

The Landscape of Viral Expression Reveals Clinically Relevant Viruses with Potential Capability of Promoting Malignancy in Lower-Grade Glioma

Zheng Wang^{1,2,3}, Yajing Hao^{4,5,6}, Chuanbao Zhang^{1,2,3}, Zhiliang Wang^{2,3}, Xing Liu^{2,3}, Guanzhang Li^{1,2,3}, Lihua Sun^{1,2,3}, Jingshan Liang^{1,2,3}, Jianjun Luo^{4,5,6}, Dabiao Zhou², Runsheng Chen^{4,5}, and Tao Jiang^{1,2,3,7,8}

Abstract

Purpose: RNA sequencing (RNA-seq) has recently proved to be effective for revealing novel virus–tumor associations. To get a thorough investigation of virus–glioma associations, we screened viruses in gliomas with RNA-seq data from the Chinese Glioma Genome Atlas project.

Experimental Design: In total, 325 samples were enrolled into this study. Reads that failed to map to the human genome were aligned to viral genomes and screened for potential virus-derived transcripts. For quantification, VPKM was calculated according to mapped reads weighted by genome sizes and sequencing depth.

Results: We observed that viruses tended to concertedly express in a certain subgroup of patients. Survival analysis revealed that individuals who were infected with Simian virus 40 (SV40) or woolly monkey sarcoma virus (WMSV) had a significantly

shorter overall survival than those uninfected. A multivariate Cox proportional hazards model, taking clinical and molecular factors into account, was applied to assess the prognostic value of SV40 and WMSV. Both SV40 and WMSV were independent prognostic factors for predicting patient's survival in lower-grade gliomas. Subsequent gene analysis demonstrated that SV40 was correlated with regulation of transcription, whereas WMSV was correlated with cell-cycle phase, which indicated frequent proliferation of tumor cells.

Conclusions: RNA-seq was sufficient to identify virus infection in glioma samples. SV40 and WMSV were identified to be prognostic markers for patients with lower-grade gliomas and showed potential values for targeting therapy. *Clin Cancer Res*; 23(9): 2177–85. ©2016 AACR.

Introduction

Glioma is the most common primary malignant brain tumor with no confirmative causes. Several human viruses have been

found to be involved in tumorigenesis in about 10% to 15% of human cancers (1), such as Epstein–Barr virus (EBV), hepatitis B virus (HBV), and human papillomavirus (HPV), causing nasopharyngeal carcinoma, hepatocellular carcinoma, and cervical cancers, respectively. Detection of expression of cytomegalovirus (CMV) in glioma has been widely reported and may be implicated in glioma genesis and speed the development or progression of glioblastomas (2, 3, 4, 5, 6), raising plenty of controversies (7, 8, 9). Different methods have been tried to uncover the existence of virus in cancers, ranging from detecting virus protein products with IHC to detecting virus nuclear acids with *in situ* hybridization (ISH).

Low-throughput methods show un-neglectable limitations when massively detecting numerous viruses in tissues are required. Previous studies about virus–tumor association are mostly based on low-throughput methodologies. Huge progress has been made in deciphering genomic code of biological creatures, enabling us to screen genomic components of viruses in tumors at genomic (8) and transcriptomic level (10). Massively parallel sequencing has been proved to be a promising method for efficiently detecting viruses in tumors.

In this study, RNA sequencing (RNA-seq) data generated within the Chinese Glioma Genome Atlas project (CGGA), encompassing 325 glioma samples, were used to validate this methodology. Genome sequences of viruses were downloaded from NCBI as reference genomes. Unmapped reads to human genome were

¹Department of Molecular Neuropathology, Beijing Neurosurgical Institute, Capital Medical University, Beijing, China. ²Department of Neurosurgery, Beijing Tiantan Hospital, Capital Medical University, Beijing, China. ³Chinese Glioma Genome Atlas Network, Beijing, China. ⁴Key Laboratory of RNA Biology, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China. ⁵Beijing Key Laboratory of Noncoding RNA, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China. ⁶University of Chinese Academy of Sciences, Beijing, China. ⁷Center of Brain Tumor, Beijing Institute for Brain Disorders, Beijing, China. ⁸China National Clinical Research Center for Neurological Diseases, Beijing, China.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (<http://clincancerres.aacrjournals.org/>).

Z. Wang, Y. Hao, and C. Zhang contributed equally to this article.

Corresponding Authors: Tao Jiang, Beijing Neurosurgical Institute, Tiantan Xili 6, Beijing 100050, China. Phone: 8610-6702-1832; Fax: 8610-6702-1832; E-mail: taojiang1964@163.com; Dabiao Zhou, Department of Neurosurgery, Beijing Tiantan Hospital, Capital Medical University, Beijing, China. E-mail: dabiao@yeah.net; Runsheng Chen, Key Laboratory of RNA Biology, Institute of Biophysics, Chinese Academy of Sciences, Beijing, 100101, China. E-mail: rschen@ibp.ac.cn

doi: 10.1158/1078-0432.CCR-16-1495

©2016 American Association for Cancer Research.

Translational Relevance

In this study, high-throughput RNA sequencing data were used to screen for viruses that may be involved in tumorigenesis of glioma. As a result, Simian virus 40 and woolly monkey sarcoma virus were revealed to be independent prognosticators for lower-grade glioma. Moreover, gene ontology analysis revealed that these two viruses were potentially oncogenic in lower-grade glioma. Thus, these two viruses can be regarded as biomarkers for lower-grade glioma in clinical practice. Furthermore, the oncogenic property of these two markers raised the possibility of using antiviral drugs for glioma treatment in the future.

used to virus genomes mapping, screening for potential elements of viruses.

Materials and Methods

Samples and RNA-seq

RNA-seq data of 325 samples were obtained from the CGGA network. Clinical specimen collection and data generating process were described in our previous article (11). Written informed consents were obtained from all the patients, and this study was approved by the Ethics Committee of Capital Medical University, Beijing, China.

