The American Journal of Human Genetics, Volume *109*

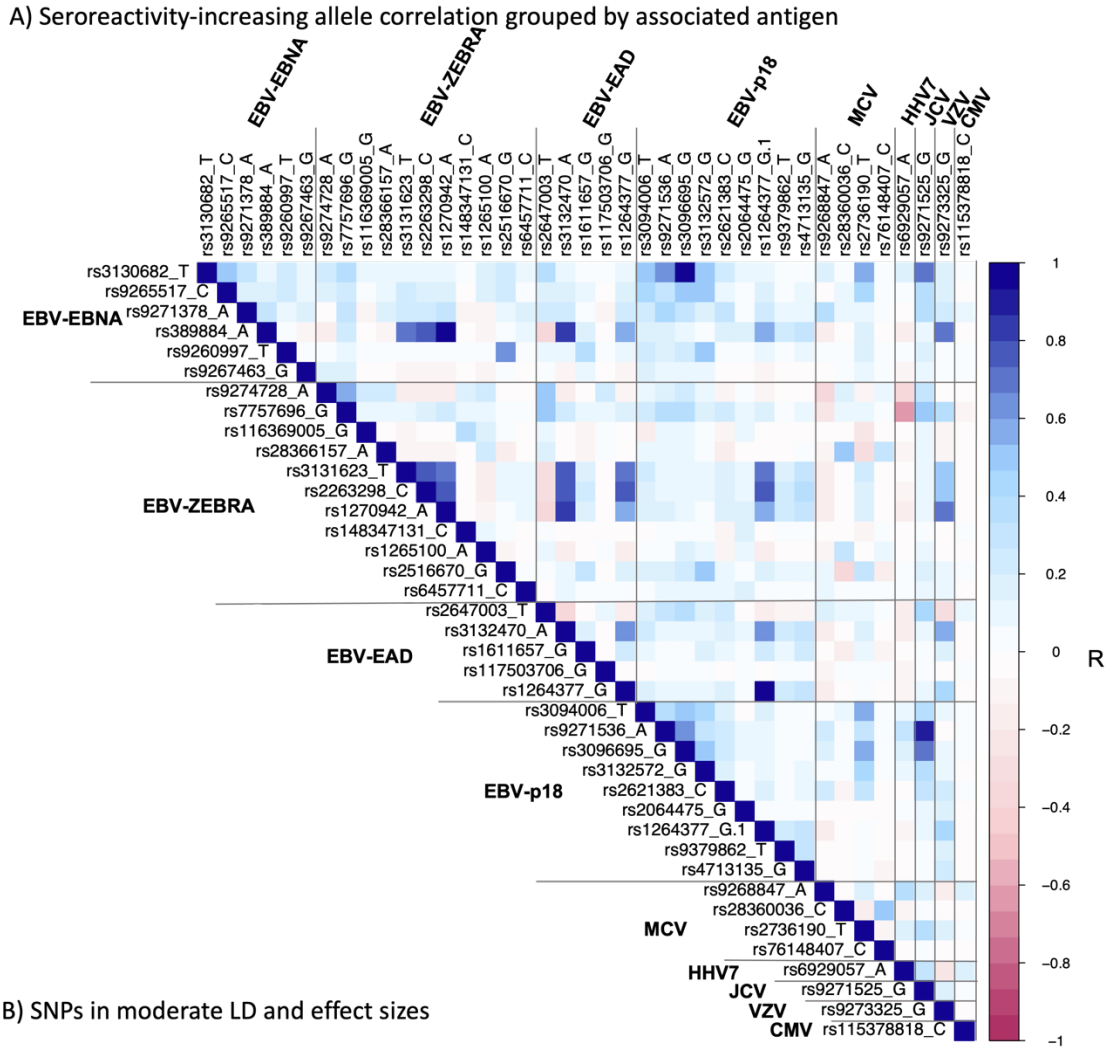
Supplemental information

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

Geno Guerra, Linda Kachuri, George Wendt, Helen M. Hansen, Steven J. Mack, Annette M. Molinaro, Terri Rice, Paige Bracci, John K. Wiencke, Nori Kasahara, Jeanette E. Eckel-Passow, Robert B. Jenkins, Margaret Wrensch, and Stephen S. Francis

Figure S1: Correlation of chromosome 6 antigen associated SNPs

A) Seroreactivity-increasing allele correlation grouped by associated antigen

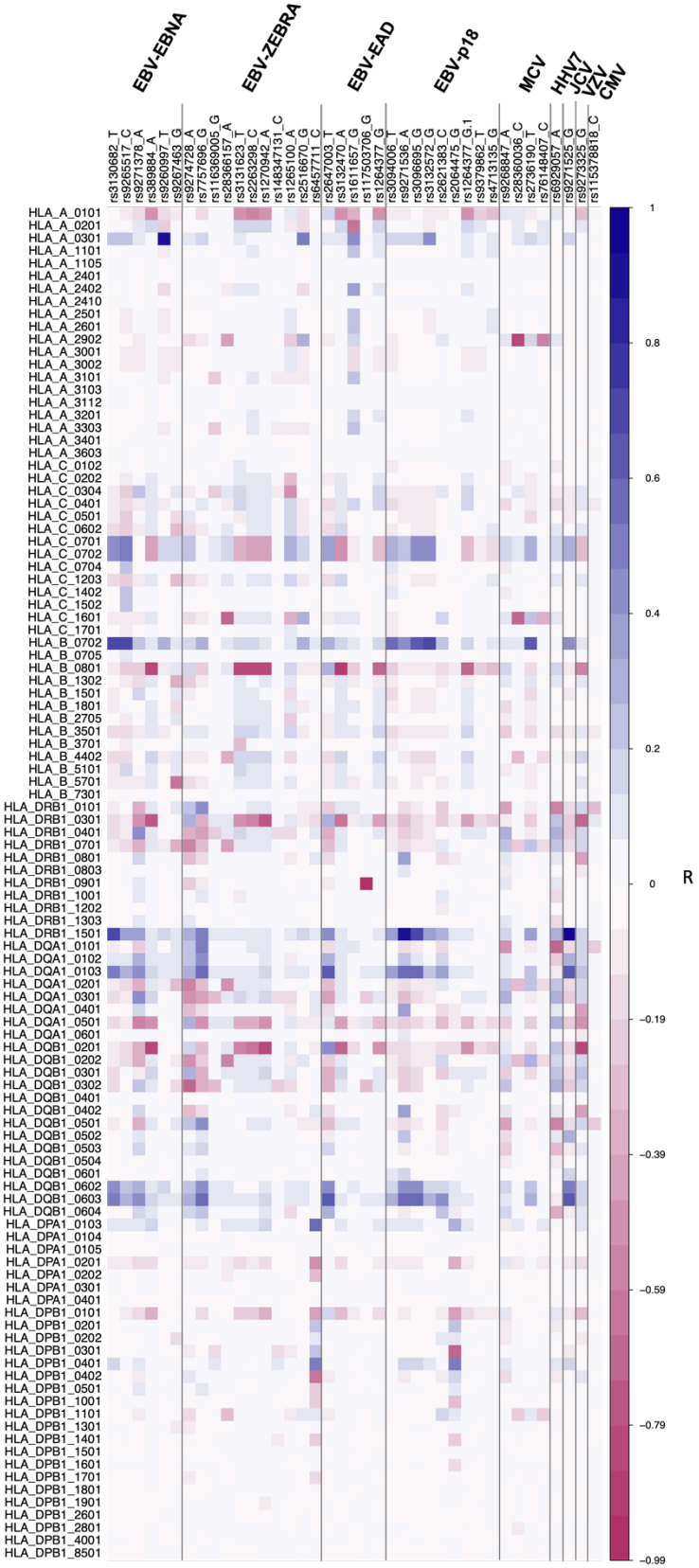


B) SNPs in moderate LD and effect sizes

Exposure 1	SNP	Effect	Correlated Allele	R^2	Correlated Allele	Effect	SNP	Exposure 2
EBV-EBNA	rs3130682	0.022	T	0.8769	G	0.207	rs3096695	EBV-p18
EBV-EBNA	rs389884	0.157	A	0.671	A	0.159	rs3132470	EBV-EAD
EBV-EBNA	rs389884	0.157	A	0.985	A	0.154	rs1270942	EBV-ZEBRA
EBV-ZEBRA	rs3131623	0.167	T	0.609	A	0.159	rs3132470	EBV-EAD
EBV-ZEBRA	rs2263298	0.154	C	0.556	G	0.119	rs1264377	EBV-EAD
EBV-ZEBRA	rs1270942	0.154	A	0.659	A	0.159	rs3132470	EBV-EAD
EBV-ZEBRA	rs2263298	0.154	C	0.556	G	0.118	rs1264377	EBV-p18
EBV-p18	rs9271536	0.200	A	0.611	G	-0.318	rs9271525	JCV

