Acta Neuropathol (2015) 130:419–434
DOI 10.1007/s00401-015-1427-y

ORIGINAL PAPER

Molecular profiling of long-term survivors identifies a subgroup of glioblastoma characterized by chromosome 19/20 co-gain

Christoph Geisenberger[1] · Andreas Mock[1,2] · Rolf Warta[1] · Carmen Rapp[1] · Christian Schwager[2,3] · Andrey Korshunov[4,5] · Ann-Katrin Nied[1] · David Capper[4,5] · Benedikt Brors[6,7,8,9] · Christine Jungk[1] · David Jones[10] · V. Peter Collins[11] · Koichi Ichimura[12] · L. Magnus Bäcklund[13] · Elena Schnabel[2,3,9] · Michel Mittelbron[14] · Bernd Lahrmann[15] · Siyuan Zheng[16] · Roel G. W. Verhaak[16] · Niels Grabe[15] · Stefan M. Pfister[10,17] · Christian Hartmann[18] · Andreas von Deimling[4,5] · Jürgen Debus[3] · Andreas Unterberg[1] · Amir Abdollahi[2,3,9] · Christel Herold-Mende[1]

Received: 13 January 2015 / Revised: 4 April 2015 / Accepted: 18 April 2015 / Published online: 1 May 2015

Abstract Glioblastoma (GBM) is a devastating tumor and few patients survive beyond 3 years. Defining the molecular determinants underlying long-term survival is essential for insights into tumor biology and biomarker identification. We therefore investigated homogeneously treated, IDH^{wt} long-term (LTS, $n = 10$) and short-term survivors (STS, $n = 6$) by microarray transcription profiling. While there was no association of clinical parameters and molecular subtypes with long-term survival, STS tumors were characterized by differential polarization of infiltrating microglia with predominance of the M2 phenotype detectable both on the mRNA and protein level. Furthermore, transcriptional signatures of LTS and STS predicted patient outcome in a large, IDH^{wt} cohort ($n = 468$). Interrogation of overlapping genomic alterations identified concurrent gain of chromosomes 19 and 20 as a favorable prognostic marker. The strong association of this co-gain with survival was validated by aCGH in a second, independent cohort ($n = 124$). Finally, FISH and gene expression data revealed gains to constitute low-amplitude, clonal events with a strong impact on transcription. In conclusion, these findings provide important insights into the manipulation of the innate immune system by particularly aggressive GBM tumors. Furthermore, we genomically characterize a previously unknown, clinically relevant subgroup of glioblastoma, which can easily be identified through modern neuropathological workup.

C. Geisenberger, A. Mock, A. Abdollahi and C. Herold-Mende have contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00401-015-1427-y) contains supplementary material, which is available to authorized users.

✉ Christel Herold-Mende
H.Mende@med.uni-heidelberg.de

1 Division of Experimental Neurosurgery, Department of Neurosurgery, Heidelberg University Hospital, Im Neuenheimer Feld 400, 69120 Heidelberg, Germany

2 Molecular and Translational Radiation Oncology, National Center for Tumor Diseases (NCT), German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany

3 Department of Radiation Oncology, National Center for Radiation Research in Oncology (NCRO), Heidelberg Ion-Beam Therapy Center (HIT), Heidelberg Institute of Radiation Oncology (HIRO), University of Heidelberg, 69120 Heidelberg, Germany

4 Clinical Cooperation Unit Neuropathology, German Cancer Research Center (DKFZ), Institute of Pathology, Heidelberg University Hospital, 69120 Heidelberg, Germany

5 Department of Neuropathology, Institute of Pathology, Heidelberg University Hospital, 69120 Heidelberg, Germany

6 National Center for Tumor Diseases (NCT), Medical Oncology, Heidelberg University Hospital, 69120 Heidelberg, Germany

7 Division of Applied Bioinformatics, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany

8 Computational Oncology, Division of Biostatistics, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany

9 German Consortium for Translational Cancer Research (DKTK), Heidelberg, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany

10 Division of Pediatric Neurooncology, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany

Keywords Glioblastoma · Molecular profiling · Microglia · Long-term survival · Short-term survival

Introduction

Glioblastoma (GBM) is the most common primary brain tumor in adults and ranks among the deadliest human cancers. Despite advances in surgical technique and improved chemoradiotherapy [11], median survival is less than 15 months [50]. Intriguingly, a small proportion (16 %) of GBM patients receiving intensified therapy survives for more than 36 months [49]. These so-called long-term survivors (LTS) are usually younger and have a higher Karnofsky Performance Score (KPS) at the time of diagnosis [31, 58]. Recently, extending prior attempts to subclassify glioblastoma [44], The Cancer Genome Atlas (TCGA) have suggested four subtypes (proneural, neural, classical and mesenchymal) based on gene expression patterns [55]. In addition, classification based on DNA methylation has been proposed [51, 52]. However, there are only marginal differences in survival between these subtypes and imperfect overlap with key genetic aberrations. The long-term clinical impact of these classifications is thus unclear. Further studies by the TCGA have revealed that a number of proneural GBMs show genome-wide hypermethylation [42]. Tumors with this so-called glioma CpG island methylator phenotype (G-CIMP) almost uniformly carry mutations in isocitrate dehydrogenase isoenzymes (*IDH1* and *IDH2*). Indeed, the vast majority of mutations (>95 %) occur in *IDH1* and among these, R132H substitutions are by far the most common [22]. It is now recognized that remodeling of the active center of *IDH* enzymes leads to production of 2-hydroxyglutarate (2-HG) [57]. Elevated levels of this metabolite in turn interfere with a number of enzymes involved in epigenetic regulation, resulting in global hypermethylation [54]. Thus, a causal link between *IDH* mutations and G-CIMP has been established. As *IDH* mutations are observed in the vast majority of lower grade gliomas (LGGs) and secondary GBMs (sGBMs) [59], it has been speculated that GBMs harboring this specific lesion are in fact derived from clinically silent precursor lesions and do not constitute true primary glioblastomas (pGBMs). This is further supported by results from the TCGA consortium and others who have shown that IDH^{mut} and G-CIMP tumors have a much better outcome and lack prototypical genetic alterations of pGBM such as *EGFR* amplification, *CDKN2A* deletions, and *PTEN* loss [2, 42, 52, 55]. Unsurprisingly, a large number of studies have found a higher proportion of IDH^{mut} [21] and G-CIMP tumors [47] among LTS. These data imply that available studies on long-term survival may have been confounded by *IDH* mutations, and further research is warranted to elucidate the mechanisms contributing to improved survival in primary, IDH^{wt} glioblastoma.

Materials and methods

Patient data

All patients received surgery at the Department of Neurosurgery (University Hospital Heidelberg, Germany) and were treated with post-operative chemoradiotherapy as well as six cycles of adjuvant temozolomide-based chemotherapy. To ensure completion of the therapy regimen, only patients who survived for more than 7 months after surgery were included. Diagnosis of glioblastoma was confirmed through histopathological review by independent, board-certified neuropathologists (AvD, DC, CH). In cases where *IDH1* data was not available as part of routine diagnostic workup, mutations were ruled out by sequencing and immunohistochemistry as described elsewhere [5, 22]. Clinical and outcome data for patients treated at the Department of Neurosurgery (University Hospital Heidelberg, Germany) were obtained through review of patient's charts. Detailed patient information is provided in Table S1.

Patient material, quality control, and RNA extraction

Tumor material was obtained following surgical resection. Tissue was snap-frozen and stored at −80 °C until further processing. RNA was extracted with TRIzol® RNA Isolation Reagents (Life Technologies) from samples

[11] Department of Pathology, University of Cambridge, Addenbrooke's Hospital, Cambridge CB2 0QQ, UK

[12] Division of Brain Tumor Translational Research, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan

[13] Department of Oncology-Pathology, Cancer Center Karolinska and Radiumhemmet Karolinska Institute and University Hospital, R8:00, 171 76 Stockholm, Sweden

[14] Neurological Institute (Edinger-Institute), Goethe University, 60528 Frankfurt, Germany

[15] Hamamatsu Tissue Imaging and Analysis Center (TIGA), BIOQUANT, University of Heidelberg, 69120 Heidelberg, Germany

[16] Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

[17] Department of Pediatric Oncology, Hematology and Immunology, Heidelberg University Hospital, 69120 Heidelberg, Germany

[18] Department of Neuropathology, Institute of Pathology, Hannover Medical School, 30625 Hannover, Germany

found eligible in terms of tumor cell content (>60 %) and necrosis (<20 %). Analyte concentration and quality were determined using the Nanodrop 2000 spectrophotometer (Thermo Scientific) and Bioanalyzer 2100 (Agilent).

Microarray experiments

All specimens were assayed on the Agilent SurePrint G3 Human Exon 2 × 400 K platform. Labeling and hybridization reactions were performed according to the manufacturer's protocols. Probe sequences were aligned to the Ensembl transcript database, filtered based on BLAST quality metrics and combined into transcript expression estimates. Median transcript expression was used as an expression estimate ($n = 21{,}389$). The resulting data were normalized using VSN [26], transformed to $\log_2$ scale and median centered. Differential expression between LTS and STS was assessed using Student's *t* test with correction for unequal variances. Exploratory data analysis (Principal Component Analysis) was conducted within the *R* statistical software environment [53]. GeneGo's Metacore software tool from Thomson Reuters was used for all pathway analyses.

Statistical analyses

GraphPad Prism was used for statistical analyses for patient data of the discovery cohort. All other analyses were conducted within the *R* software environment. Numerical data were compared using Student's *t* test with correction for unequal variances. Ordinal data were analyzed using non-parametric statistics (Mann–Whitney *U* test). Overlap of categorical variables was assessed using the χ^2 test or Fisher's exact test where applicable. Survival analysis was conducted using the "survival" *R* package. Log-rank tests and Cox' proportional hazards model were used for univariate and multivariate comparisons, respectively.

