Significant Association of Multiple Human Cytomegalovirus Genomic Loci with Glioblastoma Multiforme Samples

Padhma Ranganathan,^{a,b} Paul A. Clark,^e John S. Kuo,^{c,d,e,f} M. Shahriar Salamat,^g and Robert F. Kalejta^{a,b,c,d}

Institute for Molecular Virology,^a McArdle Laboratory for Cancer Research,^b Carbone Comprehensive Cancer Center,^c Stem Cell and Regenerative Medicine Center,^d and Departments of Neurological Surgery,^e Human Oncology,^f and Pathology and Laboratory Medicine,^g University of Wisconsin—Madison, Madison, Wisconsin, USA

Viruses are appreciated as etiological agents of certain human tumors, but the number of different cancer types induced or exacerbated by viral infections is unknown. Glioblastoma multiforme (GBM)/astrocytoma grade IV is a malignant and lethal brain cancer of unknown origin. Over the past decade, several studies have searched for the presence of a prominent herpesvirus, human cytomegalovirus (HCMV), in GBM samples. While some have detected HCMV DNA, RNA, and proteins in GBM tissues, others have not. Therefore, any purported association of HCMV with GBM remains controversial. In most of the previous studies, only one or a select few viral targets were analyzed. Thus, it remains unclear the extent to which the entire viral genome was present when detected. Here we report the results of a survey of GBM specimens for as many as 20 different regions of the HCMV genome. Our findings indicate that multiple HCMV loci are statistically more likely to be found in GBM samples than in other brain tumors or epileptic brain specimens and that the viral genome was more often detected in frozen samples than in paraffin-embedded archival tissue samples. Finally, our experimental results indicate that cellular genomes substantially outnumber viral genomes in HCMV-positive GBM specimens, likely indicating that only a minority of the cells found in such samples harbor viral DNA. These data argue for the association of HCMV with GBM, defining the virus as oncoaccessory. Furthermore, they imply that, were HCMV to enhance the growth or survival of a tumor (i.e., if it is oncomodulatory), it would likely do so through mechanisms distinct from classic tumor viruses that express transforming viral oncoproteins in the overwhelming majority of tumor cells.

Glioblastoma multiforme (GBM) is an aggressive and malignant tumor of glial origin with a grim prognosis (43). It accounts for nearly half of all central nervous system (CNS) malignancies in adults. Equivalent to a grade IV diffuse astrocytoma, GBMs consist primarily of neoplastic astrocytes, the most abundant type of glia, but tumor specimens also include other nonneoplastic cell types, including neurons, oligodendrocytes, macrophages, and glial and neural stem cells (42). The heterogeneous nature of these tumors may at least partially explain why they are refractory to current therapeutics and has led to the hypothesis that GBMs may originate from stem-like cells, perhaps in the subventricular zone of the CNS (9). The etiology of GBM is unknown, although exposure to ionizing radiation, electrical, or magnetic fields has been proposed as a risk factor (22). Recently, several reports have detected a potential association between GBMs and human cytomegalovirus (HCMV), a common betaherpesvirus.

Viruses are causative agents of human cancers (21). At least 15% of all human tumors have a viral etiology. Human cancer viruses include Epstein-Barr virus (EBV), hepatitis B virus (HBV), human T-lymphotropic virus type 1 (HTLV-1), human papillomavirus (HPV), hepatitis C virus (HCV), Kaposi's sarcoma-associated herpesvirus (KSHV), and Merkel cell polyomavirus (MCV). True viral infection (whether productive or latent) where the entire genome is maintained, as well as abortive infections where only select regions of the viral genome are present, can lead to cellular transformation and cancer development. Not surprisingly, human cancer viruses can produce within cells one or more of the molecular hallmarks of cancer (10) that promote cellular plasticity (through genomic instability, inflammation, deregulation of cellular energetics, and induction of angiogenic and metastatic processes), proliferation (by sustaining proliferative signaling, evading growth suppressors, and enabling replicative

immortality), and survival (avoidance of immune detection and inhibition of apoptosis). However, many other viruses not yet directly associated with human tumors may also initiate such molecular events, leading to speculation that more human cancers have a viral etiology or association than is currently appreciated. One such putative relationship between a virus and a cancer that is receiving increased examination is that of HCMV and GBM.

HCMV asymptotically infects the majority of the human population, and virus-induced sequelae are generally observed only under conditions of insufficient host immune function (7). Examples include birth defects in congenitally infected neonates, retinitis and blindness in AIDS patients, and graft rejection in transplant patients receiving immunosuppressive therapy. However, an emerging concept hypothesizes that chronic conditions, although perhaps not directly caused by HCMV, are likely to be exacerbated by infection with this virus (32). Disease states linked to HCMV infection include immunosenescence (40), certain cardiovascular diseases (36), and cancer (34). Either HCMV infection or the ectopic expression of individual viral proteins can produce all of the molecular hallmarks of cancer (13). For example, viral infection or HCMV-encoded proteins promote genomic instability (8, 30), inflammation (2), angiogenesis (4), and cell migration (39), modulate cellular energetics (41), proliferation (11), and life span (35), and inactivate cellular immune functions (25)

Received 22 August 2011 Accepted 3 November 2011

Published ahead of print 16 November 2011

Address correspondence to Robert F. Kalejta, rfkalejta@wisc.edu.

