Title

The immunogenetics of viral antigen response is associated with subtype-specific glioma risk and survival

Journal

American Journal of Human Genetics, 109(6)

ISSN

0002-9297

Authors

Guerra, Geno

Kachuri, Linda

Wendt, George

et al.

Publication Date

2022-06-01

DOI

10.1016/j.ajhg.2022.04.011

The immunogenetics of viral antigen response is associated with subtype-specific glioma risk and survival

Authors

Geno Guerra, Linda Kachuri, George Wendt, ..., Robert B. Jenkins, Margaret Wrensch, Stephen S. Francis

Correspondence

geno.guerra@ucsf.edu (G.G.),
stephen.francis@ucsf.edu (S.S.F.)

Guerra et al., 2022, The American Journal of Human Genetics *109*, 1105–1116
June 2, 2022
https://doi.org/10.1016/j.ajhg.2022.04.011

CellPress

The immunogenetics of viral antigen response is associated with subtype-specific glioma risk and survival

Geno Guerra,[1,*] Linda Kachuri,[2] George Wendt,[1] Helen M. Hansen,[1] Steven J. Mack,[3] Annette M. Molinaro,[1,2] Terri Rice,[1] Paige Bracci,[2] John K. Wiencke,[1,2,4] Nori Kasahara,[1,7] Jeanette E. Eckel-Passow,[5] Robert B. Jenkins,[6] Margaret Wrensch,[1,4] and Stephen S. Francis[1,2,7,*]

Summary

Glioma is a highly fatal cancer with prognostically significant molecular subtypes and few known risk factors. Multiple studies have implicated infections in glioma susceptibility, but evidence remains inconsistent. Genetic variants in the human leukocyte antigen (HLA) region modulate host response to infection and have been linked to glioma risk. In this study, we leveraged genetic predictors of antibody response to 12 viral antigens to investigate the relationship with glioma risk and survival. Genetic reactivity scores (GRSs) for each antigen were derived from genome-wide-significant ($p < 5 \times 10^{-8}$) variants associated with immunoglobulin G antibody response in the UK Biobank cohort. We conducted parallel analyses of glioma risk and survival for each GRS and HLA alleles imputed at two-field resolution by using data from 3,418 glioma-affected individuals subtyped by somatic mutations and 8,156 controls. Genetic reactivity scores to Epstein-Barr virus (EBV) ZEBRA and EBNA antigens and Merkel cell polyomavirus (MCV) VP1 antigen were associated with glioma risk and survival (Bonferroni-corrected $p < 0.01$). GRS_{ZEBRA} and GRS_{MCV} were associated in opposite directions with risk of *IDH* wild-type gliomas ($OR_{ZEBRA} = 0.91$, $p = 0.0099$/$OR_{MCV} = 1.11$, $p = 0.0054$). GRS_{EBNA} was associated with both increased risk for *IDH* mutated gliomas (OR = 1.09, p = 0.040) and improved survival (HR = 0.86, p = 0.010). *HLA-DQA1*03:01* was significantly associated with decreased risk of glioma overall ($OR = 0.85$, $p = 3.96 \times 10^{-4}$) after multiple testing adjustment. This systematic investigation of the role of genetic determinants of viral antigen reactivity in glioma risk and survival provides insight into complex immunogenomic mechanisms of glioma pathogenesis. These results may inform applications of antiviral-based therapies in glioma treatment.

Introduction

Studies linking viruses and cancer date back over 100 years[1] and laid the foundation for understanding oncogenes.[2] It has also become increasingly clear, as well evidenced by the current pandemic, that host genetics play an important role in response to viruses.[3,4] To date, seven viruses have been accepted to be tumor-initiating in humans: Epstein-Barr virus (EBV), hepatitis B virus (HBV), human papillomavirus (HPV), human T-lymphotropic virus-1 (HTLV-1), hepatitis C (HCV), Kaposi's sarcoma herpesvirus (HHV-8), and Merkel cell polyomavirus (MCV).[5] Recent analyses have shown that infections account for approximately 13% of human cancers worldwide.[6] However viruses have not been definitively implicated in the etiology of glioma despite decades of suggestive associations.[7–11]

Glioma (MIM: 137800) is a highly fatal brain cancer with a paucity of known risk factors, and exposure to ionizing radiation is an accepted causal factor.[12] A history of previous infection with Varicella-Zoster virus (VZV) has been the only infectious agent consistently linked to adult glioma, conferring an estimated 20% decrease in risk.[10,13] A suite of other viruses have been associated with the risk and grade of glioma, including EBV, MCV, John Cunningham virus (JCV), BK virus (BKV), human Cytomegalovirus (CMV), and human herpesvirus-6 (HHV-6), but with discordant results.[11,14–16] The identification of key somatic molecular alterations (e.g., *IDH* mutation [MIM: 147700], 1p/19q chromosomal arm codeletion, *TERT* mutation [MIM :187270]), which drastically affect glioma prognosis, have uncovered subtype-specific risk factors.[17,18] Recently, several studies/clinical trials have suggested a prognostic benefit of antiviral medications in the treatment of glioma.[19–21]

Genetic host response to viral infection could play a role in elucidating potential links between viruses and cancer incidence and prognosis. Studies have demonstrated significant heritable components (32%–48%) of antibody response to many viruses and have identified genetic loci within host genes related to cell entry, cytokine production, and immune response.[22–24] Genetic variants of class I and II human leukocyte antigen (HLA) genes contribute the most important identified components of genetic determinants of response to viral antigens. Class II genes each encode half of a heterodimeric class II HLA protein, which presents extra-cellularly derived peptides to $CD4^+$

[1]Department of Neurological Surgery, University of California, San Francisco, San Francisco, CA, USA; [2]Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA, USA; [3]Department of Pediatrics, University of California, San Francisco, Oakland, CA, USA; [4]Institute of Human Genetics, University of California, San Francisco, San Francisco, CA, USA; [5]Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN, USA; [6]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA; [7]Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, USA

*Correspondence: geno.guerra@ucsf.edu (G.G.), stephen.francis@ucsf.edu (S.S.F.)

https://doi.org/10.1016/j.ajhg.2022.04.011.

helper T cells. The immune response is triggered when a $CD4^+$ T cell recognizes the combination of a class II HLA protein and its bound peptide. It is well studied that $CD4^+$ T cells have an important role in creating and sustaining effective anti-tumor immunity.[25]

The HLA region of the genome is considered the most polymorphic region of the human genetic system[26] where polymorphisms have been shown to alter the risk and progression of disease in a variety of autoimmune (notably HLA class II) and malignant conditions.[27,28] Certain HLA haplotypes have shown non-additive epistatic effects on glioma risk,[29] yet no germline variants within the HLA have been directly identified as risk loci in a glioma genome-wide association study (GWAS).

In this study, we leveraged previously published genome-wide SNP associations with viral antibody response[30] to generate genetically inferred antigen reactivity profiles in glioma cases and controls and evaluated their association with risk and survival by major glioma subtypes. We further conducted imputed HLA gene association analysis with glioma risk and survival and our findings suggest a convergence of genetic mechanisms regulating host immune response to viral challenge and glioma development and progression.

Subjects and methods

Ethics

Collection of affected individual samples and associated clinico-pathological information was undertaken with written informed consent and relevant ethical review board approval at the respective study centers in accordance with the tenets of the Declaration of Helsinki. Specifically, informed consent and ethical board approval was obtained from the UCSF Committee on Human Research (USA) and the Mayo Clinic Office for Human Research Protection (USA). The diagnosis of glioma (ICDO-3 codes 9380–9480 or equivalent) was established through histology in all cases in accordance with World Health Organization guidelines.

Study populations

We analyzed three glioma case-control sets assembled on the basis of genotyping platform and study population for a total sample size of 3,418 cases and 8,156 controls (Figure 1, Table 1). The first set included 1,973 cases from the Mayo Clinic and University of California San Francisco (UCSF) Adult Glioma Study and 1,859 controls from the Glioma International Case-Control Study (GICC) who were genotyped on the Illumina OncoArray, as previously described.[18,31–34] The second dataset (AGS-i370) included 659 cases and 586 controls from the UCSF Adult Glioma study genotyped on the Illumina HumanHap370duo panel.[32] The third dataset included 786 glioma cases from The Cancer Genome Atlas (TCGA) with available molecular data genotyped on the Affymetrix 6.0 array. Cancer-free controls were assembled from two Wellcome Trust Case Control Consortium (WTCCC) studies genotyped using the Affymetrix 6.0 array: 2,917 controls from the 1958 British Birth cohort and 2,794 controls from the UK Blood Service control group.