Assessment of subtypes in discovery cohort

Centroids established by Verhaak et al. [55] for subtyping of GBM expression data were downloaded from the TCGA website (accompanying data freeze released with aforementioned publication). Pearson's *r* for each of the four centroids was calculated for all samples and all available genes ($n = 776$). Each sample was then assigned the subtype of the centroid it was most highly correlated with (Fig. S1).

Multicolor immunostaining, image analysis and evaluation

Acetone-fixed cryostat sections (5–7 μm) were used for triple immunofluorescence staining. Microglia were detected using mouse anti-CD68 (Dako). M2 Polarization of microglia cells was evaluated by mouse anti-CD163 (AbD Serotec) and rabbit anti-CD204 (Sigma-Aldrich) stainings [45]. Primary antibodies were diluted in Antibody Diluent (Dako) and incubated for 1 h at room temperature. For the detection of the primary antibodies CD163 and CD204, anti-mouse AlexaFluor647 (Invitrogen) and anti-rabbit AlexaFluor555 (Invitrogen) were used and diluted in PBS (Gibco). DAPI (Invitrogen) was used to counterstain nuclei. CD68 was detected using Zenon technology according to the manufacturers' protocols (Invitrogen).

Specificity of primary and secondary antibodies was ensured by isotype-matched controls and negative controls (without primary antibody), respectively. After application of both primary and secondary antibodies, three washing steps were performed in PBS containing 0.05 % Tween. Slides were mounted with Elvanol (Calbiochem). Evaluation was performed with the TissueQuest system (TissueGnostics) which allows for automated quantification of multiple markers in whole tissue sections. Complete tissue sections were recorded with the Olympus IX51 microscope (at 20 × resolution) equipped with a XM10 camera and the Cell Sense software (all Olympus). Tissue sections were analyzed with the TissueQuest 4.0 software (TissueGnostics) through evaluation of median staining intensity of cells in FACS-like scattergrams [34]. All analyses have been conducted by an independent, blinded researcher.

Macrophage expression signature

A gene expression signature reported by Martinez and coworkers [37] was downloaded from the supplemental files provided with the publication. Signature genes ($n = 97$) were filtered for those also available in our dataset ($n = 83$). Fold changes were provided as positive and negative values for higher expression in M1 and M2 polarized macrophages, respectively. To indicate M2 phenotype by higher correlation values, fold changes were multiplied by −1. Pearson's *r* for the correlation of these fold changes and expression data was calculated for all samples in the study cohort.

TCGA data

Level 3 gene expression data, summarized mutation data, and patient-centric copy number data released by Brennan et al. [2] were downloaded from the TCGA website (https://tcga-data.nci.nih.gov/docs/publications/gbm_2013/). Clinical data were acquired from the supplementary files provided with the publication. Gene expression microarray data were filtered for (1) non-tumor samples, (2) duplicate measurements, and (3) samples lacking clinical annotation, which were then removed ($n = 519$ remaining). Log_2 transformed gene expression data were then normalized

to z-scores and all samples classified as G-CIMP were removed ($n = 468$ remaining).

All genes with differential expression ($p < 0.01$, Student's t test; $n = 130$) between LTS and STS in the discovery cohort were used as an input for the analysis (for gene names, refer to Table S2). Pearson's r was calculated for all possible combinations of TCGA samples and tumors in the discovery cohort, resulting in a 468×16 correlation matrix.

Subsequently, each TGCA sample was classified as "LTS-like" if it was most highly correlated with a LTS tumor. Conversely, TCGA samples most highly correlated with a STS tumor were classified as "STS-like". Z-score normalized expression values were used to define a "high" and "low" expression group for each gene using the median expression as a cutoff. p values for differences in survival between the groups were calculated with the Log-rank test. Genes were classified as "rOS genes" (reduced overall survival) or "iOS genes" (increased overall survival) if median OS (overall survival) was higher in the "low" and "high" expression groups, respectively.

To correct for multiple testing, the false discovery rate for overlap of differentially expressed genes between "LTS-like" and "STS-like" samples on the one hand and survival-associated genes on the other hand was estimated through a permutation-based approach. In total, 1000 permutations of group labels were performed. Differentially expressed genes for these random groupings were identified for each permutation and compared to survival-associated transcripts. Here, the χ^2 statistic served as a parameter for the degree of enrichment. The resulting false discovery rate was calculated as n_{sig}/n_{total}. Here, n_{sig} corresponds to the number of permutations with a χ^2 statistic greater than observed in the original analysis while n_{total} denotes the total number of permutations.

To facilitate further analysis, the TCGA patient-centric CNA data were simplified as follows: all broader gains and focal amplifications (indicated by $+1$ and $+2$ in the table) were denoted by $+1$; conversely, all broader losses and focal deletions were denoted by -1. All genes with multiple entries were removed. Gains and losses of whole chromosomes were assumed to be present if >50 % of genes on a given chromosome were affected.

Genome-wide expression for co-gain tumors and the rest of the data set were estimated by first sorting for chromosomal position based on coordinates retrieved from the ENSEMBL data base and approximating mean values for each group by a *loess* regression curve (smoothing parameter $\alpha = 0.05$). Due to preprocessing (z-score normalization), expression patterns for each group represent deviations as compared to the average across the whole data set. Thus, higher expression denoted by an upwards deviation of the curve might reflect either higher proportions of gains or lower proportions of deletions, depending on the distribution of these copy number aberrations among the different groups. Accordingly, lower expression may be due to a higher proportion of deletions or a lower proportion of gains in these regions.

Validation of chromosome 19/20 co-gain

An independent cohort of GBM samples with extensive molecular annotation including *IDH1* mutation status and outcome data was collected from the Sahlgrenska University Hospital and the Karolinska Institute in Sweden. Diagnosis was performed according to the latest WHO guidelines by an experienced neuropathologist (VPC). These samples were subjected to analysis by aCGH for copy number assessment (unpublished data, platforms and processing as described [27, 38]). A lift-over to the latest release of the human genome (hg19) was performed for genomic coordinates and data for all samples were plotted along the genome. Whole chromosome gains of chromosomes 19 and 20 were assessed using visual inspection. The data set was filtered for IDH^{wt} samples with a confirmed diagnosis of primary glioblastoma. All samples which passed this filtering step ($n = 124$) were used for subsequent survival analyses.

Fluorescence in situ hybridization

Two-color interphase FISH was performed on deparaffinized sections using CEP probes for chromosome 20 (spectrum orange; Catalogue No. 06J36-030, Abbott) and chromosome 17 (spectrum aqua; Catalogue No. 06J38-027, Abbott) as a reference. Pretreatment of slides, hybridization, post-hybridization processing, and signal detection were performed as previously described [33]. Samples showing sufficient FISH efficiency (90 % nuclei with signals) were evaluated. Signals were scored in 200 non-overlapping, intact nuclei of each sample. A tumor was considered to carry a gain of chromosome 20 when the mean ratio of CEP20/CEP17 was $>1.25 \pm$ SD per sample. Moreover, the percentage of nuclei with detectable gain of chromosome 20 was calculated for each sample.

Expression data for aCGH validation cohort

Gene expression array data (Affymetrix U133A) was available for a subset ($n = 32$) of the aCGH validation cohort. These data had been deposited in the GEO database (Accession No. GSE1993) as part of a larger study on astrocytic tumors [43]. All analyses were conducted within the *R* programming environment. Series matrix files were acquired through the "GEOquery" package, available in *bioconductor* [10]. Expression estimates were transformed to $\log_2$

Table 1 Clinicopathological characteristics of LTS and STS discovery cohort

Variable	LTS $N = 10$	STS $N = 6$	p value (univariate)
OS (mo), median (range)	50 (36–83)	9 (7–9)	<0.0001
Age (years) at diagnosis, median (range)	59 (31–70)	67 (43–74)	0.19
Sex			0.63
Male (%)	5 (50)	4 (67)	
Female (%)	5 (50)	2 (33)	
KPS at diagnosis, median (range)	75 (50–100)	90 (80–100)	0.08
MGMT			0.3
Hypermethylated (%)	7 (70)	2 (33)	
Non-hypermethylated (%)	3 (30)	4 (67)	
Location			0.51
Frontal (%)	3 (30)	3 (50)	
Temporal (%)	1 (10)	0 (0)	
Parietal (%)	3 (30)	1 (17)	
Occipital (%)	1 (10)	1 (17)	
Other (%)	2 (20)	1 (17)	
Resection			0.61
Total (%)	7 (70)	3 (50)	
Subtotal (%)	3 (30)	3 (50)	
Therapy			
Extensive therapy[a]	10 (100)	6 (100)	
Subtype			0.8
Classical (%)	4 (40)	3 (50)	
Mesenchymal (%)	5 (50)	2 (33)	
Proneural (%)	1 (10)	1 (17)	

p values correspond to univariate analyses. More detailed information is provided in Table S1. Statistical tests: Log-rank test (OS), Mann–Whitney test (age, KPS), Fisher's exact test (sex, MGMT, resection), χ^2 test (location, subtype)

KPS Karnofsky Performance Score, *LTS* long-term survivor, *OS* overall survival, *STS* short-term survivor

[a] Patients who received combined radio- and chemotherapy followed by adjuvant temozolomide treatment

scale and median centered. Repeat measurements of genes were combined into a single value (mean expression). Genes with differential expression ($p < 0.01$) between LTS and STS in the discovery cohort were used to classify cases as "LTS-like" and "STS-like" as outlined above (refer to "TCGA data" within the "Materials and methods").