A) Heatmap of correlation values (R) of all LD-pruned seroreactivity-increasing alleles on chromosome 6 considered in our analysis, separated by the corresponding significantly associated antigen. Correlations were calculated using samples of European genetic descent from the UCSF-Mayo case and control dataset and the allele specified by each rsid suffix. B) Further detail into all pairs of SNPs in moderate LD (R^2>0.5 within 500kb) on chromosome 6, as estimated from European populations using LDlink, including the correlated allele pair and their effects on the associated exposure antigens, as previously calculated in Kachuri and Francis et al, 2020. Note that rs1264377_G is a shared effect increasing allele for both EBV -p18 and -EAD.

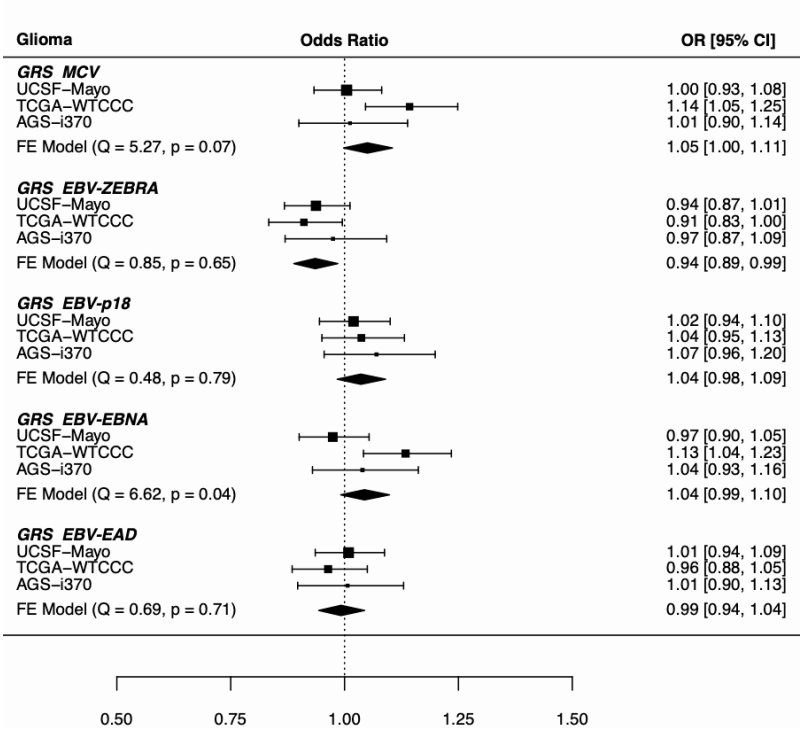
Figure S2: Correlation matrix of chr 6 GRS SNPs and two-field HLA alleles



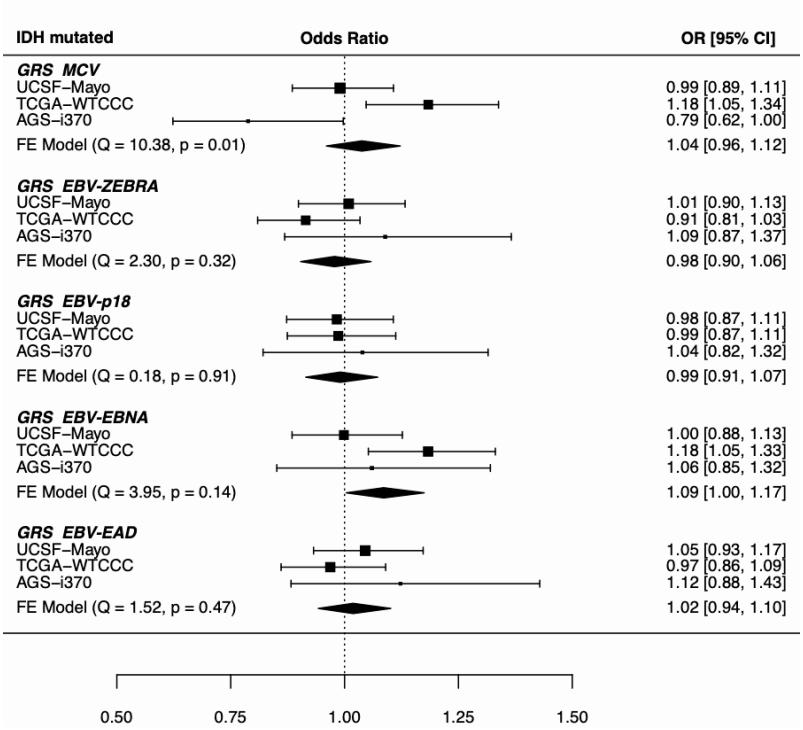
Heatmap of correlation values (R) of all LD-pruned seroreactivity-increasing alleles on chromosome 6 considered in our GRS analysis with the presence of each HLA allele. SNP alleles are partitioned by the corresponding significantly associated antigen, with imputed HLA alleles (at two-field resolution). R was calculated using European cases and controls from the UCSF-Mayo dataset. Cells with darker shades indicate higher levels of correlation.