Molecular subtype information (*IDH* mutation and 1p/19q co-deletion status) was downloaded from Ceccarelli et al.[35] Table S1 for the third dataset and was provided directly from the UCSF AGS and Mayo Clinic for the first and second datasets.

Quality control and imputation

Standard quality control procedures were implemented prior to imputation. Analyses were restricted to individuals of predominantly (>70%) European ancestry, determined with ADMIXTURE[36] and the HapMap 3 reference populations. Within each ancestral group, we removed samples with excess heterozygosity (>3 standard deviations [SD] from mean), <95% call rates, and discordant self-reported and genetically inferred sex. Relatedness checks were performed within each study with KING[37] (kinship > 0.12), filtering out up to second-degree relations and retaining the samples with higher call rate. TCGA blood samples were preferentially chosen when both blood and tumor sequencing data were available for the same affected individual (693 European samples had both blood and tumor samples available). SNPs with <95% call rate were removed, along with variants deviating from Hardy-Weinberg equilibrium ($p < 10^{-6}$) or at a low minor allele frequency (MAF < 0.005). Samples genotyped on the same platform (i.e., Affymetrix 6.0 for TCGA and WTCCC) were imputed together with the multi-ethnic TOPMed reference panel (ver. r2).

Statistical analysis

Genetically predicted viral antigen response

We calculated genetic reactivity scores (GRSs) to obtain genetically inferred viral antigen response profiles in each of the glioma datasets. For each GRS, candidate variants and corresponding effect sizes were obtained from genome-wide summary statistics for seroreactivity to 12 viral antigens previously identified in Kachuri and Francis et al. (2020)[30] in 7,895 randomly selected individuals of European descent from the UK Biobank (UKB) cohort (aged 40–69) with serological measures of immunoglobulin G (IgG) antibody response,[38] a stable biomarker of lifetime exposure to common viruses. For each antigen, we preferentially selected independent SNPs (linkage disequilibrium, LD, $r^2 < 0.01$ within 500 kb) with the lowest p value among genome-wide-significant variants ($p < 5 \times 10^{-8}$). LD proxies ($r^2 > 0.9$) were obtained for variants unavailable in the target glioma datasets. For each individual, an antigen-specific GRS was calculated as a weighted sum with weights (β) corresponding to a standard deviation increase in antibody response:

$$GRS_{\text{Antigen}} = \beta_1 \text{xSNP}_1 + \ldots + \beta_k \text{xSNP}_k.$$

Each GRS was calculated with a minimum of four variants with MAF ≥ 0.01 and imputation quality ($R^2 > 0.8$) and then standardized within each dataset.

Antigen GRS associations with glioma susceptibility. GRS associations with disease risk were examined for glioma overall and for molecular subtypes defined by the specific tumor alterations: *IDH* mutation status and 1p/19q chromosomal arm codeletion. For each GRS, odds ratios (ORs) were estimated with logistic regression models adjusted for age, sex, and the first ten genetic ancestry principal components (PCs). Models in the UCSF-Mayo dataset were further adjusted for contributing site. GRS associations from each study were meta-analyzed with a fixed-effects inverse-variance-weighted approach. Heterogeneity in study-specific GRS associations was assessed with Cochran's Q test.

Antigen GRS associations with glioma survival. The association between each antigen-GRS and overall survival was assessed with a Cox proportional hazards regression model with follow-up time calculated from the date of first surgery to either date of death or last known contact. The latter censored at that date. Analyses were conducted for glioma overall and molecular subtypes with a minimum of 50 cases and 20 events (deaths). Proportionality

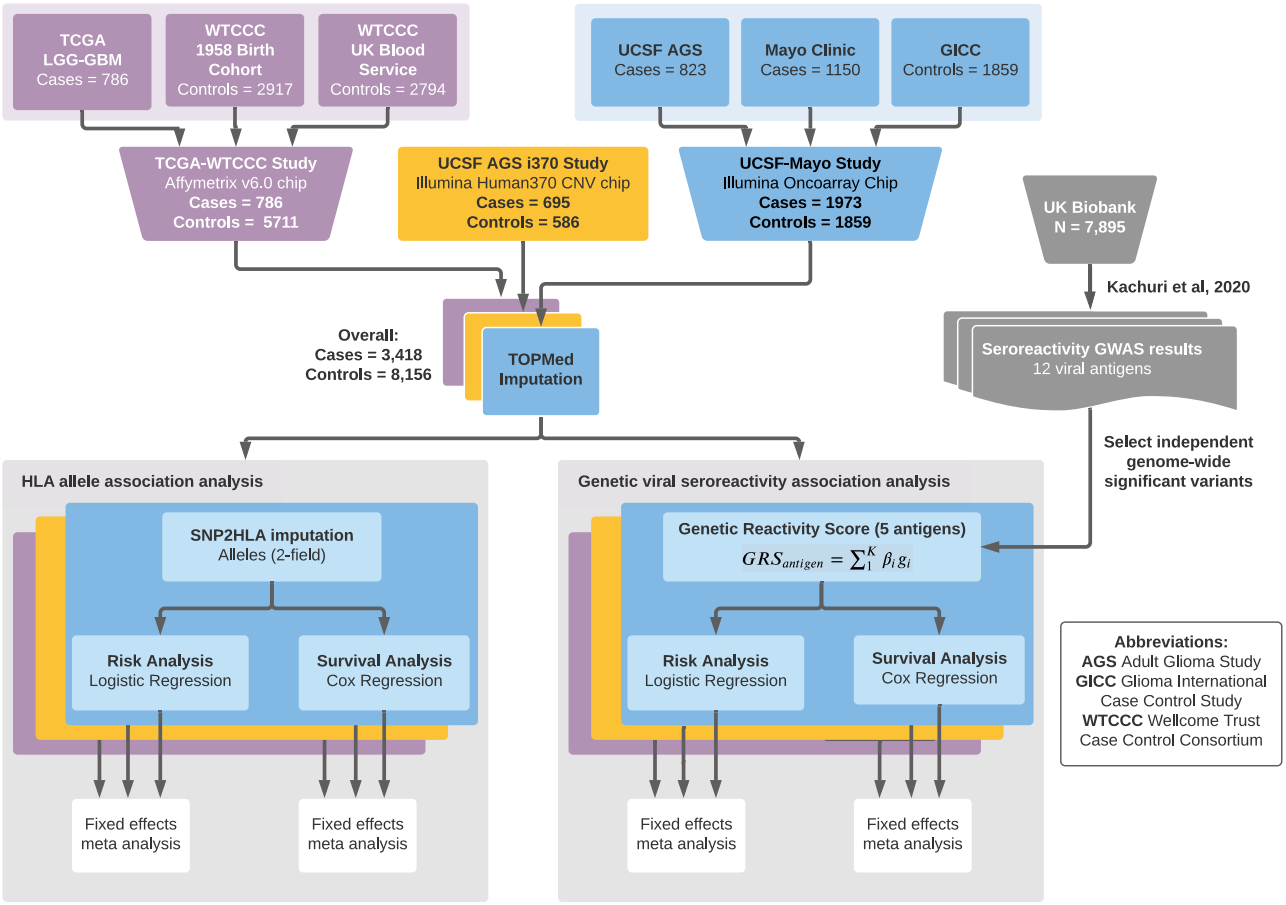

Figure 1. Summary of data processing and analysis

Our analysis consisted of three glioma case-control datasets. The first dataset (purple) included 786 glioma cases from The Cancer Genome Atlas (TCGA) genotyped on the Affymetrix 6.0 array and cancer-free controls assembled from two Wellcome Trust Case Control Consortium (WTCCC) studies genotyped with the Affymetrix 6.0 array: 2,917 controls from the 1958 British Birth cohort and 2,794 controls from the UK Blood Service control group. The second dataset (yellow) included 659 cases and 586 controls from the University of California, San Francisco (UCSF) Adult Glioma Study (AGS) genotyped on the Illumina HumanHap370duo panel.[32] The third set (blue) included 1,973 cases from the Mayo Clinic and UCSF AGS and 1,859 controls from the Glioma International Case-Control Study (GICC) who were genotyped on the Illumina OncoArray, as previously described.[18,31–34] The three resulting case-control datasets were processed through quality controls as described in the main text and imputed with the TOPMed imputation server. SNP2HLA was used to impute HLA alleles from SNP data. Risk and survival analyses were performed separately on each study's imputed HLA alleles, on multiple glioma molecular subtypes, and a fixed-effects meta-analysis was performed to aggregate results. Separately, genetic reactivity scores to five viral antigens were created with previously published GWAS data. For cases and controls across the three studies, a genetic reactivity score (GRS) for antibody response to each of the five antigens was calculated for each individual with the available sequencing data. Risk and survival analyses were performed separately on each study with each GRS as a predictor, with results aggregated via a fixed effects inverse-variance-weighted meta-analysis for each subset of glioma patients and controls based on molecular subtype.

assumptions were checked within each dataset via examination of Kaplan-Meier curves. Hazard ratios (HRs) were estimated with Cox models adjusted for age, sex, ten genetic PCs, and study site (if applicable). Associations with survival in each study were combined via fixed-effects meta-analysis. For each nominally significant ($p < 0.05$) glioma-GRS association, survival differences were further assessed with Kaplan-Meier curves by comparing mortality trajectories in affected individuals with high genetically predicted immune reactivity (top 20%) to the remainder.