Results

Clinical characteristics of GBM long- and short-term survivors

Current standard in the treatment of primary glioblastoma consists of surgery and temozolomide (TMZ) based post-operative chemoradiotherapy, followed by six cycles of adjuvant TMZ [49]. To rule out any treatment-related bias and categorically exclude secondary GBMs, we focused on IDH^{wt} GBM patients who completed the full treatment regimen. Patients with a survival of more than 36 months after diagnosis (LTS, $n = 10$) were compared to patients who survived for less than 10 months (STS, $n = 6$). Clinicopathological characteristics of the study cohort are summarized in Table 1 (for more detailed information, refer to Table S1). There were no significant differences in age ($p = 0.19$, Mann–Whitney U test), MGMT methylation status ($p = 0.3$, Fisher's exact test) or sex ($p = 0.63$, Fisher's exact test). Surprisingly, there was a trend towards better pre-operative performance (KPS) in STS ($p = 0.08$, Mann–Whitney U test). Furthermore, neither location of the tumor nor extent of resection differed between the groups ($p = 0.51$, χ^2 test and $p = 0.61$, χ^2 test, respectively). To conclude, there were no statistically significant differences in clinical parameters between LTS and STS GBM patients in our study cohort.

Transcriptional profiling of GBM long- and short-term survivors

Assessment of gene expression subtypes according to the TCGA [55] splits our study cohort into mesenchymal ($n = 7$), classical ($n = 7$), and proneural ($n = 2$) groups (Table 1; Fig. S1). We did not observe enrichment of any subtype among LTS or STS ($p = 0.8$, χ^2 test). Transcriptional profiles of LTS and STS were investigated with a number of unsupervised and supervised methods. Principal component analysis (Fig. S2) did not result in a clear separation of the groups. However, out of 21,389 genes we found 1504 to be differentially expressed ($p < 0.05$, Student's t test) between LTS and STS (Table S3). Of these, 755 showed higher expression in STS and 749 in LTS, respectively. From this, we concluded that while there are conserved differences in gene expression between STS and LTS, there is no global, uniform LTS phenotype detectable by microarray profiling of bulk tissue.

Next, pathway analysis was conducted for a better understanding of the biological processes underlying these expression differences (for detailed results, refer to Table S4 and S5). We noted that six out of seven significantly enriched pathways (FDR corrected $p < 0.01$) for genes upregulated in STS were involved in immune processes. Consequently, we found a number of genes well known for their role in tumor-promoting inflammatory processes and macrophage/microglia education. Among other genes identified through pathway analysis, *PLA2G2A*, *CHI3L2*, *CCL2*, *CCL18* and *CCL20* were overexpressed in STS (for more detailed information, refer to Tables S3 and S5).

Differential macrophage/microglia activation in LTS and STS

Recently, evidence has surfaced for divergent and even antagonistic effects of macrophages with different polarization. While the so-called M1 cells exhibit anti-tumor activity, the M2 phenotype supports tumor cell growth and invasion, a finding which also applies to brain tumors (reviewed in [18]).

We hypothesized that qualitative rather than quantitative differences regarding microglia might govern the LTS and STS phenotypes. Therefore, a previously published gene expression signature [37] was employed to investigate the activation status of microglia in our study cohort. Martinez et al. investigated transcriptional differences between macrophages with M1 and M2 polarization (reviewed in [40]). This signature consists of genes which were found to be differentially regulated between these phenotypes. Since it also includes information regarding fold changes, i.e., the quantitative amount of change, we were able to calculate correlation (Pearson's r) between signature genes and expression data for all patients in our cohort. Here, positive and negative correlations imply M2 and M1 polarization, respectively. We found correlation values among STS to be significantly higher than those for LTS ($p < 0.05$, Student's t test), indicating differential polarization and a shift towards the M2 phenotype among STS (Fig. 1a).

To examine whether these findings from bulk tissue gene expression data also translated to the protein level, we additionally performed multicolor immunofluorescence stainings. The TissueFAXS method [34] was applied to simultaneously detect and quantify CD68 (macrophage/microglia marker) as well as CD163 and CD204, two established M2 markers [30, 45]. We were thus able to assess both the amount and polarization of microglia for all cases in the study cohort. Staining for CD68 revealed no significant differences in the number of microglial cells ($p = 0.43$, Mann–Whitney U test, Fig. S3). The proportion of pro-tumorigenic M2 cells among tumor microglia, however, was significantly higher in STS as compared to LTS ($p = 0.02$, Mann–Whitney U test, Fig. 1b, d). Conversely, LTS tumors contained larger proportions of M1 cells ($p = 0.02$, Mann–Whitney U test, Fig. 1b, c). In summary, protein level data were consistent with the concept of differential polarization and, furthermore, confirmed prior transcriptomic analyses.

LTS and STS expression profiles are prognostic in non-G-CIMP GBMs

A number of recent publications on the transcriptome of GBM LTS have assessed mesenchymal, proneural, classical and neural subtypes among their patient cohorts [16, 46]. However, none of these studies validated differences in gene expression between LTS and STS in a larger number of tumors. As our study cohort was highly selected, yet small, we aimed to project the transcriptional phenotypes of LTS and STS onto a larger data set. Gene expression data were downloaded from the TCGA. As global hypermethylation is a sensitive parameter for *IDH* mutations, all G-CIMP cases were removed for further analyses ($n = 468$ remaining). Using genes which were differentially expressed between LTS and STS ($n = 130$, $p < 0.01$, Student's t test; for gene names, see Table S2), we calculated pair-wise correlations of TCGA tumors and samples in our study cohort. The resulting correlation matrix is shown in Fig. 2a. Subsequently, each TCGA tumor was classified as either "LTS-like" or "STS-like", according to the sample in the discovery cohort it was most highly correlated with.

We noted a significant difference in survival between "LTS-like" and "STS-like" glioblastomas in the TCGA data set ($p = 0.007$, Log-rank test, Fig. 2b). We found no enrichment for classical, neural, mesenchymal or proneural tumors in either group ($p = 0.67$, χ^2 test), implying our

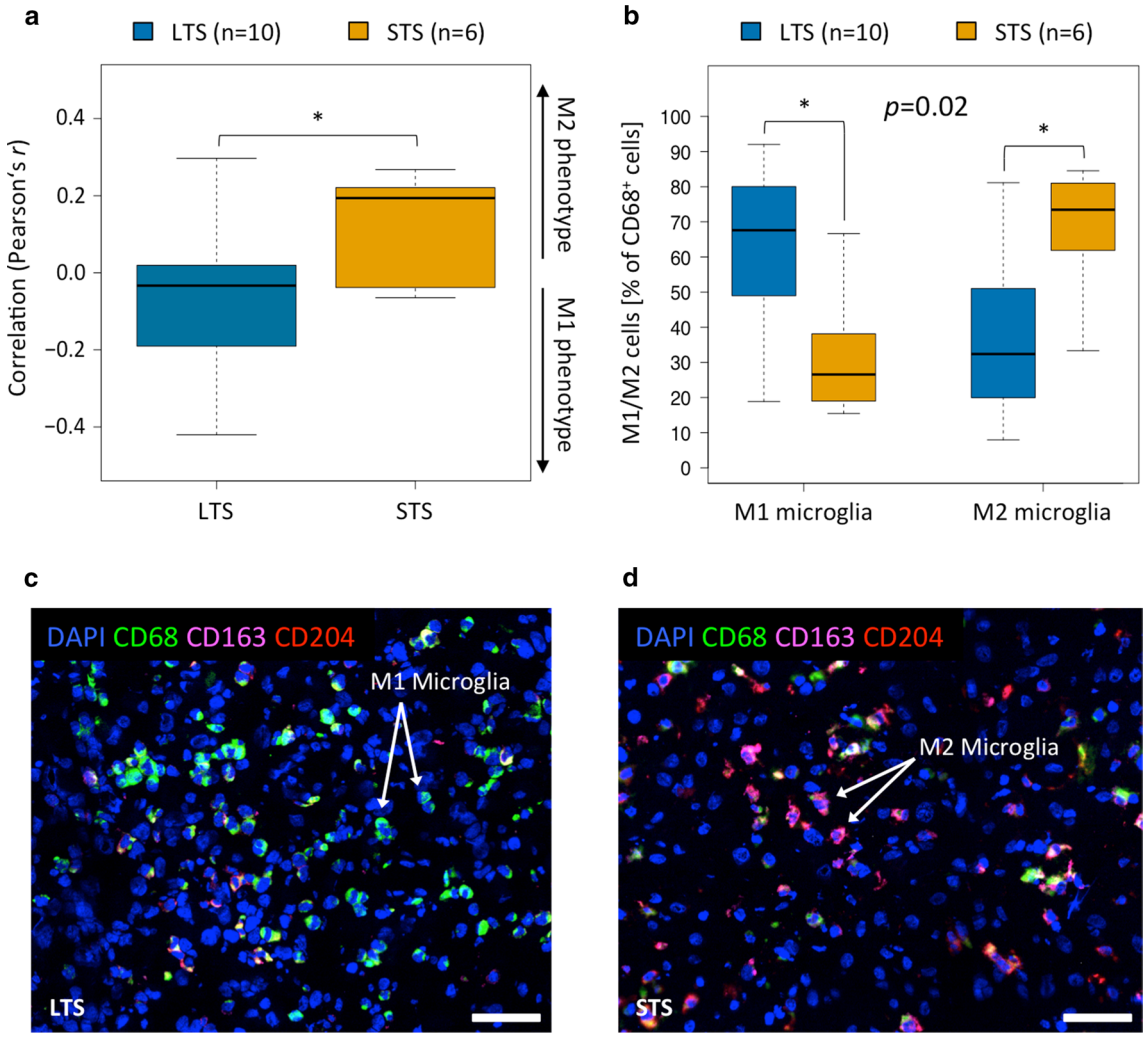


Fig. 1 Microglial infiltration and activation status in LTS and STS. A previously published gene expression signature based on transcriptional differences of M1 and M2 activated macrophages was used to assess microglial polarization in the study cohort. **a** Shows correlation with this signature for LTS and STS samples. STS tumors exhibited significantly greater correlation with the M2 signature ($p < 0.05$, Student's t test). These findings were followed up with *multicolor* immunofluorescence staining (see text). Staining for CD68, CD163 and CD204 allowed for the identification and quantification of macrophages ($CD68^+$) and their polarization. **b** Congruent with transcriptomic results, M2 microglia were more abundant in STS tumors ($p < 0.02$, Mann–Whitney U test). The number of microglial cells, however, did not differ between the groups (refer to Fig. S3). **c**, **d** Show representative stainings for a LTS and STS case, respectively. *CD* cluster of differentiation, *LTS* long-term survivor, *STS* short-term survivor