Figure S3: All GRS-Glioma risk meta-result forest plots, by antigen

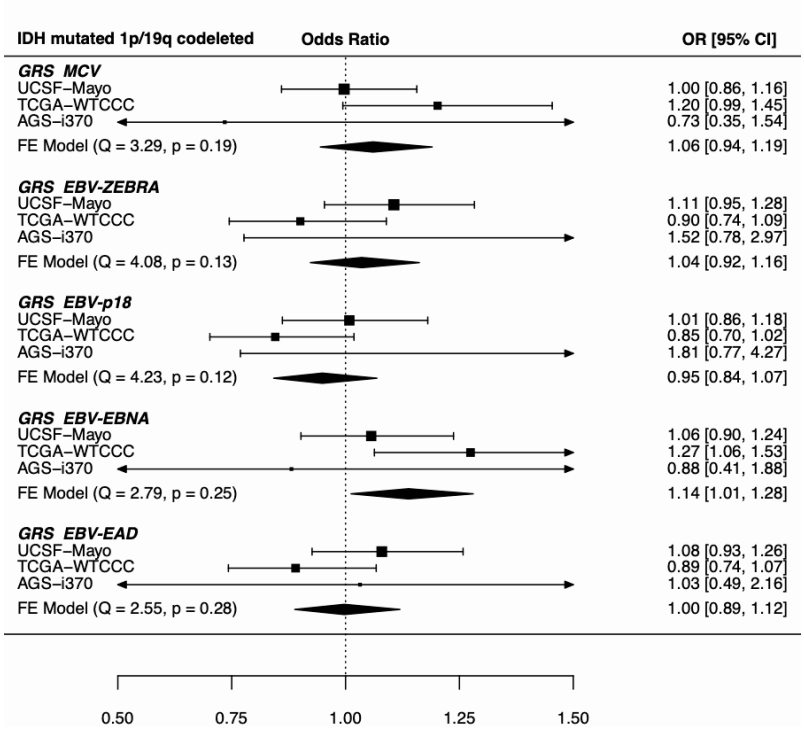
S3 A) Glioma overall



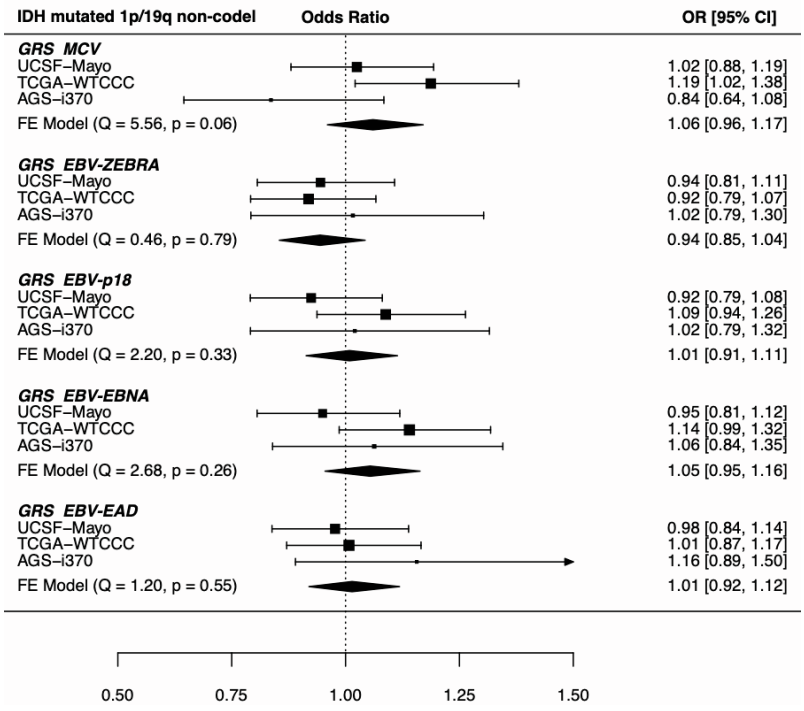
S3 B) *IDH* mutated



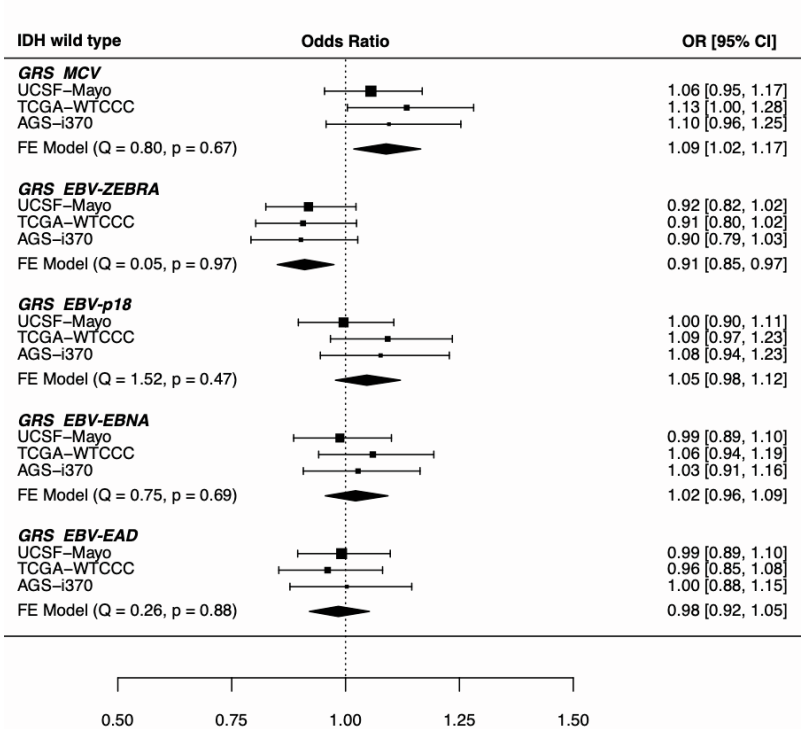
S3 C) *IDH* mutated 1p/19q codeleted



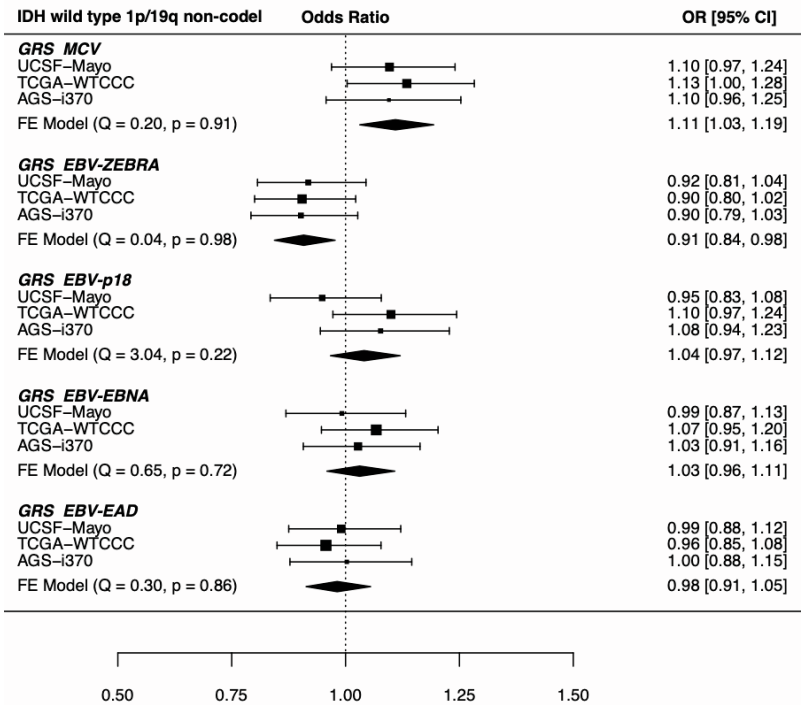
S3 D) *IDH* mutated 1p/19q non-codeleted



S3 E) *IDH* wild type



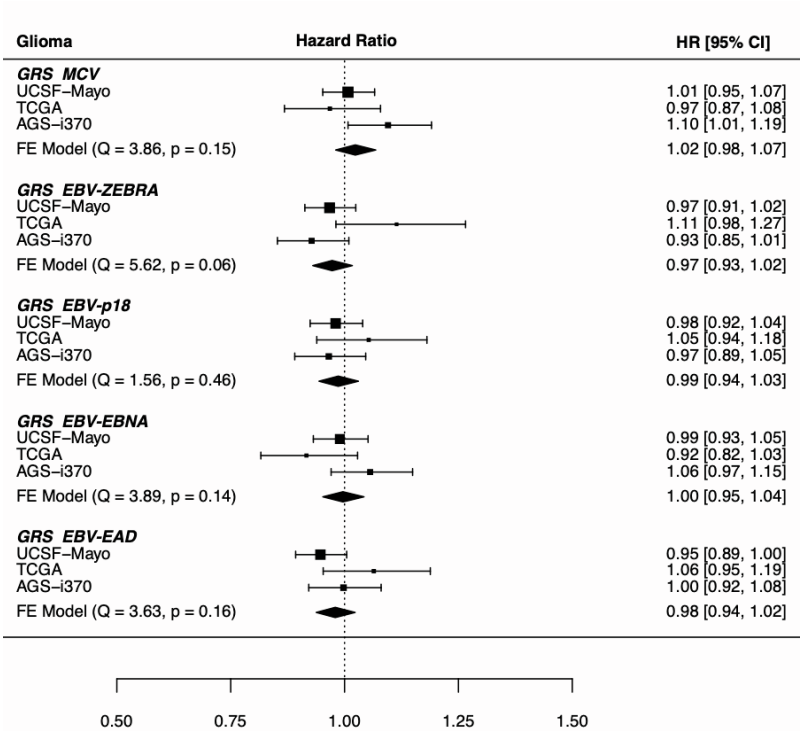
S3 F) *IDH* wild type 1p/19q non-codeleted