Regional HLA analyses of glioma risk and survival. Classical HLA alleles were imputed for samples in all cohorts at two-field resolution with SNP2HLA[39] and the Type 1 Diabetes Genetics Consortium (T1DGC) reference panel of 2,767 unrelated individuals. Associations were tested for 77 alleles (MAF ≥ 0.01) with imputation quality > 0.4 across eight genes: *HLA-A* (MIM: 142800), *HLA-B* (MIM: 142830), *HLA-C* (MIM: 142840), *HLA-DPA1* (MIM: 142880), *HLA-DPB1* (MIM: 142858), *HLA-DQA1* (MIM: 146880), *HLA-DQB1* (MIM: 604305), and *HLA-DRB1*(MIM: 142857).

Subtype-specific associations with risk were estimated with logistic regression models. HLA allele associations with mortality were assessed with SPACox,[40] an extension of the Cox model with improved type I error control in high-dimensional settings. Study-specific risk and survival associations were combined in a meta-analysis. Associations for each HLA allele were considered statistically significant if $p < 6.5 \times 10^{-4}$ (0.05/77 alleles).

Results

The creation of a GRS was attempted for all antigens with genome-wide-significant seroreactivity-associated variants ($p < 5 \times 10^{-8}$) as reported in Kachuri and Francis et al.

Table 1. Clinical and molecular summary of the three case-control datasets

	UCSF-Mayo cases N = 1,973	GICC controls N = 1,859	AGS i370 cases N = 659	AGS i370 controls N = 586	TCGA cases N = 786	WTCCC controls N = 5,711
Age						
<40	983 (50%)	538 (29%)	101 (15%)	85 (15%)	213 (27%)	0 (0%)
40–59	476 (24%)	555 (30%)	329 (50%)	252 (43%)	279 (36%)	0 (0%)
≥60	514 (26%)	766 (41%)	229 (35%)	249 (42%)	245 (31%)	0 (0%)
Missing	0 (0%)	0 (0%)	0 (0%)	0 (0%)	49 (6%)	5,711 (100%)
Sex						
Female	822 (42%)	744 (40%)	229 (35%)	280 (48%)	328 (42%)	2,819 (49%)
Male	1,151 (58%)	1,115 (60%)	430 (65%)	306 (52%)	458 (58%)	2,892 (51%)
IDH mutation status						
Mutant	588 (30%)	N/A	111 (17%)	N/A	375 (48%)	N/A
Wild type	699 (35%)	N/A	416 (63%)	N/A	364 (46%)	N/A
Missing	686 (35%)	N/A	132 (20%)	N/A	47 (6%)	N/A
Molecular subtype based on IDH mutation and 1p/19q codeletion						
MT-codel	244 (12%)	N/A	9 (1%)	N/A	143 (18%)	N/A
MT-noncodel	291 (15%)	N/A	94 (14%)	N/A	230 (29%)	N/A
WT-noncodel	507 (26%)	N/A	416 (63%)	N/A	357 (46%)	N/A
Missing/other	905 (47%)	N/A	140 (22%)	N/A	56 (7%)	N/A

N/A, not applicable to controls.

(2020).[30] This included 12 antigens: four EBV antigens (EA-D, EBNA, p18, and ZEBRA) and antigens for BKV, HHV7, HSV1, JCV, MCV, VZV, CMV, and HHV6. An antigen-specific GRS was considered successfully created if at least four independent SNPs passed LD clumping and thresholding and were present (or via proxy SNP) in the three glioma study datasets. GRSs were successfully generated for EBV EA-D, EBV EBNA, EBV p18, EBV ZEBRA, and MCV. Only one independent SNP remained for each of HSV1, BKV, JCV, and VZV after accounting for long-range LD in the HLA region, although numerous SNPs met genome-wide significance in the previous seroreactivity study. Three SNPs remained for HHV7, each on a different chromosome (6, 11, and 17). CMV and HHV6 each had only one significantly associated SNP in the previously published GWAS and therefore a GRS was not considered for either. Complete information of variants considered for all 12 viral antigens, including nearest gene, allele frequencies, and corresponding glioma risk associations, are reported in Table S1. Allele frequencies of all considered SNPs were similar between our datasets and the UK Biobank cohort (maximum frequency difference = 0.0401). LD correlations for all associated chromosome 6 variants are available in a heatmap in Figure S1.

The variants included in each of the five created GRSs were overwhelmingly located in the HLA region, and few predictors were located elsewhere across the genome: rs67886110 in 3q25.1 an eQTL for *MED12L* (MIM: 611318) and *P2RY12* (MIM: 600515) (EBV EBNA), rs7618405 in 3p24.3 (MCV), and rs7444313 in 5q31.2 near *TMEM173* (MIM: 612374) (MCV). Of the 38 SNPs included across the five GRSs, two had a significant association (Bonferroni-corrected: $p < 1.3 \times 10^{-3}$) with overall glioma risk: rs9265517 in *HLA-B* (EBV EBNA), $p = 3.08 \times 10^{-4}$ and rs9268847 near *HLA-DRB9* (MCV), $p = 2.86 \times 10^{-4}$. Figure S2 visualizes the correlation between the effect-increasing GRS alleles and each imputed HLA allele (two field resolution) from UCSF-Mayo cases and controls. Notably, the GRS for MCV was inversely correlated with GRSs for EBV ZEBRA (Pearson's $r = -0.28$, $p = 3.89 \times 10^{-70}$) and EBV EBNA ($r = -0.19$, $p = 6.28 \times 10^{-31}$) (Figure 2).

Viral antigen GRS associations with glioma risk

In the combined meta-analysis of three case-control studies, three GRSs reached at least nominal significance ($p < 0.05$), and some associations remained statistically significant after correction for the number of antigens tested (Bonferroni: $p < 0.05/5 = 0.01$) (Figure 3). Genetic predisposition to an increased serological response to EBV ZEBRA was inversely associated with risk of glioma overall (per 1 SD increase in GRS: odds ratio, $OR_{ZEBRA} = 0.94$, 95% confidence interval 0.89–0.99, $p = 0.012$, 3,418 cases). GRS_{EBNA} was associated with an increased risk of *IDH* mutated gliomas ($OR_{EBNA} = 1.09$, 1.004–1.18, $p = 0.040$, 1,074 cases) and the magnitude of this effect persisted for *IDH* mutated 1p/19q codeleted gliomas ($OR_{EBNA} = 1.14$, 1.012–1.28, $p = 0.031$, 396 cases).

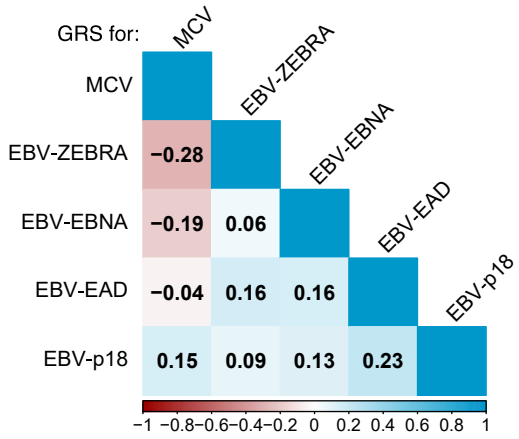

Figure 2. GRS correlations within the UCSF-Mayo cases and controls

Pearson correlations between genetically predicted antigen responses (via GRS) as computed in the UCSF-Mayo glioma cases and controls. Values were printed in each block if and only if the associated correlation test p value was less than 0.01.