classification was independent of these subtypes. There was also no significant association with established prognostic factors such as age ($p = 0.38$, χ^2 test) or extensive therapy (radiotherapy or chemoradiotherapy followed by adjuvant temozolomide treatment, $p = 0.78$, χ^2 test). We found a trend towards more frequent *MGMT* hypermethylation in the "LTS-like" group. However, this did not reach statistical significance ($p = 0.1$, χ^2 test). We further investigated how our classification performed for LTS in the TCGA data set. Out of 23 non-G-CIMP patients with survival of more than 36 months in the TCGA data set, 17 (73.9 %) were correctly classified as "LTS-like" by our approach ($p = 0.03$, χ^2 test). Pathway analysis (Table S6 and S7) again revealed inflammatory response as a major topic among genes with higher expression in "STS-like" tumors. We also found a difference in correlation with the aforementioned microglial signature in TCGA samples, similar to our discovery set. "STS-like" tumors exhibited significantly greater correlation (Pearson's r) with the M2 signature than "LTS-like" tumors ($p < 0.0001$, Student's t test; Fig. S4).

Large datasets such as those from the TCGA are often used to assess the prognostic value of certain target genes. A common approach is the median split, whereby a continuous variable is turned into a categorical one. More

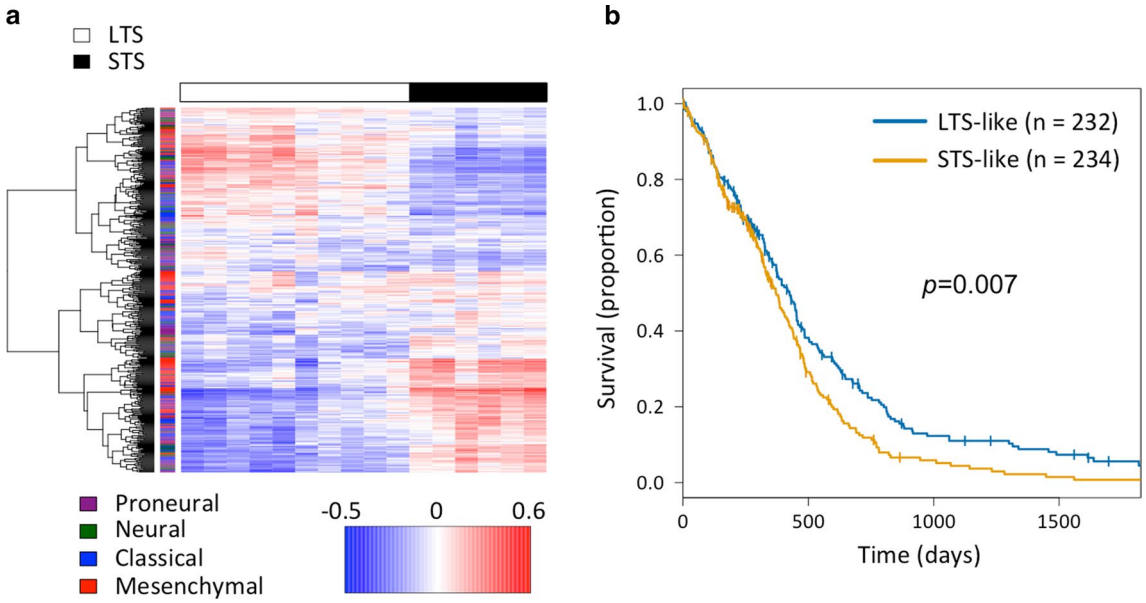


Fig. 2 Classification of TCGA samples based on LTS and STS expression profiles. **a** The heatmap shows all pair-wise correlations between LTS/STS samples of the discovery cohort (*columns*) and non-G-CIMP GBMs in the TCGA data set (*rows*) for genes with differential expression between LTS and STS ($p < 0.01$, Student's t test, $n = 130$). *Red* and *blue* indicate high and low correlation, respectively (range −0.53 to 0.63). Each TCGA tumor was classified as "LTS-like" or "STS-like", depending on the sample from the discovery cohort it was most highly correlated with. **b** Kaplan–Meier plot for non-G-CIMP TCGA samples according to the classification based on LTS/STS expression profiles. The *blue* and *orange lines* denote tumors classified as "LTS-like" and "STS-like", respectively. The p value was calculated using the Log-rank test. For illustration purposes, a cutoff of 5 years was used for the x-axis. *G-CIMP* glioma CpG island methylator phenotype, *LTS* long-term survivor, *STS* short-term survivor, *TCGA* The Cancer Genome Atlas

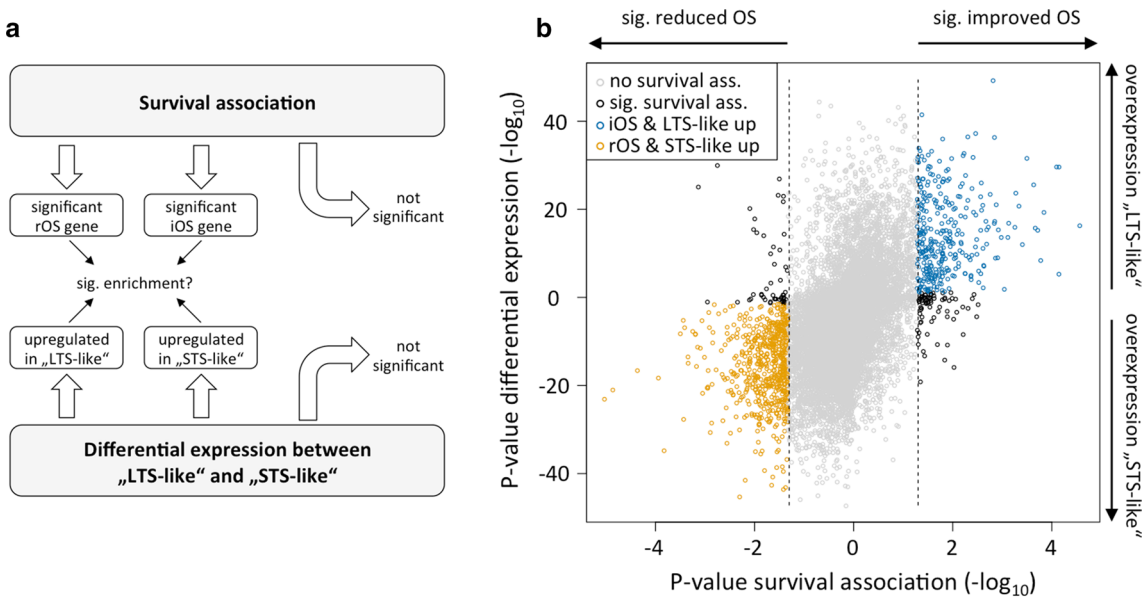
specifically, the median gene expression is used to dichotomize into a "low" and "high" expression group. Then, a statistical measure (in most cases, the Log-rank test) is used to test whether a significant difference in survival exists between these groups. Presumably, multiple mechanisms with an impact on patient outcome act in GBM tumors. Thus, various expression signatures associated with outcome might yield prognostic, yet different groupings. We therefore reasoned that a classification scheme might be considered more reliable if it has high correlation with the unbiased identification of survival-associated genes.

Thus, we extended the approach outlined above and assessed the overlap of survival-associated genes and those with differential expression between "LTS-like" and "STS-like" TCGA samples. First, the median split was used to identify genes with significant survival association in non-G-CIMP GBM ($n = 1093$; $p < 0.05$, Log-rank test). We nominally classified them as "iOS genes" (improved overall survival, $n = 404$) or "rOS genes" (reduced overall survival, $n = 689$). We hypothesized that genes with overexpression in "LTS-like" samples should be enriched for "iOS genes". Conversely, we expected transcripts with higher expression in "STS-like" to be enriched for "rOS genes" (see Fig. 3a for a brief summary of the approach). Indeed, enrichment analysis revealed a striking overlap in both cases ($p < 10^{-12}$, χ^2 test, Fig. 3b–d). Furthermore, after accounting for multiple testing by estimation of the false discovery rate (FDR), the highly significant ($q = 0.008$) interrelation of our classification as "LTS/STS-like" and single-gene prognosticators was confirmed.

Co-gain of chromosomes 19 and 20 is a prognostic factor in non-G-CIMP glioblastoma

We also explored the association of genetic lesions with LTS- or STS-like status in the TCGA data set. We identified a number of mutations ($n = 20$), which were significantly enriched ($p < 0.05$, Fisher's exact test) in one or the other group (Table S8). However, each of these mutations only occurred in a small number of samples and did not show any prognostic effect by itself.

