visual representation using the Rtsne package.[21]

In order to investigate evolutionary trajectories and branching patterns across different tumor samples, an adapted version of the TuMult algorithm was used to analyze CNV profiles inferred from the methylation data.[22] In brief, the TuMult algorithm computes a phylogeny from microarray CNV data for multiple samples from the same patient. The input data for this algorithm are copy number values at n locations in the genome for k samples. Based on these CNV profiles, TuMult divides each chromosome into m segments of similar CNV status. The input data are reduced from [n k] continuous data points to [m k] integer values, where m << n. On this reduced dataset, TuMult fits a tree topology that reconstructs the tumor lineage. Iteratively, the two nodes that are most similar to each other are joined and their common precursor is inferred from the CNV events in common, building a tree from leaves to root. For this analysis, TuMult was modified such that it can be applied to CNV profiles generated from methylation data.

RESULTS

Patient cohort

In total, 56 patients with a treatment-naive, newly diagnosed GB according to the 2016 WHO classification were included in this study. Of these, 31 (55%) were male and 25 female (45%). Mean age of the in-house cohort was 68.2 years (standard deviation 12.6 years). All patients were diagnosed with an *IDH* wild-type GB. Spatially separated biopsies (range 2-9, median = 4) were stereotactically acquired immediately before tumor resection.

Heterogeneous tumor class predictions and global heterogeneity

Of all 238 samples, 127 samples (53%) were classified as mesenchymal-like, 93 (39%) as classic-like, and 18 (8%) as PA-like. Importantly, 22 patients (39%) harbored multiple glioma subtypes within a single tumor, displaying a strong intratumoral heterogeneity (Figure 1A). Classification heterogeneity was predominantly observed with respect to two subtype combinations: classic-/mesenchymal-like, and mesenchymal-/PA-like. Patient sfb14 was the only one assigned all three subtypes associated with *IDH* wild-type GB. The intratumoral heterogeneity was also reflected in a t-SNE representation generated from the same 1300 probes dataset evaluated by the Ceccarelli classifier (Figure 1B). Whereas mostly for homogenous samples, a close clustering of samples from the same patient was observed, subtype heterogeneous samples tended to cluster distantly to the remaining samples of the same patient. To account for differing tumor purities potentially confounding subtype prediction, the influence of purity on classification probabilities was evaluated, showing no apparent correlation between classifier confidence purity estimate (Supplementary Figure S2, available at https://doi.org/10.1016/j.esmoop.2022.100566).

To investigate the influence of methylation class heterogeneity on prognosis, we compared median overall survival for patients who harbored multiple methylation subtypes ($n = 20$; 'heterogeneous') with that of patients who were subtype-homogeneous ($n = 14$). Median overall survival for subtype-heterogeneous patients was shorter than for subtype-homogeneous patients {293 days [95% confidence interval (CI) 214 days-infinite] versus 477 days (95% CI 176-1145 days)}, although this was not statistically significant ($P = 0.51$, log-rank test and Supplementary Figure S3, available at https://doi.org/10.1016/j.esmoop.2022.100566).

Intratumor MGMT promoter methylation heterogeneity

Of all 56 patients, 20 (36%) carried a homogeneously methylated *MGMT* promoter and 27 patients (48%) were unmethylated in all their samples. Nine patients (16%), however, contained both methylated and unmethylated biopsy specimens (Figure 2). *MGMT* promoter methylation heterogeneity was particularly apparent in patient sfb19, for whom three samples were classified as methylated and two as unmethylated. All other MGMT-heterogeneous patients exhibited different classifications in just one single sample. A high classifier confidence for heterogeneous MGMT methylation was observed for patient sfb41. Sample sfb41_1 was assigned a methylation probability of 0.08, whereas for biopsy sfb41_3 a probability of 0.87 was computed.

Observed CNVs

As expected in GB, large-scale genomic instability in terms of CNV was observed. Most prominent were chromosome 7 gains, and chromosome 9p and 10 losses (Figure 3A). Median log2-ratios were averaged for each chromosome arm across all samples, with chromosomes 10p and 10q exhibiting mean losses of −0.33 and −0.34, respectively, and chromosomes 7p and 7q showing average gains of +0.22 and +0.2, respectively. Also very aberrant were chromosomes 9p, 13q, 14q, and 22, with mean log2-ratios of −0.15, −0.15, −0.09, and −0.10, respectively. The largest variance of copy number values across samples was observed for chromosomes 10, 13q, 9p, 14q, 22q, and 7.

Next, discrete CNV events were assigned using a threshold of ±0.5 for gains/losses and assessed on a gene level (Figure 3B), focusing on genes prominently involved in glioma biology and/or therapeutic targets. The most frequently observed event was loss of *CDKN2A/B* genes ($n = 177$). Other prominent events include loss of *TET1* ($n = 69$), *PTEN* ($n = 61$), *MGMT* ($n = 38$), and *RB1* ($n = 27$). Copy number gains were commonly observed in *EGFR* ($n = 102$), *PDGFRA* ($n = 43$), *CDK4* ($n = 43$), *MET* ($n = 27$), and *MDM4* ($n = 26$). Chromosomes 7 and 10, in particular, contain many genes associated with GB (e.g., *EGFR*, *CDK6*, *MET*, *PTEN*).

