Microglia-Based Gene Expression Signature Highly Associated with Prognosis in Low-Grade Glioma

Evelien Schaafsma [1,†], Chongming Jiang [2,†], Thinh Nguyen [2], Kenneth Zhu [3] and Chao Cheng [2,4,5,6,*]

1   Department of Microbiology and Immunology, Dartmouth College, Hanover, NH 03755, USA
2   Department of Medicine, Baylor College of Medicine, Houston, TX 77054, USA
3   Medical School, UT Southwestern Medical Center, Dallas, TX 77054, USA
4   The Institute for Clinical and Translational Research, Baylor College of Medicine, Houston, TX 77054, USA
5   L. Duncan Comprehensive Cancer Center, Baylor College of Medicine, Houston, TX 77054, USA
6   Department of Biomedical Data Science, Geisel School of Medicine at Dartmouth, Lebanon, NH 03755, USA
*   Correspondence: chao.cheng@bcm.edu
†   These authors contributed equally to this work.

Simple Summary: Gliomas make up ~80% of malignant brain tumors in adults and are responsible for the majority of deaths from primary brain tumors. Consequently, a better understanding of the malignant features of the TME in glioma is pertinent. The aim of our study was to evaluate the expression of immune-related genes (IRGs) in glioma and their association with patient prognosis. We utilized several approaches to interrogate the glioma immune microenvironment. We found that immune genes are generally negatively associated with survival and that overall survival was significantly lower in those with a high level of microglia infiltration. The microglia abundance was significantly associated with common genomic aberrations. Lastly, we generated a 23-gene expression signature that is highly associated with patient prognosis, independent of clinical variables. These findings are relevant to investigators in the glioma field, those working in biomarker development, but also to individuals working on glioma therapeutics.

Abstract: Gliomas make up ~80% of malignant brain tumors in adults and are responsible for the majority of deaths from primary brain tumors. The glioma tumor microenvironment (TME) is a dynamic, heterogeneous mixture of extracellular matrix and malignant and non-malignant cells. Several ongoing clinical trials are evaluating the efficacy of therapies that target non-malignant cells, particularly immune cells. Consequently, a better understanding of the TME in glioma is pertinent. We utilized several gene expression datasets to evaluate the relationship between immune-related genes (IRGs) and patient prognosis. We generated microglia signatures using single-cell RNAseq data from human and mouse glioma cells to infer microglia abundance. Lastly, we built a LASSO Cox regression model that predicts patient survival. We found that 428 IRGs were negatively associated with survival in glioma patients. Overall survival was significantly lower in those with a high level of microglia infiltration. In addition, we also found that microglia abundance was significantly associated with several common genomic aberrations, including *IDH2* and *TP53* mutations. Furthermore, we found that patients with high risk scores had significantly worse overall survival than those with low risk scores in several independent datasets. Altogether, we characterized immune features predictive of overall survival in glioma and found that microglia abundance is negatively associated with survival. We developed a 23-gene risk score that can significantly stratify patients into low- and high-risk categories.

Keywords: microglia; glioma; glioblastoma; prognosis

1. Introduction

Gliomas make up ~80% of malignant brain tumors in adults and are responsible for the majority of deaths from primary brain tumors [1]. Gliomas can originate from

different types of glial cells, including astrocytes, oligodendrocytes and ependymal cells. As these cells are present throughout the nervous system, gliomas can appear in various parts of the brain and spinal cord. The 2021 World Health Organization Classification of Tumors of the Central Nervous System (WHO CNS) provides detailed tumor classification guidelines based on histology and immunohistochemistry [2]. Four grades of glioma are distinguished: grades I and II are considered low-grade gliomas (LGGs), and grades III and IV are high-grade gliomas (HGGs). In addition, grade IV gliomas are often designated as glioblastoma (GBM). Treatment is based on molecular profiling and tumor grade and generally involves surgery, adjuvant therapy and radiotherapy [3].

The glioma tumor microenvironment (TME) is a dynamic, heterogeneous mixture of extracellular matrix and malignant and non-malignant cells [4]. Half or more of the cells within the glioma TME are typically non-neoplastic and include neurons, glia, leukocytes and endothelial cells [5]. Valuable insights have been gained from both the study of T-cell biology in brain tumors [6,7] and the investigation of other cell populations contributing to immune suppression in the glioma TME [8–10]. The modulation of these cell populations in the glioma TME could improve the efficacy of immunotherapy against brain malignancies [11,12].

Immune checkpoint inhibitors, which mainly target checkpoint proteins, including PD-1 and anti-PD-L1, on T cells, have not yet significantly improved overall survival for patients with glioma. Examples of evaluated antibodies include Nivolumab, an anti-PD-1 antibody, Pembrolizumab, an anti-PD-1 antibody, and Ipilimumab, an anti-CTLA4 antibody. In a phase II clinical trial in resectable GBM, neoadjuvant Nivolumab (anti-PD-1) prior to tumor resection did not substantiate obvious clinical benefit [13]. Another phase II clinical trial compared Pembrolizumab (anti-PD-1) to Pembrolizumab with Bevacizumab (anti-VEGF) and found that Pembrolizumab alone or with Bevacizumab was well-tolerated but of limited benefit [14]. In a phase III clinical trial comparing Nivolumab (anti-PD-1) to Bevacizumab (anti-VEGF) for recurrent GBM, both treatments conferred similar median overall survival. Lastly, a phase I clinical trial showed that combination therapy of Nivolumab (anti-PD-1) and Ipilimumab (anti-CTLA-4) in patients with recurrent GBM did not improve overall survival but worsened treatment-related adverse events (AEs) [15]. Multiple clinical trials are still underway in glioma and GBM to evaluate, for example, the combination of immune checkpoint inhibitors and radiotherapy [16]. Given the extraordinary success of immune checkpoint inhibitors in other cancer types, including metastatic melanoma or lung cancer, the hope of successfully treating glioma patients with immune checkpoint inhibitors remains.

Consequently, a better understanding of the TME in glioma is pertinent. In this study, we evaluated the expression of immune-related genes (IRGs) in glioma and their association with patient prognosis. We characterized immune features predictive of overall survival and developed a 23-gene risk score that can stratify patients as having a low and high risk of death due to glioma. An overview of our study can be found in Supplementary Figure S1.

2. Materials and Methods

2.1. Data Utilized in this Study

Level 3 TCGA RNAseq data and clinical information for low-grade glioma (LGG, $n = 515$) and glioblastoma multiforme (GBM, $n = 160$) were obtained from TCGA on FireBrowse (gdac.broadinstitute.org/ (accessed on 29 July 2022)). TCGA MAF files for gene mutation analyses were obtained from https://gdc.cancer.gov/about-data/publications/pancanatlas (accessed on 29 July 2022). All genes in which non-silent mutations occurred were considered to be mutated. Total mutation burden (TMB) was represented as the sum of all non-silent mutations in a given TCGA sample. The copy number variation (CNV) data were also downloaded from Firehose, which provided DNA segments deviated from normal copy numbers (copy number = 2) in each tumor sample. The CNV burden of a sample was calculated as the total size (in bp) of genomic regions covered by those segments. Macrophage regulation scores, leukocyte and lymphocyte infiltration scores

and IFNγ response and TGFβ response scores for TCGA samples were downloaded as a supplemental file from prior work [17]. We note here that the TCGA classification of low-grade gliomas includes grades II and III, although grade III gliomas are considered high-grade gliomas clinically. The GBM dataset contains all grade IV gliomas. In addition, TCGA glioma subtypes were assigned based on the fourth WHO CNS version and may not reflect classification based on the latest WHO CNS version [2]. The transcriptomic and clinical data of glioma patients from the Rembrandt data (*n* = 580) [18] were accessed through the Rembrandt data portal (https://sites.google.com/georgetown.edu/g-doc/home (accessed on 29 July 2022)). The transcriptomic and clinical data of glioma patients from the Chinese Glioma Genome Atlas (CGGA) databases (mRNAseq_693, *n* = 693) were downloaded from cgga.org.cn (accessed on 29 July 2022).