We further assessed whether chromosomal aberrations were more prevalent in one of the groups (for detailed results, refer to Table S9). We observed that prototypical GBM lesions such as amplifications on chromosome 7 and deletions of *CDKN2A*/*CDKN2B* as well as chromosome 10 were present in both groups at similar frequencies. However, LTS-like tumors were enriched for patients with concurrent gain of chromosomes 19 and 20, defined as duplications affecting more than half of each chromosome. Here, 55 of 86 samples (64 %) with co-gain were classified as "LTS-like", whereas 31 (36 %) belonged to the



c

Enrichment analysis for iOS genes

$p < 10^{-12}$ (χ^2 test)		sig. iOS gene: no	sig. iOS gene: yes	Total
overexpressed in „LTS-like“	no	9391	75	9466
	yes	2246	**329**	2575
Total		11637	404	**12041**

d

Enrichment analysis for rOS genes

$p < 10^{-12}$ (χ^2 test)		sig. rOS gene: no	sig. rOS gene: yes	Total
overexpressed in „STS-like“	no	4142	53	4195
	yes	7210	**636**	7846
Total		11352	689	**12041**

Fig. 3 Overlap of survival-associated and differentially expressed genes in the TCGA data set. **a** Gives an overview of the approach. Survival association was investigated for all genes using a median split and Log-rank testing (see text). Genes were classified as “iOS genes” (improved OS) and “rOS genes”(reduced OS) if their expression correlated significantly with better and worse prognosis, respectively. We then investigated whether genes with significant survival association were more likely to be overexpressed in “STS-like” or “LTS-like” samples. **b** Provides a graphical representation of the results. The *x*- and *y*-axis show *p* values on a negative $\log_{10}$ scale for survival association (Log-rank test) and differential expression (Student‘s *t* test), respectively. For illustration purposes, *p* values for survival-associated “rOS genes” and transcripts overexpressed in “STS-like” samples were multiplied by −1. The *dashed lines* represent cutoffs for $p = 0.05$. The majority (81 %) of survival-associated “iOS genes” was also overexpressed in “LTS-like” tumors (*blue circles* in the *upper right part*). Conversely, “rOS genes” were also commonly (92 %) overexpressed in “STS-like” tumors (*orange circles* in *lower left part*). **c** The left contingency table provides a numerical representation for the highly significant ($p < 10^{-12}$, χ^2 test) enrichment of “iOS genes” among genes with overexpression in “LTS-like” tumors. **d** The right contingency table summarizes results for “rOS genes” and overexpression in “STS-like” tumors. Again, there was highly significant overlap ($p < 10^{-12}$, χ^2 test). *Colors* in contingency tables correspond to those in (**b**). *iOS* improved overall survival, *LTS* long-term survivor, *rOS* reduced overall survival, *STS* short-term survivor

“STS-like” group ($p = 0.004$, χ^2 test). While isolated gains exist, aberrations of chromosomes 19 and 20 co-occur frequently ($p < 0.0001$, χ^2 test) and are a common finding in non-G-CIMP GBM (18.5 %).

Notably, co-gain of chromosomes 19 and 20 is a highly significant prognostic factor ($p < 0.0001$, Log-rank test; Fig. 4a). However, it appears that only simultaneous gains are associated with a better prognosis. Cases with isolated gains of chromosomes 19 or 20 did not show improved outcome as compared to GBMs harboring neither aberration in the TCGA data set ($p = 0.2$ and $p = 0.82$, respectively, Log-rank test; Fig. 4b). Intriguingly, taken together, “LTS-like” status and chromosome 19/20 co-gain showed an even stronger association with survival ($p = 2 \times 10^{-6}$, Log-rank test) as compared to all other samples (Fig. S5).

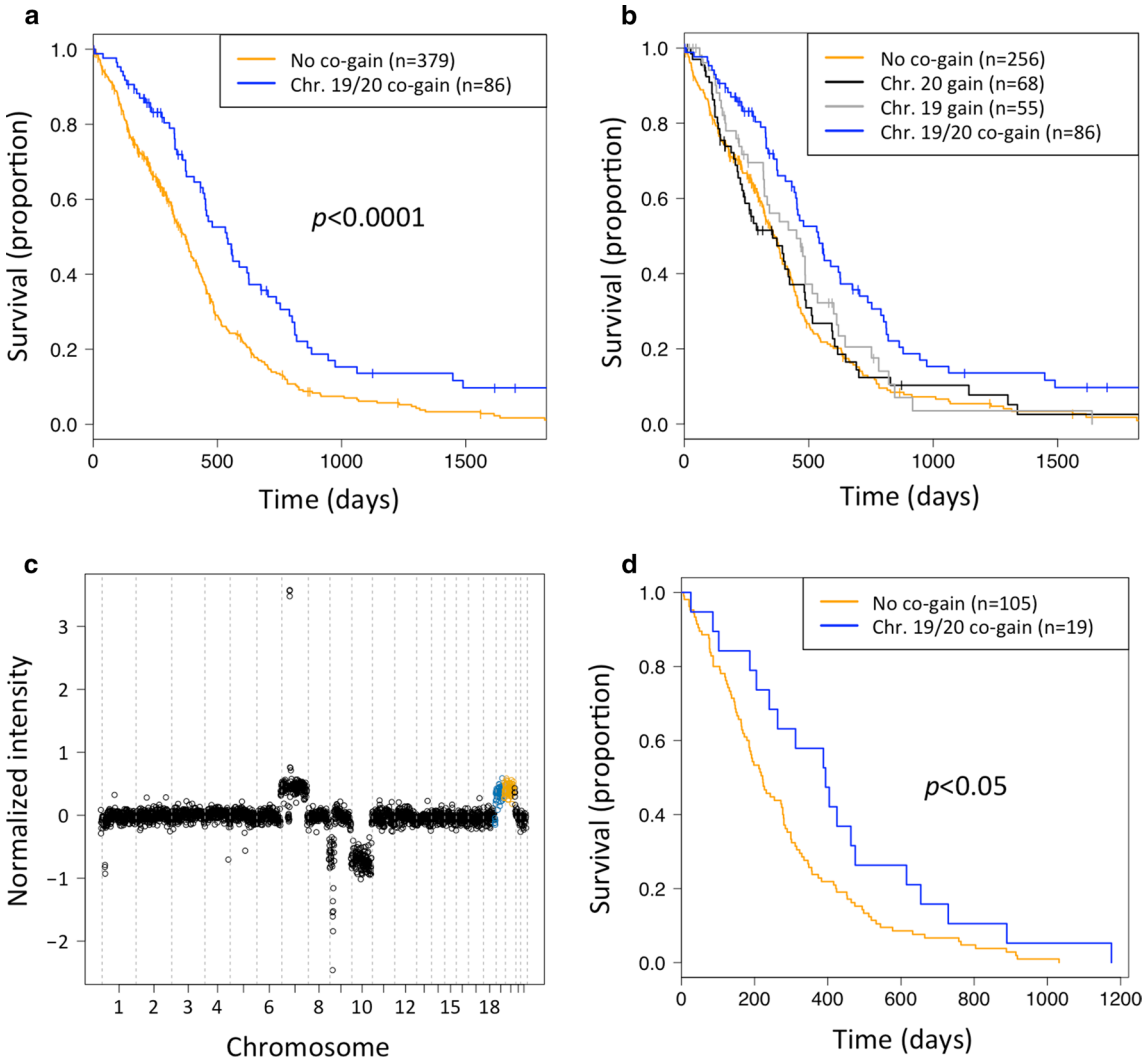


Fig. 4 Prognostic significance of chromosome 19/20 co-gain. **a** Shows a Kaplan–Meier curve for the TCGA data set ($n = 465$) comparing patients with concurrent gains of chromosome 19 and 20 to the rest of the cohort. **b** Provides a more comprehensive view comparing isolated gains of chromosomes 19 (*gray line*) or chromosome 20 (*black line*), co-gain of both chromosomes (*blue line*) and those samples with no gains on chromosomes 19 or 20 (*orange line*). Isolated gains of chromosomes 19 or 20 did not show improved outcome as compared to samples without these gains ($p = 0.2$ and $p = 0.82$, respectively, Log-rank test). Samples with co-gain show improved outcome compared to all other groups. **c** Shows aCGH data for a representative sample from a second, independent cohort of primary GBMs harboring co-gain of chromosomes 19 and 20 (with probes colored *blue* and *orange*, respectively). This tumor also exhibits additional copy number alterations typical for primary GBM, including high-level amplification of *EGFR* (chromosome 7) and homozygous deletion of *CDKN2A* on chromosome 9. **d** In congruence with our findings in the TCGA data set, co-gain of chromosomes 19 and 20 showed prognostic benefit in this second, independent validation cohort ($p < 0.05$, Log-rank test; $n = 124$). *aCGH* array comparative genomic hybridization, *CDKN2A* cyclin-dependent kinase inhibitor 2A, *EGFR* epidermal growth factor receptor, *GBM* glioblastoma multiforme, *TCGA* The Cancer Genome Atlas

Co-gain of both chromosomes also remained a highly significant prognostic factor in multivariate analysis, correcting for age and extensive therapy (Table 2). However, in concordance with findings by the TCGA [2], we did not detect significant association with survival for *MGMT* hypermethylation after correction for other known prognostic factors. In line with these findings, LTS in the TCGA data set were also more likely to carry the co-gain. Here, 8 of 22 (36 %) patients with available copy number data exhibited the combined lesion. This corresponds to a significantly higher proportion as compared to the rest of the data set (18 %; $p < 0.05$, Fisher's exact test).

Independent validation of prognostic significance of chromosome 19/20 co-gain

To further interrogate and corroborate the prognostic importance of the chromosome 19/20 co-gain, we

Table 2 Multivariate analysis of OS for co-gain of chromosomes 19 and 20

Variable	Death		*p* value
	Relative risk	95 % CI	
Age (<61 years)	0.59	0.48–0.73	1.73×10^{-6}
MGMT methylation	0.97	0.76–1.20	0.81
Extensive therapy	0.44	0.37–0.56	7×10^{-14}
Co-gain chromosomes 19 and 20	0.56	0.40–0.69	3.8×10^{-5}

Patients were dichotomized according to median age in the cohort (61 years). Extensive therapy denotes patients who received either radiotherapy alone followed by adjuvant temozolomide treatment or combined radio- and chemotherapy followed by adjuvant temozolomide treatment

OS overall survival, *CI* confidence interval

investigated a second, independent study cohort. This validation set consisted of previously unpublished aCGH data, including molecular annotation as well as clinical outcome information (for patient data, refer to Table S10). Again, we exclusively investigated *IDH*^{wt} cases with confirmed diagnosis of primary glioblastoma ($n = 124$). Figure 4c shows a representative copy number profile for a sample with typical GBM-associated alterations and co-gain of chromosomes 19 and 20. We confirmed the highly significant co-occurrence of gains on chromosomes 19 and 20 ($p < 0.0001$, χ^2 test). Also, the frequency of co-gains (15.3 %) was similar to the TCGA data set (18.5 %). Albeit the considerably smaller number of samples, co-gain of chromosomes 19 and 20 was again significantly associated with patient survival ($p < 0.05$, Log-rank test, Fig. 4d).

