Title

Adult diffuse glioma GWAS by molecular subtype identifies variants in D2HGDH and FAM20C

Journal

Neuro-Oncology, 22(11)

ISSN

1522-8517

Authors

Eckel-Passow, Jeanette E
Drucker, Kristen L
Kollmeyer, Thomas M
et al.

Publication Date

2020-11-26

DOI

10.1093/neuonc/noaa117

Peer reviewed

Neuro-Oncology

22(11), 1602–1613, 2020 | doi:10.1093/neuonc/noaa117 | Advance Access date 9 May 2020

Adult diffuse glioma GWAS by molecular subtype identifies variants in *D2HGDH* and *FAM20C*

Jeanette E. Eckel-Passow,[†] Kristen L. Drucker,[†] Thomas M. Kollmeyer,[†] Matt L. Kosel, Paul A. Decker, Annette M. Molinaro, Terri Rice, Corinne E. Praska, Lauren Clark, Alissa Caron, Alexej Abyzov, Anthony Batzler, Jun S. Song, Melike Pekmezci, Helen M. Hansen, Lucie S. McCoy, Paige M. Bracci, Joseph Wiemels, John K. Wiencke, Stephen Francis, Terry C. Burns, Caterina Giannini, Daniel H. Lachance, Margaret Wrensch,[†] and Robert B. Jenkins[†]

Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, Minnesota (J.E.E-P., M.L.K., P.A.D, A.A, A.B.); Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota (K.L.D., T.M.K., C.E.P., L.C., A.C., C.G., D.H.L., R.B.J.); Department of Neurological Surgery, University of California San Francisco (UCSF), San Francisco, California (A.M.M., T.R., H.M.H., L.S.M., J.K.W, S.F, M.W.); Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, California (A.M.M., P.M.B.,J.W., J.K.W., S.F, M.W.); Department of Physics, Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Champaign, Illinois (J.S.S.); Department of Pathology, University of California San Francisco, San Francisco, California (M.P.); Institute of Human Genetics, University of California San Francisco, San Francisco, California (J.W., J.K.W., M.W.); Department of Neurology, Mayo Clinic, Rochester, Minnesota (D.H.L.); Department of Neurologic Surgery, Mayo Clinic, Rochester, Minnesota (T.C.B.)

Corresponding Author: Robert Jenkins, M.D., Ph.D.; Mayo Clinic, 200 First Street SW, Rochester, MN 55905 (rjenkins@mayo.edu).

[†]These are co-first authors.

[†]These are co-senior authors.

Abstract

Background. Twenty-five germline variants are associated with adult diffuse glioma, and some of these variants have been shown to be associated with particular subtypes of glioma. We hypothesized that additional germline variants could be identified if a genome-wide association study (GWAS) were performed by molecular subtype.

Methods. A total of 1320 glioma cases and 1889 controls were used in the discovery set and 799 glioma cases and 808 controls in the validation set. Glioma cases were classified into molecular subtypes based on combinations of isocitrate dehydrogenase (*IDH*) mutation, telomerase reverse transcriptase (*TERT*) promoter mutation, and 1p/19q codeletion. Logistic regression was applied to the discovery and validation sets to test for associations of variants with each of the subtypes. A meta-analysis was subsequently performed using a genome-wide *P*-value threshold of 5×10^{-8}.

Results. Nine variants in or near D-2-hydroxyglutarate dehydrogenase (*D2HGDH*) on chromosome 2 were genome-wide significant in *IDH*-mutated glioma (most significant was rs5839764, meta $P = 2.82 \times 10^{-10}$). Further stratifying by 1p/19q codeletion status, one variant in *D2HGDH* was genome-wide significant in *IDH*-mutated non-codeleted glioma (rs1106639, meta $P = 4.96 \times 10^{-8}$). Further stratifying by *TERT* mutation, one variant near *FAM20C* (family with sequence similarity 20, member C) on chromosome 7 was genome-wide significant in gliomas that have *IDH* mutation, *TERT* mutation, and 1p/19q codeletion (rs111976262, meta $P = 9.56 \times 10^{-9}$). Thirty-six variants in or near *GMEB2* on chromosome 20 near regulator of telomere elongation helicase 1 (*RTEL1*) were genome-wide significant in *IDH* wild-type glioma (most significant was rs4809313, meta $P = 2.60 \times 10^{-10}$).

Conclusions. Performing a GWAS by molecular subtype identified 2 new regions and a candidate independent region near *RTEL1*, which were associated with specific glioma molecular subtypes.

Key Points

1. We performed a GWAS of molecular subtypes of adult diffuse glioma and identified 2 new regions that are associated with particular glioma subtypes.
2. One of the regions is in *D2HGDH*, a region that is also associated with allergy and asthma.