We observed some evidence of antagonistic pleiotropy between genetic determinants of antibody response to EBV ZEBRA and MCV, which generalized across glioma subtypes. GRS_{ZEBRA} and GRS_{MCV} were associated with susceptibility to *IDH* wild-type gliomas, but in opposite directions: higher genetically predicted reactivity to EBV ZEBRA was inversely associated with *IDH* wild type glioma risk (OR_{ZEBRA} = 0.91, 0.85–0.98, p = 0.0072, 1,479 cases), while increased predicted antibody response to MCV conferred an increased risk (OR_{MCV} = 1.09, 1.02–1.17, p = 0.013). This pattern persisted for *IDH* wild-type 1p/19q non-codeleted gliomas (OR_{ZEBRA} = 0.91, 0.84–0.9, p = 0.0099; OR_{MCV} = 1.11, 1.03–1.19, p = 0.0054, 1,280 cases). When GRS_{ZEBRA} and GRS_{MCV} were tested together along with interaction term in a single logistic model, we observed no significant interaction (interaction term p = 0.34 in *IDH* wild-type risk). The general correlation between GRS_{ZEBRA} and GRS_{MCV} (Pearson's r = −0.28, p = 3.9×10^{-70} in UCSF-Mayo, Figure 2) suggests a possible shared underlying genetic mechanism between the response to the two antigens. This genetic correlation is most likely driven by the significant inverse relationship of the effect increasing allele of rs9268847 near *HLA-DRB9* (MCV) with both rs9274728 near *HLA-DQB1* (EBV ZEBRA) (Pearson's r = −0.38, p = 2.6×10^{-134}) and rs7757696 near *HLA-DQA1* (EBV ZEBRA) (r = −0.23, p = 3.1×10^{-45}). The odds ratios were robust to the level of genetic ancestry controlled for, as measured by the number of genetic principal components included in the model (Table S6). Results reported here were adjusted for the first ten PCs, but analysis was conducted in parallel with 15 and 20 PCs.

Viral antigen GRS-survival associations

The number of available cases and deaths across the three studies is available in Table 2. Associations between genetically predicted viral antigen response profiles and survival were restricted to *IDH* mutated gliomas (Figure 4). GRS_{EBNA} was associated with survival in 1,074 *IDH* mutated glioma cases (per 1 SD increase: hazard ratio, HR = 0.86, 0.76–0.96, p = 0.010, 325 events), suggesting that a higher genetically predicted reactivity to EBV EBNA improved duration of survival.

GRS for two EBV antigens were nominally associated with survival amongst 244 *IDH* mutated 1p/19q codeleted glioma cases, but in opposite directions (HR_{EBNA} = 0.75, 0.57–0.99, p = 0.048/HR_{ZEBRA} = 1.27, 1.01–1.6, p = 0.042, 64 events). This subtype-specific result was limited to the UCSF-Mayo dataset due to insufficient number of reported events (deaths) in *IDH* mutated 1p/19q codeleted glioma cases from TCGA and AGS-i370.

GRS_{EBNA} was also associated with improved survival outcomes in 614 *IDH* mutated 1p/19q non-codeleted glioma cases (HR = 0.75, 0.74–0.997, p = 0.045, 222 events), suggesting the association of GRS_{EBNA} is independent of 1p/19q status. Figure 5 visualizes the significant GRS_{EBNA} associations with Kaplan-Meier survival curves.

We did not detect any significant GRS associations in the analysis of glioma overall, which is consistent with the strong prognostic significance of molecular glioma subtypes.[17,33] Also, as the prognosis in *IDH* wild-type gliomas is the poorest, we suspect our GRS instruments are underpowered to detect significant deviations over such short intervals. Full survival results are available in Figures S3 and S4.

HLA allele-glioma associations

Of the 77 HLA alleles imputed at two-field resolution, only *HLA-DQA1*03:01* reached Bonferroni-corrected significance ($p < 6.5 \times 10^{-4}$ after correcting for 77 HLA alleles tested) for association with glioma risk in a meta-analysis. The presence of an *HLA-DQA1*03:01* allele was associated with a decrease in overall glioma risk (OR = 0.84, p = 3.96×10^{-4}). This allele was also nominally associated with risk of *IDH* wild-type (OR = 0.82, p = 1.4×10^{-3}) and *IDH* wild-type 1p/19q non-codeleted gliomas (OR = 0.81, p = 2.6×10^{-3}). These subtype-specific associations and direction of effect mirror those of GRS_{ZEBRA} presented above, suggesting *HLA-DQA1*03:01* as a potential shared marker for both glioma risk and EBV ZEBRA seroreactivity. Results of all HLA allele (one and two field resolution) associations with glioma risk tested in a meta-analysis are available in Table S2. Odds ratios for the association of *HLA-DQA1*03:01* were slightly attenuated when adding further genetic principal components to the model (Table S6).

Discussion

We investigated genetically predicted antibody response to twelve antigens for nine viruses in relation to glioma risk and prognosis in a meta-analysis of three patient cohorts with molecular subtype information. Our analysis discovered evidence that genetic predisposition to reactivity to specific viral antigens associated with susceptibility to glioma in a subtype-specific manner. This pattern of effect

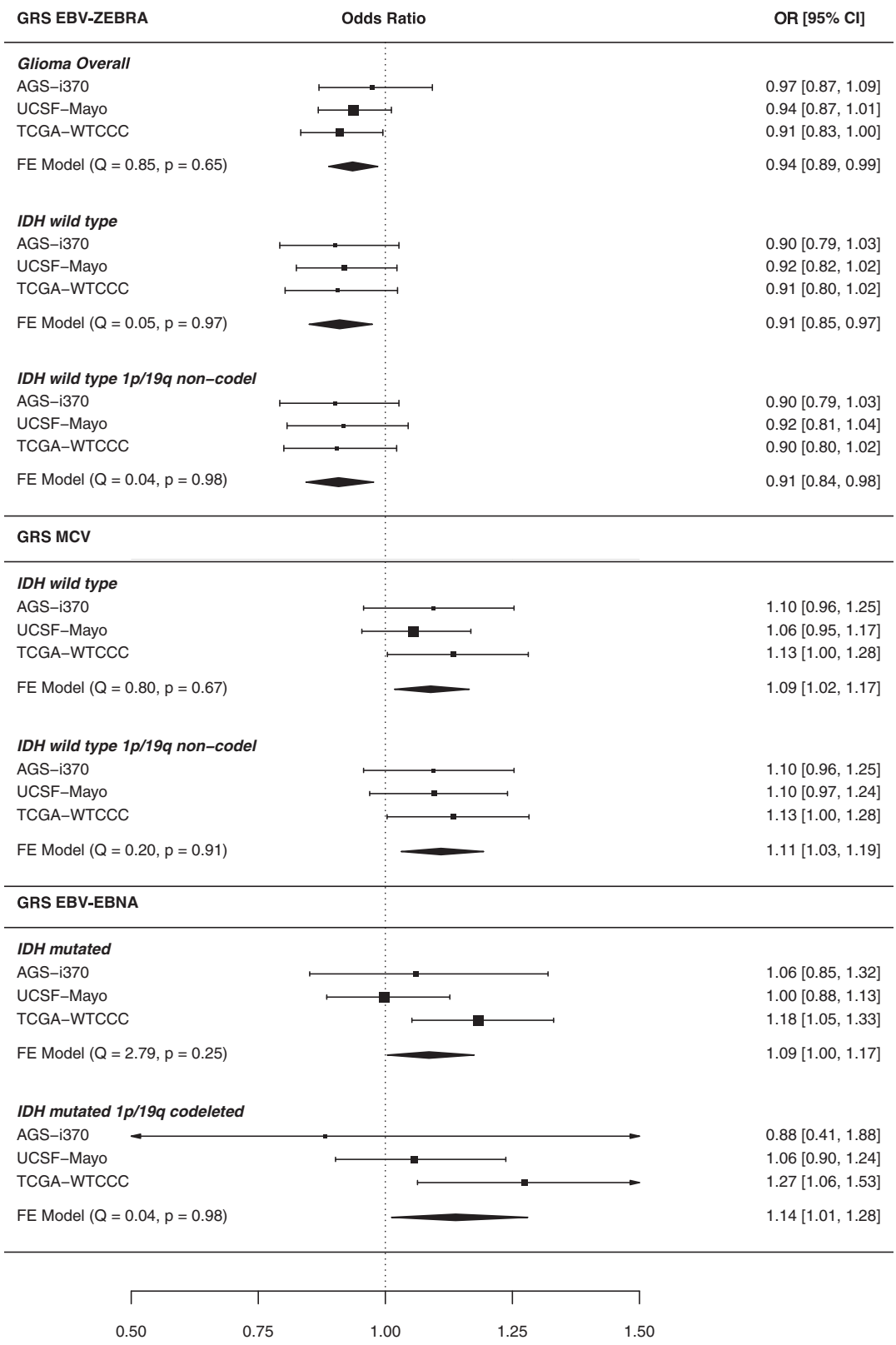

Figure 3. Significant GRS-glioma subtype risk association meta-analysis forest plots

Forest plot meta-analysis results of GRS-glioma risk associations that were at least nominally statistically significant ($p < 0.05$). Response to antigens EBV ZEBRA, MCV, and EBV EBNA had associations that reached this threshold. Results are reported as odds ratios along with 95% confidence intervals. Briefly, each header indicates the studied viral antigen GRS, within are its association with molecular glioma subtypes reported with $p < 0.05$ and the 95% confidence interval of each study-specific effect. The diamond visualizes the 95% confidence interval for the fixed effect (FE) meta-analysis across all three studies. Each meta-analysis was tested for between-study heterogeneity (Q statistic), and $p < 0.05$ indicates evidence of study-specific associations.