Intratumoral copy number heterogeneity

To assess intratumoral heterogeneity of CNV events, patient-wise standard deviations (SD) of the purity-corrected median log2-ratio in chromosomes and selected genes were calculated. *CDKN2A/B* loss was the most prominent and the most variant gene-level CNV among the genes investigated. In 26 patients, it was the CNV with the

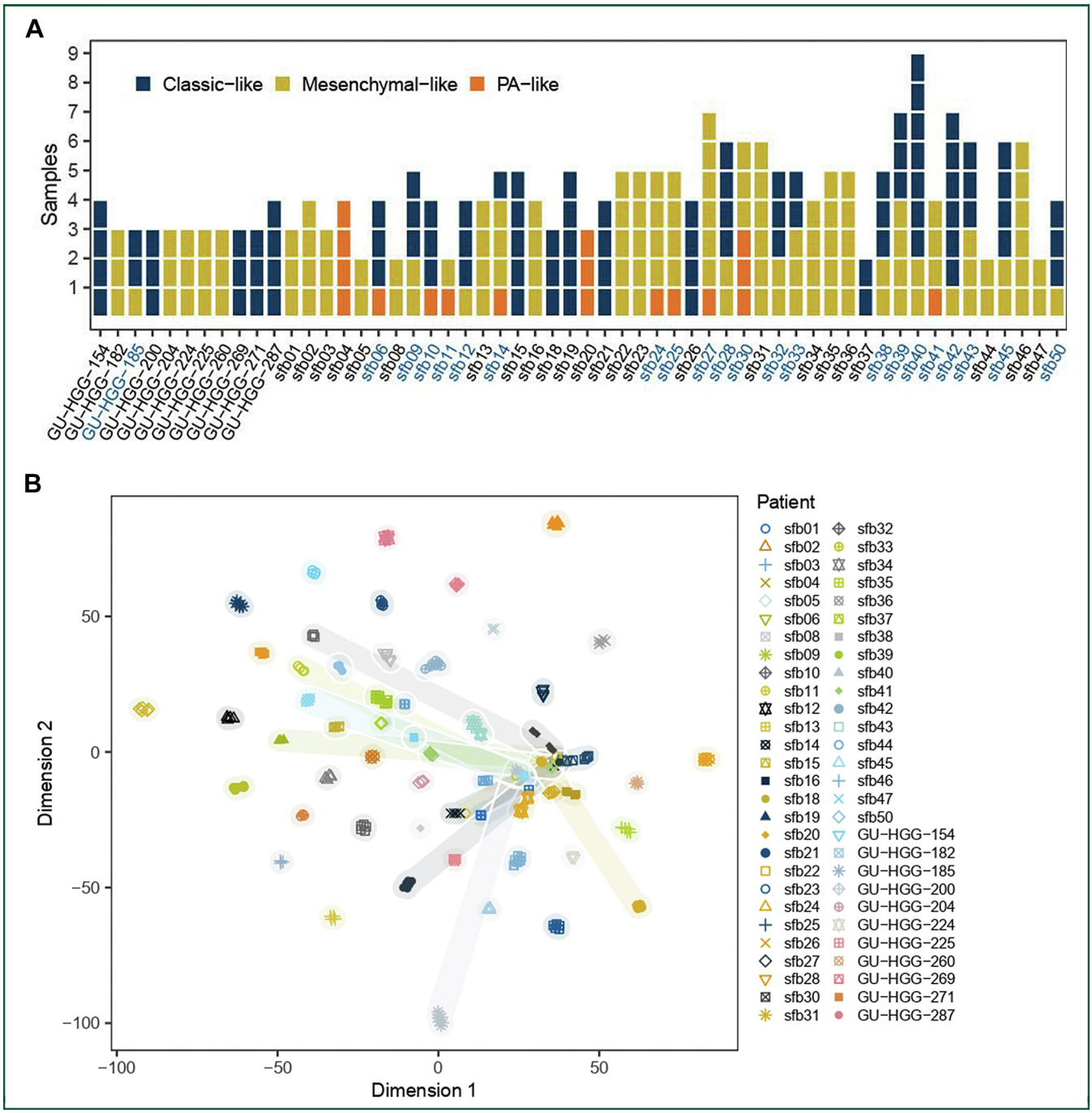

Figure 1. Methylation heterogeneity on an epigenome-wide level.
For each patient and biopsy, the respective classification from Ceccarelli et al.[7] is displayed and subtype-heterogeneous patients are labeled in blue (A). t-Distributed stochastic neighbor embedding clustering of biopsies reveals how samples from patients with heterogeneous methylation labels tend to disperse, highlighting their heterogeneity (B).
PA, pilocytic astrocytoma.

highest standard deviation and was the most intratumorally heterogeneous gene-level CNV (mean SD = 0.4). Second in terms of standard deviation was the *EGFR* gene, with an average intratumor mean SD = 0.19, followed by *PDGFRA* (mean SD = 0.14), *PTEN* (mean SD = 0.14), and *MDM4* (mean SD = 0.14) (Figure 4A).