VPKM generating process

Reads were first mapped to the human genome (RNA sequences, GRCh37, hg19), which was downloaded from the UCSC Genome Browser (<http://genome.ucsc.edu>) using TopHat (12) with `-no-mixed` parameter. Unaligned (nonhuman) reads were extracted using the `bamtofastq` from BedTools (13) with the default parameters. Next, unmapped sequences were aligned to the NCBI collection of viral genomes ($N = 6,402$), downloaded on July 22, 2015, using TopHat with default paired-end parameters. Then, we obtained reliable alignments using SAMtools (14) by setting a threshold on mapping quality (`-bq 1`). To quantify the expression of transcripts derived from virus in different individuals, we defined VPKM, which was calculated according to mapped reads weighted by viral genome sizes and sequencing depth:

$$\text{VPKM} = \frac{n \times 10^6}{L \times \text{sum1} \times \text{sum2}}$$

Here, n , L , sum1 , and sum2 denote the number of mapped reads, the length of the viral genome, the number of reads mapped to the human genome in the individual, and the number of reads mapped to the all viral genome in the individual, respectively. The L is used to adjust the expression with viral genome length. The sum1 was used to adjust the expression according to the whole sequencing depth, and the sum2 was used to adjust the expression according to the sequencing depth influenced by the viral genomes. The reads number will be counted twice if the reads can be aligned to more than one different viral genomes, but the reads will be counted once if the reads can be aligned to different locations in one viral genome.

IDH mutation detection

Among 325 samples, *IDH1/2* mutation status of 300 samples was determined by pyrosequencing as described in our previous study (15), whereas for the other 25 samples, *IDH* mutation status was called from RNA-seq data, which showed a satisfying concordance with *IDH* status called from DNA samples (Spearman correlation, $R = 0.8935$, $P < 2.2 \times 10^{-16}$).

TCGA subtype annotation

For The Cancer Genome Atlas (TCGA) subtype annotation, the samples were classified according to the proneural–neural–classical–mesenchymal classes using the signatures published and single-sample gene set enrichment analysis algorithm (16).

Detection of SV40 and WMSV with PCR

RNA of available frozen samples was isolated using the RNA-prep Pure Tissue Kit (Tiangen Biotech) followed by reverse transcription to cDNA (RevertAid RT Kit, Thermo Fisher Scientific) according to the manufacturer's instructions. Primers for SV40 (forward: 5'-GGGTCTTCTACCTTTCTCTCTTT-3', reverse: 5'-GCAGTGGTGAATGCCCTT-3') and for WMSV (forward: 5'-TGAAGGAAGTGTCTTTCAAGCTAG-3', reverse: 5'-AGAACT-CCGACACAGAGGCTTAT-3') were used to clone the expression of these two viruses.

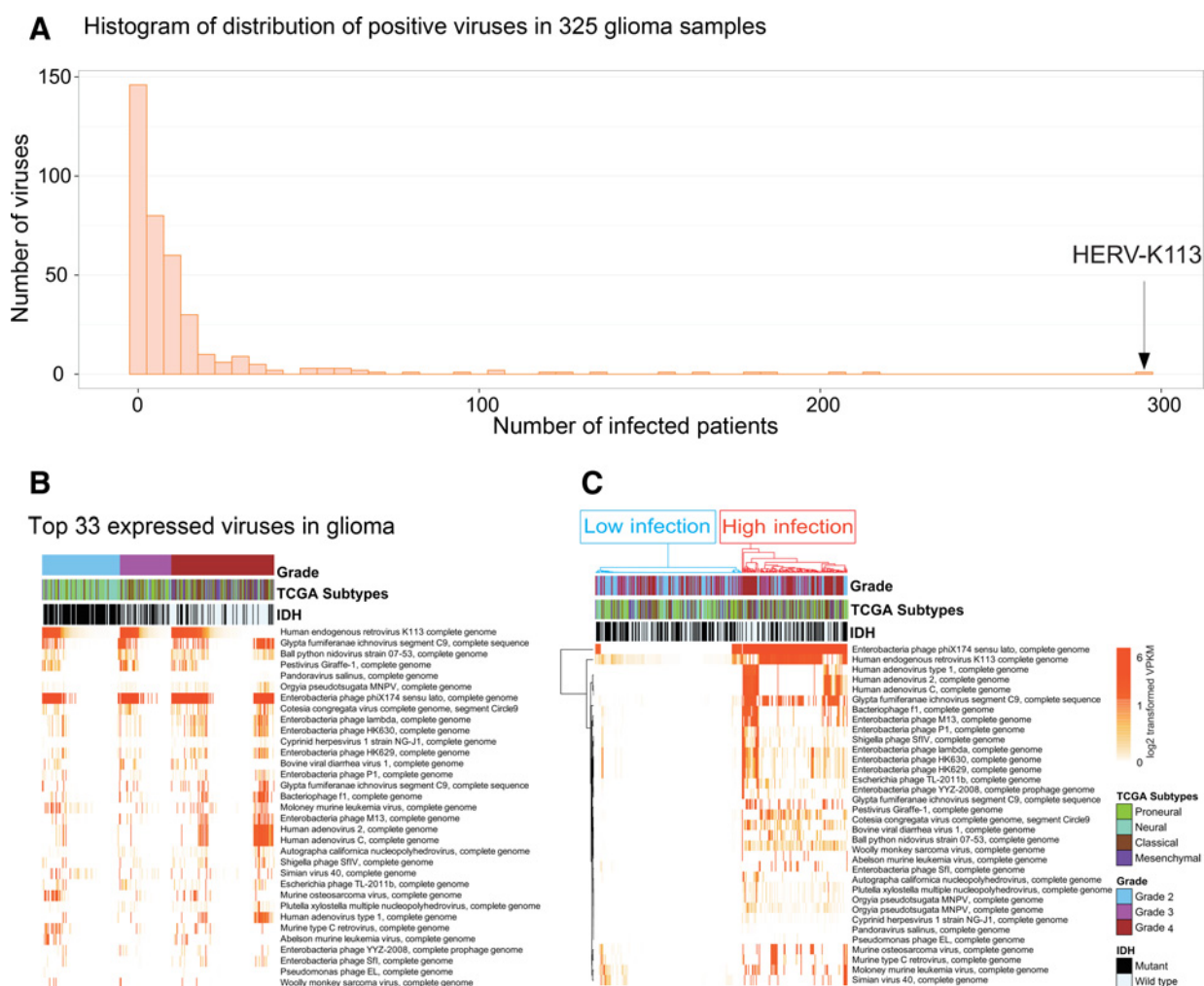
Statistical analysis

Fisher exact test was performed with R language. Survival analysis was performed using survival package of R, and log-rank test was applied to detect the difference of survival time. Graphical works were achieved by R language with a range of publicly available packages. Gene functions were annotated by DAVID online (<https://david.ncifcrf.gov/>). A P value less than 0.05 was considered to be significant.