Table 2. Summary of available cases and events for survival analyses

	UCSF-Mayo		TCGA		AGS i370	
Molecular subtype	**cases**	**events**	**cases**	**events**	**cases**	**events**
Glioma	1,973	1,218	786	310	659	592
IDH mutated[a]	588	201	375	50	111	74
1p/19q codeleted[b]	244	64	143	13	9	1
1p/19q non-codeleted[b]	291	117	230	36	94	69
IDH wild type[a]	699	594	364	228	416	402
1p/19q non-codeleted[c]	507	458	357	224	416	402

Study/subtype combinations with less than 50 cases or 20 events were not utilized in the meta-analysis of survival associations.
[a]Further subset of Glioma.
[b]Further subset of IDH mutated.
[c]Further subset of IDH wild type.

modification by glioma subtype also extends into prognosis, where we observed associations between viral antigen GRSs and survival among specific glioma subtypes.

One of our main findings is that genetic predisposition to increased seroreactivity to EBV ZEBRA was associated with a decreased overall glioma risk, with a significant decrease in *IDH* wild-type subtypes. Predicted reactivity to the MCV VP1 antigen mirrored the same *IDH* wild-type associations but in the opposite direction, where higher reactivity was associated with increased glioma risk. The significant inverse relationship between predicted reactivity to EBV ZEBRA and MCV VP1 highlights the possibility of shared genetic components of antibody response to the two antigens. We saw evidence that this relationship is driven by SNPs near class II HLA alleles. Understanding underlying mechanisms guided by these genetics is left as an open question. One possible link is that the class II HLA allele *DQA1*03:01* was associated with decreased glioma risk in the same subtypes associated with GRSs for EBV ZEBRA. In our previous UKB analysis,[30] the presence of *HLA-DQA1*03:01* was positively associated with EBV ZEBRA seroreactivity measurements ($\beta = 0.168$, $p = 1.3 \times 10^{-16}$). Taken together, the associations between *HLA-DQA1*03:01*, EBV ZEBRA, and glioma risk suggest possible shared underlying immunogenetic architecture. As HLA class II proteins present potentially antigenic peptides, functional genetic alterations can result in changes to the binding affinity of specific antigens, modulating the potential for recognition by $CD4^+$ helper T cells. It is possible that the heterodimeric DQ proteins half-encoded by *DQA1*03:01* have improved binding or recognition of peptides presented by both EBV ZEBRA proteins and glioma (specifically *IDH* wild type), allowing for an increased immune response in both cases. As *IDH* wild type has generally been shown to serve as a marker for more severe glioma cases, the exact somatic tumor alterations leading to recognized peptide variants in these tumors is not clear and is an area that warrants future investigation. Further analysis of viral-tumor protein homology is needed to understand if this could be a possible connection.

Interestingly, a higher genetically predicted response to EBV EBNA was nominally associated with increased risk in *IDH* mutated/1p/19q codeleted gliomas. Still, reactivity toward EBV EBNA was more strongly associated with improved survival in *IDH* mutated gliomas. This discrepancy between the disease-promoting and pro-survival associations of GRS_{EBNA} may suggest the presence of different biologic pathways after initiation of disease. It may also point toward EBV latency-mediated treatment effects. Previous studies show that the EBV latency type predicts the relative amount of induced reactivation generated by cytotoxic chemotherapy drugs.[41] Therefore, individuals who react strongly to EBV EBNA antigens may exhibit a different pattern of reactivation when treated with temozolomide. Further research is required to elucidate this putative association, yet it is clear that EBV lytic/latent cycling is a critical aspect of other germline interactions and disease risk.

Epstein-Barr virus was the first recognized human oncovirus[42] and exists in two distinct life cycles within a host: a lytic phase of active infection where new viruses are produced and a latent phase where the virus remains dormant to avoid detection by the host. Distinct sets of antigens are produced during these two phases, two of such are EBNA, during the latent (inactive) phase of infection, and the ZEBRA protein, which initiates a change from latent to lytic gene expression. Past evidence suggests processes during the latent stage are responsible for the virus' oncogenic properties via mechanisms promoting cell growth and preventing cell apoptosis (reviewed in Akhtar et al., 2018[43]). Recent work has also implicated the ZEBRA protein as oncogenic with evidence demonstrating its ability to deregulate immune surveillance and promote immune escape.[44]

In contrast, MCV is the most recently recognized human oncovirus, first described in 2008 and later accepted as a causal agent for Merkel cell carcinoma,[45] a neuroendocrine carcinoma. Previous studies have identified the presence of MCV DNA in glioma patients and have drawn an association with infection and increased risk of glioma.[15,16] It has