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JVI.06097-11

TABLE 1 PCR primer sequences used in this study^a

Gene	Forward primer	Reverse primer	Expected product size (bp)
UL1	AGCTTCTGGAACATATGGCGA	CGTAGTCTGTGTCTCCGGGT	416
UL17	ACCCCGTCACTGTGCGTTGGA	TTCGGGAGACCGTGCCGGAT	268
UL18	ATAGCGAGCCTCAATG	GTTAGCTGTCGGGTGATCAGGG	196
UL19	GCCCCCGTCAACACGGTCAT	GCGTCCCTGCGACTCCCATC	245
UL36	TGTCCCTTCTTTCCCATGAC	TTCTGGACCCACAACACAGA	302
UL44	TCGCGTCGTTGAAGTAATTG	TAGTTACGAGACACGACACG	498
UL55	TCCAACACCCACAGTACCCGT	CGGAAACGATGGTGTAGTTCG	270
UL69	CTCGTCGTGTGAACAGCAGGATG	GAACTACAGCAACTCAGCCGTTTGA	200
UL82	CGTTCATTTGGAACCCGAC	CTCTTCCTCTTCTTCTC	222
UL96	ACAGCTCTTAAAGGACGTGATGCG	ACCGTGTCTTCAGCTCGGTTAAA	288
UL97	CTTCGGACAAGGAGAACCAG	AGCTCCACGTCGTTTTCTTC	325
UL98	AGTCGCGCAACTCGGTGTGG	GAGCACGCCAAGCACAGGT	320
UL111A	GTACGGTGGTCAAAGGCTGT	CCAACTCGCTGAGACCTTTC	350
UL122	AAGCGGCCTCTGATAACCAAGCC	AGCACCATCCTCCTCTTCTCTCTGG	431
UL144	TACGTGTTGTCCCCCATGTA	GTCGTTGCTGTGATGTTGG	248
US2	CCCGGAATGGAGGGTAAT	CTAGCCCTGCGTCTGGTACT	183
US11	ACCACTGGTCCGAAAACATC	GTTCTACTTTTCCCCCTGCC	269
US28	AGCGTGCCGTGTACGTTAC	ATAAAGACAAGCACGACC	395
IE1	CGACGTTCTCGAGACTATG	TCTCGGTCACTTGTTCAAA	117

^a Nucleotide sequences for primer pairs used to amplify segments of the indicated HCMV gene are shown. The expected size of the amplified product is listed in base pairs. Also, the sequences of the primers (named "IE1") employed in the comparative PCR (Fig. 1) and copy number (Fig. 10) studies are indicated.

and apoptotic pathways (3). Therefore, HCMV represents an intriguing candidate human cancer virus.

The major experimental method identifying HCMV in GBM samples has been immunohistochemical (IHC) detection of the 72-kDa viral immediate-early 1 (IE1) protein. Individual studies have detected IE1 in the following number of GBM specimens: 27 out of 27 (5), 42 out of 45 (20), 21 out of 21 (29), 10 out of 10 (35), 20 out of 21 (31), or 8 out of 49 (16). The viral phosphoprotein of 65 kDa (pp65) has also been detected by flow cytometry in 5 out of 5 GBM samples (6) and by IHC in 8 out of 8 (5), 30 out of 33 (20), or 25 out of 59 (16) GBM samples. Furthermore, the protein product of the 28th gene in the unique short region of the genome (US28) was found in 20 out of 21 GBM samples (31). *In situ* hybridization (ISH) for either the HCMV IE locus (5), total viral genomic DNA (5, 20), or an undisclosed region(s) of the viral genome (29) has also been used to score for the presence of HCMV, obtaining positive results in every (29 out of 29) GBM sample examined. Finally, PCRs have detected the 55th gene in the unique long section of the genome (UL55 that encodes glycoprotein B [gB]) in 7 out of 9 (5) or 21 out of 34 (20) GBM specimens. Many interpret these results as solid evidence indicating that HCMV is present in GBMs.

However, other studies using similar approaches have failed to detect HCMV in GBM samples. PCR and IHC for IE1 or pp65 failed to detect the presence of HCMV in 22 GBM samples (24). Likewise, an independent study (15) that used IHC for pp65, ISH for IE1 or pp65, and PCR for gB failed to detect HCMV in 8 GBM samples. An additional investigation (28) detected IE1 by IHC in only 9 out of 81 GBMs and by ISH in only 7 out of 81 GBMs, although those 7 were also positive by IHC. Furthermore, complicating issues have created uncertainty about the results of the above studies that detected a strong association of HCMV with GBM tumors. For example, the IHC images presented are very difficult for all but a trained pathologist to decode, and thus, it could be argued that the conclusions drawn from such experi-

ments are more subjective than objective. Furthermore, the viral antigens are detected only when ultrasensitive IHC methods are employed. Perhaps most troubling, the normally nuclear IE1 and pp65 proteins detected with this technique are almost invariably cytoplasmic when detected in GBM tissue, calling into question the assay's specificity. Thus, any association of HCMV with GBM is considered with a healthy skepticism.

The previous studies discussed above often suffer from small sample sizes, low numbers of HCMV loci analyzed, subjective assays, and insufficient quantitative analysis. Thus, we sought to independently determine whether HCMV was statistically more likely to be present in GBM samples as opposed to other cancerous or noncancerous brain tissues by testing for the presence of multiple viral genomic loci by PCR in GBM specimens. Our results lead us to conclude that all regions of the HCMV genome are present in the vast majority of GBM samples but that only a small minority of cells in any individual sample harbors HCMV DNA.

MATERIALS AND METHODS

Sample collection and DNA extraction. This retrospective study was conducted in agreement with the terms of a University of Wisconsin—Madison (UW-Madison) institutional review board (IRB) protocol (M-2009-1420). Tumor samples were obtained with prior patient consent. The samples were deidentified and are untraceable to individual patients. Only the date of collection of the sample and diagnosis information were available. The samples were destroyed during the course of the analysis. Sample preparation was performed in clean laboratories never before used for HCMV research in order to prevent spurious contamination. Paraffin-embedded archived astrocytoma grade 4 (GBM), meningioma, schwannoma, oligodendroglioma, or nonneoplastic epileptic brain samples that were of sufficient size (>3 mm) were pulled, and histopathology slides were examined. Sections containing the highest percentage of tumor cells and the least amount of necrosis were chosen. The 75 GBM samples analyzed here represent essentially the entire collection of usable samples collected from 1994 through 2009 at UW-Madison. To purify DNA from paraffin-embedded tissues, 10-μm sections were extracted