2.2. Curation of Immune-Related Genes (IRGs)

Immune-related genes (IRGs) were obtained from Supplementary Table S6 from Charoentong et al. [19], Supplementary Table S1 from Bindea et al. [20] and Supplementary Table S1A from Xu et al. [21]. All genes from immune cells were collected; i.e., marker genes attributed to cancer cells were excluded and combined into a single list of IRGs totaling 831 genes.

2.3. Immune Cell Inference

Immune infiltration scores of six immune cells were calculated using Binding Association with Sorted Expression (BASE) [22], a rank-based gene set enrichment method. Immune cell infiltration using this method has been detailed and validated in previous publications [23,24]. Briefly, BASE uses immune-cell-specific weight profiles and patient gene expression data to infer immune cell infiltration for each patient and immune cell type. Full details on the calculation and validation of the immune infiltration scores can be found in [23,24]. Similarly, BASE was used to calculate single-cell-based microglia scores using microglia signatures (see next section).

2.4. Generation of Microglia Signatures

Glioma single-cell-RNAseq datasets from human [25] and mouse gliomas [26] were obtained from previous publications. Cluster annotations were obtained from these works, and the biological interpretation of each cluster is extensively described in the corresponding publications. For each human/mouse cluster, a list of marker genes was provided by identifying genes that were overexpressed in the corresponding cluster as compared to all other clusters. These cluster-specific marker gene sets were used as microglia signatures [27]. In total, 9 human and 20 mouse microglia signatures were defined. Given an LGG/GBM gene expression dataset, the BASE algorithm was used to calculate sample-specific microglia scores for each signature. For all mouse microglia signature genes, we queried for homologous human genes using biomaRt [28] and used these homologs as signature genes. Of note, microglia signatures were represented as gene sets without assigning weights to genes. In this case, the BASE algorithm degenerated into a method like the single-sample GSEA analysis [29]. A higher microglia score indicates that the corresponding subtype of microglia cells is more abundant in the tumor.

2.5. Lasso Cox Regression

The TCGA LGG dataset was randomly divided into a training and testing set with a 1:1 ratio. The training set was analyzed to identify potential prognostic genes, and both the testing set and the entire set were used for validation. First, univariate Cox proportional hazards regression analysis was used to evaluate the association between the expression of the 831 IRGs and overall survival. Genes with a *p*-value of <0.05 based on the log-rank test were selected as candidate genes. Second, least absolute shrinkage and selection operator (LASSO) Cox regression analysis from the R glmnet package was employed to screen the IRGs most associated with overall survival in a multivariate model, which resulted in

23 genes (*ADAMTSL2, APOBEC3C, ARHGAP12, CARHSP1, CBFB, CD274, EVL, FAM161A, HIST2H2BE, HSPB1, ITGAV, MYBL1, PLAT, RCCD1, SAR1A, SLC25A45, SMC4, STAP1, TFDP1, TGIF1, TMEM55B, TNFRSF11B, WEE1*). These 23 genes composed the final risk score, which is described as follows:

$$iskscore = \sum_{i=0}^{n} \beta_i x_i$$

where β_i refers to the coefficients of each gene and x_i represents the expression value of the gene.

2.6. Survival Analysis

For univariate and multivariate survival analyses, Cox proportional hazards models were calculated using the "coxph" function from the R "survival" package. Survival curves were visualized using Kaplan–Meier curves using the "survfit" function from the R "survival" package. Median immune cell infiltration scores were used to stratify patients into "high" and "low" groups for univariate analyses. For multivariate analyses, an infiltration score of 0 was used as a separator to stratify patients into "high" and "low" groups. Differences in survival distributions in each Kaplan–Meier plot were calculated using a log-rank test using the "survdiff" function from the R "survival" package.

2.7. Statistical Analyses

The Spearman correlation coefficient (SCC) was reported for all correlation analyses as the assumptions underlying the Pearson correlation (i.e., normal distribution, homoscedasticity or linearity) were not met. The SCC was calculated using the R function cor, and significance was assessed using cor.test. Principal component analysis (PCA) was performed using the prcomp R function. Principal component coordinates for each sample were extracted using the factoextra R package (https://github.com/kassambara/factoextra (accessed on 29 July 2022)). Principal component 1 (PC1) was used to represent microglia infiltration. The sensitivity and specificity of the diagnostic and prognostic prediction models were analyzed by the ROC curve and quantified based on the area under the ROC curve (AUC). All statistical tests were two-sided, and *p*-values < 0.05 were considered statistically significant. All statistical analyses were performed using R software (Version 3.5.2).

2.8. Data Availability

All data available in this study are publicly available. These data can be found at: gdac.broadinstitute.org/ (accessed on 29 July 2022), https://gdc.cancer.gov/about-data/publications/pancanatlas (accessed on 29 July 2022), https://sites.google.com/georgetown.edu/g-doc/home (accessed on 29 July 2022), http://www.cgga.org.cn/ (accessed on 29 July 2022).

3. Results

3.1. Immune-Related Genes (IRGs) Are Negatively Associated with Prognosis in Glioma

The glioma TME is heterogeneous and incorporates many different cell types. To focus in on infiltrating immune cells, we compiled a list of 831 immune-related genes (IRGs) from several publications [19–21] and evaluated the expression of these genes in LGGs and GBM. We observed that the majority of IRGs were negatively associated with survival in both LGGs and GBM (Figure 1A,B). This was especially prominent in LGGs, where 428 of the 520 significant IRGs were negatively associated with survival, indicating that higher expression of these IRGs conferred shorter survival. Comparing the expression of IRGs between LGGs and GBM, we found that GBM expressed the highest levels of IRGs, with 294 IRGs overexpressed in GBM and only 71 IRGs overexpressed in LGGs (Figure 1C). Several immune checkpoints were among the IRGs associated with poor survival in LGGs, including *HAVCR2* (log-rank *p* = 0.001), *CD274* (log-rank *p* = 0.001), *CD276* (log-rank *p* = 4×10^{-5}) and *CTLA-4* (log-rank *p* = 0.005) (Figure 1D). Thus, IRGs are overexpressed

in LGGs and GBM tumors with poor prognosis, and these include immune checkpoint molecules. Since many more IRGs were differentially expressed in LGGs as compared to GBM, we focused the remainder of the study on LGGs.
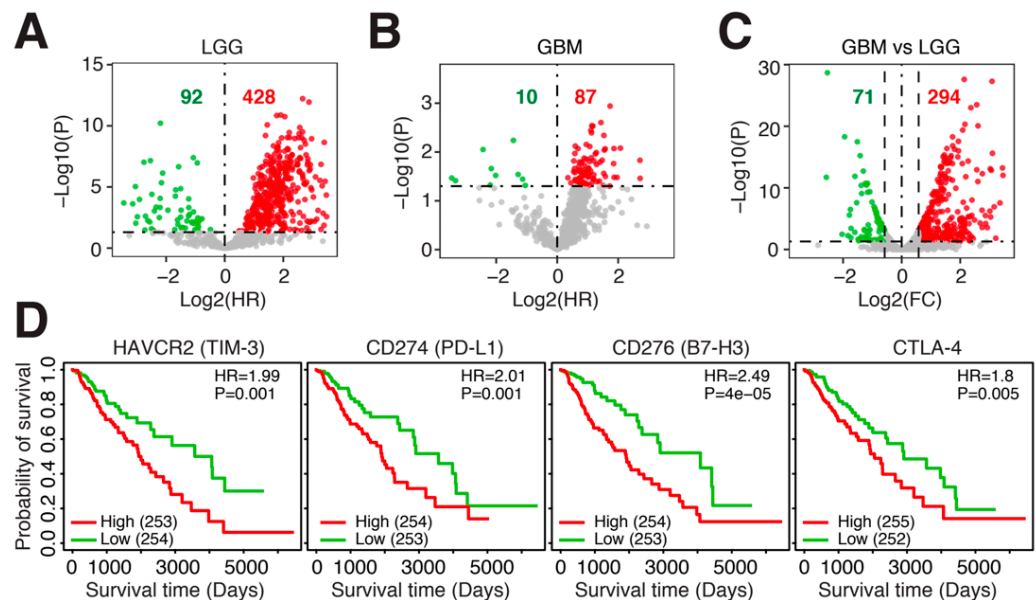

Figure 1. Immune-related genes (IRGs) are negatively associated with prognosis in glioma. Log2 hazard ratio of univariate Cox regression models evaluating the association between overall survival and IRG expression in (**A**). TCGA-LGG (*n* = 515) and (**B**). TCGA-GBM (*n* = 160). (**C**). Log2 fold change between IRG expression in TCGA-LGG (*n* = 515) and TCGA-GBM (*n* = 160). (**D**). Kaplan–Meier plots of four immune checkpoint genes in TCGA-LGG (*n* = 515). HR = hazard ratio. *p* = log-rank *p*-value.