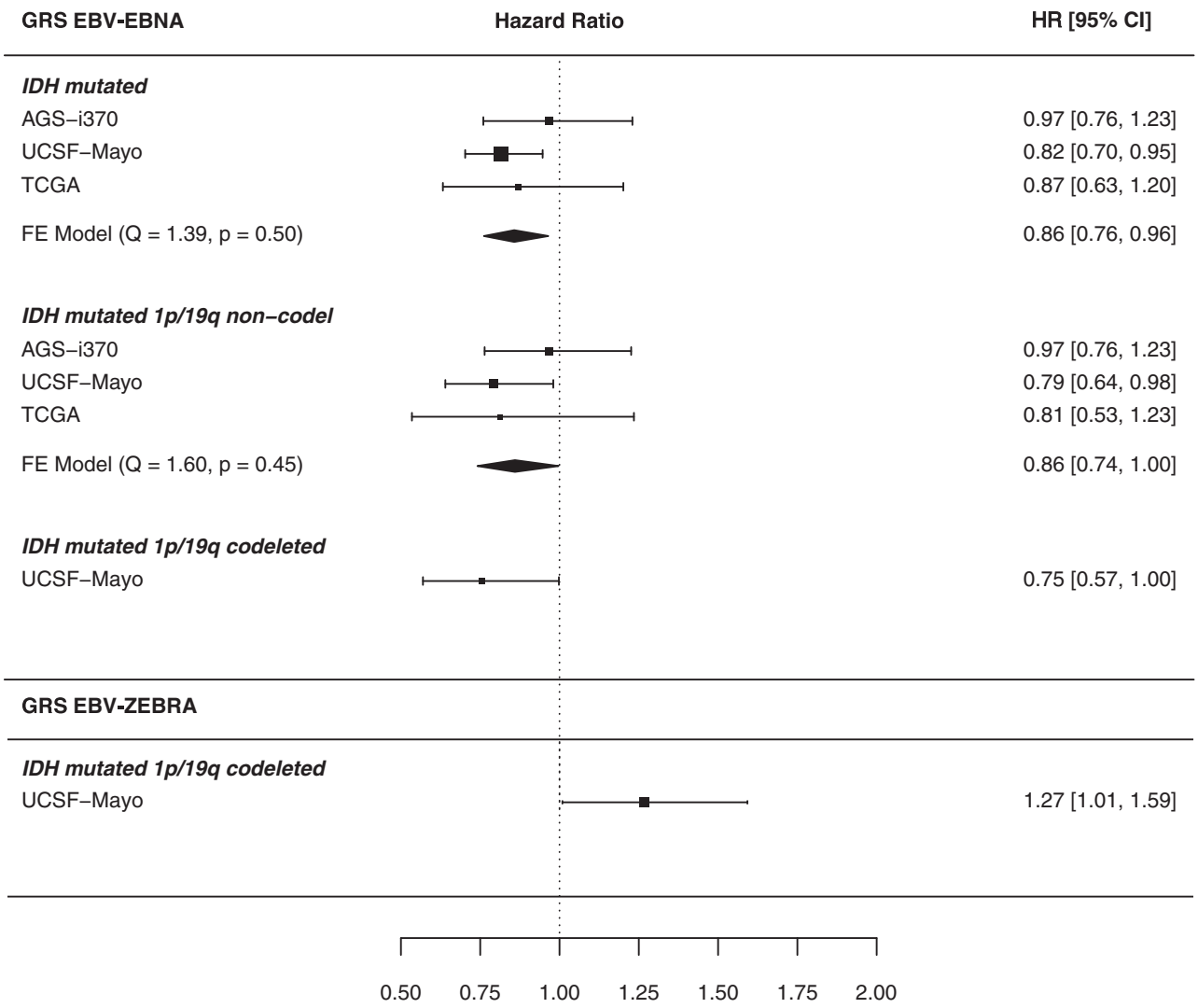

Figure 4. Nominal GRS-glioma subtype survival association meta-analysis forest plots
Forest plot meta-analysis results of GRS-glioma survival associations that were at least nominally statistically significant ($p < 0.05$). Genetically inferred response to antigens EBV ZEBRA, MCV, and EBV EBNA had associations that reached this threshold. Results are reported as hazard ratios along with 95% confidence intervals. Briefly, each header indicates the studied viral antigen GRS, within are its association with molecular glioma subtypes reported with $p < 0.05$ and the 95% confidence interval of each study-specific effect. The diamond visualizes the 95% confidence interval for the fixed effect (FE) meta-analysis across included studies. Studies that had an insufficient number of cases/events in a subtype were not included in the meta-analysis. Each meta-analysis (where more than one study was included) was tested for between-study heterogeneity (Q statistic), and $p < 0.05$ indicates evidence of study-specific associations.

been shown that both the large T and small T antigens of MCV are oncoproteins that target tumor suppressor proteins such as pRB. Although the exact latency mechanism in polyomaviruses is unknown, it has been suggested that complex protein-mediated latency may be critical to the MCV life cycle.[46]

In interpreting our findings, several limitations should be acknowledged. The UKB GWAS used to develop antigen GRSs was restricted to individuals of predominantly European ancestry and we would expect low portability to other populations considering the high level of polymorphism in the HLA region. Therefore, our analyses of glioma susceptibility were also restricted to European ancestry subjects. Furthermore, sample size in other ancestry groups was insufficient for a meta-analysis (African: 73 cases/29 controls, Asian: 48/49, American: 24/24, admixed: 131/68).

The shared genetic architecture between the viruses studied here, as seen in the correlations in Figures 2 and S1, may result in a lack of specificity. This limitation in our study may be applicable in other viruses that share the same underlying genetic programming of antigen response, particularly VZV, which is associated with a unique pattern of LD spanning a large region of the HLA.[30] Furthermore, the clumping and thresholding approach to GRS development may not be optimal for regions with complex LD structure, such as HLA. An approach that can appropriately account for the long-range correlation structure and capture non-linear interactions, such as haplotype effects, may improve future

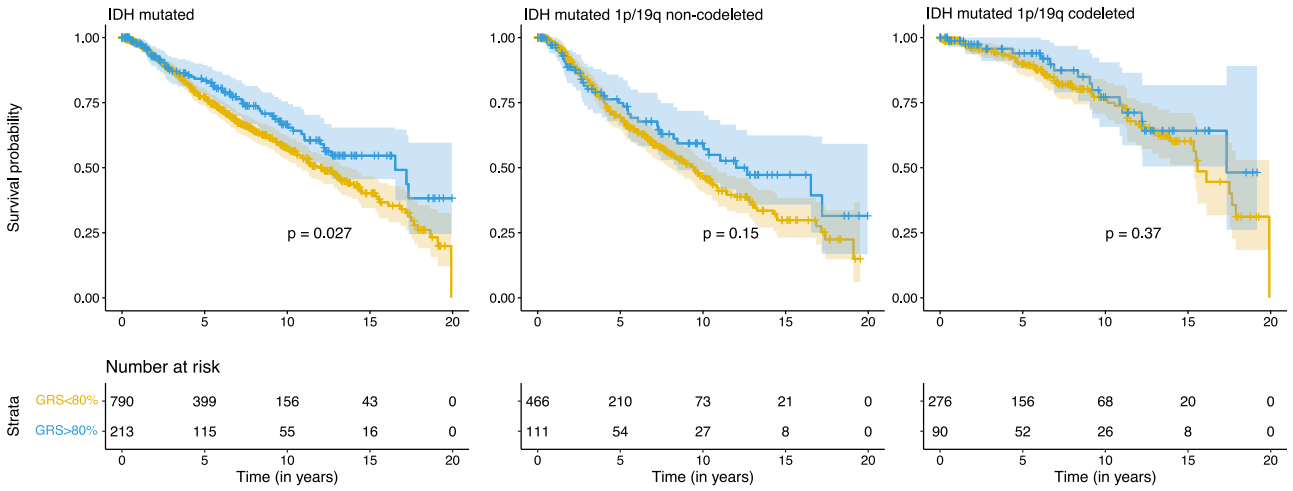

Figure 5. Kaplan Meier curves for significant GRS_{EBNA}-glioma molecular subtype associations
Kaplan-Meier curves for subtypes where GRS_{EBNA} was nominally associated with subtype-specific glioma survival differences ($p < 0.05$ via Cox regression). The second and third plots are distinct partitions of the IDH-mutated subgroup. To visualize, unnormalized GRSs across the included studies were binned on the basis of the case-specific 80^{th} percentile score in the UCSF-Mayo dataset. p values included on each plot are results of a log-rank test for difference between the two curves. Below each set of curves provides the number of cases surviving beyond that time point for each of the two GRS groups. In all cases, the glioma cases with higher GRS for EBV EBNA had visually improved survival outcomes compared to the bottom 80%.

studies that use genetically inferred immune responses. Similarly, the complex LD patterns present challenges in disentangling the specific function of individual SNPs located within the HLA region.

The magnitude and direction of associations with glioma risk observed for GRS and HLA alleles were robust to the number of genetic ancestry principal components included in the models. Including up to 20 principal components did not result in an appreciable change in effect size for GRS associations. The OR for *HLA-DQA1*03:01* increased slightly from 0.82 to 0.89 (Table S6) and become less precise. Analyses of the HLA region are complicated by the highly polymorphic HLA variation that is driven by demographic factors associated with fine-scale geographic variation.[47] Additional studies leveraging targeted HLA sequencing approaches can provide further clarity into the genetic determinants of antigen response and glioma risk/survival.

We could only study the single SNP-glioma relationships (Table S1) for variants associated with reactivity to several antigens of interest (BKV, HSV-1, JCV, VZV, HHV-7, HHV-6, CMV) because of the limited number reaching genome-wide significance.[30] Although the GRS-glioma association results presented here require further replication in independent affected individual populations, our findings are intriguing and suggest previous inconsistencies in observational epidemiologic viral association studies may be partially due to individual differences in genetic factors that affect antigen reactivity. Further, our approach focused on the genetic underpinnings of antibody response in the adaptive immune response to common viral infections as measured by IgG levels, which is only one pathway utilized by the adaptive immune system. Another arm of the adaptive immune system, T cell-mediated immunity, is not accounted for in this study. The construction of genetic instruments that capture the role of this arm in viral response would complement the current study well. Differences in the innate immune system, such as the efficiency of natural killer cells, could also provide additional associations between viral infections and glioma. Lastly, there is inherent noise in the seroreactivity GWAS used here, as the time between infection and IgG measurement was unknown in the UK Biobank cohort. As we restricted our analysis to common viruses, it is however a safe assumption that most people were exposed early in life (before the age of 20), and all viruses we considered set up some form of permanent latency in their hosts.

This study directly examined the underlying genetic architecture of antigen reactivity to common viruses and glioma risk and survival. We observed important associations between programmed reactivity to viruses and glioma etiology and prognosis. This methodology is not limited to the study of glioma, as the GRS instruments proposed here can readily be applied to cohorts of any cancer type. This offers a unique approach for future studies to reinvestigate the genetic contributions of long-running epidemiologic associations between viruses and cancer and possibly clarify effects of viral therapies in their treatment.

Data and code availability

Genotype data of control samples from the 1958 British Birth Cohort and UK Blood Service Control Group were made available from the Wellcome Trust Case Control Consortium (WTCCC) and downloaded from the European Genotype Archive under accession numbers WTCCC2: EGAD00000000021 and WTCCC2: EGAD00000000023, respectively. Genotype data of glioma cases from The Cancer Genome Atlas (TCGA) were obtained from

Database of Genotypes and Phenotypes (dbGaP) (dbGaP: phs000178). Genotype data of glioma samples from Mayo Clinic and control samples from the Glioma International Case Control Study (GICC) are available from dbGaP under accession dbGaP: phs001319.v1.p1. Genotype data from the University of California San Francisco Adult Glioma Study (AGS) are available under dbGap: phs001497.v2.p1. The code supporting the current study have not been deposited in a public repository because no novel methods have been developed but are available from the corresponding author on request.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.ajhg.2022.04.011.