FIG 1 Comparison of Q-PCR and conventional PCR for the detection of an HCMV genomic sequence. (A) Quantitative real-time PCR detection of the HCMV UL123 (IE1) genomic locus from human fibroblasts infected at MOIs of 0.1, 0.01, 0.001, and 0.0001 for 18 h. (B) Linear plot of threshold values from panel A. (C) Agarose gel detection of the HCMV UL123 (IE1) genomic locus amplified by conventional PCR using the same primers and templates as in panel A. (D) Linear plot of Image J quantification of the gel shown in panel C.

twice with xylene and treated with DNA lysis buffer (50 mM KCl, 10 mM Tris-HCl [pH 8.3], 2.5 mM MgCl₂, 100 μg/ml gelatin, 0.45% IGEPAL [octylphenoxypolyethoxyethanol], 0.45% Tween 20, and 60 μg/ml proteinase) at 55°C for 3 h and then at 95°C for 10 min. DNA was subsequently precipitated from supernatants with sodium acetate and ethanol and finally resuspended in deionized/distilled water. For nonarchived tissue samples, ~1 mg of frozen but not fixed tissue was minced and DNA was extracted with the Qiagen DNeasy kit (catalog no. 69506) according to the manufacturer's protocol. Positive-control samples were generated by infecting primary human foreskin fibroblasts with HCMV strain TB40/E at multiplicities of infection (MOIs) of 10⁻⁵, 10⁻⁴, 10⁻³, 10⁻², or 10⁻¹ PFU/cell for 18 h and preparing DNA with the Qiagen DNeasy kit. It is suspected that little to no viral DNA replication takes place within this time frame, so the amount of viral DNA present reflects input viral genomes. DNA was quantitated by absorbance at 260 nm.

Experimental and statistical analyses. Samples were analyzed with PCRs and agarose-ethidium bromide gel electrophoresis in clean laboratories never before used for HCMV research in order to prevent spurious contamination. For conventional PCRs, 100 ng template DNA and *Taq* polymerase (New England BioLabs) along with nucleotides and buffers

TABLE 2 *P* values from Wilcoxon rank sum test

Gene	<i>P</i> value for comparison ^a		No. of replicates
	GBM vs epilepsy	GBM vs control tumors	
UL17	0.0001	0.18	2
UL27	0.015	0.83	1
UL55	0.53	<0.0001	1
UL69	0.012	0.14	1
UL82	0.03	0.24	2
UL96	0.007	0.013	1
UL111A	0.75	0.37	1
UL122	0.0001	0.18	2
UL144	0.19	0.54	1
US2	0.75	0.38	2
US11	0.02	0.14	2
US28	0.0001	0.0002	1

^a *P* values for the indicated comparisons of the levels of PCR amplification products for the indicated viral genes are displayed. *P* values that indicate statistical significance are shown in boldface type.

provided by the polymerase manufacturer were utilized in 25-μl reaction mixtures that were run for 35 cycles. Primer sequences are presented in Table 1. Primer pair fidelity was confirmed by sequencing positive-control and paraffin-embedded GBM tumor DNA reaction products. Sequences were analyzed by nucleotide BLAST (NCBI). PCR products were separated by electrophoresis, and bands of the expected size were quantitated with densitometry using Image J software. Quantitative real-time PCR (Q-PCR) methods have been previously described (12). Statistical analyses were performed with Graphpad prism. Linear graphs were created with Microsoft Excel, and heat maps were created with Heatmap Builder (14).

RESULTS

Semiquantitative PCR as an objective assay for HCMV sequences in GBM samples. In previous studies where PCR was used to probe for HCMV genomes in GBM samples (5, 20, 29), a positive or negative scoring method was employed based, presumably, on visual examination of reaction products separated on agarose gels. We sought a more quantitative and objective method, and thus, we compared the efficiency with which quantitative real-time PCR (Q-PCR) and semiquantitative (conventional) PCR could differentiate between differing levels of HCMV genomes in complex samples. Templates consisted of a series of DNA preparations from primary human fibroblasts (HFs) infected with 10-fold dilutions of HCMV strain TB40/E for 18 h. These samples also served as positive controls during the analysis of the samples described below. At this time point, nuclear entry of viral genomes should be complete but viral DNA replication not yet initiated. Thus, the level of DNA detected should be directly

FIG 2 HCMV genomic regions surveyed. Schematic representation of the genomic loci amplified by the PCR primers used in this study. Primer pair locations are approximate, and sizes are not drawn to scale. TR, terminal repeats; UL, unique long; IR, internal repeat; US, unique short. Numbers indicate the specific gene (e.g., UL82) amplified by the indicated primer pair.

FIG 3 Mann-Whitney U/Wilcoxon rank sum test of paraffin-embedded samples. DNAs from tissue samples were used as templates for conventional PCR amplification with primers detecting the indicated HCMV genomic locus. Bands on agarose gels were quantitated by Image J software, and arbitrary units are plotted. Samples from individual classes (epilepsy, GBM, and other brain tumors [OBT]) were compared with the Mann-Whitney U/Wilcoxon rank sum test. Statistically significant differences are indicated with asterisks as follows: ***, P value < 0.0001 ; *, $P < 0.05$.

related to input viral genomes. For this test comparison, primers (Table 1) that could be used for both Q-PCR and traditional PCR amplification of the IE1 locus were employed.

Q-PCR was performed (Fig. 1A), and threshold values were determined and plotted (Fig. 1B) as a function of the multiplicity of infection (MOI). Regression analysis demonstrated the linearity of the assay (coefficient of determination [R^2] = 0.93). Con-

ventional PCR was also performed, and reaction products were separated by agarose gel electrophoresis (Fig. 1C), quantitated by Image J software, and plotted (Fig. 1D) as a function of MOI. Regression analysis demonstrated the linearity of this assay (R^2 = 0.97). Thus, for this primer set, conventional PCR (yielding the larger R^2 value) would have comparable (in fact, superior) predictive value for future samples than Q-PCR. Therefore, we

judged that conventional PCR would be sufficient to obtain the semiquantitative data that would allow for an objective evaluation of the presence of HCMV DNA in a given sample. Due to the large number of samples to be examined and primer pairs to be employed, we chose to use conventional PCR for most of our analysis.