Several types of immune cells infiltrate into the glioma TME. For example, T cells, tumor-associated macrophages (TAMs)/monocytes, NK cells, B cells, neutrophils and dendritic cells (DCs) have all been observed in glioma tumors [10]. To further elucidate which immune cell types confer a poor prognosis in glioma, we estimated the abundance of six immune cell types commonly present in the TME: memory B cells, naïve B cells, CD4+ T cells, CD8+ T cells, NK cells and monocytes. We then clustered patients based on their immune infiltration profile and observed two groups of patients: patients with high infiltration of monocytes but no other immune cell types (group 1) and patients with high infiltration of all immune cell types except monocytes (group 2) (Figure 2A). When comparing overall survival, patients with high monocyte infiltration (group 1) had significantly shorter survival as compared to patients with a broader immune cell infiltration pattern (group 2) (Figure 2B, log-rank $p = 9 \times 10^{-5}$).

We observed a similar pattern when investigating the relationship between overall survival and individual immune cell types; higher infiltration of naïve B cells (log-rank $p = 0.001$), memory B cells (log-rank $p = 8 \times 10^{-4}$) and CD4+ T cells (log-rank $p = 0.001$) was associated with longer overall survival, whereas higher infiltration of monocytes was associated with shorter overall survival in LGGs (log-rank $p = 6 \times 10^{-4}$) (Figure 2C). We confirmed the significant association of naïve B cells and monocytes with survival in two independent datasets (Supplementary Tables S1 and S2). Although CD8+ T cell infiltration was not significantly associated with survival in the TCGA dataset, the two independent LGG datasets both showed a significant association between higher CD8+ T cell infiltration and longer overall survival (Supplementary Tables S1 and S2). Lastly, higher infiltration of both naïve and memory B cells was significantly associated with longer survival in GBM (Supplementary Tables S3 and S4). Thus, immune cell infiltration is associated with prognosis in glioma. In LGGs, high monocyte infiltration is associated with shorter survival, and the infiltration of other immune cells is associated with a better prognosis.
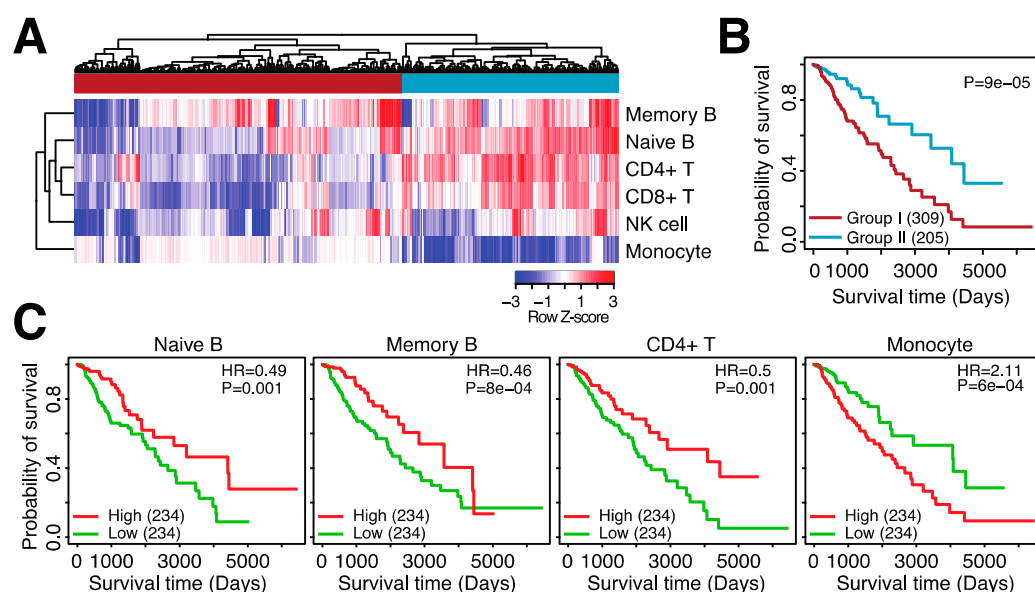
Figure 2. Patterns of immune infiltration in low-grade glioma. (**A**). Heatmap depicting the infiltration scores for six immune cell types in TCGA-LGG. Top sidebar indicates the grouping of the samples based on hierarchical clustering. (**B**). Kaplan–Meier plot depicting differential overall survival between group I and group II. (**C**). Kaplan–Meier plots comparing low and high infiltration of four immune cell types. HR = hazard ratio. p = log-rank p-value.

3.2. Microglia Abundance Is Negatively Associated with Prognosis in Glioma

Microglia are the resident macrophages of the brain [1], and the above infiltration of monocytes likely resembles, at least partially, the infiltration of microglia into the tumor. Indeed, we observed that several genes prominently expressed in microglia, *CCR5* and *TREM2*, were also negatively associated with patient survival (Figure 3A), mirroring the negative association between monocytes and overall survival (Figure 2C). We thus sought to improve the inferred monocyte profile to resemble microglia infiltration more closely. We obtained two single-cell RNA-seq (scRNA-seq) datasets, one from human [25] and one from murine gliomas [26], and isolated all cell clusters that were designated as microglia cells by the study authors. This resulted in 9 clusters for human and 20 clusters for mouse microglia cells. For each microglia cluster, we generated a gene expression signature (see Methods) and inferred the abundance of these signatures in LGG patients.

We observed that almost all microglia signatures were strongly associated with monocyte infiltration (Supplementary Figure S2), again suggesting that our monocyte signature likely captures signals from infiltrating microglia cells. In addition, inferred microglia abundance was highly positively correlated with several markers expressed in microglia cells, including ITGAM, CCR5 and FCGR2A (Supplementary Figure S3A). Lastly, the majority of IRGs (Supplementary Figure S3B) and immune checkpoint genes (Supplementary Figure S3C) were positively associated with microglia signature scores. It thus seems that our inferred microglia scores represent intratumoral microglia abundance and are positively associated with IRG expression, including checkpoint genes.