Results

Viruses concertedly expressed in certain individuals

To get a landscape of virus expression, VPKM value, representing abundance of viruses, was calculated for each sample. A total of 374 viruses were detected to be positive in at least one sample, with most viruses only detected in a very small number of individuals (Fig. 1A). On the other hand, we noted that one virus was detected in almost all the samples (293/325, 90.2%), human endogenous retrovirus K113 (HERV-K113), which has been known for many years (17, 18, 19). Although many human endogenous retroviruses have been reported to potentially induce cancers (20), very few evidence emerged to support this kind of capability of HERV-K113. Moreover, we failed to find any influence that HERV-K113 might exert on patient's survival (Supplementary Fig. S1). As 325 samples contained World Health Organization (WHO) grades 2, 3, and 4 gliomas, we asked the distribution of viruses across histopathologic grades with top 10% (33/325, Supplementary Table S1) positively infected viruses. To our surprise, viruses seemed to concertedly express in certain individuals (Fig. 1B). This was the same case when we tested the distribution of viruses across glioma expression subtypes (Supplementary Fig. S2A). To further validate what we found, unsupervised clustering analysis was performed. As shown in Fig. 1C, two groups of patients were identified, with few individuals

**Figure 1.**

Virus reads expression in glioma. **A**, Number of positive viruses in 325 patients. HERV-K113 is the most prevalent virus in glioma, as the arrow indicates. **B**, Distribution of top 10% (33/325) positively infected viruses across all histopathologic grades. **C**, Unsupervised cluster of top 10% (33/325) positively infected viruses identified a subgroup of patients that seemed to be susceptible to various virus infections.

infected with viruses in one group, and most of patients got infected with a number of viruses in the other. These two groups were designated as low-infection group or high-infection group according to the infection status, respectively. This result indicated that immune status of patients might account for the huge difference of virus infection among individuals. We compared mutation status of IDH, which is the most common mutation in glioma, but failed to find significant correlation between the two groups (χ^2 test, $P = 0.0888$). Different gene expression and gene ontology analysis showed that low virus infection-related genes were more involved in differentiated biological processes such as chemical synaptic transmission and learning, while high virus infection-related genes were more associated with DNA repair and extracellular matrix organization, although no statistical significance was observed after Bonferroni correction (Supplementary Fig. S3). Moreover, we compared the clinical outcome of the two groups, but no significant difference was revealed (Supplementary Fig. S2B).

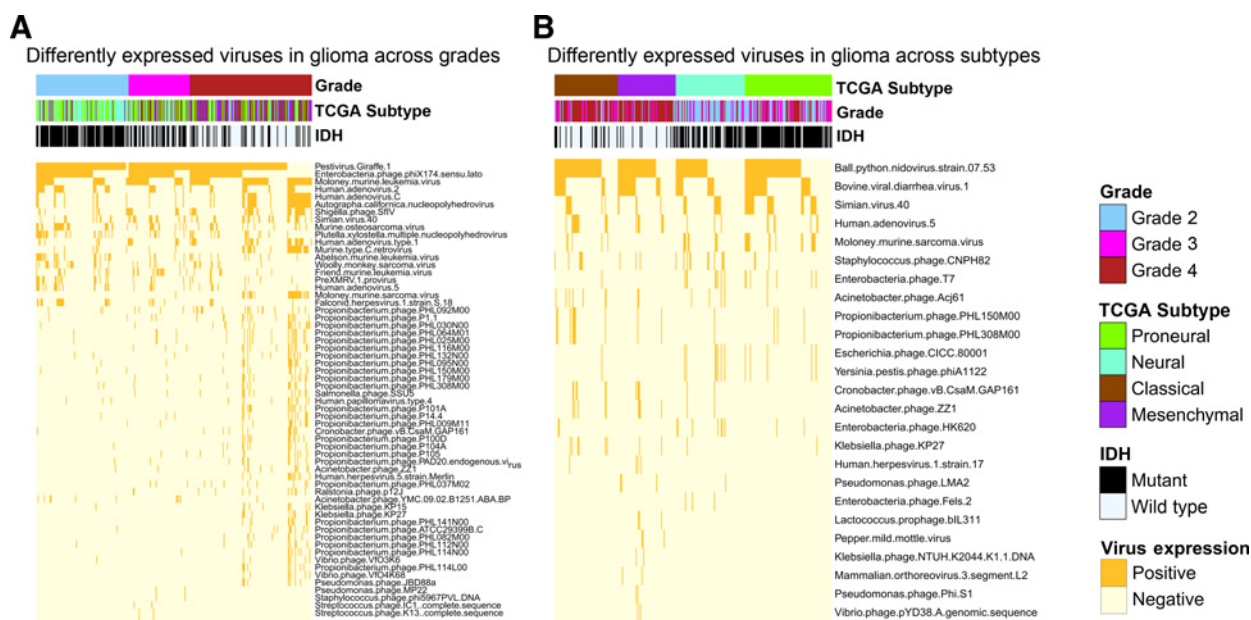
Differently expressed viruses across WHO grade and subtypes

To our knowledge, different grades or subtypes of glioma may have different origins. Virus infection can be one of the potential causes for different gliomas. In this analysis, we first investigated the differently expressed viruses across different WHO grade or subtypes with Fisher exact test. Infection of virus was dichotomized as positive or negative based on whether virus genome-derived reads were detected or not. As a result, we found 61 and 25 differently infected viruses across WHO grade and expression subtypes, respectively (Fig. 2; Supplementary Table S2).