Next, copy number gain and loss events were classified as common (in all samples for a given patient), shared (more than once in a patient), and unique (in only one sample for a given patient). Figure 4B illustrates how many of these events were detected at each location (on chromosome and gene level, respectively). From the detected gene-level CNV, on average, 34% were classified as common, 35% as shared, and 31% as unique. In some patients, none of the observed events were classified as common, whereas in other patients all events were common. Among those gene-level CNVs were *PDGFRA* gain, *EGFR* gain, *CDKN2A/B* deletion, and *CDK4* gain. The largest proportion of common events

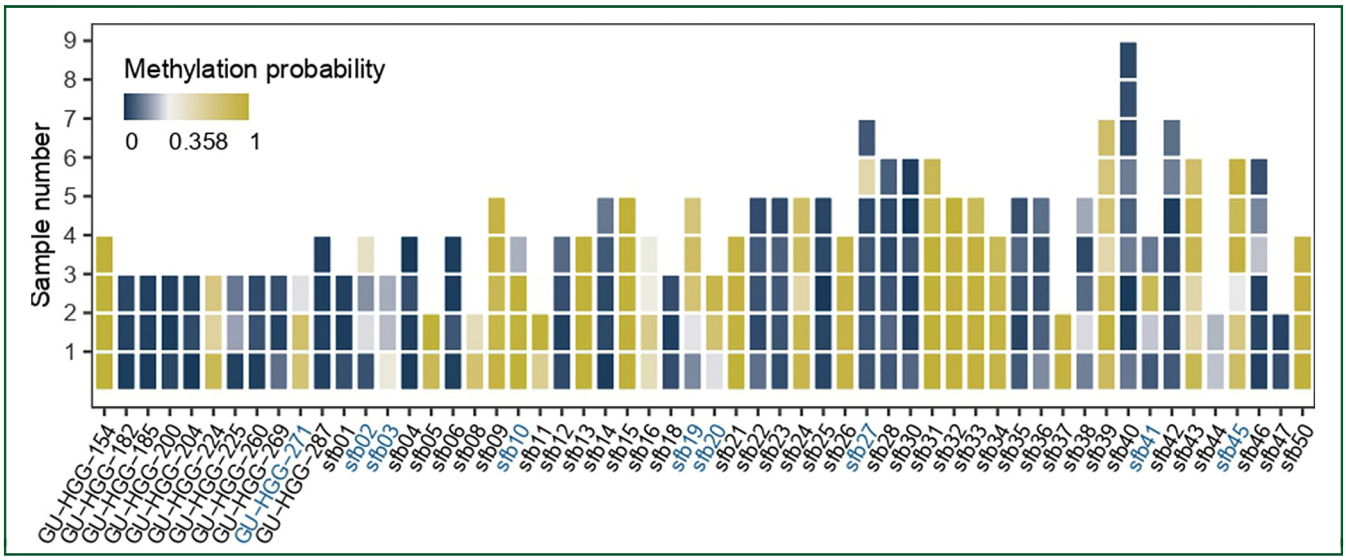

Figure 2. **Heterogeneity of *MGMT* promoter methylation. In 9/56 patients, areas with and without *MGMT* promoter methylation co-exist.**