We next assessed the relationship between microglia infiltration and overall survival. Almost all 29 signatures, except for 1 human and 3 mouse signatures, were negatively associated with overall survival (Figure 3B). For example, higher inferred infiltration of hC1_MG1, resembling microglia with homeostatic functions [25], and hC9_MG9, resembling proliferating microglial cells [25], was associated with shorter overall survival (Figure 3C). In addition, higher abundance of mC3_Ube2c, resembling proliferating microglial cells [26], and mC9_center, resembling microglia with interferon expression (26), was also associated with shorter overall survival (Figure 3C). When correlating the inferred levels of all signatures with each other, we noticed that all signatures were highly positively correlated with one another (Figure 3D), suggesting high overlap in the signals captured by

these different signatures. We thus aimed to unify these signatures and generate a single signature that resembles all microglia signatures. We performed dimensionality reduction (principle component analysis (PCA)) on the inferred human microglia signature scores and observed that the first principle component (PC) was highly positively correlated with all microglia signatures (Supplementary Figure S3D) and captured 72% of the variation among patients (Figure 3E). This suggested that PC1 can represent all human microglia signatures, and we thus utilized PC1 to represent overall microglia infiltration. When comparing overall survival between patients with high microglia abundance (PC1-high) and low microglia abundance (PC1-low), we observed that patients with high microglia abundance had significantly shorter survival (log-rank $p = 4 \times 10^{-5}$, Figure 3F). In conclusion, high microglial infiltration can be captured by gene-expression-based microglia abundance inference and is associated with shorter survival.
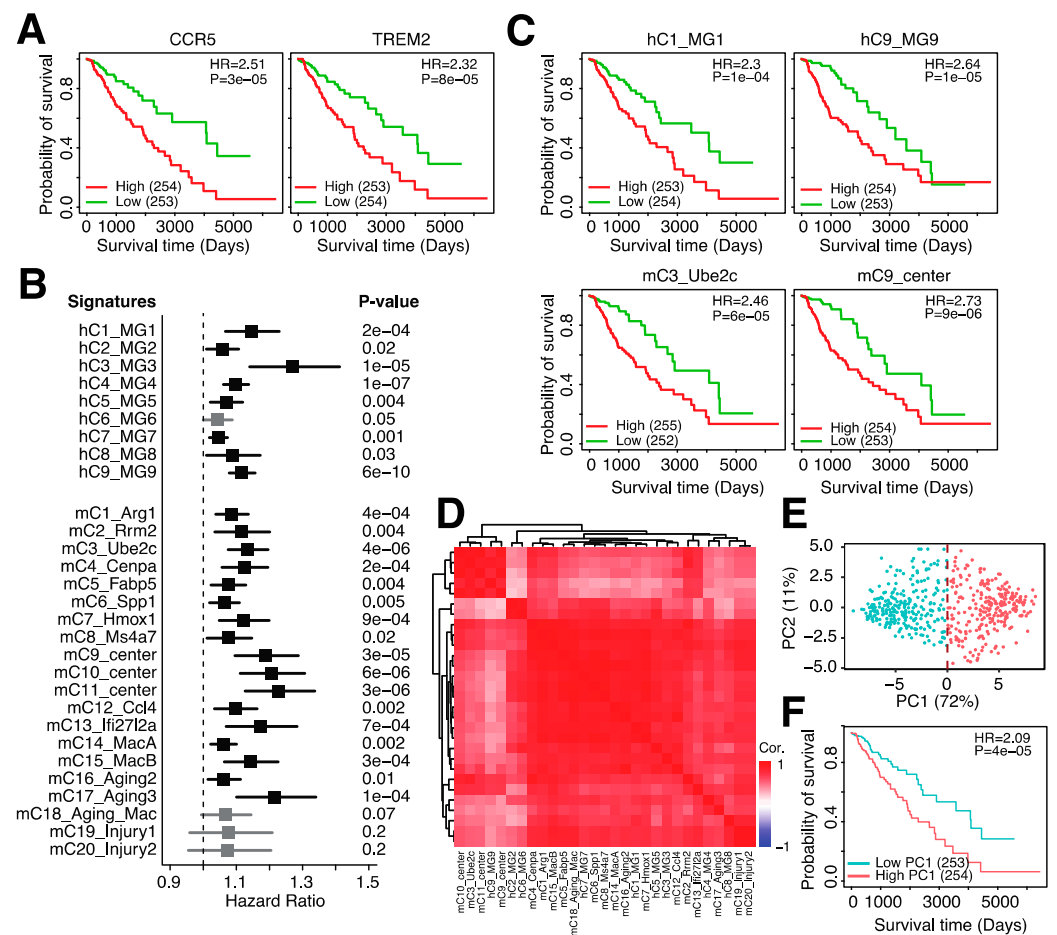

Figure 3. Microglia abundance is negatively associated with prognosis in low-grade gliomas. (**A**). Kaplan–Meier plot comparing overall survival between patients with low and high expression of *CCR5* (left) and *TREM2* (right). (**B**). Hazard ratios of univariate Cox regression models evaluating the association between overall survival and 9 human microglia signatures ("h") and 20 mouse microglia signatures ("m"). (**C**). Kaplan–Meier plots depicting the association between overall survival and four microglia signatures. (**D**). Spearman correlation between the 29 microglia gene expression signatures. (**E**). Principle component analysis (PCA) on the expression of the 29 microglia gene expression signature in TCGA-LGG patients. (**F**). Kaplan–Meier plot showing the association between overall survival and principle component 1 (PC1) in TCGA-LGG.

We next sought to validate our findings in the TCGA-LGG dataset in independent gene expression datasets. Since both independents datasets are based on human glioma biopsies, we focused on the 9 human microglia signatures. We indeed confirmed that almost all human signatures were significantly associated with overall survival in the Rem-

brandt and CGGA datasets and that PC1-high patients had significantly shorter survival (Figure 4A–D), confirming our earlier findings. Several subtypes of glioma exist, with astrocytoma, oligoastrocytoma and oligodendroglioma being the most common subtypes [1]. As each subtype originates from a different cell type and consequently confers different morphologies, we investigated what would occur if microglia abundance were different as well. We indeed observed a difference among subtypes: astrocytomas contained the highest levels of microglia infiltration, whereas oligodendrogliomas had the lowest levels of microglia infiltration (Figure 4E). When assessing overall survival, no difference in overall survival in astrocytomas was observed when comparing high vs. low microglia infiltration, but patients with lower microglia infiltration had significantly longer survival than patients with oligoastrocytoma and oligodendroglioma (Figure 4F). Lastly, we evaluated the relationship between overall survival and microglial abundance while adjusting for several important clinical variables known to be associated with prognosis. We observed that microglia infiltration was still associated with prognosis in multivariate regression analysis (Figure 4G), suggesting that microglia abundance is an independent prognosis marker. In conclusion, microglia infiltration is associated with survival, varies among glioma subtypes and can be used as a prognostic marker independent of clinical variables.
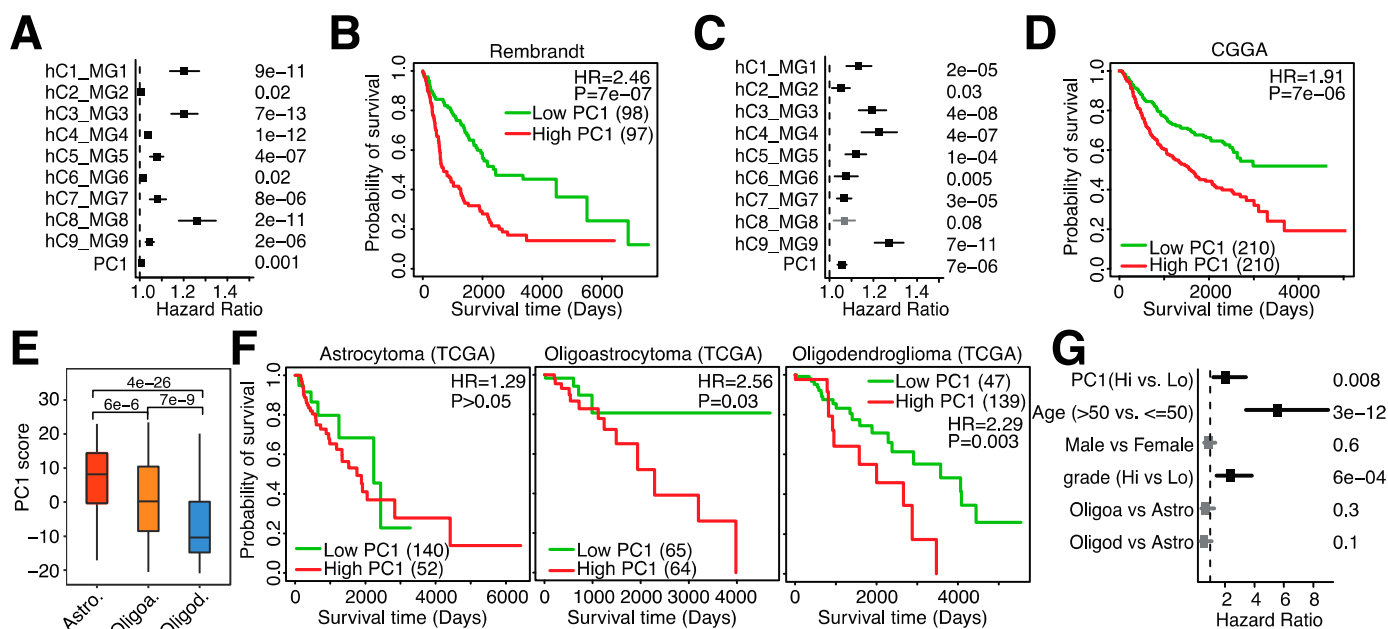