For a subset of these samples, mRNA microarray data had been released as part of a prior study (GSE1993, [43]). This allowed us to also infer "LTS-like" and "STS-like" status for a limited number of samples ($n = 32$) in the aCGH validation cohort. We found that LTS-like tumors ($n = 13$) included all samples with co-gain ($n = 3$), whereas none belonged to the STS-like group ($n = 19$). Despite the very small sample numbers, this trend almost reached significance ($p = 0.057$, Fisher's exact test).

Extent and transcriptional effect of chromosome 19/20 co-gain

We further assessed the extent of gains on chromosomes 19 and 20 in the TCGA dataset. Here, we noted a nearly bimodal distribution in both cases, suggesting that the majority of samples show either no aberration or gain of the whole chromosome (Fig. 5a–b).

Recent studies have highlighted the extensive genetic heterogeneity within tumors of the same patient [17]. The current model of tumor evolution states that early, initiating drivers are present in all cancer cells whereas later events only occur in subpopulations of the tumor (reviewed in [19]). We thus interrogated the extent of the co-gain and the proportion of affected cells by FISH in four *IDH*^{wt} GBMs harboring this alteration and two controls.

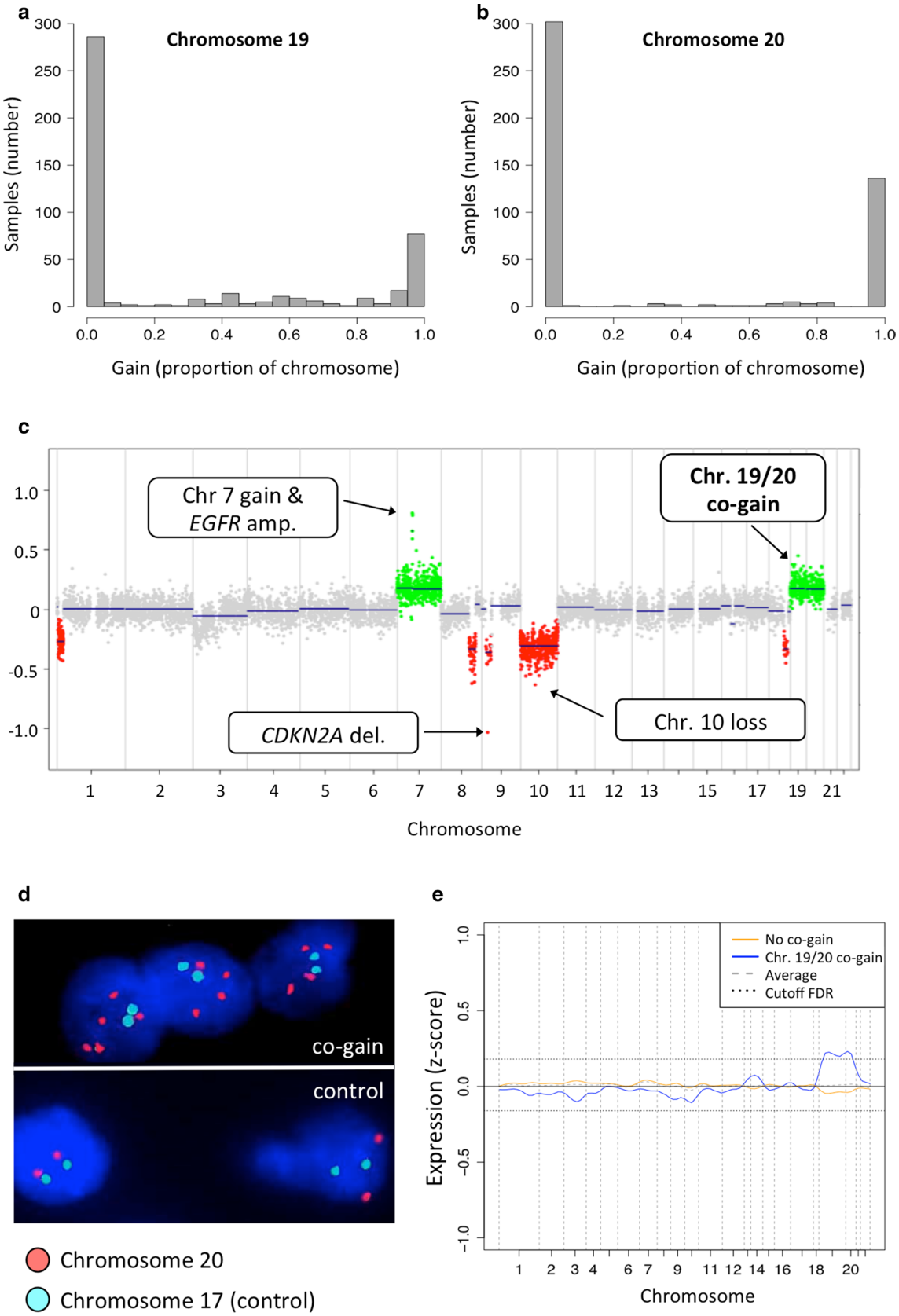
These samples were selected based on copy number data generated from DNA methylation microarrays [25, 52]. A copy number profile for a representative co-gain sample of this cohort is shown in Fig. 5c (for other cases, refer to Fig. S6). Of note, all samples were *IDH*^{wt} and exhibited genomic landscapes typical for primary GBM. As these tumors show almost identical signal intensities for chromosomes 19 and 20, similar dynamics regarding their distribution among tumor cells is likely. For this reason and technical limitations, we used chromosome 20 as a surrogate parameter for the co-gain. Our results confirmed gains in the majority of tumor cells in all selected samples, while controls were uniformly negative (data for two representative cases are shown in Fig. 5d). For most affected cells, 2–3 additional copies of chromosome 20 were observed per nucleus.

Upon analysis of gene expression across the genome, in addition to low amplitude variations in several regions of the genome, we found a strong upregulation of genes located on chromosomes 19 and 20 for tumors harboring the co-gain, indicating a gene dosage effect (Fig. 5e). Gene expression also showed smaller changes across other chromosomes. Indeed, certain copy number aberrations such as deletions on chromosomes 13, 14 and 15 seem to be almost mutually exclusive with chromosome 19/20 co-gain. However, these copy number changes were themselves not associated with patient survival (data not shown). Yet, upon FDR correction, only expression changes on chromosomes 19 and 20 remained significant ($q < 0.05$).

We conclude that co-gains of chromosomes 19 and 20 (1) frequently involve the whole chromosome, (2) likely affect the majority of tumor cells and (3) lead to upregulation of genes located in these regions.

FISH analysis in the discovery cohort

Co-gain of chromosome 19 and 20 was identified in the TCGA data set based on a transcriptional signature learned from our initial STS and LTS cohort. We thus hypothesized the co-gain might also be present in LTS cases of the discovery cohort. Therefore, we again performed FISH analysis for chromosome 20. Intriguingly, the experiment revealed gains among eight out of ten (80 %) LTS samples (Table S11). Due to the lack of additional data, we cannot preclude that some of these cases might harbor isolated gains of chromosome 20. However, even when considering the cumulative frequency (33 %) for co-gains and isolated chromosome 20 gains, this finding represents highly significant enrichment ($p = 0.003$, Binomial test).


a
Chromosome 19
Samples (number)
Gain (proportion of chromosome)
b
Chromosome 20
Samples (number)
Gain (proportion of chromosome)
c
Chr 7 gain & EGFR amp.
Chr. 19/20 co-gain
CDKN2A del.
Chr. 10 loss
Chromosome
d
co-gain
control
Chromosome 20
Chromosome 17 (control)
e
No co-gain
Chr. 19/20 co-gain
Average
Cutoff FDR
Expression (z-score)
Chromosome

◀ **Fig. 5** Extent, clonality and transcriptional impact of gains on chromosomes 19 and 20. **a**, **b** Give an overview of the extent of gains on chromosomes 19 and 20, respectively. Histograms show the number of samples (*y*-axis) in which gains of different extent (*x*-axis) occur. *Zero* and *one* correspond to normal gene dose and whole chromosome gain, respectively. These data indicate a nearly bimodal distribution in both cases. **c** Depicts the genomic landscape in a typical co-gain sample. Comparable signals for gains on chromosome 19 and 20 suggest similar dynamics regarding their distribution among tumor cells. FISH was used to further differentiate between two models which might account for this observation (high amplitude gains in a subset of malignant cells vs. low amplitude gains in the majority of the tumor cell population). **d** Shows representative cases for FISH in a cohort of co-gain cases ($n = 4$, *upper panel*) and controls without gains on chromosome 19 or 20 ($n = 2$, *lower panel*). The majority of tumor cells in co-gain samples exhibited chromosome 20 gains, while controls were uniformly negative. **e** Shows normalized expression in the TCGA data set, approximated by a *loess* regression curve (smoothing parameter $\alpha = 0.05$), plotted along the genome for samples harboring co-gain of chromosomes 19 and 20 (*blue line*) and the rest of the cohort (*orange line*). The *dashed gray line* represents average expression across the whole dataset. While minor changes were observed in different regions, only those for chromosome 19 and 20 reached statistical significance after correction for multiple testing (*horizontal*, *dotted lines* correspond to a FDR with $q < 0.05$). *FDR* false discovery rate, *FISH* fluorescence in situ hybridization, *TCGA* The Cancer Genome Atlas

Discussion

Due to the grim prognosis of GBM, LTS have intrigued researchers for decades and reports about these patients have been published as early as 1950. However, most results are heavily biased by *IDH* mutations. In the present study, we accounted for this confounding variable and identified a subgroup of GBM, characterized by co-gain of chromosomes 19 and 20 and markedly better survival. Moreover, we provided considerable evidence for differential polarization of microglia with higher abundance of the M2 phenotype in STS. Most importantly, these findings have been confirmed in an independent patient cohort and by different experimental approaches.