Detection of HCMV sequences in paraffin-embedded samples. Twelve primer pairs for the examination of paraffin-embedded samples (Table 1) were developed based on the sequence of the cloned TB40/E strain (GenBank accession no. EF999921.1) that spanned 229 kb of the HCMV genome at approximately 20-kb intervals (Fig. 2). Primers within genes encoding proteins with functional activities relevant to cancer (13) were used whenever possible. Primer specificity was confirmed by sequencing PCR amplification products generated with positive-control templates (data not shown). Paraffin-embedded tissue sections were obtained from the Pathology archives of the Hospital & Clinics at the University of Wisconsin—Madison and consisted of samples from patients diagnosed with GBM ($n = 75$), meningioma ($n = 6$), schwannoma ($n = 5$), oligodendroglioma ($n = 5$), or epilepsy ($n = 15$). Healthy (nondiseased) brain samples were rare in the collection, so nonneoplastic brain samples from patients with epilepsy were used as a surrogate. For the GBM samples, the number analyzed represented essentially every usable sample from the collection. DNA preparation and PCR analysis were carried out in clean rooms not previously used for HCMV research to avoid spurious contamination of samples and reaction mixtures. The total reaction products were separated by agarose gel electrophoresis, and bands of the appropriate size were quantitated with Image J software (data not shown). Background values were subtracted from individual band intensities. Where indicated (Table 2), technical replicates were averaged. Data and statistical analyses are described below.

Mann-Whitney U/Wilcoxon rank sum test. The levels of HCMV sequences in GBM, epilepsy, or other brain tumor paraffin-embedded samples were compared using the Mann-Whitney U test (also called the Wilcoxon rank sum test), a non-parametric statistical examination applicable to arbitrary sample sizes (33). We considered comparisons with $P < 0.05$ (denoted by an asterisk in Fig. 3) to indicate that GBM samples were more likely to contain larger amounts of HCMV DNA than the comparison cohort and comparisons with $P < 0.0001$ (denoted by three asterisks in Fig. 3) to indicate that GBM samples were much more likely to contain HCMV DNA than the comparison cohort.

Our analysis indicated that 8 out of 12 HCMV genomic loci tested were statistically more likely to be present in GBM samples than in brain samples from patients with epilepsy (Fig. 3 and Table 2). The eight loci were UL17, UL27, UL69, UL82, UL96, UL122, US11, and US28. Of these loci, three (UL96, US11, and US28) were also statistically overrepresented in GBM samples compared to other brain tumors. The HCMV UL55 locus was statistically more likely to be present in GBM samples compared to other brain tumors, but not compared to brain samples from patients with epilepsy. In no case was the HCMV genome statistically overrepresented in control samples (epilepsy or other brain tumors) compared to GBM samples. We also corrected for multiple comparisons by estimating false discovery rates (FDR) by calculating Q values (1). We considered comparisons with $Q < 0.05$ to indicate that HCMV DNA was more likely to be found in GBM samples than in the comparison cohort and comparisons with $Q <$

FIG 4 False discovery rate analysis. P values (Table 2) were converted to Q values to account for the multiple comparisons performed. The negative log of the Q value for each primer pair is plotted for two comparisons, GBM versus epilepsy and GBM versus OBT. Lines are drawn at P values of 0.05 and 0.001. Values between the lines are considered to represent significant differences between the compared cohorts. Values above both lines are considered highly significant. A schematic diagram of the viral genome labeled with the genomic loci surveyed is shown for illustrative purposes (not to scale).

0.0001 to indicate that HCMV DNA was much more likely to be found in GBM samples than in the comparison cohort. The results of FDR analysis echoed our initial analysis (Fig. 4 and Table 2). From these two tests, we conclude that multiple HCMV genomic loci, and very likely the entire genome, are found in GBM tumors.

y -intercept analysis. We extended the analysis of the PCR data to objectively classify each sample as positive or negative for the viral genomic region amplified by each primer pair. To obtain objectivity for the complete data set (Fig. 5), we used a set of positive-control samples and linear regression analysis to generate a standard curve with which we could calculate slope, y -axis intercept, and R^2 values (Table 3). Note that the R^2 values indicate that while some primer sets (e.g., UL17) can be used to predict unknown values very efficiently, others (e.g., UL55) do so less effectively. At the y intercept, the value of x (the MOI of HCMV) is equal to zero, setting the theoretical limit of detection in this assay. Therefore, samples with band intensities greater than the y -intercept value were considered positive, and those with band intensities equal to or less than the y -intercept value were considered negative. Heat maps depicting either positive (red) or negative (green) results for each sample and each primer pair were generated (Fig. 6).

Two observations emerged from this analysis. First, the amplification of HCMV DNA sequences from paraffin-embedded tissue was less likely in brain samples from patients with epilepsy (Fig. 6A) and other brain tumors (Fig. 6B) than from GBM samples (Fig. 6C). By using this method of data analysis, which is independent of the statistical one described above, we reached a similar conclusion, namely, that multiple HCMV genomic loci, and very likely the entire genome, are found in GBM tumors.

FIG 5 y -intercept analysis of individual paraffin-embedded samples. PCR data from GBM (gray squares), epilepsy (black triangles), meningioma (crosses), schwannoma (large gray asterisks), and oligodendroglioma (gray circles) samples (Fig. 3) are plotted along with values from positive-control samples (black diamonds) that consisted of human fibroblasts infected with HCMV at MOIs of 0.01, 0.001, 0.0001, and 0.00001. Linear regression analysis with the positive-control samples allowed for extrapolation to the y -axis intercept, denoted as a horizontal line.