Figure 4. Microglia abundance is negatively associated with prognosis in independent datasets and subtypes of low-grade glioma. (**A**). Hazard ratios of univariate Cox regression models evaluating the association between overall survival and 9 human microglia signatures and PC1 in the Rembrandt dataset. (**B**). Kaplan–Meier plot depicting the association between overall survival and PC1 in the Rembrandt dataset. (**C**). Hazard ratios of univariate Cox regression models evaluating the association between overall survival and 9 human microglia signatures and PC1 in the CGGA dataset. (**D**). Kaplan–Meier plot depicting the association between overall survival and PC1 in the CGGA dataset (mRNAseq_693). (**E**). Boxplot comparing PC1 scores among glioma subtypes. Astro. = astrocytoma, Oligoa. = oligoastrocytoma and Oligod. = oligodendroglioma. (**F**). Kaplan–Meier plots depicting the association between overall survival and PC1 in different glioma subtypes in the TCGA dataset. (**G**). Forest plot depicting hazard ratios of univariate Cox regression models evaluating the association between overall survival and several clinical parameters.

3.3. Microglia Infiltration Is Associated with the mRNA Expression of Immune Checkpoint Genes and Immune Regulatory Pathways

The glioma TME is often highly infiltrated by non-neoplastic cells, with half or more of the cells within the glioma TME being non-neoplastic [5]. We thus wanted to evaluate

the relationship between microglial abundance and non-neoplastic cells, particularly immune cells, in more detail. We first correlated microglial abundance with our curated IRGs and observed that a substantial fraction of IRGs was positively associated with microglial abundance, including immune checkpoint genes such as C10orf54 (VISTA), HAVCR2 (TIM-3) and CD274 (PD-L1) (Figure 5A). In addition, genes expressed predominantly in microglia were strongly positively associated with microglial abundance (Figure 5B), including CD68 and MSR1 (Figure 5C), suggesting again that PC1 represents microglial infiltration. Similarly, a macrophage regulation score was also positively associated with microglia abundance (Figure 5D). This score represents a colony-stimulating factor-1 (CSF1) response [30], which is the primary regulator of tissue macrophages and induces proliferation, differentiation and survival of macrophages and microglia [31]. Lastly, microglia abundance was more highly correlated with M2 macrophages (SCC = 0.50) as compared to M1 macrophages (SCC = 0.17) (Supplementary Figure S4), suggesting that microglia abundance is associated with a suppressive TME.
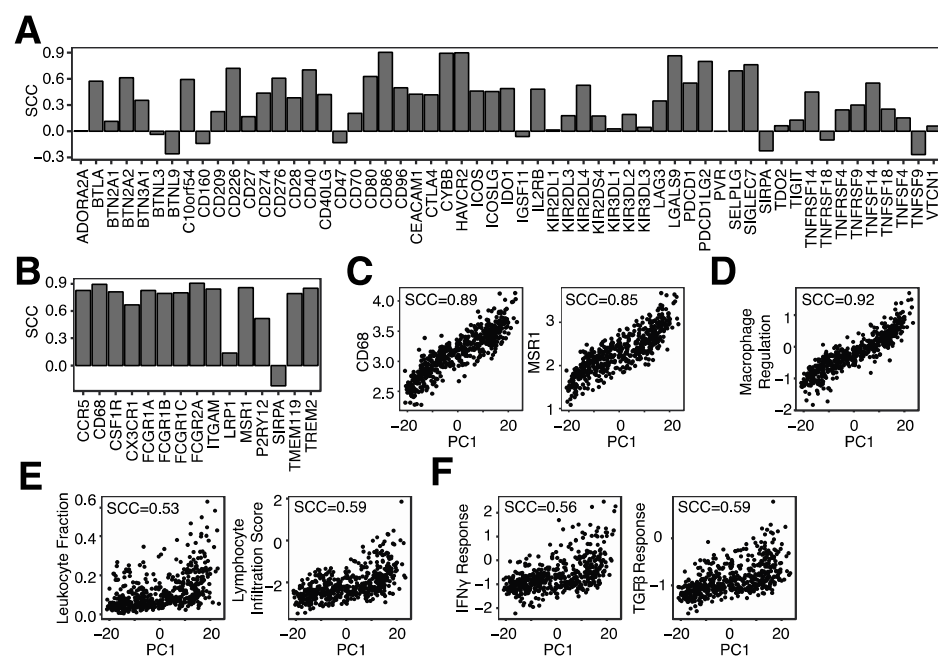

Figure 5. Microglia infiltration is associated with the mRNA expression of immune checkpoint genes and immune regulatory pathways. (**A**). Spearman correlation coefficient (SCC) between PC1 and the mRNA expression of several immune checkpoint genes. (**B**). SCC between PC1 and mRNA genes expressed in microglia. (**C**). SCC between PC1 and CD86 and MSR1 mRNA levels. (**D**). SCC between PC1 and a macrophage regulation score. (**E**). SCC between PC1 and leukocyte and lymphocyte infiltration. (**F**). SCC between PC1 and IFNγ response and TGFβ response.

When assessing the relationship between microglia infiltration and general immune cell infiltration, we observed a positive association with both leukocyte and lymphocyte infiltration (Figure 5E). We previously noticed a positive correlation between single-cell-based microglial signatures and B-cell infiltration (Supplementary Figure S2) and hypothesize that the positive correlation observed in Figure 5E was at least in part due to B-cell infiltration. Lastly, we inferred pathway activity of several immune pathways, including IFNγ and TGFβ pathways. Both pathways were positively correlated with microglia infiltration (Figure 5F). In conclusion, microglia infiltration is associated with the expression of many immune checkpoint proteins and immune-related pathways.