As of now, cytogenetic studies and genome-wide screens in GBM have focused mostly on common copy number alterations (CNAs) such as amplifications of *EGFR*, losses on chromosome 10 (including the tumor suppressor *PTEN*) and homozygous deletions of *CDKN2A* (reviewed in [51]). Furthermore, these studies have produced conflicting results. For example, EGFR amplification has been associated with both improved [24] and poor [48] outcome, while some reports indicated no prognostic effect at all [23]. Consequently, none of these CNAs are assessed in routine diagnostics due to their inconclusive nature. Isolated gains of chromosomes 19 or 20 are also common CNAs in GBM [2]. Despite their frequency, there is only anecdotal evidence of their co-occurrence [32]. Here, we not only consolidate this finding, but also provide compelling evidence for the strong prognostic benefit for tumors with concurrent gain of chromosomes 19 and 20. Also, this co-gain could easily be assessed as part of a routine neuropathological workup in the near future as high-throughput methods are finding their way into diagnostics. For instance, array-based methylation profiling allows for simultaneous detection of G-CIMP tumors and MGMT hypermethylation while generating copy number profiles at the same time, including information about chromosome 19/20 co-gain [52]. Thus, identification of primary GBMs and stratification according to prognostic factors such as the co-gain could soon become a useful complement to histopathological evaluation.

While many prognostic classifications have been proposed for GBM [7, 15, 29, 36, 44], none have adequately corrected for *IDH* mutations. This is also supported by results from Reifenberger et al., which indicate that identification of IDH^{mut} tumors is a hallmark feature of previously published signatures [46]. By accounting for both *IDH* status and treatment, we are among the first to provide a classification scheme capable of predicting prognosis in purely IDH^{wt} tumors.

Combined gains of chromosome 19 and 20 had been identified through a transcriptional signature learned from LTS and STS. As the same samples from which the signature was initially derived were later shown to also harbor gains, these findings are most likely interrelated. This is also supported by the cumulative prognostic effect of "LTS-like" status and presence of the co-gain. Possible causative effects, however, should be addressed in future studies.

Chromosome 19 has the highest gene density among the genome and contains several large gene families, including zinc finger transcription factors and cytochrome P450 enzymes [20]. Both chromosomes also contain genes implicated in diseases of the nervous system [12, 20] such as *PRNP* which is linked to Creutzfeldt–Jakob disease. Furthermore, different alterations of these chromosomes have been implicated in cancer. For instance, loss of 1p and 19q is a hallmark CNA of oligodendroglioma [35], while gains have been described for leukemia [1]. As the observed co-gains involve all of chromosomes 19 and 20, the prognostic effect is unlikely caused by single transcripts. As a consequence, our study clearly delineates the urgent need for further research addressing the functional consequences of whole chromosome gains rather than single genes.

Pathway analysis of differentially regulated genes in our study cohort implicated tumor-promoting, microglia-driven inflammatory processes in STS. Active anti-tumor response by the adaptive immune system constitutes a favorable prognostic factor in GBM [34] and has also been identified as a feature of LTS tumors [13]. Yet, the ambivalent role of the innate immune system in glioma is well documented (reviewed in [18]). Recruited macrophages/microglia and myeloid progenitors can exert pro-tumorigenic effects such

as angiogenesis [41], invasion [56], and proliferation [9, 28].

These cancer-enabling inflammatory processes have previously been described as a salient feature of the *mesenchymal* subgroup. Differentiation towards this phenotype is driven by a conserved transcriptional network [6] and was shown to correlate with poorer survival [44] in early high-throughput studies. Subsequent efforts have also provided evidence for higher numbers of infiltrating macrophages and greater extent of necrosis in this subtype [8, 14]. In addition to these findings, abundance of M2 macrophages has been found to correlate with glioma grade [45].

However, after exclusion of IDH^{mut}/G-CIMP samples, the largest study in GBM to date [2] has not confirmed inferior prognosis for *mesenchymal* tumors. Indeed, there even was a trend towards better outcome as compared to IDH^{wt} *proneural* samples, highlighting the confounding effect of *IDH* mutations. It is important to realize the imminent impact of these results, as a considerable number of studies have made use of the very same data for validation purposes.

Recently, a pro-tumorigenic inflammation signature has been described for IDH^{wt} STS [16]. In addition to confirming this finding, we have resolved these expression differences further. Transcriptomic analyses and multicolor immunofluorescence stainings have provided considerable evidence for divergent polarization with higher numbers of M2 microglia in STS tumors. Thus, our multilevel approach provides important insights into the association of innate immune response and survival in IDH^{wt} GBM.

The substantial impact of unrecognized IDH mutations on research in GBM is also one of the most important shortcomings of previous studies on LTS. As only recently became evident, these tumors have an utterly different biology despite histological similarities with GBM. Those differences include, but are not limited to, (1) younger age at diagnosis, (2) a different pattern of CNAs, (3) an almost uniform proneural expression subtype and (4) G-CIMP. Their confounding effect on gene expression signatures has been outlined above. At the same time, studies on CNAs usually detected hallmark lesions of primary GBM among STS and common aberrations of lower grade lesions among LTS [3, 4]. This bias extends across all levels of cell biology including epigenetics, which is why methylation screening found G-CIMP tumors to be overrepresented among LTS [47]. Additional proof comes from the reportedly high histopathological misclassification rates [39] in LTS and the high prevalence of *IDH* mutations among them [21]. Therefore, most studies on LTS to date have identified prototypical features of lower grade gliomas for which better survival is well documented, whereas we have taken great care to avoid this bias.

In summary, we provide strong evidence for M2 polarization of microglia as a feature of particularly aggressive tumors. Furthermore, we have shown co-gain of chromosomes 19 and 20 to constitute a bona fide marker for a subgroup of primary GBM with better outcome. It is detectable both on the single-cell and tissue level with different technical platforms and robustly associated with patient survival. In addition, we utilized extensive validation cohorts as part of our study. As clinical studies necessitate appropriate stratification according to known prognostic biomarkers, our findings will have important clinical implications and can, at the same time, be easily integrated into modern molecular diagnostics.

Acknowledgments We thank Steffen Dettling and Saskia Rösch for proofreading of the manuscript as well as Anja Metzner and Daniela Zito for review of patient data. Furthermore, we thank Farzaneh Kashfi, Hildegard Göltzer, Ilka Hearn and Melanie Greibich as well as Barbara Schwager and Claudia Rittmüller for their excellent technical assistance. This study has been funded in part by a grant from the Anni Hofmann Stiftung to CHM and the German Federal Ministry of Research and Education (BMBF) to BB (01GS0883).

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with animals performed by any of the authors. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

References

1. Alvarez S, MacGrogan D, Calasanz MJ, Nimer SD, Jhanwar SC (2001) Frequent gain of chromosome 19 in megakaryoblastic leukemias detected by comparative genomic hybridization. Genes Chromosom Cancer 32:285–293
2. Brennan CW, Verhaak RGW, McKenna A, Campos B, Noushmehr H, Salama SR et al (2013) The somatic genomic landscape of glioblastoma. Cell 155:462–477
3. Burton EC, Lamborn KR, Feuerstein BG, Prados M, Scott J, Forsyth P et al (2002) Genetic aberrations defined by comparative genomic hybridization distinguish long-term from typical survivors of glioblastoma. Cancer Res 62:6205–6210
4. Burton EC, Lamborn KR, Forsyth P, Scott J, O'Campo J, Uyehara-Lock J et al (2002) Aberrant p53, mdm2, and proliferation differ in glioblastomas from long-term compared with typical survivors. Clin Cancer Res 8:180–187
5. Capper D, Zentgraf H, Balss J, Hartmann C, Von Deimling A (2009) Monoclonal antibody specific for IDH1 R132H mutation. Acta Neuropathol 118:599–601
6. Carro MS, Lim WK, Alvarez MJ, Bollo RJ, Zhao X, Snyder EY et al (2010) The transcriptional network for mesenchymal transformation of brain tumours. Nature 463:318–325
7. Colman H, Zhang L, Sulman EP, McDonald JM, Shooshtari NL, Rivera A et al (2010) A multigene predictor of outcome in glioblastoma. Neurooncology 12:49–57