Second, GBM samples archived more recently were more likely to be identified as HCMV positive than older samples. Therefore, we repeated our analysis restricting samples to those collected within the same time frame (after 2007), which allowed us to examine 35 GBM samples, 8 brain samples from patients with epilepsy, and all 16 non-GBM brain tumors. Counting all PCRs, we found the GBM samples to be 55% positive, non-GBM brain tumors to be 42% positive, and brain samples from patients with epilepsy to be 34% positive. The Mann-Whitney U/Wilcoxon rank sum test and

the false discovery rate statistical test applied to these samples (Fig. 7) agreed well with the previous analysis (Fig. 3 and Fig. 4), indicating that multiple regions of the genome are statistically more likely to be found in GBMs than in other brain tumors or brain samples from patients with epilepsy.

Sequencing confirms PCR product identity. Amplified PCR products were quantitated and served as the diagnostic criteria for labeling a sample as HCMV positive or HCMV negative. To confirm that they actually represented HCMV DNA, selected PCR

TABLE 3 Estimated MOIs and parameters of semiquantitative analysis^a

Gene	Slope	y-axis intercept	R ²	Estimated MOI
UL17	34	39	0.99	4.5 × 10 ⁻⁷
UL27	28	42	0.82	2.7 × 10 ⁻⁵
UL55	16	85	0.5	3.6 × 10 ⁻⁵
UL69	26	25	0.88	1.06 × 10 ⁻⁶
UL82	13	34	0.87	9.7 × 10 ⁻⁴
UL96	43	-4	0.96	1.2 × 10 ⁻⁶
UL111A	35	73	0.84	4.7 × 10 ⁻⁷
UL122	31	80	0.98	5.3 × 10 ⁻⁷
UL144	21	4.5	0.77	4.7 × 10 ⁻⁶
US2	12	39	0.89	3.3 × 10 ⁻⁵
US11	33	50	0.89	7.6 × 10 ⁻⁶
US28	26	57	0.8	8.9 × 10 ⁻⁶

^a Slope, y-axis intercept, and correlation coefficient (R²) values for standard curve analysis with the indicated PCR primer set are shown. These parameters were used to estimate the multiplicity of infection (MOI) with HCMV within tumor samples based on the MOIs of the standards.

products resulting from amplifications using template DNA derived from paraffin-embedded GBM samples were sequenced (Table 4). Seventy-three individual PCR bands representing 10 distinct HCMV loci were sequenced. Internal regions were compared to sequences derived from the two most commonly utilized HCMV strains in our laboratory, AD169 (GenBank accession no. [AC146999.1](#)) and TB40/E (GenBank accession no. [EF999921.1](#)). Seventy-one (97%) of the sequences matched the predicted HCMV locus (Table 4). One sample amplified with UL144 primers generated a product with no similarity to any known sequence, and one sample amplified with the US2 primers generated a product with limited identity (82% over 47 nucleotides) to the human SLC25A16 gene (GenBank accession no. [AL713888.10](#)) and various human and primate bacterial artificial chromosome (BAC) clones. Only 11 of the HCMV sequences (15%) were 100% matches to the strains used in our laboratory (9 complete matches to TB40/E and 2 complete matches to AD169), indicating that the bands were unlikely to result from sample contamination. Sequence confirmation strengthens our conclusion that HCMV is associated with GBM tumors.

Detection of HCMV sequences in frozen samples. The differences between newer and older paraffin-embedded archival tissue samples prompted us to examine newer, nonarchival GBM samples for the presence of HCMV genomes. DNA was prepared from 12 frozen GBM specimens and analyzed by conventional PCR with 19 primer pairs (Fig. 2) that include the same 12 primer pairs employed above and seven novel primer pairs (Table 1) that detect other viral genomic loci (UL1, UL18, UL19, UL36, UL44, UL97, and UL98). The positive control was HFs infected with HCMV strain TB40/E (MOI of 0.1), and negative controls included uninfected HFs and water in place of a DNA template. A portion of the PCR was separated by agarose gel electrophoresis (Fig. 8), and bands of the appropriate size were quantitated with Image J software. Bands with intensities more than twice that of the water control were considered positive for HCMV DNA, and along with negative reactions, are displayed in heat map format (Fig. 9). Overall, 70% of the individual PCRs with GBM sample templates were positive for HCMV. None of the uninfected HF samples were positive. These data support our previous conclusion that multiple HCMV genomic loci, and very likely the entire genome, are

found in GBM tumors. Unfortunately, the small sample number and close collection dates (2006 to 2008) prevent us from determining whether HCMV sequences are more likely to be detected in newer frozen samples than in older ones, as they are for the paraffin-embedded samples.

One primer pair (UL18) produced only negative PCRs of the GBM samples but was able to amplify the positive control. This could indicate that this region of the genome may be missing in otherwise HCMV-positive GBM samples. However, adjacent primer sets (UL17 and UL19) showed overwhelmingly positive reactions (Fig. 8). Furthermore, when we amplified five randomly selected GBM samples with a primer pair designed to detect a 3-kb fragment spanning UL17, UL18, and UL19, all were positive for the full-length fragment (Fig. 10). Therefore, we conclude that the negative results with the UL18 primer pair likely represents inefficient amplification rather than the absence of those genomic regions in HCMV-positive GBM specimens.

Cellular genomes outnumber viral genomes in HCMV-positive GBM samples. A qualitative examination of the data (Fig. 5, 8, and 10) seems to indicate that while the majority of GBM samples are positive for multiple regions of the viral genomes, the amount of viral DNA within any individual sample mimics what is observed in low-MOI positive controls. Thus, one might conclude that not all of the cells in a positive tissue sample are infected with, or are carrying viral DNA sequences. If this were true, cellular genomes might outnumber viral genomes in GBM samples.