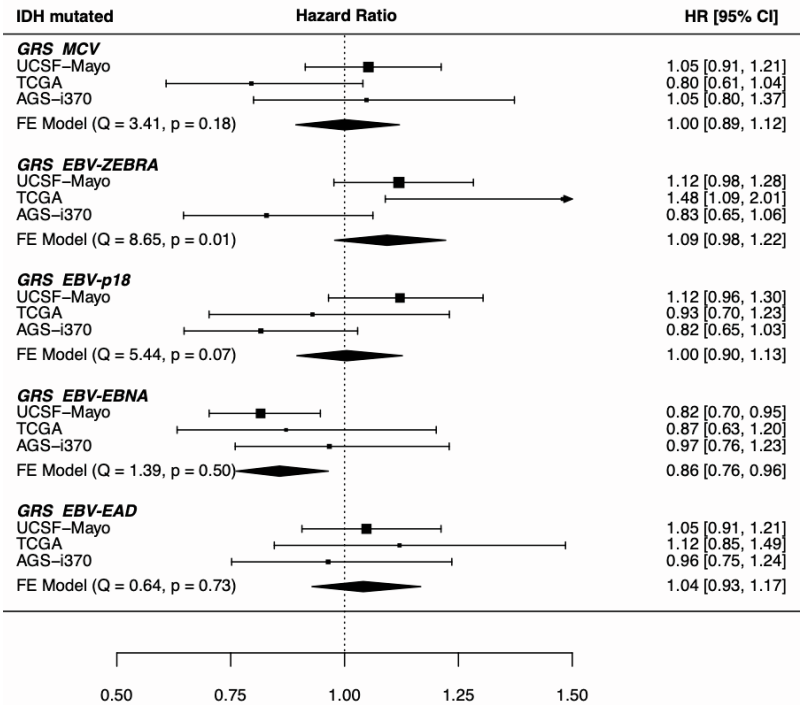
A-F) Forest plot meta-analysis results of all tested GRS-glioma risk associations for the specific primary subtype. Results are reported as odds ratios, along with 95% confidence intervals. Briefly, each header indicates the studied glioma molecular subtype, within are each GRSs associations with the subtype, 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 studies. Each meta-analysis was tested for between-study heterogeneity (Q statistic), with $p<0.05$ indicating evidence of study-specific associations. GRS-subtype tests with evidence of significant heterogeneity were re-analyzed using a random-effects meta-analysis (results not reported). No new suggestive associations were found after re-analysis.