3.4. Microglia Infiltration Is Associated with Specific Genomic Alterations

After evaluating the relationship between immune-related characteristics and microglia infiltration, we next investigated the relationship between genomic alterations and

microglia infiltration. Several genomic alterations are prominent in glioma, including isocitrate dehydrogenase (*IDH*) mutations and *CDKNA2A* deletions. We first evaluated somatic mutations and observed that several common gene mutations were associated with microglia infiltration (Figure 6A). For example, patients with *TP53* mutations tended to have higher levels of microglia in their tumor, whereas patients with *CIC* or *IDH2* mutations tended to have lower microglia infiltration (Figure 6B). A large number of copy number variations (CNVs) were associated with microglia infiltration, most notably EGFR amplifications and CDKNA2A deletions; both of these genomic abnormalities were associated with increased microglia infiltration (Figure 6C). Lastly, neither overall mutation counts nor overall CNVs were not associated with microglia infiltration (Figure 6D,E). These results suggest that specific genomic alterations are involved in microglia infiltration and not the general level of genomic abnormalities.
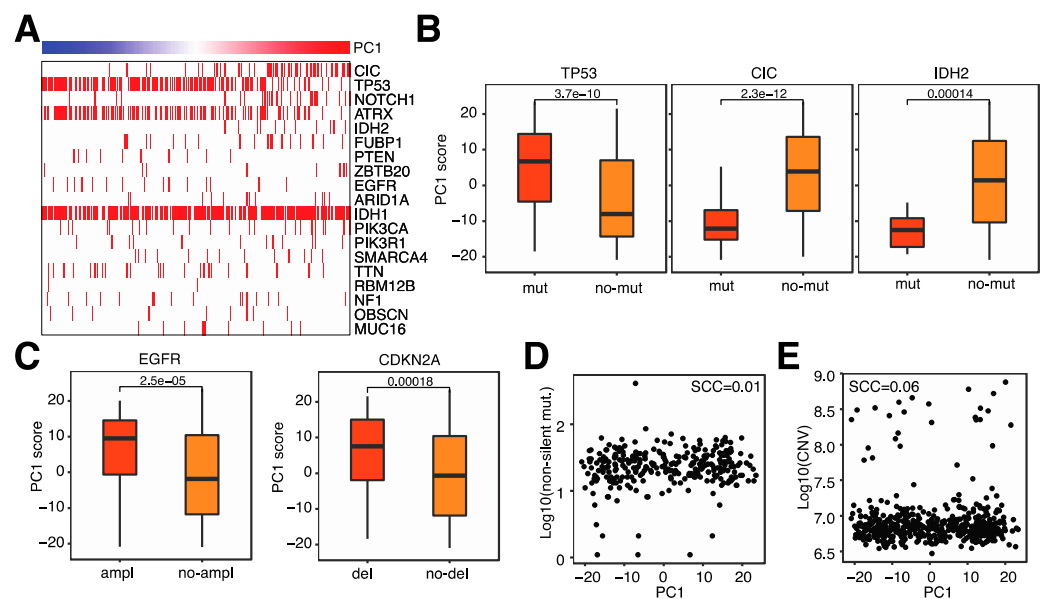

Figure 6. Microglia infiltration is associated with specific genomic alterations. (**A**). Heatmap indicating the incidence of coding mutations (horizontal red bars) and the level of microglia infiltration (top bar). (**B**). Box plot of microglia infiltration in patients with or without coding mutations in the specified genes. (**C**). Box plot of microglia infiltration in patients with or without copy number variations (CNVs) in the specified genes. (**D**). Relationship between microglia infiltration and the number of non-silent mutations. (**E**). Relationship between microglia infiltration and total CNVs.

3.5. A 23-Gene Risk Score Is Highly Associated with Overall Survival in Glioma

As we have shown previously, microglia infiltration is associated with survival independently of clinical variables (Figure 4). However, the inference of microglia abundance throughout our studies has been based on a large number of genes, each contributing to the final infiltration score. In order to generate a potentially clinically useful tool, we thus aimed to significantly reduce the number of genes in a signature that can be used to risk-stratify glioma patients. We used Lasso Cox regression for feature (gene) selection and ended up with a 23-gene risk score (Methods, Supplementary Figure S5A). The following genes were included in the risk score: *ADAMTSL2, APOBEC3C, ARHGAP12, CARHSP1, CBFB, CD274, EVL, FAM161A, HIST2H2BE, HSPB1, ITGAV, MYBL1, PLAT, RCCD1, SAR1A, SLC25A45, SMC4, STAP1, TFDP1, TGIF1, TMEM55B, TNFRSF11B* and *WEE1*. The expression of these genes was weighted and summed to generate a final risk score ($riskscore = \sum_{i=0}^{n} \beta_i x_i$ where β_i refers to the coefficients of each gene and x_i represents the expression value of the gene). This signature was significantly associated with glioma survival in the TCGA training dataset (Figure 7A) and in two independent datasets (Figure 7B,C), with higher risk scores being associated with shorter overall survival. This signature is specific to glioma and not

GBM as only one of the three datasets showed a significant relationship between overall survival and the risk score in GBM (Figure 7D, Supplementary Figure S5B). Notably, the risk score was highly associated with survival when evaluated in a multivariate model with several clinical variables (Figure 7E). In this model, the hazard risk (HR) of our risk score was much larger than that of tumor grade, a commonly used risk stratification measure in glioma [3]. In conclusion, we developed a 23-gene risk score that is highly associated with prognosis in glioma and can be used in addition to other clinical variables.
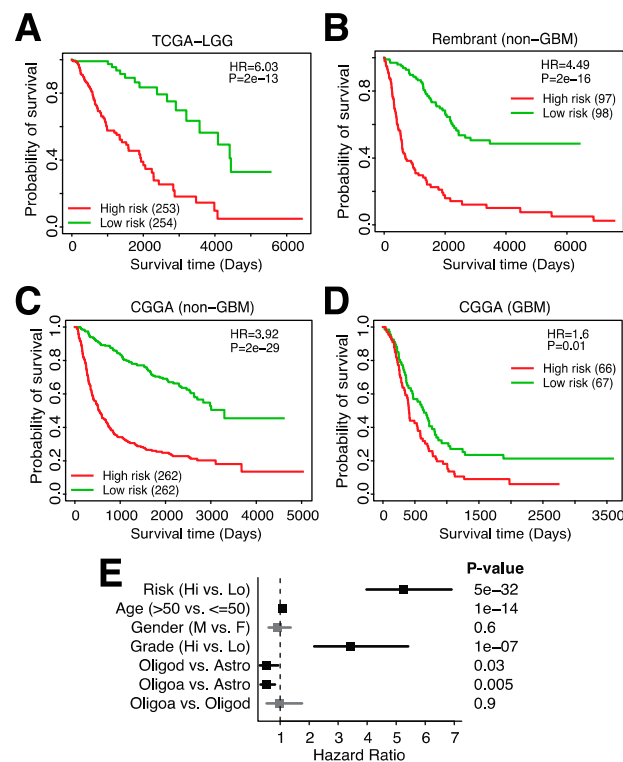

Figure 7. A 23-gene risk score is highly associated with overall survival in low-grade glioma. Kaplan–Meier plot showing the association between overall survival and the 23-gene risk score in (**A**). TCGA-LGG, (**B**). the Rembrandt dataset (mRNAseq_693, non-GBM), (**C**). CGGA (mRNAseq_693, non-GBM) and (**D**). CGGA (GBM). (**E**). Forest plot of hazard ratios derived from multivariable Cox regression analysis including the listed variables.

4. Discussion

Glioma is a heterogeneous disease in which many non-malignant cells infiltrate into the TME. Several clinical trials are underway to assess the efficacy of targeting these non-malignant cells [5], especially immune cells. To better understand the relationship between the immune system and patient survival in gliomas, we first started out by investigating the relationship between IRG expression and patient prognosis. We found that the overexpression of many IRGs is associated with shorter overall survival. However, this is contrary to the expression of immune genes in several other cancer types; expression of IRGs is associated with better survival in, for example, melanoma and kidney cancer [19]. Notably, several immune checkpoint genes were among the IRGs negatively associated with survival. Thus, this led us to investigate the effect of immune checkpoint inhibitors as a possible mechanism for these unique effects.

Further investigation showed that the abundance of monocytes, but not other immune cell types, was negatively associated with prognosis. As microglia are the resident macrophages of the brain [5], we reasoned that this monocyte profile is likely in part reflective of microglia abundance. We used scRNAseq-based microglia signatures to evaluate the infiltration of microglia in a more efficient manner. These signatures, as well as a compiled

microglia signature, were positively associated with the majority of IRGs and immune checkpoint genes. We hypothesize that microglia express several immune checkpoint genes and in turn negatively regulate T-cell activation. In addition, microglia potentially act as antigen-presenting cells (APCs), given their ability to upregulate the expression of MHC class II and costimulatory molecules [32]. Thus, in the presence of immune checkpoint expression, T-cell activation will be attenuated, and no effective antitumor response would be initiated. Alternatively, in pathological states, microglia produce proinflammatory factors such as cytokines and chemokines [33] that promote tumor invasiveness [34].