Acknowledgments

Work at University of California, San Francisco was supported by the National Institutes of Health (grant numbers T32CA151022, K99CA246076, R01CA52689, P50CA097257, R01CA126831, R01CA139020, R01AI128775, and R25CA112355), the National Brain Tumor Foundation, the Stanley D. Lewis and Virginia S. Lewis Endowed Chair in Brain Tumor Research, the Robert Magnin Newman Endowed Chair in Neuro-oncology, and by donations from families and friends of John Berardi, Helen Glaser, Elvera Olsen, Raymond E. Cooper, and William Martinusen. This publication was supported by the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through UCSF-CTSI grant number UL1 RR024131. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. The collection of cancer incidence data used in this study was supported by the California Department of Public Health pursuant to California Health and Safety Code Section 103885; Centers for Disease Control and Prevention's (CDC) National Program of Cancer Registries, under cooperative agreement 5NU58DP006344; the National Cancer Institute's Surveillance, Epidemiology, and End Results Program under contract HHSN261201800032I awarded to the University of California, San Francisco, contract HHSN261201800015I awarded to the University of Southern California, and contract HHSN261201800009I awarded to the Public Health Institute, Cancer Registry of Greater California. The ideas and opinions expressed herein are those of the author(s) and do not necessarily reflect the opinions of the State of California, Department of Public Health, the National Cancer Institute, and the Centers for Disease Control and Prevention or their contractors and subcontractors.

Author contributions

G.G., L.K., and S.S.F. conceived of the study and wrote main drafts of the manuscript. G.G., L.K., G.W., S.J.M., N.K., and S.S.F. conducted and advised on informatic and statistical analyses along with result interpretations. H.M.H., A.M.M., T.R., P.B., J.K.W., J.E.E., R.B.J., and M.W. were involved in primary data collection. All authors contributed and reviewed the final manuscript.

Declaration of interests

The authors declare no competing interests.

Received: November 29, 2021
Accepted: April 18, 2022
Published: May 11, 2022

References

1. Rous, P. (1910). A transmissible avian neoplasm. (Sarcoma of the common fowl.). J. Exp. Med. *12*, 696–705. https://doi.org/10.1084/jem.12.5.696.
2. Duesberg, P.H., and Vogt, P.K. (1970). Differences between the ribonucleic acids of transforming and nontransforming avian tumor viruses. Proc. Natl. Acad. Sci. U S A *67*, 1673–1680. https://doi.org/10.1073/pnas.67.4.1673.
3. COVID-19 Host Genetics Initiative. Mapping the human genetic architecture of COVID-19. Nature. ;600(7889):472-477Published online July 8, 2021. doi:10.1038/s41586-021-03767-x
4. Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., et al. (2015). UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. *12*, e1001779. https://doi.org/10.1371/journal.pmed.1001779.
5. Mui, U.N., Haley, C.T., and Tyring, S.K. (2017). Viral oncology: molecular biology and pathogenesis. J. Clin. Med. *6*, E111. https://doi.org/10.3390/jcm6120111.
6. de Martel, C., Georges, D., Bray, F., Ferlay, J., and Clifford, G.M. (2020). Global burden of cancer attributable to infections in 2018: a worldwide incidence analysis. Lancet Glob. Health *8*, e180–e190. https://doi.org/10.1016/S2214-109X(19)30488-7.
7. Neves, A.M., Thompson, G., Carvalheira, J., Trindade, J.C., Rueff, J., Caetano, J.M., Casey, J.W., and Hermouet, S. (2008). Detection and quantitative analysis of human herpesvirus in pilocytic astrocytoma. Brain Res. *1221*, 108–114. https://doi.org/10.1016/j.brainres.2008.05.009.
8. Chauvin, C., Suh, M., Remy, C., and Benabid, A.L. (1990). Failure to detect viral genomic sequences of three viruses (herpes simplex, simian virus 40 and adenovirus) in human and rat brain tumors. Ital. J. Neurol. Sci. *11*, 345–357. https://doi.org/10.1007/BF02335937.
9. Cobbs, C.S., Harkins, L., Samanta, M., Gillespie, G.Y., Bharara, S., King, P.H., Nabors, L.B., and Britt, W.J. (2002). Human cytomegalovirus infection and expression in human malignant glioma. Cancer Res. *62*, 3347–3350.
10. Wrensch, M., Weinberg, A., Wiencke, J., Masters, H., Miike, R., Barger, G., and Lee, M. (1997). Does prior infection with varicella-zoster virus influence risk of adult glioma? Am. J. Epidemiol. *145*, 594–597. https://doi.org/10.1093/oxfordjournals.aje.a009155.
11. Strojnik, T., Duh, D., and Lah, T.T. (2017). Prevalence of neurotropic viruses in malignant glioma and their oncomodulatory potential. In Vivo *31*, 221–230. https://doi.org/10.21873/invivo.11049.
12. Bondy, M.L., Scheurer, M.E., Malmer, B., Barnholtz-Sloan, J.S., Davis, F.G., Il'yasova, D., Kruchko, C., McCarthy, B.J., Rajaraman, P., Schwartzbaum, J.A., et al.; On behalf of the Brain Tumor Epidemiology Consortium (2008). Brain tumor epidemiology: consensus from the brain tumor epidemiology Consortium. Cancer *113*, 1953–1968. https://doi.org/10.1002/cncr.23741.
13. Amirian, E.S., Scheurer, M.E., Zhou, R., Wrensch, M.R., Armstrong, G., Lachance, D., Olson, S.H., Lau, C.C., Claus, E.B., Barnholtz-Sloan, J., et al. (2016). History of chickenpox in