Figure S4: All GRS-Glioma survival meta result forest plots, by antigen

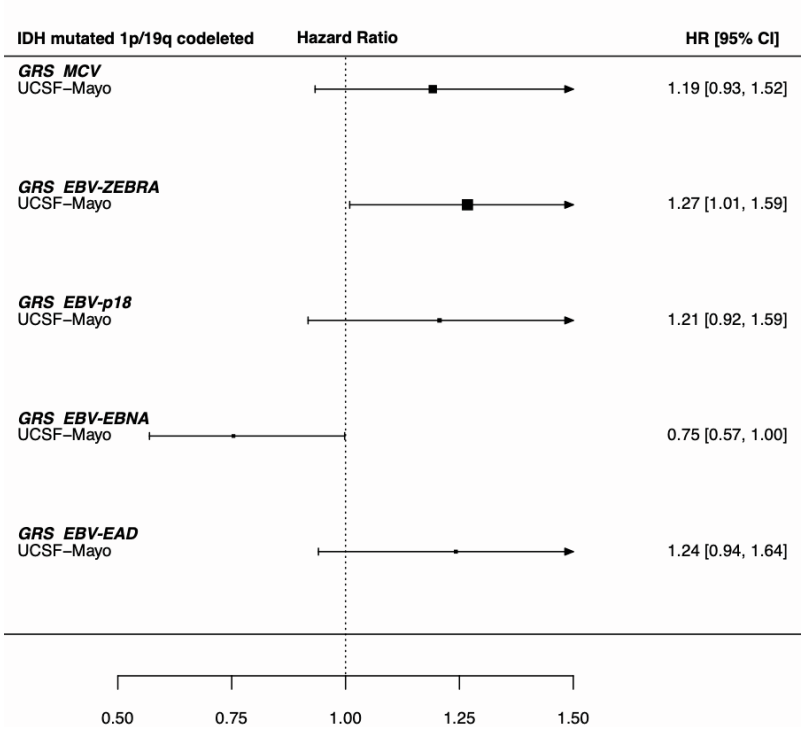
S4 A) Glioma overall



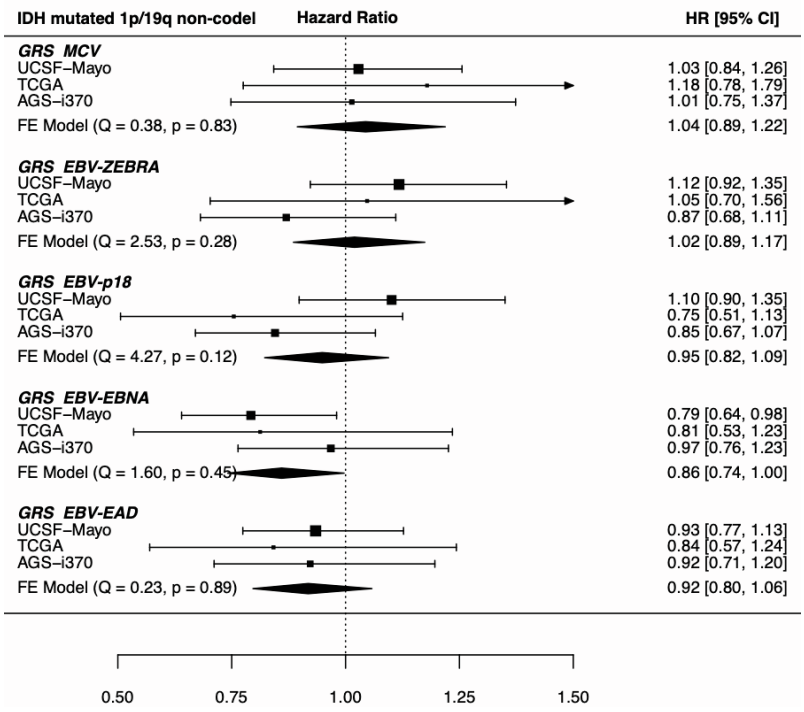
S4 B) *IDH* mutated



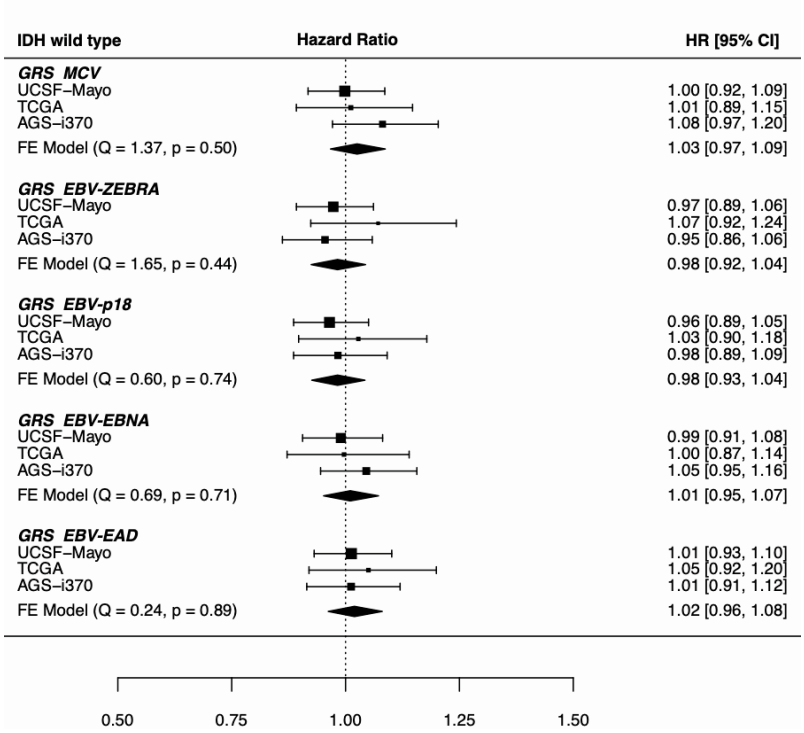
S4 C) *IDH* mutated 1p/19q codeleted



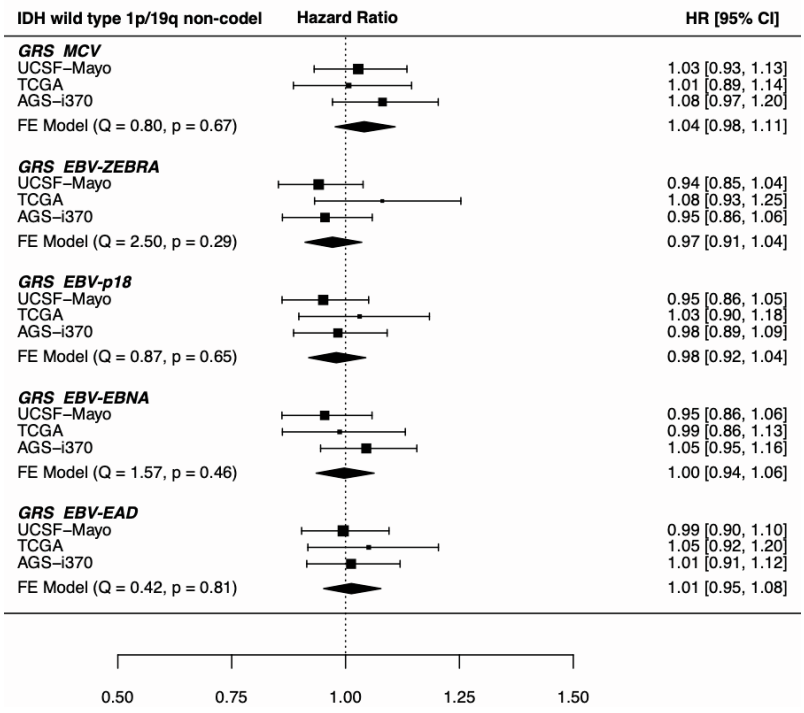
S4 D) *IDH* mutated 1p19 non-codeleted



S4 E) *IDH* wild type



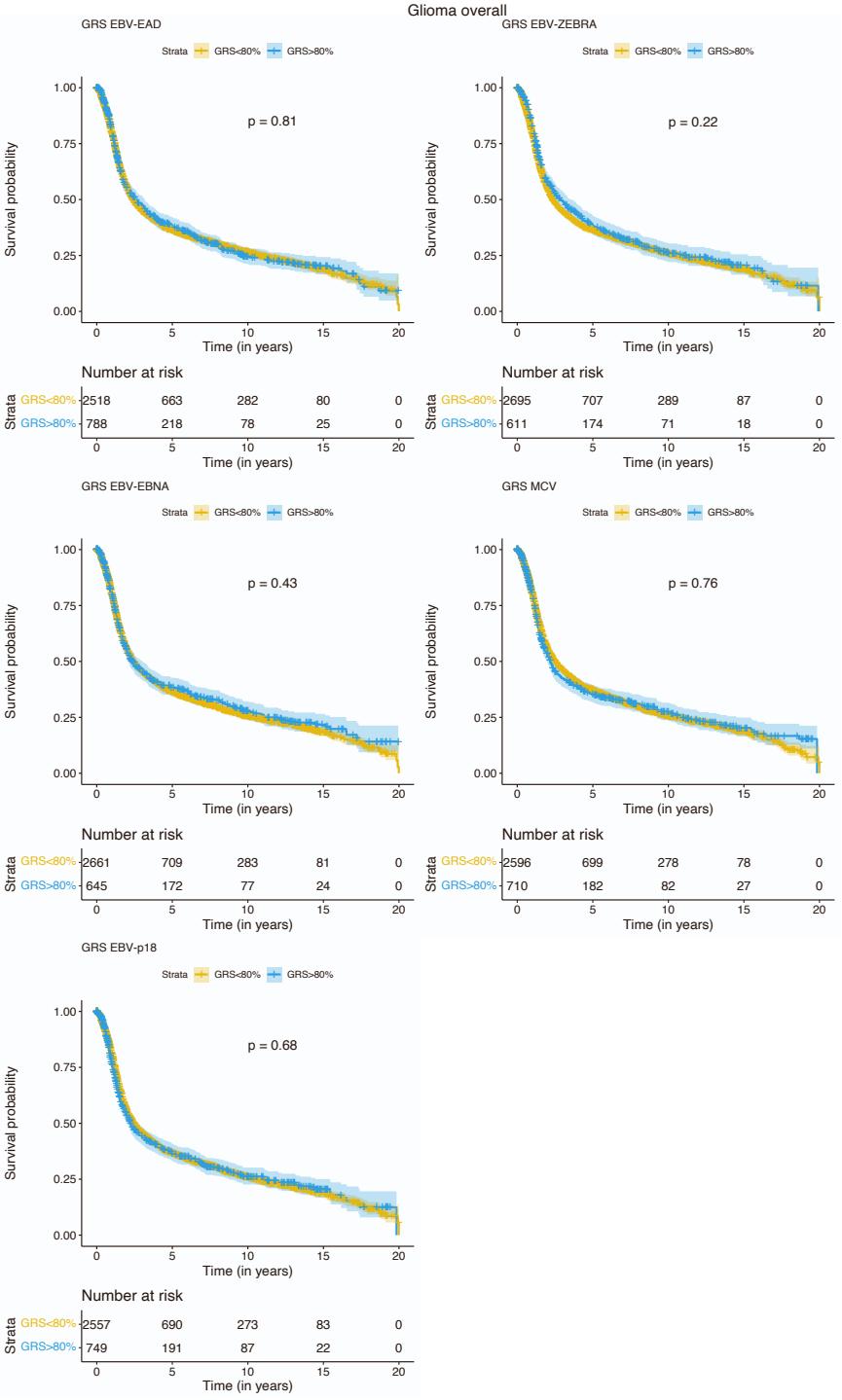
S4 F) *IDH* wild type 1p/19q non-codeleted