Identification of clinically relevant viruses in lower-grade glioma

Survival analysis was performed separately in lower-grade gliomas and glioblastomas. SV40 was one of the viruses that differently expressed across WHO grades according to Fisher exact analysis. A larger proportion of lower-grade gliomas, proneural gliomas as well as IDH wild-type tumors were SV40 positive (Fig. 3A–D). Previous studies showed that SV40 demonstrated

Wang et al.

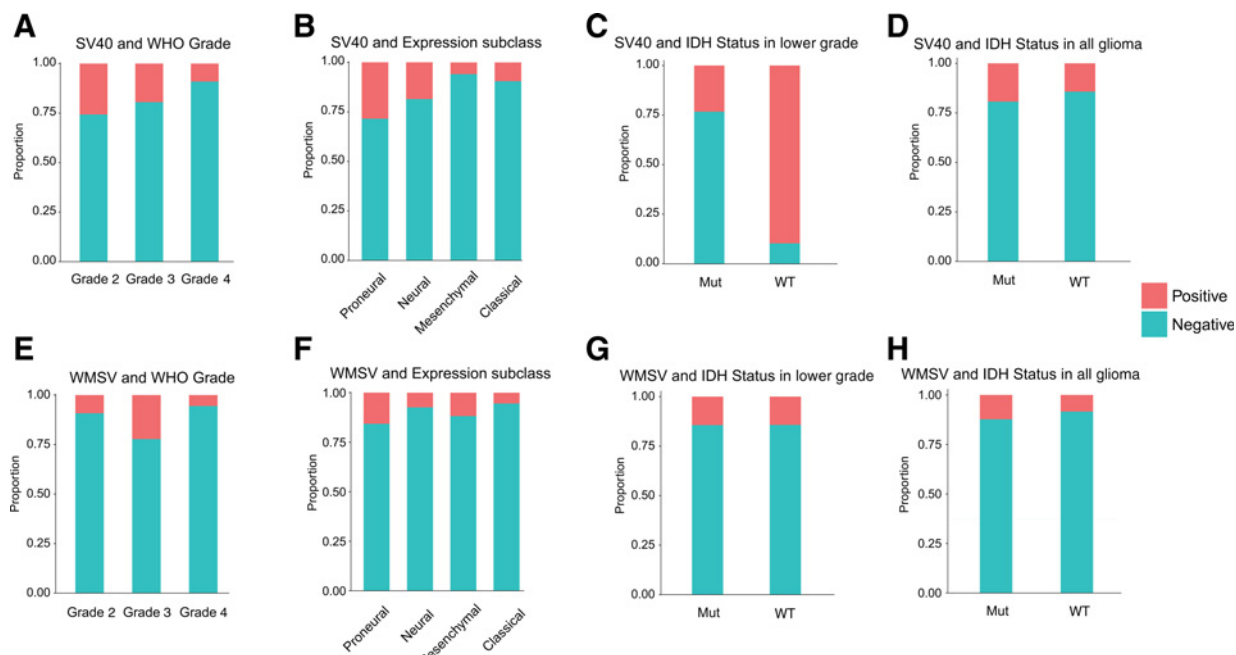
**Figure 2.**

Differently expressed viruses in glioma across grades and TCGA subtypes. **A**, Differently expressed virus across WHO grades. **B**, Differently expressed virus across TCGA expression subtypes.

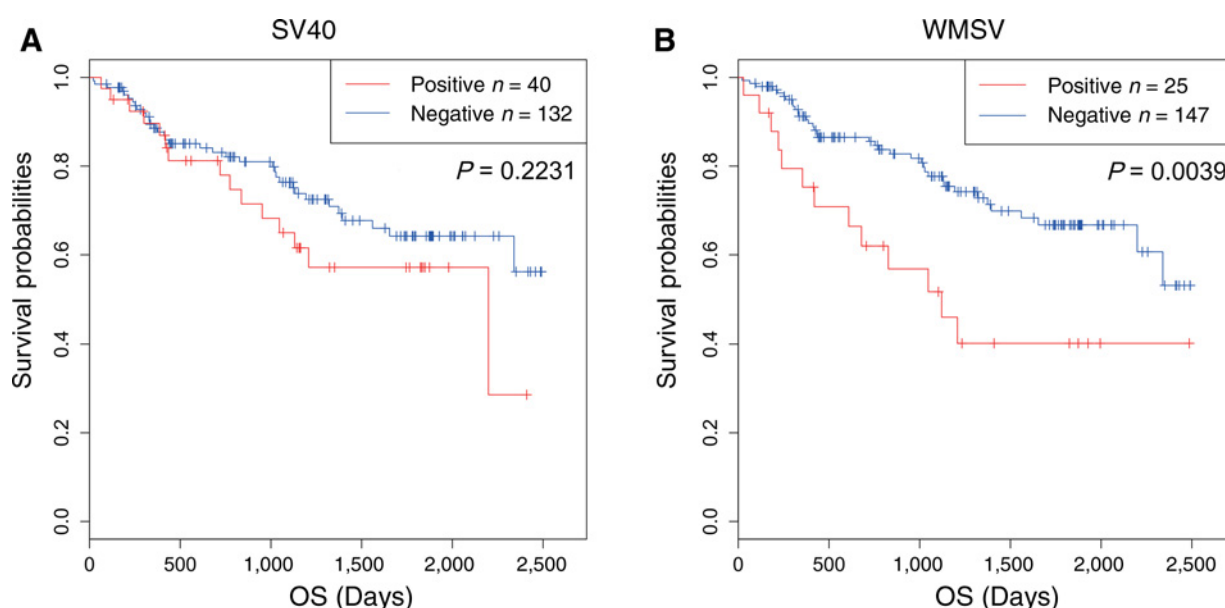
oncogenic property in the central nervous system (21, 22, 23). Kaplan–Meier analysis indicated that in lower-grade gliomas, SV40 seemed to be a negative prognosticator. Patients who were infected with SV40 in their tumors suffered relatively shorter

survival than those who were not infected, although the *P* value was not significant (Fig. 4A).