Microglia abundance was associated with several genomic alterations. For example, tumors with IDH2 mutations had significantly lower infiltration of microglia as compared to tumors without IDH2 mutations. It is known that IDH wildtype gliomas have a poorer prognosis and are molecularly similar to GBM [35]. This is consistent with our finding that higher microglia infiltration is associated with shorter survival. CNVs in EGFR and CDKNA2A often occur in the absence of IDH mutations [35]. These alterations were consequently associated with higher microglia infiltration. Overall, TMB and CNV burden were not associated with microglia infiltration, suggesting that specific genomic alterations affect microglia infiltration and not the magnitude of genomic aberrations.

Lastly, we developed a 23-gene risk score that can stratify high- and low-risk patients irrespective of the predictive ability of clinical variables. This risk score was validated in independent datasets and was strongly associated with prognosis. Interestingly, prognosis could only be predicted in gliomas but not high-grade gliomas, or glioblastoma (GBM). It was also notable that a much smaller number of IRGs were significantly and differentially expressed in GBM as compared to LGGs. This is likely due to the more aggressive phenotype of GBM as compared to a glioma, which confers a high risk itself. We do note that the TCGA-LGG dataset contains includes both grades II and III, although grade III gliomas are considered high-grade gliomas clinically. Thus, this observation seems specific to high-grade grade IV gliomas. The 23-gene signature could be clinically relevant as only a small number of genes need to be measured to establish a patient-specific risk score.

Previous studies have aimed at identifying TME-related factors associated with prognosis. For example, Ni et al. (2020) used the ESTIMATE algorithm [36] in combination with network analysis to identify a signature of 25 immune genes associated with prognosis [37]. Similar to our study, high signature scores, i.e., high immune infiltration, was associated with shorter overall survival. While this study identified a non-specific immune signature that was validated in one dataset, we developed a microglia-specific signature and validated our signature in multiple independent datasets. The increased specificity of our signature adds biological relevance and could aid in patient selection for microglia-specific interventions in the future.

Although our study provides valuable insights into the potential role of microglia in LGGs, we note a few limitations of our study. First, we used immune inference methods to infer the abundance of immune cells and microglia. While these methods have been thoroughly validated, a more precise quantification using, for example, immunohistochemistry, would complement our inference estimates. Second, the expression of immune checkpoints was based on mRNA data, which might not reflect protein levels. Lastly, our 23-gene signature needs to be validated in a prospective study to assess clinical efficacy. The potential difference between risk score quantification in FFPE or fresh glioma samples would also need to be evaluated to establish clinical utility.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/cancers14194802/s1, Figure S1: Overview of our study; Figure S2: Correlation between immune cell infiltration and microglia abundance; Figure S3: Correlation of IRGs with microglia infiltration; Figure S4: Spearman correlation between PC1 and (A) M1 macrophage and (B) M2 macrophage scores; Figure S5: Development of a microglia risk score for glioma by LASSO-cox regression analysis. Table S1: Immune cell infiltration in glioma (Rembrandt); Table S2: Immune cell infiltration in glioma (CGGA); Table S3: Immune cell infiltration in GBM (Rembrandt); Table S4: Immune cell infiltration in GBM (CGGA).

References

1. Weller, M.; Wick, W.; Aldape, K. Glioma. *Nat. Rev. Dis. Primers* **2015**, *1*, 15017. [CrossRef] [PubMed]
2. Louis, D.N.; Perry, A.; Wesseling, P.; Brat, D.J.; Cree, I.A.; Figarella-Branger, D.; Hawkins, C.; Ng, H.K.; Pfister, S.M.; Reifenberger, G.; et al. The 2021 WHO Classification of Tumors of the Central Nervous System: A summary. *Neuro-Oncol.* **2021**, *23*, 1231–1251. [CrossRef] [PubMed]
3. Nabors, L.B.; Portnow, J.; Ahluwalia, M.; Baehring, J.; Brem, H.; Brem, S.; Butowski, N.; Campian, J.L.; Clark, S.W.; Fabiano, A.J.; et al. Central Nervous System Cancers, Version 3.2020, NCCN Clinical Practice Guidelines in Oncology. *Natl. Compr. Cancer Netw.* **2020**, *18*, 1537–1570. [CrossRef] [PubMed]
4. Varn, F.S.; Johnson, K.C.; Martinek, J.; Huse, J.T.; Nasrallah, M.P.; Wesseling, P.; Cooper, L.A.; Malta, T.M.; Wade, T.E.; Sabedot, T.S.; et al. Glioma progression is shaped by genetic evolution and microenvironment interactions. *Cell* **2022**, *185*, 2184–2199.e16. [CrossRef]
5. Andersen, B.M.; Akl, C.F.; Wheeler, M.A.; Chiocca, E.A.; Reardon, D.A.; Quintana, F.J. Glial and myeloid heterogeneity in the brain tumour microenvironment. *Nat. Cancer* **2021**, *21*, 786–802. [CrossRef]
6. Kwok, D.; Okada, H. T-Cell based therapies for overcoming neuroanatomical and immunosuppressive challenges within the glioma microenvironment. *Neuro-Oncol.* **2020**, *147*, 281–295. [CrossRef]
7. Mathewson, N.D.; Ashenberg, O.; Tirosh, I.; Gritsch, S.; Perez, E.M.; Marx, S.; Jerby-Arnon, L.; Chanoch-Myers, R.; Hara, T.; Richman, A.R.; et al. Inhibitory CD161 receptor identified in glioma-infiltrating T cells by single-cell analysis. *Cell* **2021**, *184*, 1281–1298.e26. [CrossRef]
8. Klemm, F.; Maas, R.R.; Bowman, R.L.; Kornete, M.; Soukup, K.; Nassiri, S.; Brouland, J.-P.; Iacobuzio-Donahue, C.A.; Brennan, C.; Tabar, V.; et al. Interrogation of the Microenvironmental Landscape in Brain Tumors Reveals Disease-Specific Alterations of Immune Cells. *Cell* **2020**, *181*, 1643–1660.e17. [CrossRef]
9. Goswami, S.; Walle, T.; Cornish, A.E.; Basu, S.; Anandhan, S.; Fernandez, I.; Vence, L.; Blando, J.; Zhao, H.; Yadav, S.S.; et al. Immune profiling of human tumors identifies CD73 as a combinatorial target in glioblastoma. *Nat. Med.* **2019**, *26*, 39–46. [CrossRef]
10. Friebel, E.; Kapolou, K.; Unger, S.; Núñez, N.G.; Utz, S.; Rushing, E.J.; Regli, L.; Weller, M.; Greter, M.; Tugues, S.; et al. Single-Cell Mapping of Human Brain Cancer Reveals Tumor-Specific Instruction of Tissue-Invading Leukocytes. *Cell* **2020**, *181*, 1626–1642.e20. [CrossRef]
11. Pyonteck, S.M.; Akkari, L.; Schuhmacher, A.J.; Bowman, R.L.; Sevenich, L.; Quail, D.F.; Olson, O.C.; Quick, M.L.; Huse, J.T.; Teijeiro, V.; et al. CSF-1R inhibition alters macrophage polarization and blocks glioma progression. *Nat. Med.* **2013**, *19*, 1264–1272. [CrossRef] [PubMed]
12. Fujita, M.; Kohanbash, G.; Fellows-Mayle, W.; Hamilton, R.L.; Komohara, Y.; Decker, S.A.; Ohlfest, J.R.; Okada, H. COX-2 Blockade Suppresses Gliomagenesis by Inhibiting Myeloid-Derived Suppressor Cells. *Cancer Res.* **2011**, *71*, 2664–2674. [CrossRef] [PubMed]
13. Schalper, K.A.; Rodriguez-Ruiz, M.E.; Diez-Valle, R.; López-Janeiro, A.; Porciuncula, A.; Idoate, M.A.; Inogés, S.; De Andrea, C.; De Cerio, A.L.-D.; Tejada, S.; et al. Neoadjuvant nivolumab modifies the tumor immune microenvironment in resectable glioblastoma. *Nat. Med.* **2019**, *25*, 470–476. [CrossRef]
14. Nayak, L.; Molinaro, A.M.; Peters, K.B.; Clarke, J.L.; Jordan, J.T.; de Groot, J.F.; Nghiemphu, P.L.; Kaley, T.J.; Colman, H.; McCluskey, C.; et al. Randomized Phase II and Biomarker Study of Pembrolizumab plus Bevacizumab versus Pembrolizumab Alone for Patients with Recurrent Glioblastoma. *Clin. Cancer Res.* **2021**, *27*, 1048–1057. [CrossRef]
15. Omuro, A.; Vlahovic, G.; Lim, M.; Sahebjam, S.; Baehring, J.; Cloughesy, T.; Voloschin, A.; Ramkissoon, S.H.; Ligon, K.L.; Latek, R.; et al. Nivolumab with or without ipilimumab in patients with recurrent glioblastoma: Results from exploratory phase I cohorts of CheckMate 143. *Neuro Oncol.* **2018**, *20*, 674–686. [CrossRef]