glioma risk: a report from the glioma international case–control study (GICC). Cancer Med. *5*, 1352–1358. https://doi.org/10.1002/cam4.682.
14. Kofman, A., Marcinkiewicz, L., Dupart, E., Lyshchev, A., Martynov, B., Ryndin, A., Kotelevskaya, E., Brown, J., Schiff, D., and Abounader, R. (2011). The roles of viruses in brain tumor initiation and oncomodulation. J. Neurooncol. *105*, 451–466. https://doi.org/10.1007/s11060-011-0658-6.
15. Limam, S., Missaoui, N., Bdioui, A., Yacoubi, M.T., Krifa, H., Mokni, M., and Selmi, B. (2020). Investigation of simian virus 40 (SV40) and human JC, BK, MC, KI, and Wu polyomaviruses in glioma. J. Neurovirol. *26*, 347–357. https://doi.org/10.1007/s13365-020-00833-4.
16. Egan, K.M., Kim, Y., Bender, N., Hodge, J.M., Coghill, A.E., Smith-Warner, S.A., Rollison, D.E., Teras, L.R., Grimsrud, T.K., and Waterboer, T. (2021). Prospective investigation of polyomavirus infection and the risk of adult glioma. Sci. Rep. *11*, 9642. https://doi.org/10.1038/s41598-021-89133-3.
17. Molinaro, A.M., Taylor, J.W., Wiencke, J.K., and Wrensch, M.R. (2019). Genetic and molecular epidemiology of adult diffuse glioma. Nat. Rev. Neurol. *15*, 405–417. https://doi.org/10.1038/s41582-019-0220-2.
18. Eckel-Passow, J.E., Drucker, K.L., Kollmeyer, T.M., Kosel, M.L., Decker, P.A., Molinaro, A.M., Rice, T., Praska, C.E., Clark, L., Caron, A., et al. (2020). Adult diffuse glioma GWAS by molecular subtype identifies variants in *D2HGDH* and *FAM20C*. Neuro-Oncol. *22*, 1602–1613. https://doi.org/10.1093/neuonc/noaa117.
19. Volpin, F., Casaos, J., Sesen, J., Mangraviti, A., Choi, J., Gorelick, N., Frikeche, J., Lott, T., Felder, R., Scotland, S.J., et al. (2017). Use of an anti-viral drug, Ribavirin, as an anti-glioblastoma therapeutic. Oncogene *36*, 3037–3047. https://doi.org/10.1038/onc.2016.457.
20. Stragliotto, G., Pantalone, M.R., Rahbar, A., Bartek, J., and Söderberg-Naucler, C. (2020). Valganciclovir as add-on to standard therapy in glioblastoma patients. Clin. Cancer Res. *26*, 4031–4039. https://doi.org/10.1158/1078-0432.CCR-20-0369.
21. Piper, K., Foster, H., Gabel, B., Nabors, B., and Cobbs, C. (2019). Glioblastoma mimicking viral encephalitis responds to acyclovir: a case series and literature review. Front. Oncol. *9*, 8. https://doi.org/10.3389/fonc.2019.00008.
22. Besson, C., Amiel, C., Le-Pendeven, C., Plancoulaine, S., Bonnardel, C., Ranque, B., Abbed, K., Brice, P., Ferme, C., Carde, P., et al. (2009). Strong correlations of anti–viral capsid antigen antibody levels in first-degree relatives from families with epstein-barr virus–related lymphomas. J. Infect. Dis. *199*, 1121–1127. https://doi.org/10.1086/597424.
23. Kenney, A.D., Dowdle, J.A., Bozzacco, L., McMichael, T.M., St Gelais, C., Panfil, A.R., Sun, Y., Schlesinger, L.S., Anderson, M.Z., Green, P.L., et al. (2017). Human genetic determinants of viral diseases. Annu. Rev. Genet. *51*, 241–263. https://doi.org/10.1146/annurev-genet-120116-023425.
24. Venkataraman, T., Valencia, C., Mangino, M., Morgenlande, W., Clipman, S.J., Liechti, T., Valencia, A., Christofidou, P., Spector, T., Roederer, M., et al. (2021). Antiviral antibody epitope selection is a heritable trait. Genetics. https://doi.org/10.1101/2021.03.25.436790.
25. Tay, R.E., Richardson, E.K., and Toh, H.C. (2021). Revisiting the role of CD4+ T cells in cancer immunotherapy—new insights into old paradigms. Cancer Gene Ther. *28*, 5–17. https://doi.org/10.1038/s41417-020-0183-x.
26. Shiina, T., Hosomichi, K., Inoko, H., and Kulski, J.K. (2009). The HLA genomic loci map: expression, interaction, diversity and disease. J. Hum. Genet. *54*, 15–39. https://doi.org/10.1038/jhg.2008.5.
27. Bateman, A.C., and Howell, W.M. (1999). Human leukocyte antigens and cancer: is it in our genes? J. Pathol. *188*, 231–236. https://doi.org/10.1002/(SICI)1096-9896(199907)188:3<231::AID-PATH325>3.0.CO;2-A.
28. Gebe, J.A., Swanson, E., and Kwok, W.W. (2002). HLA Class II peptide-binding and autoimmunity: gebe et al : HLA Class II peptide-binding and autoimmunity. Tissue Antigens *59*, 78–87. https://doi.org/10.1034/j.1399-0039.2002.590202.x.
29. Zhang, C., de Smith, A.J., Smirnov, I.V., Wiencke, J.K., Wiemels, J.L., Witte, J.S., and Walsh, K.M. (2017). Non-additive and epistatic effects of HLA polymorphisms contributing to risk of adult glioma. J. Neurooncol. *135*, 237–244. https://doi.org/10.1007/s11060-017-2569-7.
30. Kachuri, L., Francis, S.S., Morrison, M.L., Wendt, G.A., Bosse, Y., Cavazos, T.B., Rashkin, S.R., Ziv, E., and Witte, J.S. (2020). The landscape of host genetic factors involved in immune response to common viral infections. Genome Med. *12*, 93. https://doi.org/10.1186/s13073-020-00790-x.
31. Melin, B.S., Barnholtz-Sloan, J.S., Wrensch, M.R., Johansen, C., Il'yasova, D., Kinnersley, B., Ostrom, Q.T., Labreche, K., Chen, Y., Armstrong, G., et al.; GliomaScan Consortium (2017). Genome-wide association study of glioma subtypes identifies specific differences in genetic susceptibility to glioblastoma and non-glioblastoma tumors. Nat. Genet. *49*, 789–794. https://doi.org/10.1038/ng.3823.
32. Wrensch, M., Jenkins, R.B., Chang, J.S., Yeh, R.F., Xiao, Y., Decker, P.A., Ballman, K.V., Berger, M., Buckner, J.C., Chang, S., et al. (2009). Variants in the CDKN2B and RTEL1 regions are associated with high-grade glioma susceptibility. Nat. Genet. *41*, 905–908. https://doi.org/10.1038/ng.408.
33. Eckel-Passow, J.E., Lachance, D.H., Molinaro, A.M., Walsh, K.M., Decker, P.A., Sicotte, H., Pekmezci, M., Rice, T., Kosel, M.L., Smirnov, I.V., et al. (2015). Glioma groups based on 1p/19q, *IDH* , and *TERT* promoter mutations in tumors. N. Engl. J. Med. *372*, 2499–2508. https://doi.org/10.1056/NEJMoa1407279.
34. Jenkins, R.B., Wrensch, M.R., Johnson, D., Fridley, B.L., Decker, P.A., Xiao, Y., Kollmeyer, T.M., Rynearson, A.L., Fink, S., Rice, T., et al. (2011). Distinct germ line polymorphisms underlie glioma morphologic heterogeneity. Cancer Genet. *204*, 13–18. https://doi.org/10.1016/j.cancergencyto.2010.10.002.
35. Ceccarelli, M., Barthel, F.P., Malta, T.M., Sabedot, T.S., Salama, S.R., Murray, B.A., Morozova, O., Newton, Y., Radenbaugh, A., Pagnotta, S.M., et al. (2016). Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. Cell *164*, 550–563. https://doi.org/10.1016/j.cell.2015.12.028.
36. Alexander, D.H., and Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. BMC Bioinformatics *12*, 246. https://doi.org/10.1186/1471-2105-12-246.
37. Manichaikul, A., Mychaleckyj, J.C., Rich, S.S., Daly, K., Sale, M., and Chen, W.M. (2010). Robust relationship inference in genome-wide association studies. Bioinforma Oxf Engl. *26*, 2867–2873. https://doi.org/10.1093/bioinformatics/btq559.
38. Mentzer, A.J., Brenner, N., Allen, N., Littlejohns, T.J., Chong, A.Y., Cortes, A., Almond, R., Hill, M., Sheard, S., McVean, G., et al. (2019). Identification of Host-Pathogen-Disease Relationships Using a Scalable Multiplex Serology Platform in

UK Biobank (Infectious Diseases (except HIV/AIDS)). https://doi.org/10.1101/19004960.

39. Jia, X., Han, B., Onengut-Gumuscu, S., Chen, W.M., Concannon, P.J., Rich, S.S., Raychaudhuri, S., and de Bakker, P.I. (2013). Imputing amino acid polymorphisms in human leukocyte antigens. J. Tang, ed. *8*, e64683. https://doi.org/10.1371/journal.pone.0064683.
40. Bi, W., Fritsche, L.G., Mukherjee, B., Kim, S., and Lee, S. (2020). A fast and accurate method for genome-wide time-to-event data analysis and its application to UK Biobank. Am. J. Hum. Genet. *107*, 222–233. https://doi.org/10.1016/j.ajhg.2020.06.003.
41. Phan, A.T., Fernandez, S.G., Somberg, J.J., Keck, K.M., and Miranda, J.L. (2016). Epstein–Barr virus latency type and spontaneous reactivation predict lytic induction levels. Biochem. Biophys. Res. Commun. *474*, 71–75. https://doi.org/10.1016/j.bbrc.2016.04.070.
42. Epstein, M.A., Achong, B.G., and Barr, Y.M. (1964). Virus particles in cultured lymphoblasts from burkitt's lymphoma. The Lancet *283*, 702–703. https://doi.org/10.1016/S0140-6736(64)91524-7.
43. Akhtar, S., Vranic, S., Cyprian, F.S., and Al Moustafa, A.E. (2018). Epstein–barr virus in gliomas: cause, association, or artifact? Front. Oncol. *8*, 123. https://doi.org/10.3389/fonc.2018.00123.
44. Germini, D., Sall, F.B., Shmakova, A., Wiels, J., Dokudovskaya, S., Drouet, E., and Vassetzky, Y. (2020). Oncogenic properties of the EBV ZEBRA protein. Cancers *12*, 1479. https://doi.org/10.3390/cancers12061479.
45. Chang, Y., and Moore, P.S. (2012). Merkel cell carcinoma: a virus-induced human cancer. Annu. Rev. Pathol. *7*, 123–144. https://doi.org/10.1146/annurev-pathol-011110-130227.
46. Kwun, H.J., Chang, Y., and Moore, P.S. (2017). Protein-mediated viral latency is a novel mechanism for Merkel cell polyomavirus persistence. Proc. Natl. Acad. Sci. U S A *114*, E4040–E4047. https://doi.org/10.1073/pnas.1703879114.
47. Sanchez-Mazas, A., Buhler, S., and Nunes, J.M. (2013). A new HLA map of europe: regional genetic variation and its implication for peopling history, disease-association studies and tissue transplantation. Hum. Hered. *76*, 162–177. https://doi.org/10.1159/000360855.