In Kaplan–Meier analysis of 181 lower-grade gliomas, WMSV was the only virus that showed a significantly prognostic value for

**Figure 3.**

Infection of SV40 and WMSV in different WHO grades, TCGA expression subclass, and *IDH* mutation status. **A–D**, Infection of SV40 in different WHO grades, TCGA expression subclass, and *IDH* mutation status. **E–H**, Infection of WMSV in different WHO grades, TCGA expression subclass, and *IDH* mutation status. Mut, mutant; WT, wild type.

**Figure 4.**

Survival analysis for SV40 and MMSV in lower-grade glioma. **A**, Kaplan-Meier analysis for SV40 in lower-grade glioma. **B**, Kaplan-Meier analysis for WMSV in lower-grade glioma. OS, overall survival.

patients' outcome. Patients who were infected with WMSV in their tumors had significantly shorter survival than those who were not infected (Fig. 4B, median survival: 1,121 days vs. not reached, HR: 3.434). Distribution of WMSV in different grade, subtypes, and correlation with IDH mutation status was presented in Fig. 3E–H.

To further validate the prognostic value of SV40 and WMSV for patients with lower-grade gliomas, multivariate analysis by Cox proportional hazards model was performed with survival package of R, taking vital clinical and molecular factors into account. As shown in Table 1, after being adjusted by WHO grade, IDH mutation status, radiotherapy, and chemotherapy, both SV40 and WMSV remained to be independent prognostic factors for patients with lower-grade gliomas ($P = 0.0019$ and 0.032 , respectively). This raised the possibility of potential impacts that SV40 and WMSV might exert on glioma malignancy. PCR was performed to validate the expression of positive samples of which frozen samples were available. For validation of SV40, 11 positive samples in RNA-seq were available, and two negative samples of RNA-seq were used as negative control (Supplementary Fig. S4A). For validation of WMSV, only eight positive samples were available (Supplementary Fig. S4B), and two negative samples were used as control as well. Because of low abundance of these two viruses, we performed PCR at 50 cycles to improve the detectable rate. As a result, we found the expression of virus mRNA of most

samples was in concordance with RNA-seq except for the two negative samples in WMSV validation, which also showed minimal expression of WMSV but at much lower abundance than positive samples (Supplementary Fig. S4).

In glioblastoma, survival analysis was performed for each virus individually. A couple of viruses seemed to be prognostic in the GMB cohort, but most of them turned out to be phages, or too few positive samples to be acknowledged as statistically significant (Supplementary Fig. S5).

Function annotation of SV40 and WMSV

To further investigate the potential impact that SV40 and WMSV might exert on glioma, we compared the expression difference and calculated fold change with VPKM value between positive and negative samples. Genes with a P value of Student t test < 0.05 and fold change > 2 were considered to be statistically significant. Significantly altered genes of SV40 and WMSV were shown in volcano plots, respectively (Fig. 5A and B).

Moreover, when taking expression level into account, Spearman correlation analysis was performed between virus expression and human gene expression profile, screening for significantly altered genes. Top 400 correlated genes, including 200 positively related and 200 negatively related with SV40 or WMSV, were enrolled for further analysis. Genes with 0 expression in more

Table 1. Cox proportional hazards regression model for SV40 and WMSV

Terms	Coef	exp(coef)	se(coef)	z	P
Gender	0.1016	1.107	0.3224	0.315	0.75
Age	0.0211	1.021	0.0174	1.217	0.22
Grade	1.6957	5.45	0.3819	4.44	0.000009
IDH	-1.2494	0.287	0.3563	-3.507	0.00045
Radiotherapy	-1.1132	0.329	0.3742	-2.974	0.0029
Chemotherapy	-0.0476	0.954	0.3773	-0.126	0.9
SV40	1.0768	2.935	0.3461	3.112	0.0019
WMSV	0.7463	2.109	0.348	2.144	0.032

Wang et al.