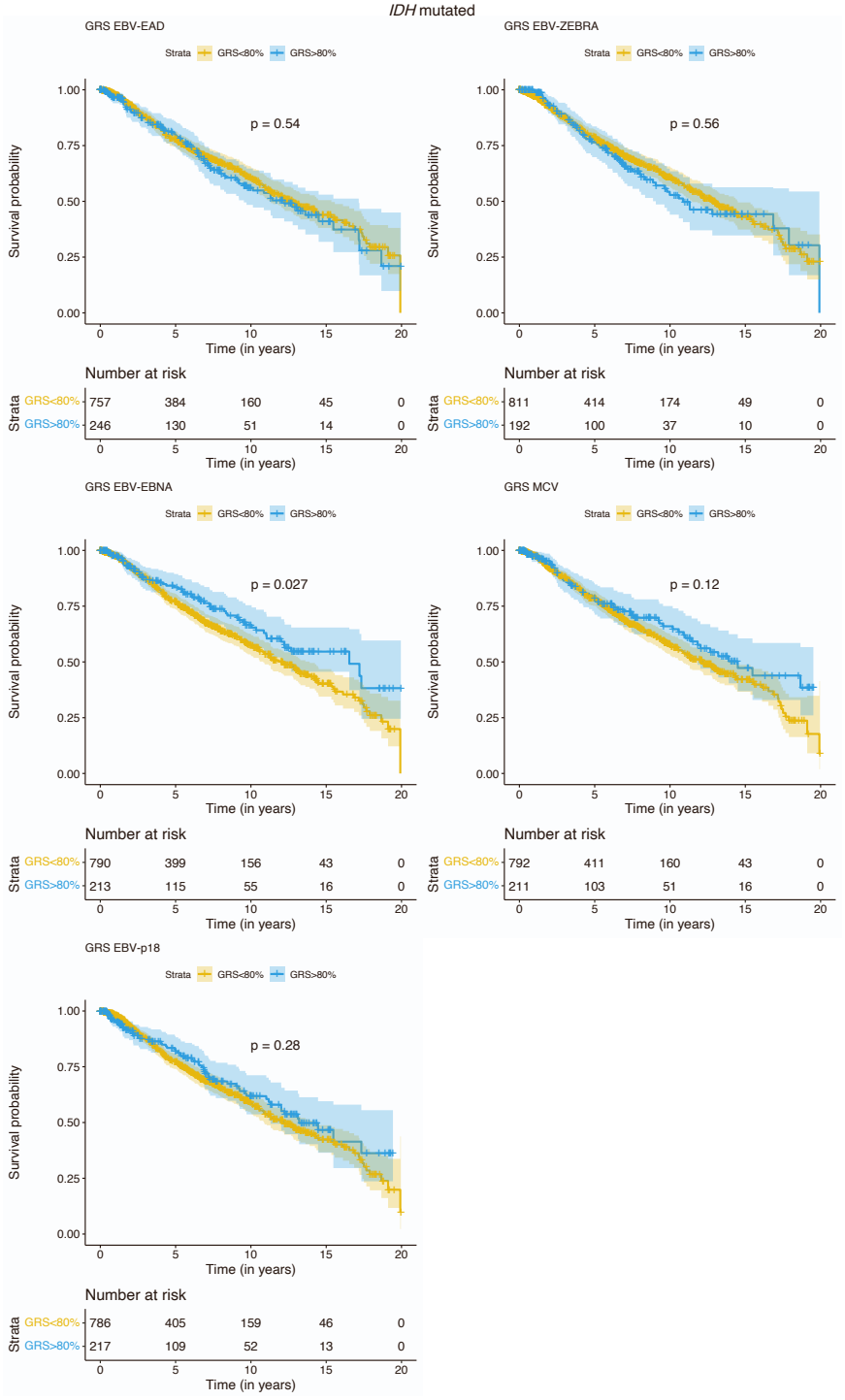
A-F) Forest plot meta-analysis results of all tested GRS-glioma survival associations. Results are reported as hazard ratios, along with 95% confidence intervals. Briefly, each header indicates the studied glioma molecular subtype, within are each GRSs associations with the clinical outcomes of that subtype, 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 which had an insufficient number of cases/events in a subtype were not included in the meta-analysis and no meta-analysis was performed if only one study was included. Each meta-analysis was tested for between-study heterogeneity (Q statistic), with $p<0.05$ indicating evidence of study-specific associations. GRS-subtype tests with evidence of significant heterogeneity were re-analyzed using a random-effects meta-analysis (results not reported). No new suggestive associations were found after re-analysis.