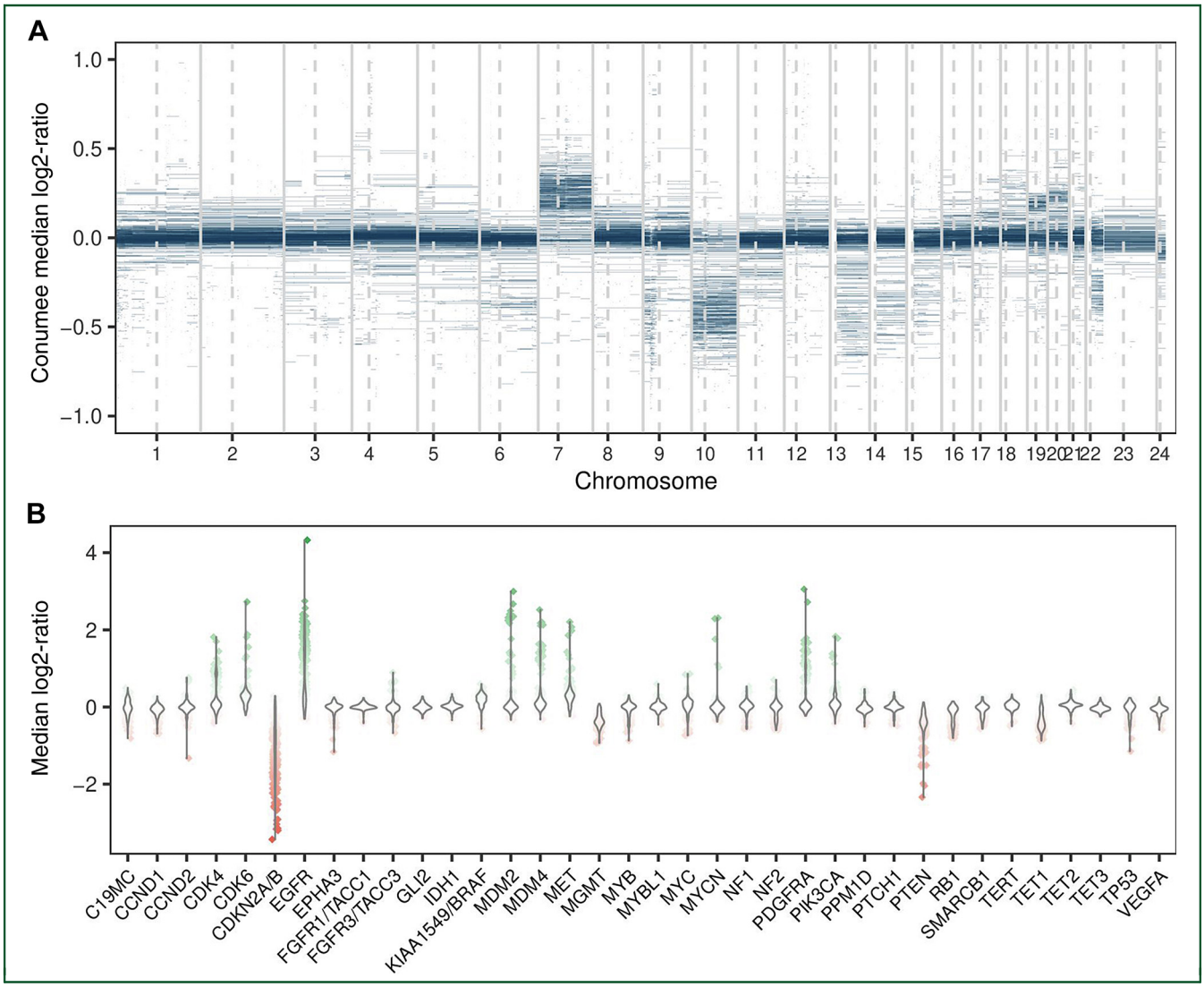

Figure 3. **Overview of copy number status for all biopsies on segment-wise (A) and gene-wise (B) level, showing 'typical' glioblastoma profiles such as (A) +7/−10 or (B) gain of *EGFR* and loss of *CDKN2A/B*, but also relevant variability.**

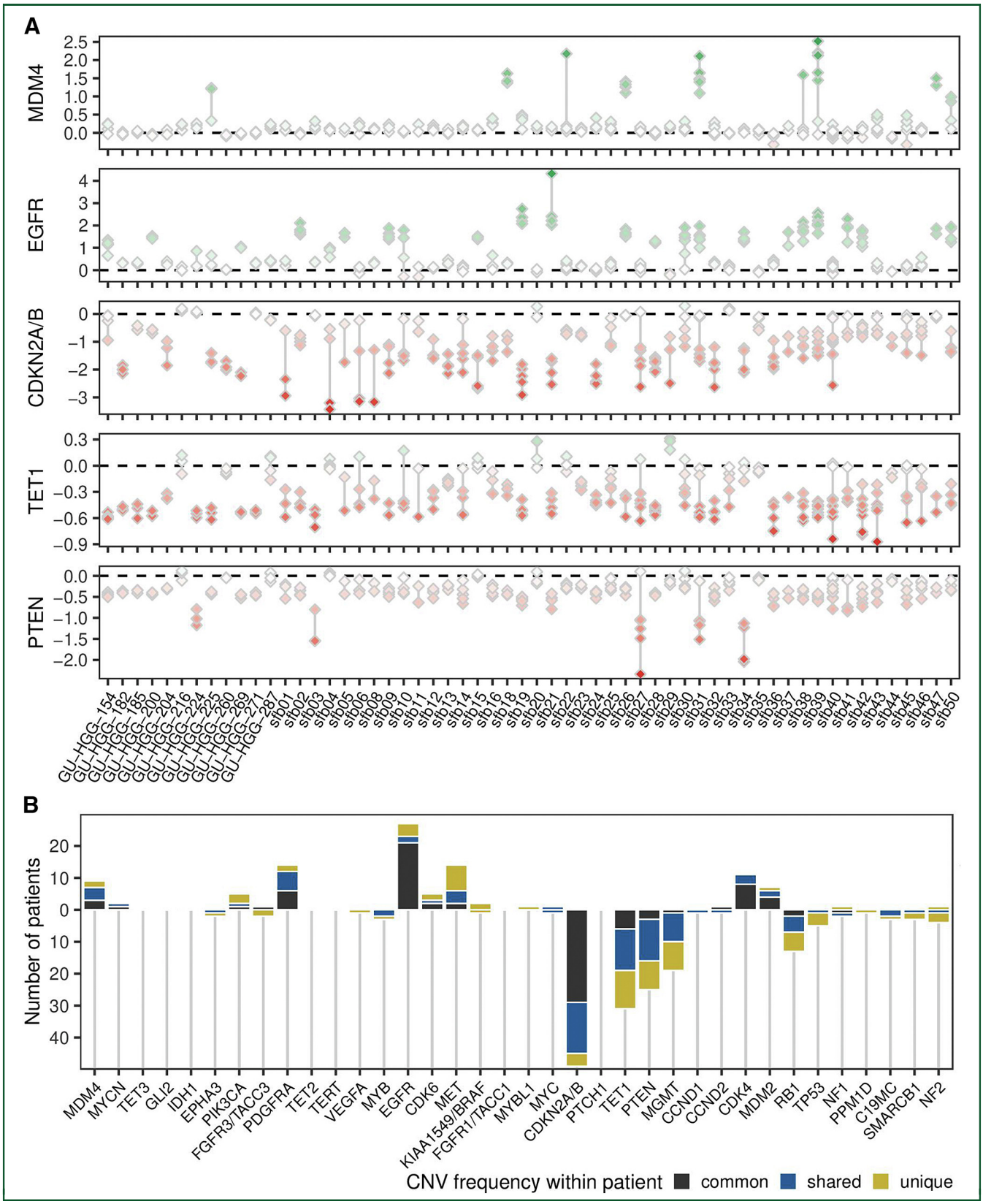

Figure 4. Individual CNV plots per sample for selected genes (A). In (B), genes were grouped as common (in all samples for a given patient), shared (more than once in a patient), and unique (in only one sample for a given patient).
CNV, copy number variation.