Importance of the Study

Twenty-five germline variants have been associated with adult diffuse glioma, and some of these variants have subsequently been shown to be associated with particular subtypes of glioma. By performing a GWAS by molecular subtype, we identified 2 new regions that are associated with specific molecular subtypes of glioma. Variants in *D2HGDH* on chromosome 2 were associated with *IDH*-mutated glioma. A variant near *FAM20C* on chromosome 7 was associated with gliomas that have *IDH* mutation, *TERT* mutation, and 1p/19q codeletion. One of the regions, *D2HGDH*, is a region that is also associated with allergy and asthma. The identification of additional novel germline variants will help to further understand the etiology of adult diffuse glioma.

Genome-wide association studies (GWAS) identified variants in 25 regions that are associated with development of adult diffuse glioma.[1–9] Recently, GWAS by histological subtype identified novel germline variants that were associated specifically with glioblastoma (GBM; World Health Organization [WHO] grade IV) and non-GBM (grades II–III).[2,3] Importantly, these newly identified GBM and non-GBM germline variants did not reach genome-wide significance when a GWAS was performed on overall glioma; they only reached genome-wide significance when a GWAS was performed within a more homogeneous subgroup of glioma. As highlighted by the 2016 WHO classification criteria,[10] which include genetic tumor testing of isocitrate dehydrogenase (*IDH*) mutation and 1p/19q codeletion, glioma can more accurately be subtyped by somatic alterations. Telomerase reverse transcriptase (*TERT*) promoter mutation has also been reproducibly shown to be associated with age at diagnosis, patient outcome, and specific germline associations.[11–14] However, a GWAS by these molecular subtypes has not been reported. Thus, we hypothesized that novel germline variants might be identified by GWAS performed by clinically relevant molecular subtypes. Here, we describe the results of performing a GWAS within molecular subtypes defined by combinations of *IDH* mutation, *TERT* promoter mutation, and 1p/19q codeletion.

Methods

Subjects

Mayo Clinic glioma cases.—Mayo Clinic glioma cases have been previously described.[3,8,11,15] The study was approved by the Mayo Clinic Office for Human Research Protection. Histologically confirmed grades II, III, and IV glioma cases were identified at diagnosis (at Mayo Clinic) or at the time of pathologic confirmation (diagnosed elsewhere and treated at Mayo Clinic). Included subjects were at least 18 years of age and had a surgical resection or biopsy between 1973 and 2014. A total of 653 cases that were run on the OncoArray genotyping assay had necessary molecular data.[3]

UCSF glioma cases.—UCSF glioma cases include participants of the San Francisco Bay Area Adult Glioma Study (AGS). This study was approved by the UCSF Committee on Human Research. Details of subject recruitment for AGS have been reported previously.[1,3,8,11,16–18] Cases were adults (>18 y of age) with newly diagnosed, histologically confirmed grade II, III, or IV glioma. Population-based cases diagnosed between 1991 and 2009 and residing in the 6 San Francisco Bay Area counties were ascertained using the Cancer Prevention Institute of California's early case ascertainment system. Clinic-based cases diagnosed between 2002 and 2012 were recruited from the UCSF Neuro-oncology Clinic, regardless of place of residence. A total of 667 cases that were run on the OncoArray genotyping assay had necessary molecular data.[3]

TCGA glioma cases.—A total of 799 glioma cases from The Cancer Genome Atlas (TCGA) obtained from Database of Genotypes and Phenotypes (dbGaP) (phs000178) had necessary molecular data. All samples were run on the Affymetrix 6.0 genotyping array. *IDH* mutation, *TERT* promoter mutation, and 1p/19q codeletion status were obtained from supplementary table 1 in Ceccarelli et al,[19] and D-2-hydroxyglutarate dehydrogenase (*D2HGDH*) copy number calls were downloaded from cBioPortal on February 19, 2019.

International Glioma Case-Control (GICC) Study controls.—A total of 1907 GICC controls that were approved for "General Research Use" and "Brain Tumors"

were obtained from dbGaP (phs001319); 1889 passed quality control metrics and were included in the analyses. All samples were run on the OncoArray genotyping assay.[3]

Mayo Clinic Biobank controls.—A total of 808 Mayo Clinic Biobank controls were utilized.[20] All samples were run on the Illumina Omni Express genotyping array.