Figure S5: Glioma survival Kaplan-Meier curves by antigen GRS and subtype

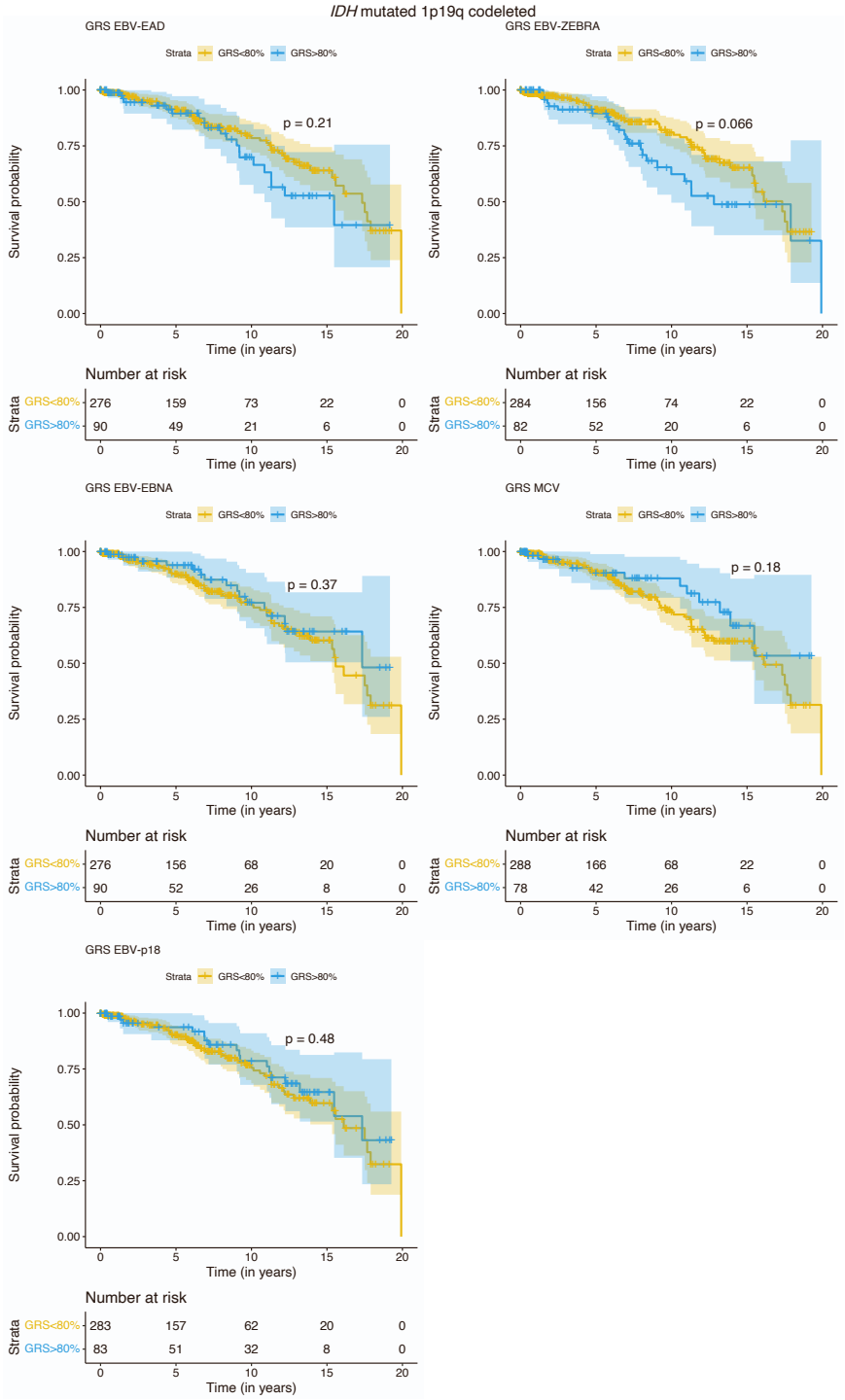
S5 A) Glioma overall



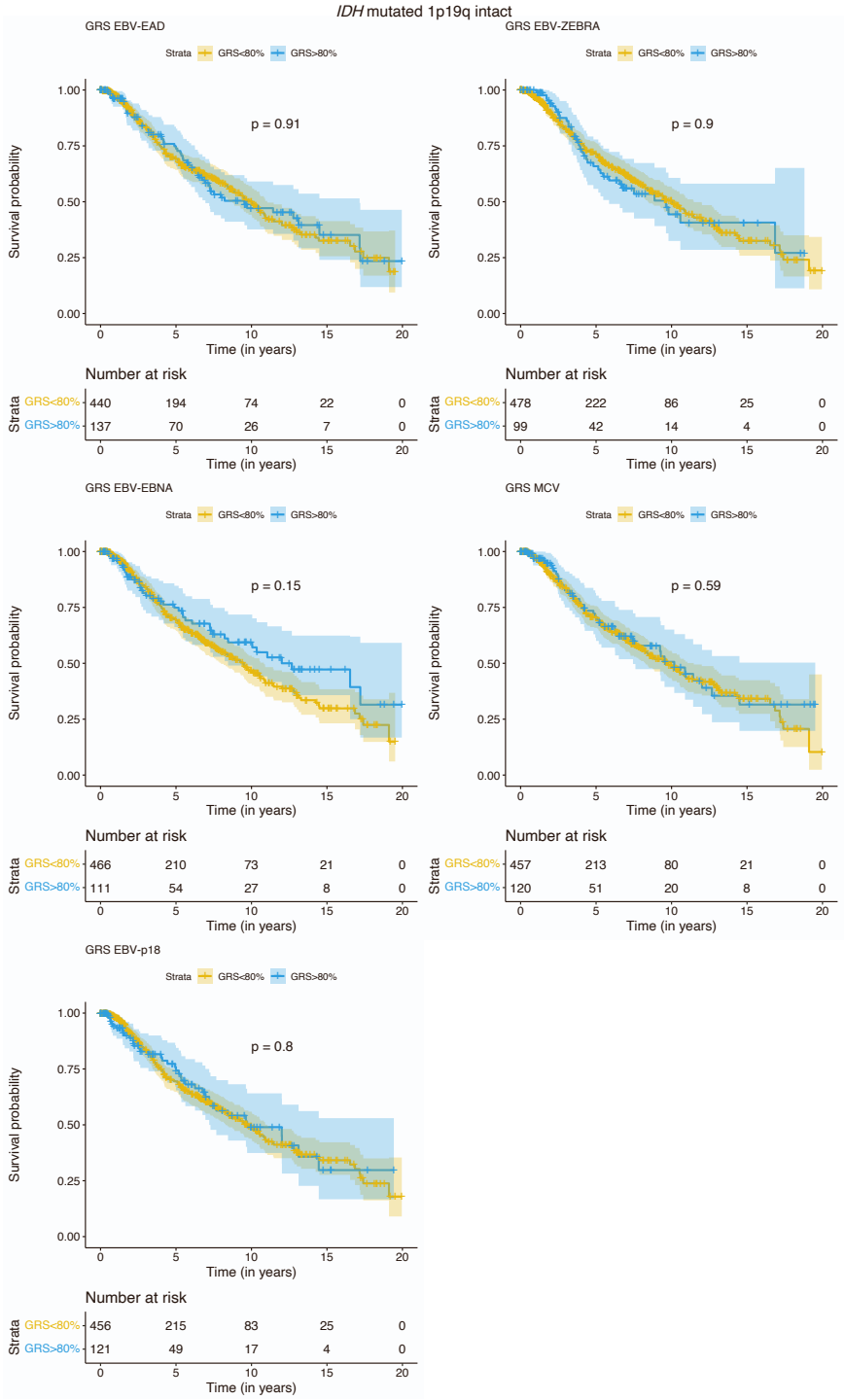
S5 B) *IDH* mutated



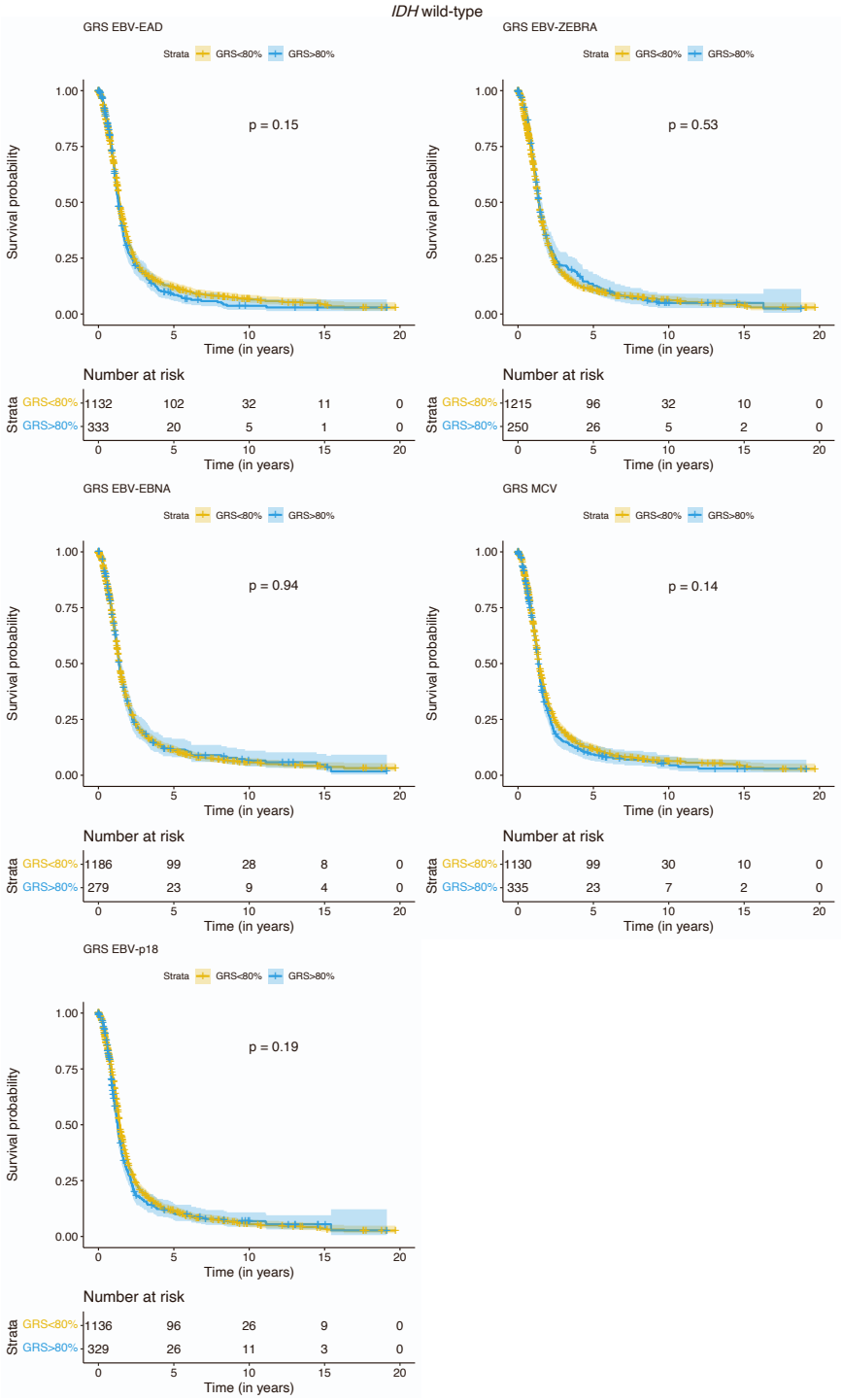
S5 C) *IDH* mutated 1p/19q codeleted



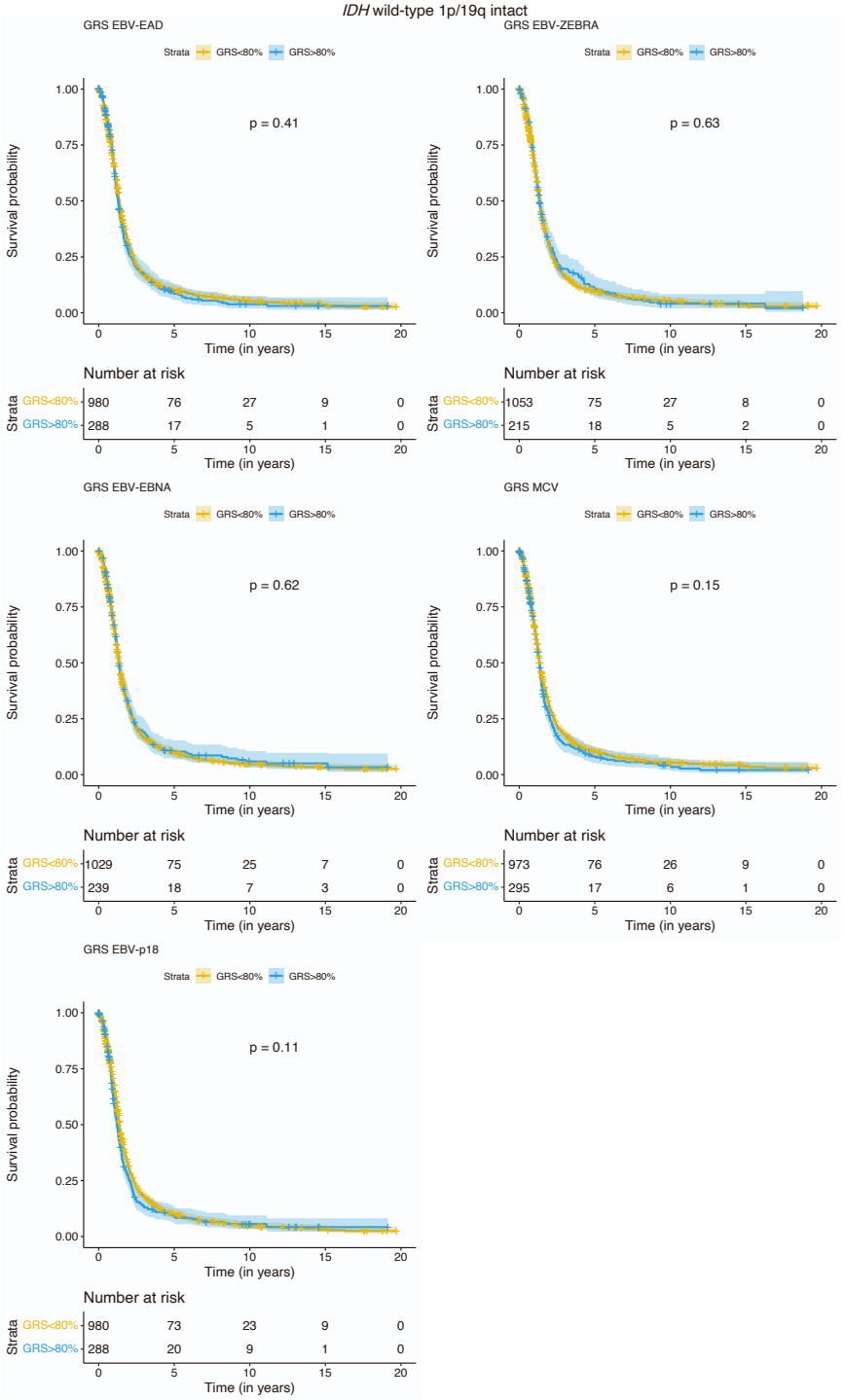
S5 D) *IDH* mutated 1p19 non-codeleted



S5 E) *IDH* wild type



S5 F) *IDH* wild type 1p/19q non-codeleted



A-F) Kaplan-Meier curves for GRS associations with the survival in the specified molecular glioma subtype. Each plot represents a visualization of the association of a specific viral antigen GRS with clinical outcomes in the molecular subtype. To visualize, each antigen's unnormalized GRS scores across the included studies were binned based on 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.

Supplemental Tables

Table S1: GWAS glioma risk meta results by subtype for the GRS variants

Separate excel file. All antigen-associated SNPs considered, basic variant information and beta/SE/p-value for a meta-GWAS for the association of the SNPs for each glioma subtype.

Table S2: SNP2HLA HLA variant glioma risk results by subtype

Separate excel file. All HLA alleles/genes which were imputed via SNP2HLA, beta/SE/p-value for associations in a meta-analysis of risk analysis of each of the HLA alleles on each glioma subtype.

Table S3: Summary of GRS-glioma risk significant results*

Subtype	Antigen GRS	Meta OR (95% CI)	Meta P-value
Glioma	EBV ZEBRA	0.936 [0.888, 0.985]	0.0116
IDH mutated	EBV EBNA	1.086 [1.004, 1.175]	0.0402
1p/19q codeleted	EBV EBNA	1.138 [1.012, 1.280]	0.0308
IDH wild type	EBV ZEBRA	0.910 [0.850, 0.975]	0.0072
	MCV	1.089 [1.018, 1.165]	0.0131
1p/19q non-codeleted	EBV ZEBRA	0.908 [0.844, 0.977]	0.0099
	MCV	1.11 [1.031, 1.194]	0.0054

*These results mirror those of Figure 3 in the main text