To directly examine whether cellular genomes outnumber viral genomes, we used quantitative real-time PCR to determine exact amounts of viral IE1 and cellular actin sequences present in HCMV-positive frozen GBM specimens and used those values to develop a ratio between cellular and viral DNA sequences. We found that, on average, there were approximately 160 copies of actin present for every molecule of IE1 DNA (Fig. 11). The diploid nature of the HFs means that if every infected cell retained only a single copy of the HCMV genome, only 1 in 80 cells within an average HCMV-positive GBM specimen would carry viral sequences. As many virus-infected cancer cells harbor more than one genome per cell (27, 38), it is likely that even fewer cells within an HCMV-positive GBM sample are actually HCMV positive. Additionally, by comparison to a standard curve (R² = 0.65), we were able to establish the approximate MOIs of the infected samples. The average MOI of the frozen GBM samples, 0.1, generally agrees with the number of viral genomes/actin copy analysis (Fig. 11) and as expected is substantially higher than those predicted from the analysis of paraffin-embedded samples (Table 3). In total, the results of our analysis indicate that the majority of the HCMV genome is present in the majority of GBM samples, but only in a minority of the cells of any individual tumor specimen.

DISCUSSION

Data we present here clearly demonstrate that HCMV genomes are associated with GBM tumors (Fig. 3 and 9), in agreement with the majority of studies that have investigated this potential interaction (5, 16, 20, 29, 31, 35). Our results also indicate that primer selection and sample quality (age or method of preservation) can have substantial effects on the outcome of the assay (Fig. 6 and 10). Interestingly, recent evidence indicates that with time, herpesviral genomes inside cells become refractory to PCR amplification due to physical damage (19) and that HCMV genomes *in vivo* show sequence variability that mirrors that of RNA viruses (26). All of

FIG 6 Heat map representation of HCMV-positive and HCMV-negative PCRs from paraffin-embedded samples. HCMV genomic loci are indicated across the top of the heat maps. Unique samples (including the year in which they were collected) are displayed on individual rows. (A) Brain samples from patients with epilepsy with values greater than the y -intercept value (Fig. 5) for each individual PCR are considered positive and denoted as a red square. Values equal to or less than the y -intercept value are considered negative and denoted as a green square. (B and C) OBT samples (B) and GBM samples (C) are displayed as in panel A.

these factors may contribute to the few negative studies that have been reported (24, 28) and may in fact lead to an underestimation of the association of HCMV with GBMs.

Multiple loci that span the viral genome were detected in asso-

ciation with GBM (Fig. 3 and 8), making it unlikely that only select genomic fragments are present, and it is possible that entire, intact genomes exist in these tumors. Deep sequencing analysis should give a more definitive answer as to the extent of viral sequences

FIG 7 False discovery rate analysis of recently collected paraffin-embedded samples. *P* values from Mann-Whitney U/Wilcoxon rank sum tests performed only on samples collected after 2007 were converted to *Q* values to account for the multiple comparisons performed. The negative log of the *Q* value for each primer pair is plotted for two comparisons, GBM versus epilepsy and GBM versus OBT. Lines are drawn at *P* values of 0.05 and 0.001. Values between the lines are considered to represent significant differences between the compared cohorts. Values above both lines are considered highly significant. A schematic diagram of the viral genome labeled with the genomic loci surveyed is shown below the graph for illustrative purposes (not to scale).

present in GBMs. Our data would strongly argue for the use of recently sampled tumors for such experiments. Should entire genomes be observed in GBMs, determining their structure (likely circular episomes or integrated, linear units) could provide clues as to whether or how such genomes are replicated and maintained in these tumors.

The results of our quantitative analysis indicated that cellular genomes greatly outnumber viral genomes in the GBMs we analyzed (Fig. 11). This likely means that not every cell within the

FIG 8 PCR analysis of frozen GBM tumors. DNA extracted from 12 individual frozen GBM tumors (lanes 1 to 12) was analyzed by PCR with primers detecting the indicated HCMV genomic locus. Reaction products were separated by agarose gel electrophoresis and visualized with ethidium bromide staining. DNA extracted from mock-infected (M) or HCMV-infected (MOI of 0.1) human fibroblasts (HCMV strain TB40/E), as well as water (W) served as template controls.

sample is HCMV positive and probably explains why viral genomes were not detected in a genomic sequencing study (23) of GBM tumors (Charles A. Whittaker, personal communication). Therefore, if HCMV is contributing to the oncogenicity of the GBM tumor (see below), it likely does so through mechanisms distinct from other classical cancer viruses (such as EBV and HPV), where full genomes or genomic fragments are found in essentially every cell within a virally induced tumor.

Because it appears that all cells within an HCMV-positive GBM are not infected, it will be important to determine whether specific cell types within the tumor are preferentially (or exclusively) infected with HCMV. A recent analysis (6) detected pp65 antigen in CD11b-positive cells of the monocyte lineage, but not in CD11b-negative lymphocytes. Another intriguing possibility is that neural stem cells harbor the virus. The theory that pluripotent stem cells or stem-like cells are uniquely able to maintain and propagate at least a subset of human tumors is gaining widespread acceptance. HCMV infection of only the neural stem cells within a GBM could explain our results indicating that very few cells within a tumor

TABLE 4 Sequence confirmations^a

Locus	No. of samples analyzed	No. of samples homologous to HCMV	Region analyzed (nt) (TB40/E genome)	Avg % identity to TB40/E	Avg % identity to AD169
UL17	9	9	200 (56128–56328)	96.4	95.7
UL27	5	5	180 (67287–67467)	98.8	98
UL55	5	5	225 (116334–116559)	95.6	93
UL69	10	10	160 (133375–133531)	94.5	94
UL82	5	5	100 (151643–151743)	96.4	97.8
UL96	12	12	210 (173778–173988)	98.3	98.2
UL144	5	4	100 (214774–214874)	95	NA
US2	5	4	75 (228975–229050)	96.7	94
US11	9	9	210 (3728–3938)	97	95.5
US28	8	8	360 (23505–23865)	99	98.5

^a The indicated loci were amplified by PCR and sequenced, and the indicated internal region (in nucleotides [nt]) was compared to the NCBI nucleotide database. Average percent identities for each locus compared to clinical (TB40/E) and laboratory (AD169) strains of HCMV are shown. AD169 lacks the UL144 locus, so this comparison is not applicable (NA).