compared with shared and unique events was found in the *EGFR* gene. *CDKN2A/B* deletion also tended to occur homogeneously (common in 28 patients), but showed heterogeneity in 16 patients and was a unique event in only 4 patients. *CDK4* gain was observed as common in eight and shared in three patients, respectively, and *PDGFRA* gain was

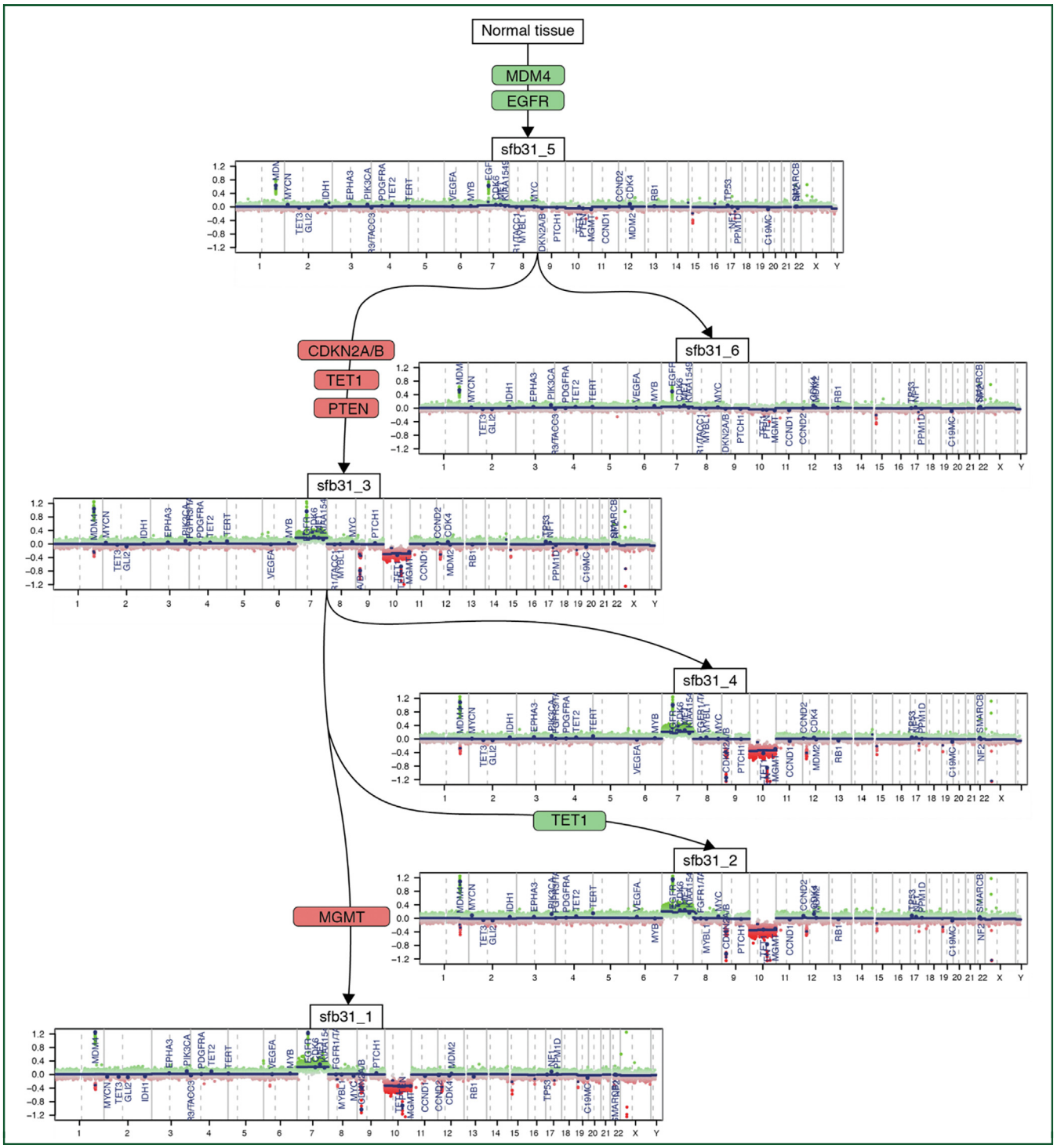

Figure 5. Exemplary phylogenetic tree for patient sfb31.
Here, *EGFR* gain and *CDKN2A/B* loss are precursor alterations occurring near the root, followed by a branching out of the tree.

common in five, shared in six, and unique in two cases. Copy number aberrations frequently classified as unique involved genes *MET* (8/14), *TET1* (12/31), *PTEN* (9/25), and *NF2* (4/5).

Evolutionary trajectory via phylogenetic trees

To investigate the evolution of intratumoral heterogeneity, phylogenetic trees were generated with a modified version of the TuMult algorithm for all patients with three or more samples ($n = 59$). An example of such an evolutionary tree, generated for patient sfb31, is shown in Figure 5. Eleven patients were assigned a strictly linear evolutionary trajectory.