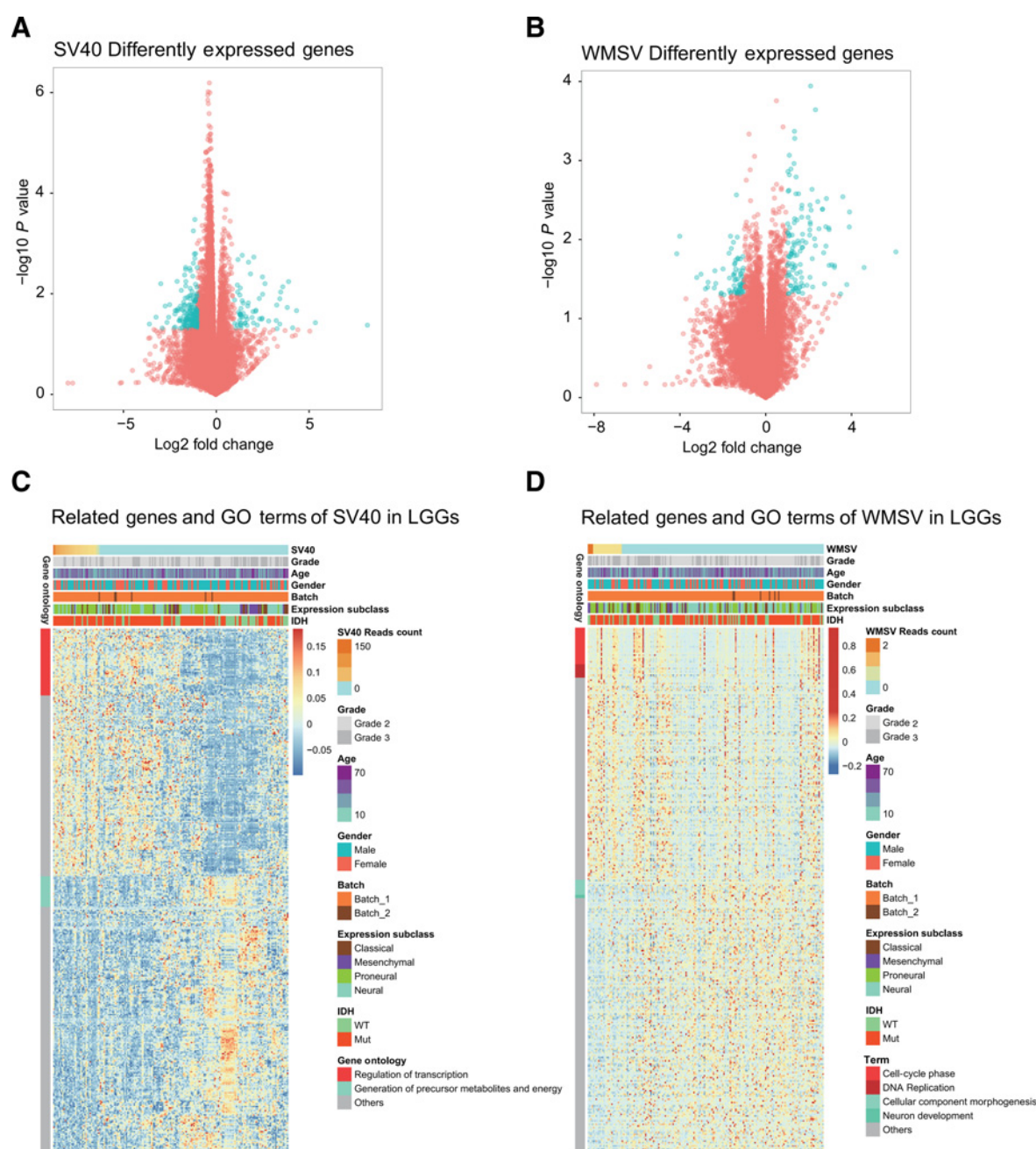


Figure 5. Gene ontology analysis for SV40 and WMSV in lower-grade glioma (LGG). **A** and **B**, Volcano plots showed differentially expressed genes between SV40/WMSV-positive and negative patients. **C** and **D**, Gene ontology (GO) analysis for SV40 and WMSV. WT, wild type.

than 30% samples were discarded (Supplementary Table S3). Gene ontology annotation was analyzed with DAVID online tools (<https://david.ncicrf.gov/summary.jsp>). For SV40, positively related genes turned to focus on regulation of transcription. Negatively related genes with SV40 mostly focused on activities of differentiated cells, such as generation of precursor metabolites and energy and ATP synthesis-coupled electron transport (Fig. 5C, Supplementary Fig. S6A). For WMSV, cell cycle and mitosis dominated the most positively related terms, while the negatively

related processes were mainly involved in biological processes of highly differentiated cells, such as neuron development and cell morphogenesis involved in differentiation (Fig. 5D, Supplementary Fig. S6B).

Virus infection in paired samples

Paired samples of seven patients (samples from primary lesion and reoperated after tumor recurrence) were available, which allowed us to infer the relationship between viral

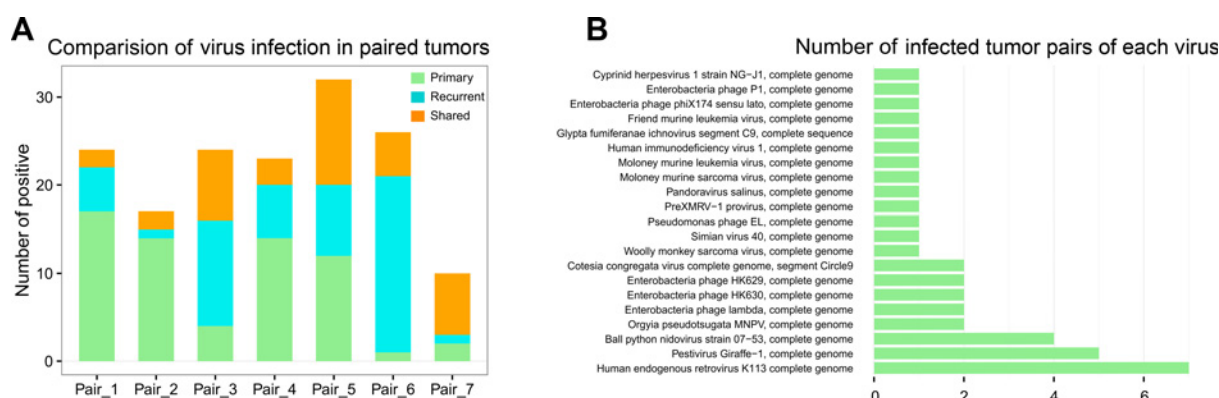


Figure 6. Virus infection status in paired samples. **A**, Virus infection status in seven paired glioma samples. **B**, Positive viruses in paired samples.

infection and glioma progression. This might convey some information about the oncogenesis effect or promotion of progression effect of a certain virus. In the seven pairs of samples, most viruses were found to be private to primary tumors or recurrent counterparts, except for one pair (Fig. 6A). When comparing the number of viruses that revealed to be positive in the paired sample, we failed to find a statistically significant difference between primary tumors and recurrent tumors (Student *t* test for paired samples, $P = 0.7314$). Only three viruses were identified to be positive in more than three pairs (Fig. 6B): HERV-K113, Ball python nidovirus strain 07-53 and Pestivirus Giraffe-1, none of which was reported to be oncogenic. This result indicated that viruses were not likely to promote recurrent or malignant progression of glioma, but the results need further validation in a larger dataset.