FIG 9 Heat map representation of HCMV-positive and -negative PCRs from frozen GBMs. Amplified bands (Fig. 8) were quantitated with Image J software. Values greater than twice the water control for each individual PCR are considered positive and denoted as a red square. Values equal to or less than twice the water control are considered negative and denoted as a green square. Genomic loci are listed across the top of the heat map, and unique samples are found on each individual row along with their year of collection.

sample are infected and why cultured GBM cell lines (that are clearly not neural stem cells) are invariably HCMV negative. Interestingly, HCMV antigens were recently found to be expressed in cultured neural stem cells (6). The primary location of neural stem cells is the subventricular zone in the CNS (9), which is also one of the major sites of mouse cytomegalovirus reactivation from latency (37). Thus, a more thorough examination of natural HCMV infections of neural stem cells appears warranted.

The question remains as to what, if anything, HCMV is doing within virus-positive GBMs. It is certainly possible that the microenvironment created by the tumor may provide a favorable location for either viral replication or simply for the persistence of HCMV-infected cells. Were this the case, the virus might serve as a useful biomarker for GBM tumors, and studying the interactions between HCMV and GBMs might provide some insight into virus biology, but it would be unlikely that a virus-based treatment would be an effective cancer therapy. Clinical data for the usefulness of anti-HCMV therapy for the treatment of GBM patients is currently being collected, and while positive results have been reported anecdotally (32), they have yet to be published. Thus, while HCMV has been described as oncomodulatory (17, 18), at this time, it is perhaps more appropriate to describe it as oncoaccessory, since the current evidence supports its presence in GBM tumors without directly addressing whether natural infections actually modulate tumor cell biology *in vivo* or *in vitro*.

However, if antiviral treatments prove effective therapies for GBMs, it may indicate that viral infection is providing some growth or survival advantage to the tumor. Clearly, HCMV en-

FIG 10 Long-range PCR. Randomly selected DNA templates prepared from the indicated GBM and control samples (from Fig. 8) were amplified with primers that span the UL17 to UL19 loci. Reaction products were visualized by agarose-ethidium bromide gel electrophoresis. HFF, human foreskin fibroblasts.

FIG 11 Q-PCR calculation of viral DNA levels. Randomly selected DNA templates prepared from the indicated GBM and control samples (from Fig. 8) were analyzed by quantitative real-time PCR for the viral UL123 (IE1) locus and cellular actin. The amount of actin DNA compared to the amount of viral DNA in each sample is presented. Error bars represent the standard errors of the means.

codes activities capable of eliciting all of the hallmarks of human cancers (10, 13). As the functions of many viral proteins and RNAs have been identified, determining which portions of the viral genome are expressed in GBM tumors should provide clues as to how viral infection may be associated with oncogenesis.

That viruses can cause cancer is a well-accepted paradigm. However, the heterogeneity of human tumors and the substantially different ways in which known tumor viruses are oncogenic make it challenging to establish a set of diagnostic experiments that can prove or disprove a causative role for a virus in a neoplasm. A major confounding issue is that most infections even with known cancer viruses do not result in oncogenic events. Despite these difficulties, there is a growing awareness that infectious pathogens, including viruses, may be oncogenic agents in a greater number of human tumors than is currently appreciated. Previously published results, supported by the new data we report here, indicate that GBMs, at the very least, are virus-associated tumors.

ACKNOWLEDGMENTS

We thank Phil Balandyk for expert technical assistance, Norman Drinkwater for guidance during the statistical analysis, Charlie Whittaker (David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology) for unpublished bioinformatics analysis, and the members of our laboratories for helpful comments.

This work was supported by NIH grants AI080675 to R.F.K., T32AG027566 (P.A.C.) to the University of Wisconsin—Madison Stem Cell Training Program, and CA014520 (J.S.K.) to the University of Wisconsin—Madison Carbone Cancer Center. J.S.K. was partially supported by the HEADRUSH Brain Tumor Research Professorship and the Roger Loff Memorial Fund for GBM Research. R.F.K. is a Burroughs Wellcome Fund Investigator in the Pathogenesis of Infectious Disease.

P.R. and R.F.K. designed the experiments. P.R. performed the experiments. P.A.C., J.S.K., and M.S.S. provided materials. P.R. and R.F.K. analyzed the data. P.R. and R.F.K. wrote the manuscript with input from all authors.

REFERENCES

- Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57:289–300.
- Britt W. 2008. Manifestations of human cytomegalovirus infection: pro-