Most prominent was *CDKN2A/B* deletion as the first event (31/48). In the remaining 17 patients exhibiting *CDKN2A/B* loss, this event was always observed in very early stages of the phylogeny. In addition, *EGFR* amplification is among

https://doi.org/10.1016/j.esmoop.2022.100566

those events occurring at early stages of tumor development. Of all 27 patients exhibiting *EGFR* gain, this event appeared at the root of tumor diversification in 22 cases. *PDGFRA* gain also occurred early during tumorigenesis in 8/13 cases. *CDK4* gain was also frequently observed as an initial event (9/11). *PTEN* loss and *MET* gain, however, were found to occur during the middle and late stages of tumor evolution. *TET1* loss, *MGMT* loss, and *MDM4* gain were placed heterogeneously in the phylogenies.

In summary, evolutionary trajectory of GB as inferred by phylogenetic analysis could be described as follows: *CDKN2A/B* loss and gain of *EGFR*, *PDGFRA*, and *CDK4* genes occurred in the early stages of tumor development. This was accompanied by chromosome 7 gain and chromosome 10 loss. At later phases of disease progression, subclone populations developed individual variations involving loss of *TET1*, *PTEN*, and *MGMT* genes as well as *MDM4* and *MET* gain.

DISCUSSION

Our study revealed extensive intratumoral heterogeneity on methylation and copy number levels, affecting both molecular classification of tumors and therapeutically relevant molecular markers. In addition, reconstruction of intratumoral evolution revealed a clear pattern of *CDKN2A/B* loss and gain of *EGFR*, *PDGFRA*, and *CDK4* during early stages of tumor development. With increasing insights into the biology of gliomas, stemming largely from high-throughput studies, molecular-based classifications of gliomas have been proposed.[23-25] Importantly, by explicitly incorporating the molecular basis of tumors, such classifications promise an improved prognostication of tumors and identification of potential therapeutic targets. Individual studies have indeed identified molecular subgroups with benefit from a specific therapy, such as patients with a proneural GB potentially having improved survival with the addition of bevacizumab to temozolomide.[26] This enthusiasm has not been met with success in targeted phase II and III trials, however, where targeted therapies, such as EGFR inhibition, failed to improve the survival of patients. Our study contributes to an emerging picture[4] of intratumoral heterogeneity as a main challenge to a successful implementation of targeted therapies in clinical routine. We observe methylation subgroup heterogeneity in almost half of the patients (40%). This is corroborated by a two-dimensional t-SNE embedding of methylation data (Figure 1B). Although many samples of a patient tend to form clusters, this analysis also demonstrated a clear spatial shift of samples away from the other samples of the respective patient. In view of reports attributing differential effects of therapies on different molecular subtypes, this finding clearly highlights the importance of considering intratumoral heterogeneity for treatment decisions. Further, we observed a relevantly longer median overall survival for subtype-homogeneous patients (477 versus 293 days, Supplementary Figure S3, available at https://doi.org/10.1016/j.esmoop.2022.100566). Although this difference was not statistically significant (likely due to sample size), this matches another recent report on the influence of intratumoral heterogeneity on outcome[27] and underlines the importance of future studies in this direction.

Predicting the *MGMT* promoter methylation status revealed a heterogeneous methylation status in nine patients. Intratumoral *MGMT* methylation heterogeneity was previously documented,[8] and could be confirmed by us. Given the importance of *MGMT* promoter methylation for therapy stratification,[28] this is another clear indication of the role of intratumoral heterogeneity for therapy resistance in GB. Besides methylation, we also report major heterogeneity in copy number events. We found genome-wide chromosomal instability affecting chromosomes 7, 9p, 10, 13, 14, and 15. The observed frequencies of genetic CNVs accurately coincide with previous reported studies.[5] *EPHA3* overexpression, which is investigated as a target for novel therapies, was not reflected in CNV status as it was never significantly amplified and in some cases was reduced in copy number. We also report a high intratumoral variability of *CDKN2A/B* loss. In view of the role of *CDKN2A/B* loss for prognostication[29] and classification,[1,30] this again cautions against a potential sampling bias with (false-negative) assessment of *CDKN2A/B* status in neuropathological routine assessment of gliomas.

Lastly, we investigated intratumoral evolutionary trajectories from copy number data. We were able to reconstruct a general pattern of *CDKN2A/B* loss and gain of *EGFR*, *PDGFRA*, and *CDK4* during early stages of tumor development, followed by observation of individual variation in *TET1*, *PTEN*, and *MGMT* genes in later phases of disease progression. This is partially in line with other findings, which also identified *EGFR* and *CDKN2A/B* as tumor-initiating events.[31] In contrast to that report, we identified *PDGFRA* gain as another early event in patient-wise tumor phylogenies. 'RTK I PDGFRA', a methylation subgroup defined by Sturm et al.[23] is indeed characterized by *PDGFRA* amplification, supporting our result of *PDGFRA* gain as an early (and thus common) event in intratumoral evolution. Nonetheless, future, large-scale studies will be necessary to provide further insight. It remains to be investigated if the evolutionary trajectory constructed through CNV-based phylogenetic trees is also represented in the spatial positions of the tissue biopsies and if it is consistent with the local dynamics of tumor growth.