CMV infection in glioma

Detection of expression of CMV in glioma has been widely reported and may be implicated in glioma genesis and speed the development or progression of glioblastomas, raising plenty of controversies (2, 24). Different methods have been tried to uncover the existence of virus in cancers (such as IHC and ISH). In our current cohort, only 11 samples were detected to be infected with CMV, most of which were detected at a relatively low abundance (ranging from 1 read to 69 reads). Most CMV-positive samples (8/11) were glioblastoma (Supplementary Fig. S7A). Moreover, no significant difference was observed in survival analysis between patients who were infected with CMV and those were not in glioblastoma (Supplementary Fig. S7B). This finding further diminished the possibility of carcinogenic capability of CMV.

Discussion

Glioma is the most lethal primary brain tumor. Some viruses were reported to be potentially carcinogenic in many malignant tumors, such as EBV, HBV, and HPV, etc. With RNA-seq profiling data of 325 samples, we aligned reads that were unmapped to human genome to viruses' genomes, looking for clues of virus infection in glioma.

Only a handful of viruses were detected to be expressed in more than half of the cohort. Taking the top 10% most infected viruses

(33 viruses) into account, unsupervised clustering analysis was performed. Surprisingly, we observed that patients who were infected with viruses tended to be classified into one group regardless of WHO grade and expression subtypes. Thus, viruses concurredly expressed in one subset of the cohort. This encouraged us to consider basic health status of the patients, which may lead to different clinical outcomes. However, survival analysis demonstrated no significant difference between low-infection group and high-infection group. Immune status emerged to be one of the possible reasons yet to be explored.

SV40 was detected to be differently expressed across WHO grades and expression subtypes. Previous studies showed that SV40 demonstrated oncogenic property in the central nervous system (21, 22, 23, 25), urging us to explore its clinical influence on patients. Kaplan-Meier analysis showed that patients infected with SV40 were prone to suffer shorter survival, although the *P* value was not significant. However, when Cox proportional hazards model was applied, taking multiple clinical and molecular factors into analysis, both SV40 and WMSV showed independent prognostic value for lower-grade gliomas (Table 1). The discrepancy about SV40 between Kaplan-Meier analysis and Cox proportional hazards model may be accounted for by sample volume and relatively short follow-up for lower-grade glioma.

About the function annotation of SV40, only regulation of transcription was identified as the most relevant biological process. No robust GO terms were identified for SV40. Thus, SV40 may serve as a prognosticator for lower-grade glioma. For WMSV, cell cycle and mitosis-related terms were identified to be positively related to WMSV infection. This finding raised the possibility that WMSV may be involved in the malignant process of lower-grade glioma.

Seven paired samples showed that no robust viruses were identified as candidates to promote recurrence or progression of glioma, but in this study, only very limited samples were analyzed. More paired samples are needed for exploring the contribution of viruses in progression or recurrence of lower-grade glioma.

Detection of CMV in glioma has been controversial for many years, giving rise to a hot debate in its existence and biological impacts on glioma (2, 4, 24, 26). Some therapies have even been proposed on the basis of CMV infection in glioma and may

provide a new way to treat individuals with glioma (27, 28). In this study, only 11 samples were detected to be infected with CMV, and in most samples, CMVs were detected at a relatively low abundance. These results showed high concordance with previous studies that detected CMV with next-generation sequencing data of glioma (10), instead of studies with ISH, IHC (3), or qPCR (5). This result indicated that CMV expression was relatively low to detect with RNA-seq, but was high enough for qPCR or ISH detection. Meanwhile, more efforts are needed to explore the biological function of viruses that express at such a low abundance.

Tang and colleagues (10) quantified viral mRNA expression by computing the fraction of viral reads (FVR), presented as parts per million (ppm) of total library size. Only viruses with FVR (viral expression) >2 ppm were regarded as positive. In another study performed by Khoury and colleagues (26), the expression level of each viral transcript was quantified by the normalized depth of coverage within each viral transcript. The cutoff of viral gene expression detection was determined by profiling the distribution of viral gene expression levels across multiple cancer-associated viruses. Any viral expression level below the cutoff was treated as no expression. Both studies used data from TCGA project. They have done a great job in detecting viral expression and host gene fusion in human cancer. Both studies treated low expression of virus as negative. As we all know, for detecting low-abundance mRNAs, RNA-seq is not as good as PCR or qPCR. We do not think it is appropriate to neglect viruses with low-abundance in RNA-seq, because low abundance does not necessarily mean that these viruses exert little influence on the process of development and progression of glioma. In the current analysis, we included the viruses even with extreme low abundance. Correlation analysis and survival analysis indicated that SV40 and WMSV were associated with malignant biological process and worse survival.

Taken together, through RNA-seq profiling data of 325 samples from CGGA, we reestablished the landscape of virus infection in glioma. Two viruses, SV40 and WMSV, were revealed to be correlated with the clinical outcome of lower-grade glioma patients and showed significant biological functions of promoting malignancy, especially for WMSV. CMVs were detected at a very low abundance in RNA-seq data and showed no clinical relevance with the survival of glioblastoma patients.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: Y. Hao, Z. Wang, D. Zhou, R. Chen, T. Jiang

Development of methodology: Y. Hao, C. Zhang

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): Z. Wang, C. Zhang, G. Li, L. Sun

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): Z. Wang, Y. Hao, C. Zhang, X. Liu, L. Sun, J. Liang, R. Chen

Writing, review, and/or revision of the manuscript: Z. Wang, Y. Hao, C. Zhang, T. Jiang

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): T. Jiang

Study supervision: J. Luo, D. Zhou, R. Chen, T. Jiang

Grant Support

This study was supported by the Research Special Fund for Public Welfare Industry of Health (nr. 201402008) and the Capital Medical Development Research Fund (2016-1-1072).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received June 17, 2016; revised October 31, 2016; accepted November 8, 2016; published OnlineFirst November 18, 2016.