8. Cooper LAD, Gutman DA, Chisolm C, Appin C, Kong J, Rong Y et al (2012) The tumor microenvironment strongly impacts master transcriptional regulators and gene expression class of glioblastoma. Am J Pathol 180:2108–2119
9. Da Fonseca ACC, Romão L, Amaral RF, Assad Kahn S, Lobo D, Martins S et al (2012) Microglial stress inducible protein 1 promotes proliferation and migration in human glioblastoma cells. Neuroscience 200:130–141
10. Davis S, Meltzer PS (2007) GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics 23:1846–1847
11. Debus J, Abdollahi A (2014) For the next trick: new discoveries in radiobiology applied to glioblastoma. Am Soc Clin Oncol Educ Book 34:95–99
12. Deloukas P, Matthews LH, Ashurst J, Burton J, Gilbert JG, Jones M et al (2001) The DNA sequence and comparative analysis of human chromosome 20. Nature 414:865–871
13. Donson AM, Birks DK, Schittone SA, Kleinschmidt-DeMasters BK, Sun DY, Hemenway MF et al (2012) Increased immune gene expression and immune cell infiltration in high-grade astrocytoma distinguish long-term from short-term survivors. J Immunol 189:1920–1927
14. Engler JR, Robinson AE, Smirnov I, Hodgson JG, Berger MS, Gupta N et al (2012) Increased microglia/macrophage gene expression in a subset of adult and pediatric astrocytomas. PLoS One 7:e43339
15. Freije WA, Castro-Vargas FE, Fang Z, Horvath S, Cloughesy T, Liau LM et al (2004) Gene expression profiling of gliomas strongly predicts survival. Cancer Res 64:6503–6510
16. Gerber NK, Goenka A, Turcan S, Reyngold M, Makarov V, Kannan K et al (2014) Transcriptional diversity of long-term glioblastoma survivors. Neurooncology 16:1186–1195
17. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E et al (2012) Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. N Engl J Med 366:883–892
18. Glass R, Synowitz M (2014) CNS macrophages and peripheral myeloid cells in brain tumours. Acta Neuropathol 128:347–362
19. Greaves M, Maley CC (2012) Clonal evolution in cancer. Nature 481:306–313
20. Grimwood J, Gordon LA, Olsen A, Terry A, Schmutz J, Lamerdin J et al (2004) The DNA sequence and biology of human chromosome 19. Nature 428:529–535
21. Hartmann C, Hentschel B, Simon M, Westphal M, Schackert G, Tonn JC et al (2013) Long-term survival in primary glioblastoma with versus without isocitrate dehydrogenase mutations. Clin Cancer Res 19:5146–5157
22. Hartmann C, Meyer J, Balss J, Capper D, Mueller W, Christians A et al (2009) Type and frequency of IDH1 and IDH2 mutations are related to astrocytic and oligodendroglial differentiation and age: a study of 1,010 diffuse gliomas. Acta Neuropathol 118:469–474
23. Heimberger AB, Hlatky R, Suki D, Yang D, Weinberg J, Gilbert M et al (2005) Prognostic effect of epidermal growth factor receptor and EGFRvIII in glioblastoma multiforme patients. Clin Cancer Res 11:1462–1466
24. Houillier C, Lejeune J, Benouaich-Amiel A, Laigle-Donadey F, Criniere E, Mokhtari K et al (2006) Prognostic impact of molecular markers in a series of 220 primary glioblastomas. Cancer 106:2218–2223
25. Hovestadt V, Remke M, Kool M, Pietsch T, Northcott PA, Fischer R et al (2013) Robust molecular subgrouping and copy-number profiling of medulloblastoma from small amounts of archival tumour material using high-density DNA methylation arrays. Acta Neuropathol 125:913–916
26. Huber W, von Heydebreck A, Sültmann H, Poustka A, Vingron M (2002) Variance stabilization applied to microarray data calibration and to the quantification of differential expression. Bioinformatics 18(Suppl 1):S96–104
27. Jones DTW, Ichimura K, Liu L, Pearson DM, Plant K, Collins VP (2006) Genomic analysis of pilocytic astrocytomas at 0.97 Mb resolution shows an increasing tendency toward chromosomal copy number change with age. J Neuropathol Exp Neurol 65:1049–1058
28. Kees T, Lohr J, Noack J, Mora R, Gdynia G, Tödt G et al (2012) Microglia isolated from patients with glioma gain antitumor activities on poly (I:C) stimulation. Neurooncology 14:64–78
29. Kim Y-W, Koul D, Kim SH, Lucio-Eterovic AK, Freire PR, Yao J et al (2013) Identification of prognostic gene signatures of glioblastoma: a study based on TCGA data analysis. Neurooncology 15:829–839
30. Komohara Y, Ohnishi K, Kuratsu J, Takeya M (2008) Possible involvement of the M2 anti-inflammatory macrophage phenotype in growth of human gliomas. J Pathol 216:15–24
31. Krex D, Klink B, Hartmann C, von Deimling A, Pietsch T, Simon M et al (2007) Long-term survival with glioblastoma multiforme. Brain 130:2596–2606
32. Li B, Senbabaoglu Y, Peng W, Yang M-L, Xu J, Li JZ (2012) Genomic estimates of aneuploid content in glioblastoma multiforme and improved classification. Clin Cancer Res 18:5595–5605
33. Lichter P, Bentz M, Joos S (1995) Detection of chromosomal aberrations by means of molecular cytogenetics: painting of chromosomes and chromosomal subregions and comparative genomic hybridization. Methods Enzymol 254:334–359
34. Lohr J, Ratliff T, Huppertz A, Ge Y, Dictus C, Ahmadi R et al (2011) Effector T-cell infiltration positively impacts survival of glioblastoma patients and is impaired by tumor-derived TGF-β. Clin Cancer Res 17:4296–4308
35. Louis DN, Ohgaki H, Wiestler OD, Cavenee WK, Burger PC, Jouvet A et al (2007) The 2007 WHO classification of tumours of the central nervous system. Acta Neuropathol 114:97–109
36. Marko NF, Toms SA, Barnett GH, Weil R (2008) Genomic expression patterns distinguish long-term from short-term glioblastoma survivors: a preliminary feasibility study. Genomics 91:395–406
37. Martinez FO, Gordon S, Locati M, Mantovani A (2006) Transcriptional profiling of the human monocyte-to-macrophage differentiation and polarization: new molecules and patterns of gene expression. J Immunol 177:7303–7311
38. McCabe MG, Ichimura K, Liu L, Plant K, Bäcklund LM, Pearson DM et al (2006) High-resolution array-based comparative genomic hybridization of medulloblastomas and supratentorial primitive neuroectodermal tumors. J Neuropathol Exp Neurol 65:549–561
39. McLendon RE, Halperin EC (2003) Is the long-term survival of patients with intracranial glioblastoma multiforme overstated? Cancer 98:1745–1748
40. Murray PJ, Wynn TA (2011) Protective and pathogenic functions of macrophage subsets. Nat Rev Immunol 11:723–737
41. Nishie A, Ono M, Shono T, Fukushi J, Otsubo M, Onoue H et al (1999) Macrophage infiltration and heme oxygenase-1 expression correlate with angiogenesis in human gliomas. Clin Cancer Res 5:1107–1113
42. Noushmehr H, Weisenberger DJ, Diefes K, Phillips HS, Pujara K, Berman BP et al (2010) Identification of a CpG Island Methylator Phenotype that Defines a Distinct Subgroup of Glioma. Cancer Cell 17:510–522
43. Petalidis LP, Oulas A, Backlund M, Wayland MT, Liu L, Plant K et al (2008) Improved grading and survival prediction of human

astrocytic brain tumors by artificial neural network analysis of gene expression microarray data. Mol Cancer Ther 7:1013–1024
44. Phillips HS, Kharbanda S, Chen R, Forrest WF, Soriano RH, Wu TD et al (2006) Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. Cancer Cell 9:157–173
45. Prosniak M, Harshyne LA, Andrews DW, Kenyon LC, Bedelbaeva K, Apanasovich TV et al (2013) Glioma grade is associated with the accumulation and activity of cells bearing M2 monocyte markers. Clin Cancer Res 19:3776–3786
46. Reifenberger G, Weber RG, Riehmer V, Kaulich K, Willscher E, Wirth H et al (2014) Molecular characterization of long-term survivors of glioblastoma using genome- and transcriptome-wide profiling. Int J Cancer 135:1822–1831
47. Shinawi T, Hill VK, Krex D, Schackert G, Gentle D, Morris MR et al (2013) DNA methylation profiles of long- and short-term glioblastoma survivors. Epigenetics 8:149–156
48. Shinojima N, Tada K, Shiraishi S, Kamiryo T, Kochi M, Nakamura H et al (2003) Prognostic value of epidermal growth factor receptor in patients with glioblastoma multiforme. Cancer Res 63:6962–6970
49. Stupp R, Hegi ME, Mason WP, van den Bent MJ, Taphoorn MJB, Janzer RC et al (2009) Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. Lancet Oncol 10:459–466
50. Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJB et al (2005) Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. N Engl J Med 352:987–996
51. Sturm D, Bender S, Jones DTW, Lichter P, Grill J, Becher O et al (2014) Paediatric and adult glioblastoma: multiform (epi) genomic culprits emerge. Nat Rev Cancer 14:92–107
52. Sturm D, Witt H, Hovestadt V, Khuong-Quang D-A, Jones DTW, Konermann C et al (2012) Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. Cancer Cell 22:425–437
53. The R Development Core Team (2008) R: a language and environment for statistical computing
54. Turcan S, Rohle D, Goenka A, Walsh LA, Fang F, Yilmaz E et al (2012) IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. Nature 483:479–483
55. Verhaak RGW, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD et al (2010) Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell 17:98–110
56. Wesolowska A, Kwiatkowska A, Slomnicki L, Dembinski M, Master A, Sliwa M et al (2008) Microglia-derived TGF-beta as an important regulator of glioblastoma invasion–an inhibition of TGF-beta-dependent effects by shRNA against human TGF-beta type II receptor. Oncogene 27:918–930
57. Xu W, Yang H, Liu Y, Yang Y, Wang P, Kim S-H et al (2011) Oncometabolite 2-Hydroxyglutarate Is a Competitive Inhibitor of α-Ketoglutarate-Dependent Dioxygenases. Cancer Cell 19:17–30
58. Yamanaka R, Arao T, Yajima N, Tsuchiya N, Homma J, Tanaka R et al (2006) Identification of expressed genes characterizing long-term survival in malignant glioma patients. Oncogene 25:5994–6002
59. Yan H, Parsons DW, Jin G, McLendon R, Rasheed BA, Yuan W et al (2009) IDH1 and IDH2 mutations in gliomas. N Engl J Med 360:765–773