A recently published study by Verburg et al.[9] investigated the methylation landscape of 16 adult patients with diffuse gliomas. In contrast to our results, they find that methylation subtypes are conserved in space and much of the heterogeneity seen can be explained by differences in tumor purity. We investigated the relationship between methylation class probability and tumor purity and found no correlation (Supplementary Figure S2, available at https://doi.org/10.1016/j.esmoop.2022.100566; Pearson's $r = -0.039$). Further, we also found methylation class heterogeneity between samples from the same patient when both samples had high (>60%) tumor cell content, clearly indicating that even when considering purity as a potential confounder,

intratumoral heterogeneity is a defining trait of a GB. Along this line, in the purity-corrected CNV data we also observed evident signs of heterogeneity.

Classifiers evaluating genome-wide molecular data like DNA methylation that use machine learning algorithms have shown promising results and are currently being tested in clinical settings.[32,33] One problem of this glioma subtyping is widespread heterogeneity, as demonstrated here. Each tissue biopsy can only capture tumor properties at that single point in space and time in a complex and dynamical system of evolving subclone populations. As microarray- and sequencing-based genotyping technologies become cheaper, algorithmic tumor classification from multiple tissue samples might present a feasible way of making routine analysis cost and time effective while explicitly incorporating heterogeneity in this assessment, making a comprehensive diagnosis more accessible.

Another potential way forward lies in advanced, computational assessment of magnetic resonance imaging.[34] Radiomics is a rapidly advancing field, linking the imaging phenotype to the underlying genotype. Previous studies have, for example, demonstrated how assessment of perfusion imaging can predict *IDH* status and the activity of angiogenic pathways in gliomas.[35] Given its 'bird's eye view' of the entire tumor, imaging and advanced image analysis provide a fascinating opportunity to better characterize and unravel intratumoral heterogeneity, as well as its development over time.

Study limitations

This work benefited from a reasonably large study cohort (56 patients, 238 samples) compared with similar studies investigating GB intratumor methylation heterogeneity.[31] There are, however, some limitations. First of all, no distinction between 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC) was made. Although the role of 5hmC in the genome is receiving increasing attention, particularly in the context of cerebral tissue,[36] methods involving standard bisulfite treatment are incapable of differentiating between the two. It remains to be investigated if patterns and dynamics of 5hmC can be distinguished from 5mC profiles in the GB methylome. In addition, the predefined selection of CpG sites on the EPIC array should be kept in mind, as it queries only ∼850 000 of the overall 30 million CpG dinucleotides in the human genome. As such, the effect of open-sea non-gene-related DNA methylation might be underrepresented. Since most of the probes found to be differentially methylated within tumors were located in open-sea regions, their role remains to be investigated. Approaches to 'bridge this gap', however, are being developed.[37]

The range of biopsies varies from two to nine per patient and was dictated by size and resectability of tumors. Therefore, it is possible that our study even underestimates the true extent of intratumoral heterogeneity (for example missing subtype heterogeneity in patients with only two samples).

Furthermore, part of the observed intratumor heterogeneity might be related to differing concentrations of tumor cells in the samples. Through manual assessment by a neuropathologist, however, a minimum tumor content of 70% for each sample was ensured, which somewhat reduces the possible variability of sample-wise tumor cell concentration.

The process of identifying CNV events using absolute thresholds on conumee log2-ratios causes a loss of information and disregards the complex nature of the measurement process that is affected by differing concentrations of tumor cells and subclone populations in each sample, thus many CNV events that are actually present in the data are ignored.

Conclusion

GB exhibits extensive heterogeneity at both the DNA methylation and copy number levels, affecting molecular classification of tumors as well as therapeutic targets. The observed heterogeneity within tumors is therefore of utmost clinical importance, such that multiple tumor biopsies may be mandatory in future to effectively characterize this disease. Phylogenetic analysis of tumors revealed patterns of intratumoral evolution, with a general pattern of *CDKN2A/B* loss and gain of *EGFR*, *PDGFRA*, and *CDK4* during early stages of tumor development.

ACKNOWLEDGEMENTS

We thank Dr. Martin Sill (DKFZ Heidelberg, Germany) for providing a reference cohort for CNV annotation.

FUNDING

This work was supported by the Deutsche Forschungsgemeinschaft (DFG), SFB-824, subproject B12 (no grant number).

DISCLOSURE

The authors have declared no conflicts of interests.

REFERENCES

1. Louis DN, Perry A, Wesseling P, et al. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Neuro Oncol*. 2021;23(8):1231-1251.
2. Miller KD, Ostrom QT, Kruchko C, et al. Brain and other central nervous system tumor statistics, 2021. *CA Cancer J Clin*. 2021;71(5):381-406.
3. Lan X, Jörg DJ, Cavalli FMG, et al. Fate mapping of human glioblastoma reveals an invariant stem cell hierarchy. *Nature*. 2017;549(7671):227-232.
4. Lee JK, Wang J, Sa JK, et al. Spatiotemporal genomic architecture informs precision oncology in glioblastoma. *Nat Genet*. 2017;49(4):594-599.
5. Taylor OG, Brzozowski JS, Skelding KA. Glioblastoma multiforme: an overview of emerging therapeutic targets. *Front Oncol*. 2019;9:963.
6. Brennan CW, Verhaak RGW, McKenna A, et al. The somatic genomic landscape of glioblastoma. *Cell*. 2013;155(2):462-477.
7. Ceccarelli M, Barthel FP, Malta TM, et al. Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. *Cell*. 2016;164(3):550-563.
8. Parker NR, Hudson AL, Khong P, et al. Intratumoral heterogeneity identified at the epigenetic, genetic and transcriptional level in glioblastoma. *Sci Rep*. 2016;6:22477.