Indentation indicates subtype is further subset of the above result

Table S4: Summary of GRS-glioma survival significant results*

Subtype	Antigen GRS	Meta HR (95% CI)	Meta P-value
IDH mutated	EBV EBNA	0.857 [0.762, 0.964]	0.010
1p/19q codeleted	EBV EBNA	0.754 [0.569, 0.998]	0.048
	EBV ZEBRA	1.268 [1.009, 1.592]	0.042
1p/19q non-codeleted	EBV EBNA	0.860 [0.741, 0.997]	0.045

*These results mirror those of Figure 4 in the main text
Indentation indicates subtype is further subset of the above result

Table S5: Number of SNPs used in each antigen GRS

	EBV-ZEBRA	EBV-p18	EBV-EBNA	MCV	EBV-EAD	HHV7	BKV	HSV1	JCV	VZV	CMV	HHV6
#SNPs	**11**	**9**	**7**	**6**	**5**	**3**	**1**	**1**	**1**	**1**	**1**	**1**

Antigens with >4 SNPs were further used in polygenic analyses

Table S6: Effect of genetic ancestry on significant glioma risk results

		Including 10 PCs	Including 15 PCs	Including 20 PCs
Subtype	**Antigen GRS/ HLA Allele**	**Meta OR (95% CI)**	**Meta OR (95% CI)**	**Meta OR (95% CI)**
Glioma overall	EBV ZEBRA	0.936 [0.888, 0.985]	0.930 [0.883, 0.981]	0.939 [0.891, 0.991]
	HLA DQA1*03:01	0.845 [0.769, 0.927]	0.887 [0.801, 0.982]	0.896 [0.808,0.993]
IDH mut	EBV EBNA	1.086 [1.004, 1.175]	1.091 [1.005, 1.185]	1.098 [1.010, 1.194]
IDH mut 1p/19q codel	EBV EBNA	1.138 [1.012, 1.280]	1.144 [1.009, 1.296]	1.158 [1.019, 1.316]
IDH WT	EBV ZEBRA	0.910 [0.850, 0.975]	0.905 [0.844, 0.970]	0.910 [0.848, 0.977]
	MCV	1.089 [1.018, 1.165]	1.086 [1.015, 1.162]	1.083 [1.012, 1.159]
	HLA DQA1*03:01	0.817 [0.721, 0.925]	0.850 [0.733, 0.986]	0.891 [0.766, 1.037]
IDH WT 1p/19q non-codel	EBV ZEBRA	0.908 [0.844, 0.977]	0.950 [0.858, 1.051]	0.906 [0.840, 0.977]
	MCV	1.11 [1.031, 1.194]	1.045 [0.946, 1.155]	1.101 [1.022, 1.185]

Glioma risk meta-analyses were conducted including 10, 15, or 20 genetic principal components to control for fine-scale population substructure.