16. Ghouzlani, A.; Kandoussi, S.; Tall, M.; Reddy, K.P.; Rafii, S.; Badou, A. Immune Checkpoint Inhibitors in Human Glioma Microenvironment. *Front. Immunol.* **2021**, *12*, 679425. [CrossRef]

17. Thorsson, V.; Gibbs, D.L.; Brown, S.D.; Wolf, D.; Bortone, D.S.; Ou Yang, T.-H.; Porta-Pardo, E.; Gao, G.F.; Plaisier, C.L.; Eddy, J.A.; et al. The Immune Landscape of Cancer. *Immunity* **2018**, *48*, 812–830.e14. [CrossRef]

18. Madhavan, S.; Zenklusen, J.-C.; Kotliarov, Y.; Sahni, H.; Fine, H.A.; Buetow, K. Rembrandt: Helping Personalized Medicine Become a Reality through Integrative Translational Research. *Mol. Cancer Res.* **2009**, *7*, 157–167. [CrossRef]

19. Charoentong, P.; Finotello, F.; Angelova, M.; Mayer, C.; Efremova, M.; Rieder, D.; Hackl, H.; Trajanoski, Z. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep.* **2017**, *18*, 248–262. [CrossRef]

20. Bindea, G.; Mlecnik, B.; Tosolini, M.; Kirilovsky, A.; Waldner, M.; Obenauf, A.C.; Angell, H.; Fredriksen, T.; Lafontaine, L.; Berger, A.; et al. Spatiotemporal Dynamics of Intratumoral Immune Cells Reveal the Immune Landscape in Human Cancer. *Immunity* **2013**, *39*, 782–795. [CrossRef]

21. Xu, L.; Deng, C.; Pang, B.; Zhang, X.; Liu, W.; Liao, G.; Yuan, H.; Cheng, P.; Li, F.; Long, Z.; et al. TIP: A Web Server for Resolving Tumor Immunophenotype Profiling. *Cancer Res.* **2018**, *78*, 6575–6580. [CrossRef] [PubMed]

22. Cheng, C.; Yan, X.; Sun, F.; Li, L.M. Inferring activity changes of transcription factors by binding association with sorted expression profiles. *BMC Bioinform.* **2007**, *8*, 452. [CrossRef] [PubMed]

23. Varn, F.S.; Tafe, L.J.; Amos, C.I.; Cheng, C. Computational immune profiling in lung adenocarcinoma reveals reproducible prognostic associations with implications for immunotherapy. *Oncoimmunology* **2018**, *7*, e1431084. [CrossRef]

24. Varn, F.S.; Wang, Y.; Mullins, D.W.; Fiering, S.; Cheng, C. Systematic Pan-Cancer Analysis Reveals Immune Cell Interactions in the Tumor Microenvironment. *Cancer Res.* **2017**, *77*, 1271–1282. [CrossRef] [PubMed]

25. Olah, M.; Menon, V.; Habib, N.; Taga, M.F.; Ma, Y.; Yung, C.J.; Cimpean, M.; Khairallah, A.; Coronas-Samano, G.; Sankowski, R.; et al. Single cell RNA sequencing of human microglia uncovers a subset associated with Alzheimer's disease. *Nat. Commun.* **2020**, *11*, 1–18. [CrossRef]

26. Hammond, T.R.; Dufort, C.; Dissing-Olesen, L.; Giera, S.; Young, A.; Wysoker, A.; Walker, A.J.; Gergits, F.; Segel, M.; Nemesh, J.; et al. Single-Cell RNA Sequencing of Microglia throughout the Mouse Lifespan and in the Injured Brain Reveals Complex Cell-State Changes. *Immunity* **2019**, *50*, 253–271.E6. [CrossRef]

27. De Boeck, A.; Ahn, B.Y.; D'Mello, C.; Lun, X.; Menon, S.V.; Alshehri, M.M.; Szulzewsky, F.; Shen, Y.; Khan, L.; Dang, N.H.; et al. Glioma-derived IL-33 orchestrates an inflammatory brain tumor microenvironment that accelerates glioma progression. *Nat. Commun.* **2020**, *11*, 4997. [CrossRef]

28. Durinck, S.; Spellman, P.T.; Birney, E.; Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **2009**, *4*, 1184–1191. [CrossRef]

29. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15545–15550. [CrossRef]

30. Beck, A.H.; Espinosa, I.; Edris, B.; Li, R.; Montgomery, K.; Zhu, S.; Varma, S.; Marinelli, R.J.; van de Rijn, M.; West, R.B. The Macrophage Colony-Stimulating Factor 1 Response Signature in Breast Carcinoma. *Clin. Cancer Res.* **2009**, *15*, 778–787. [CrossRef]

31. Stanley, E.R.; Chitu, V. CSF-1 Receptor Signaling in Myeloid Cells. *Cold Spring Harb. Perspect. Biol.* **2014**, *6*, a021857. [CrossRef] [PubMed]

32. Sankowski, R.; Böttcher, C.; Masuda, T.; Geirsdottir, L.; Sagar; Sindram, E.; Seredenina, T.; Muhs, A.; Scheiwe, C.; Shah, M.J.; et al. Mapping microglia states in the human brain through the integration of high-dimensional techniques. *Nat. Neurosci.* **2019**, *22*, 2098–2110. [CrossRef] [PubMed]

33. Hanisch, U.-K.; Kettenmann, H. Microglia: Active sensor and versatile effector cells in the normal and pathologic brain. *Nat. Neurosci.* **2007**, *10*, 1387–1394. [CrossRef]

34. Zhang, J.; Sarkar, S.; Cua, R.; Zhou, Y.; Hader, W.; Yong, V.W. A dialog between glioma and microglia that promotes tumor invasiveness through the CCL2/CCR2/interleukin-6 axis. *Carcinogenesis* **2012**, *33*, 312–319. [CrossRef] [PubMed]

35. Cancer Genome Atlas Research Network. Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. *N. Engl. J. Med.* **2015**, *372*, 2481–2498. [CrossRef] [PubMed]

36. Yoshihara, K.; Shahmoradgoli, M.; Martínez, E.; Vegesna, R.; Kim, H.; Torres-Garcia, W.; Trevino, V.; Shen, H.; Laird, P.W.; Levine, D.A.; et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* **2013**, *4*, 2612. [CrossRef]

37. Ni, J.; Liu, S.; Qi, F.; Li, X.; Yu, S.; Feng, J.; Zheng, Y. Screening TCGA database for prognostic genes in lower grade glioma microenvironment. *Ann. Transl. Med.* **2020**, *8*, 209. [CrossRef]