9. Verburg N, Barthel FP, Anderson KJ, et al. Spatial concordance of DNA methylation classification in diffuse glioma. *Neuro Oncol*. 2021;23(12):2054-2065.
10. Wenger A, Ferreyra Vega S, Kling T, Bontell TO, Jakola AS, Carén H. Intratumor DNA methylation heterogeneity in glioblastoma: implications for DNA methylation-based classification. *Neuro Oncol*. 2019;21(5):616-627.
11. Louis DN, Perry A, Reifenberger G, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol*. 2016;131(6):803-820.
12. Huber W, Carey VJ, Gentleman R, et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods*. 2015;12(2):115-121.
13. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. 2020. Available at https://www.R-project.org. Accessed February 15, 2022.
14. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*. 2014;30(10):1363-1369.
15. Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F. Evaluation of the Infinium Methylation 450K technology. *Epigenomics*. 2011;3(6):771-784.
16. Maksimovic J, Gordon L, Oshlack ASWAN. Subset-quantile Within Array Normalization for Illumina Infinium HumanMethylation450 BeadChips. *Genome Biol*. 2012;13(6):R44.
17. Qin Y, Feng H, Chen M, Wu H, Zheng X. InfiniumPurify: An R package for estimating and accounting for tumor purity in cancer methylation research. *Genes Dis*. 2018;5(1):43-45.
18. Johann PD, Jäger N, Pfister SM, Sill M. RF_Purify: a novel tool for comprehensive analysis of tumor-purity in methylation array data based on random forest regression. *BMC Bioinformatics*. 2019;20(1):428.
19. Hovestadt V, Zapatka M. Conumee: Enhanced Copy-Number Variation Analysis Using Illumina DNA Methylation Arrays. Available at http://bioconductor.org/packages/conumee/. Accessed February 15, 2022.
20. Bady P, Sciuscio D, Diserens AC, et al. MGMT methylation analysis of glioblastoma on the Infinium methylation BeadChip identifies two distinct CpG regions associated with gene silencing and outcome, yielding a prediction model for comparisons across datasets, tumor grades, and CIMP-status. *Acta Neuropathol*. 2012;124(4):547-560.
21. Maaten LJP van der, Hinton GE. Visualizing high-dimensional data using t-SNE. *J Mach Learn Res*. 2008;9:2579-2605.
22. Letouzé E, Allory Y, Bollet MA, Radvanyi F, Guyon F. Analysis of the copy number profiles of several tumor samples from the same patient reveals the successive steps in tumorigenesis. *Genome Biol*. 2010;11(7):R76.
23. Sturm D, Witt H, Hovestadt V, et al. Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. *Cancer Cell*. 2012;22(4):425-437.
24. Verhaak RGW, Hoadley KA, Purdom E, et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell*. 2010;17(1):98-110.
25. Wiestler B, Capper D, Sill M, et al. Integrated DNA methylation and copy-number profiling identify three clinically and biologically relevant groups of anaplastic glioma. *Acta Neuropathol*. 2014;128(4):561-571.
26. Sandmann T, Bourgon R, Garcia J, et al. Patients with proneural glioblastoma may derive overall survival benefit from the addition of bevacizumab to first-line radiotherapy and temozolomide: retrospective analysis of the AVAglio trial. *J Clin Oncol*. 2015;33(25):2735-2744.
27. Liesche-Starnecker F, Mayer K, Kofler F, et al. Immunohistochemically characterized intratumoral heterogeneity is a prognostic marker in human glioblastoma. *Cancers*. 2020;12(10):E2964.
28. Hegi ME, Diserens AC, Gorlia T, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med*. 2005;352(10):997-1003.
29. Shirahata M, Ono T, Stichel D, et al. Novel, improved grading system(s) for IDH-mutant astrocytic gliomas. *Acta Neuropathol*. 2018;136(1):153-166.
30. Brat DJ, Aldape K, Colman H, et al. cIMPACT-NOW update 5: recommended grading criteria and terminologies for IDH-mutant astrocytomas. *Acta Neuropathol*. 2020;139(3):603-608.
31. Sottoriva A, Spiteri I, Piccirillo SGM, et al. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc Natl Acad Sci U S A*. 2013;110(10):4009-4014.
32. Capper D, Jones DTW, Sill M, et al. DNA methylation-based classification of central nervous system tumours. *Nature*. 2018;555(7697):469-474.
33. Capper D, Stichel D, Sahm F, et al. Practical implementation of DNA methylation and copy-number-based CNS tumor diagnostics: the Heidelberg experience. *Acta Neuropathol*. 2018;136(2):181-210.
34. Wiestler B, Menze B. Deep learning for medical image analysis: a brief introduction. *Neuro Oncol Adv*. 2021;2(suppl 4):iv35-iv41.
35. Kickingereder P, Sahm F, Radbruch A, et al. IDH mutation status is associated with a distinct hypoxia/angiogenesis transcriptome signature which is non-invasively predictable with rCBV imaging in human glioma. *Sci Rep*. 2015;5:16238.
36. Sun W, Zang L, Shu Q, Li X. From development to diseases: the role of 5hmC in brain. *Genomics*. 2014;104(5):347-351.
37. Huang Y, Sun X, Jiang H, et al. A machine learning approach to brain epigenetic analysis reveals kinases associated with Alzheimer's disease. *Nat Commun*. 2021;12(1):4472.