- posed mechanisms of acute and chronic disease. *Curr. Top. Microbiol. Immunol.* 325:417–470.
3. Brune W. 2011. Inhibition of programmed cell death by cytomegaloviruses. *Virus Res.* 157:144–150.
 4. Caposio P, Orloff SL, Streblow DN. 2011. The role of cytomegalovirus in angiogenesis. *Virus Res.* 157:204–211.
 5. Cobbs CS, et al. 2002. Human cytomegalovirus infection and expression in human malignant glioma. *Cancer Res.* 62:3347–3350.
 6. Dziurzynski K, et al. 2011. Glioma-associated cytomegalovirus mediates subversion of the monocyte lineage to a tumor propagating phenotype. *Clin. Cancer Res.* 17:4642–4649.
 7. Fields BN, Knipe DM, Howley PM, Griffin DE. 2007. *Fields virology*, 5th ed. Lippincott Williams & Wilkins, Philadelphia, PA.
 8. Fortunato EA, Dell'Aquila ML, Spector DH. 2000. Specific chromosome 1 breaks induced by human cytomegalovirus. *Proc. Natl. Acad. Sci. U. S. A.* 97:853–858.
 9. Glantz M, Kesari S, Recht L, Fleischhack G, Van Horn A. 2009. Understanding the origins of gliomas and developing novel therapies: cerebrospinal fluid and subventricular zone interplay. *Semin. Oncol.* 36: S17–S24.
 10. Hanahan D, Weinberg RA. 2011. Hallmarks of cancer: the next generation. *Cell* 144:646–674.
 11. Hume AJ, Kalejta RF. 2009. Regulation of the retinoblastoma proteins by the human herpesviruses. *Cell Div.* 4:1.
 12. Hwang J, Saffert RT, Kalejta RF. 2011. Elongin B-mediated epigenetic alteration of viral chromatin correlates with efficient human cytomegalovirus gene expression and replication. *mBio* 2(2):e00023–11.
 13. Hwang J, Winkler L, Kalejta RF. 2011. Ubiquitin-independent proteasomal degradation during oncogenic viral infections. *Biochim. Biophys. Acta* 1816:147–157.
 14. King JY, et al. 2005. Pathway analysis of coronary atherosclerosis. *Physiol. Genomics* 23:103–118.
 15. Lau SK, et al. 2005. Lack of association of cytomegalovirus with human brain tumors. *Mod. Pathol.* 18:838–843.
 16. Lucas KG, Bao L, Bruggeman R, Dunham K, Specht C. 2011. The detection of CMV pp65 and IE1 in glioblastoma multiforme. *J. Neurooncol.* 103:231–238.
 17. Michaelis M, Doerr HW, Cinatl J. 2009. The story of human cytomegalovirus and cancer: increasing evidence and open questions. *Neoplasia* 11:1–9.
 18. Michaelis M, Doerr HW, Cinatl J Jr. 2009. Oncomodulation by human cytomegalovirus: evidence becomes stronger. *Med. Microbiol. Immunol.* 198:79–81.
 19. Millhouse S, et al. 2010. Evidence that herpes simplex virus DNA derived from quiescently infected cells in vitro, and latently infected cells in vivo, is physically damaged. *J. Neurovirol.* 16:384–398.
 20. Mitchell DA, et al. 2008. Sensitive detection of human cytomegalovirus in tumors and peripheral blood of patients diagnosed with glioblastoma. *Neuro-Oncology* 10:10–18.
 21. Moore PS, Chang Y. 2010. Why do viruses cause cancer? Highlights of the first century of human tumour virology. *Nat. Rev. Cancer* 10:878–889.
 22. Ohgaki H. 2009. Epidemiology of brain tumors. *Methods Mol. Biol.* 472: 323–342.
 23. Parsons DW, et al. 2008. An integrated genomic analysis of human glioblastoma multiforme. *Science* 321:1807–1812.
 24. Poltermann S, et al. 2006. Lack of association of herpesviruses with brain tumors. *J. Neurovirol.* 12:90–99.
 25. Powers C, DeFilippis V, Malouli D, Fruh K. 2008. Cytomegalovirus immune evasion. *Curr. Top. Microbiol. Immunol.* 325:333–359.
 26. Renzette N, Bhattacharjee B, Jensen JD, Gibson L, Kowalik TF. 2011. Extensive genome-wide variability of human cytomegalovirus in congenitally infected infants. *PLoS Pathog.* 7:e1001344.
 27. Rodriguez-Inigo E, et al. 2005. Percentage of hepatitis C virus-infected hepatocytes is a better predictor of response than serum viremia levels. *J. Mol. Diagn.* 7:535–543.
 28. Sabatier J, et al. 2005. Detection of human cytomegalovirus genome and gene products in central nervous system tumours. *Br. J. Cancer* 92: 747–750.
 29. Scheurer ME, Bondy ML, Aldape KD, Albrecht T, El-Zein R. 2008. Detection of human cytomegalovirus in different histological types of gliomas. *Acta Neuropathol.* 116:79–86.
 30. Shen Y, Zhu H, Shenk T. 1997. Human cytomegalovirus IE1 and IE2 proteins are mutagenic and mediate “hit-and-run” oncogenic transformation in cooperation with the adenovirus E1A proteins. *Proc. Natl. Acad. Sci. U. S. A.* 94:3341–3345.
 31. Slinger E, et al. 2010. HCMV-encoded chemokine receptor US28 mediates proliferative signaling through the IL-6-STAT3 axis. *Sci. Signal.* 3:ra58.
 32. Soderberg-Naucler C. 2008. HCMV microinfections in inflammatory diseases and cancer. *J. Clin. Virol.* 41:218–223.
 33. Sokal RR, Rohlf FJ. 1995. *Biometry: the principles and practice of statistics in biological research*, 3rd ed. Freeman, New York, NY.
 34. Soroceanu L, Cobbs CS. 2011. Is HCMV a tumor promoter? *Virus Res.* 157:193–203.
 35. Straat K, et al. 2009. Activation of telomerase by human cytomegalovirus. *J. Natl. Cancer Inst.* 101:488–497.
 36. Streblow DN, Dumortier J, Moses AV, Orloff SL, Nelson JA. 2008. Mechanisms of cytomegalovirus-accelerated vascular disease: induction of paracrine factors that promote angiogenesis and wound healing. *Curr. Top. Microbiol. Immunol.* 325:397–415.
 37. Tsutsui Y, Kawasaki H, Kosugi I. 2002. Reactivation of latent cytomegalovirus infection in mouse brain cells detected after transfer to brain slice cultures. *J. Virol.* 76:7247–7254.
 38. Vereide DT, Sugden B. 2011. Lymphomas differ in their dependence on Epstein-Barr virus. *Blood* 117:1977–1985.
 39. Vomaske J, et al. 2010. HCMV pUS28 initiates pro-migratory signaling via activation of Pyk2 kinase. *Herpesviridae* 1:2.
 40. Waller EC, Day E, Sissons JG, Wills MR. 2008. Dynamics of T cell memory in human cytomegalovirus infection. *Med. Microbiol. Immunol.* 197:83–96.
 41. Yu Y, Clippinger AJ, Alwine JC. 2011. Viral effects on metabolism: changes in glucose and glutamine utilization during human cytomegalovirus infection. *Trends Microbiol.* 19:360–367.
 42. Yuan X, et al. 2004. Isolation of cancer stem cells from adult glioblastoma multiforme. *Oncogene* 23:9392–9400.
 43. Zhu Y, Parada LF. 2002. The molecular and genetic basis of neurological tumours. *Nat. Rev. Cancer.* 2:616–626.