References

- Moore PS, Chang Y. Why do viruses cause cancer? Highlights of the first century of human tumour virology. *Nat Rev Cancer* 2010;10: 878–89.
- Dziurzynski K, Chang SM, Heimberger AB, Kalejta RF, McGregor Dallas SR, Smit M, et al. Consensus on the role of human cytomegalovirus in glioblastoma. *Neuro Oncol* 2012;14:246–55.
- Scheurer ME, Bondy ML, Aldape KD, Albrecht T, El-Zein R. Detection of human cytomegalovirus in different histological types of gliomas. *Acta Neuropathol* 2008;116:79–86.
- Price RL, Song J, Bingmer K, Kim TH, Yi J-Y, Nowicki MO, et al. Cytomegalovirus contributes to glioblastoma in the context of tumor suppressor mutations. *Cancer Res* 2013;73:3441–50.
- Bhattacharjee B, Renzette N, Kowalik TF. Genetic analysis of cytomegalovirus in malignant gliomas. *J Virol* 2012;86:6815–24.
- Cobbs CS, Harkins L, Samanta M, Gillespie GY, Bharara S, King PH, et al. Human cytomegalovirus infection and expression in human malignant glioma. *Cancer Res* 2002;62:3347–50.
- Sabatier J, Uro-Coste E, Pommepuy I, Labrousse F, Allart S, Trémolet M, et al. Detection of human cytomegalovirus genome and gene products in central nervous system tumours. *Br J Cancer* 2005;92: 747–50.
- Tang KW, Hellstrand K, Larsson E. Absence of cytomegalovirus in high-coverage DNA sequencing of human glioblastoma multiforme. *Int J Cancer* 2015;136:977–81.
- Wick W, Platten M. CMV infection and glioma, a highly controversial concept struggling in the clinical arena. *Neuro Oncol* 2014;16:332–3.
- Tang KW, Alaei-Mahabadi B, Samuelsson T, Lindh M, Larsson E. The landscape of viral expression and host gene fusion and adaptation in human cancer. *Nat Commun* 2013;4:2513.
- Bao ZS, Chen HM, Yang MY, Zhang CB, Yu K, Ye WL, et al. RNA-seq of 272 gliomas revealed a novel, recurrent PTPRZ1-MET fusion transcript in secondary glioblastomas. *Genome Res* 2014;24:1765–73.
- Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 2009;25:1105–11.
- Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010;26:841–2.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009;25: 2078–9.
- Yan W, Zhang W, You G, Bao Z, Wang Y, Liu Y, et al. Correlation of IDH1 mutation with clinicopathologic factors and prognosis in primary glioblastoma: a report of 118 patients from China. *PLoS One* 2012;7: e30339.
- Brennan CW, Verhaak RG, McKenna A, Campos B, Nounshmehr H, Salama SR, et al. The somatic genomic landscape of glioblastoma. *Cell* 2013; 155:462–77.
- Moyes D, Martina A, Sawcerb S, Temperton N, Worthington J, Griffiths DJ, et al. The distribution of the endogenous retroviruses HERV-K113 and HERV-K115 in health and disease. *Genomics* 2005;86:337–41.
- Boller K, Schönfeld K, Lischer S, Fischer N, Hoffmann A, Kurth R, et al. Human endogenous retrovirus HERV-K113 is capable of producing intact viral particles. *J Gen Virol* 2008;89:567–72.

19. Griffiths DJ. Endogenous retroviruses in the human genome sequence. *Genome Biol* 2001;2:reviews1017.1–1017.5.
20. Lower R. The pathogenic potential of endogenous retroviruses: facts and fantasies. *Trends Microbiol* 1999;7:350–6.
21. Vilchez RA, Kozinetz CA, Arrington AS, Madden CR, Butel JS. Simian virus 40 in human cancers. *Am J Med* 2003;114:675–84.
22. Suzuki SO, Mizoguchi M, Iwaki T. Detection of SV40 T antigen genome in human gliomas. *Brain Tumor Pathol* 1997;14:125–9.
23. Weggen S, Bayer TA, von Deimling A, Reifenberger G, von Schweinitz D, Wiestler OD, et al. Low frequency of SV40, JC and BK polyomavirus sequences in human medulloblastomas, meningiomas and ependymomas. *Brain Pathol* 2000;10:85–92.
24. Michaelis M, Doerr HW, Cinatl J. The story of human cytomegalovirus and cancer: increasing evidence and open questions. *Neoplasia* 2009;11:1–9.
25. Wong-Staal F, Favera RD, Gelmann EP, Manzari V, Szala S, Josephs SF, et al. The v-sis transforming gene of simian sarcoma virus is a new onc gene of primate origin. *Nature* 1981;294:273–5.
26. Khoury JD, Tannirb NM, Williams MD, Chenb Y, Yaod H, Zhang J, et al. Landscape of DNA virus associations across human malignant cancers: analysis of 3,775 cases using RNA-Seq. *J Virol* 2013;87: 8916–26.
27. Nair SK, De Leon G, Boczkowski D, Schmittling R, Xie W, Staats J, et al. Recognition and killing of autologous, primary glioblastoma tumor cells by human cytomegalovirus pp65-specific cytotoxic T cells. *Clin Cancer Res* 2014;20:2684–94.
28. Soderberg-Naucler C, Rahbar A, Stragliotto G. Survival in patients with glioblastoma receiving valganciclovir. *New Engl J Med* 2013;369: 985–6.

Clinical Cancer Research

The Landscape of Viral Expression Reveals Clinically Relevant Viruses with Potential Capability of Promoting Malignancy in Lower-Grade Glioma

Zheng Wang, Yajing Hao, Chuanbao Zhang, et al.

Clin Cancer Res 2017;23:2177-2185. Published OnlineFirst November 18, 2016.

Updated version Access the most recent version of this article at:
doi:[10.1158/1078-0432.CCR-16-1495](https://doi.org/10.1158/1078-0432.CCR-16-1495)

Supplementary Material Access the most recent supplemental material at:
<http://clincancerres.aacrjournals.org/content/suppl/2016/11/18/1078-0432.CCR-16-1495.DC1>

Cited articles This article cites 27 articles, 11 of which you can access for free at:
<http://clincancerres.aacrjournals.org/content/23/9/2177.full.html#ref-list-1>

E-mail alerts [Sign up to receive free email-alerts](#) related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.