Statistical Methods

GWAS analyses.—Prior to imputation, the following quality control procedures were applied to the discovery and validation sets: tests of Hardy–Weinberg equilibrium ($P < 10^{-6}$), duplicate and relatedness checks, sex checks, variant call rates (>95%), and subject call rates (>95%). For the validation set, because cases and controls were run on different genotyping platforms, quality control was performed on cases and controls separately. Imputation was performed using the Michigan Imputation Server, utilizing the 1000 Genome V3 data and the Haplotype Reference Consortium. Structure[21] was applied to determine racial groups, using 1000 Genome samples to anchor the racial groups. Population stratification was assessed using Eigenstrat and principal components (Supplementary Figure 1).[22] Principal components that were significantly associated with overall glioma cases versus control ($P < 0.05$) were included as covariates in each of the subtype-specific GWAS analyses. Subtype-specific GWAS analyses were performed with cases defined by somatic alterations and corresponding cases being compared with controls. GWAS was first performed stratified by *IDH* mutation: *IDH*-mutated glioma and *IDH* wild-type glioma were each compared with controls. *IDH*-mutated cases were subsequently stratified by 1p/19q codeletion into *IDH*-mutated 1p/19q codeleted and *IDH*-mutated 1p/19q non-codeleted. GWAS were then conducted for subgroups further stratified by *TERT* promoter mutation into triple-positive (*IDH*-mutated, *TERT* mutated, and 1p/19q codeleted), *IDH* and *TERT* mutations, *IDH* mutation only, *TERT* mutation only, and triple-negative (*IDH* wild-type, *TERT* wild-type, 1p/19q non-codeleted). Logistic regression was utilized comparing subtype-specific cases with controls, with genotype coded as 0, 1, or 2 copies (or dosage if imputed) of the alternate allele, and age, sex, and principal components included as covariates. Q-Q plots and lambda values were used to evaluate the excess false-positive rate. Imputation R^2 value for the discovery set was required to be larger than 0.80 and the *P*-value threshold was 5×10^{-6}. Variants that passed the imputation and *P*-value thresholds in the discovery set were meta-analyzed across the discovery and validation sets. Variants that had the same direction of effect in the discovery and validation sets, genome-wide significant meta *P*-value ($P < 5 \times 10^{-8}$), and *P*-value less than 0.05 in the validation set were further evaluated. For a subset of the imputed variants that passed genome-wide significance in the meta-analysis, 93 glioma patient samples were custom genotyped using TaqMan and compared with imputed genotypes (Supplementary Table 1).

Expression quantitative trait loci, Hi-C and ChromHMM analyses.—Expression quantitative trait loci (eQTL) analyses were performed using data obtained from the GTEx Portal on 10/22/2019. Analyses were performed for the most significant variant at each locus. All genes within 1 Mb of the variant were evaluated in normal brain tissues available in GTEx. Hi-C analyses were performed using the 3D-genome Interaction Viewer and database.[23] Interactions were examined for a 200 kb window surrounding the variant of interest in dorsolateral prefrontal cortex, hippocampus and H1-derived neural progenitor cells. ChromHMM was also evaluated for the dorsolateral prefrontal cortex and hippocampus using the UC Santa Cruz genome browser and the Roadmap Epigenomics project.[24]

Results

GWAS by Glioma Subtypes: Subtypes Defined by *IDH* Mutation and 1p/19q Codeletion

Of the 2119 glioma cases analyzed in the meta-analysis, 1012 were *IDH*-mutated and 1107 *IDH* wild-type (Table 1). Further stratifying *IDH*-mutated glioma according to 1p/19q codeletion, in the meta-analysis 390 patients had *IDH*-mutated 1p/19q-codeleted tumors and 561 patients had *IDH*-mutated non-codeleted tumors (Table 1).

IDH-mutated GWAS.—Ninety-three variants passed genome-wide significance in *IDH*-mutated glioma versus controls (Fig. 1A; Supplementary Figure 2A; Supplementary Table 2). Most of the variants were in regions that have previously been reported: *CCDC26*, *PHLDB1*, *AKT3*, and *IDH1*.[1,3,5] Nine variants in or near *D2HGDH* on chromosome region 2q27 were genome-wide significant (Supplementary Table 2; Fig. 2A). The most significant variant in the *D2HGDH* region was rs5839764 (discovery odds ratio [OR] = 1.51, meta *P*-value = 2.82×10^{-10}; Table 2), and this variant remained significant after adjustment for the known *IDH1* variant rs7572263 on chromosome 2 ($P = 5.46 \times 10^{-7}$; Supplementary Table 3).[3] TCGA reported that the 2q37 region was commonly deleted in *IDH*-mutated gliomas that do not have 1p/19q codeletion.[19] In TCGA data for *IDH*-mutated non-codeleted glioma, we observed that rs5839764 was inversely associated with tumor deletions of *D2HGDH* (OR = 0.57, 95% CI: 0.36–0.90, $P = 0.015$), with deletions more likely to occur in patients who carry the reference allele C versus the alternative allele G. Hi-C DNA interactions were observed between rs5839764 and nearby regions, including the 5′ end of *D2HGDH*, in the hippocampus and H1-derived neural progenitor cells (Fig. 2A). GTEx did not have data on rs5839764, and thus the second most significant variant was evaluated as a surrogate (rs71430382). The rs71430382 alternate allele T was associated with decreased expression of *D2HGDH* in normal brain tissues ($P = 4.9 \times 10^{-14}$, 1.7×10^{-18}, and 2.2×10^{-11} for frontal cortex, cortex, and hippocampus, respectively; Fig. 2B).

Table 1. Clinical, pathology, and molecular characteristics of discovery and validation sets

	Discovery		Validation	
	Mayo and UCSF Cases, N = 1320 (%)	GICC Controls, N = 1889 (%)	TCGA Cases, N = 799 (%)	Mayo Biobank Controls, N = 808 (%)
Age, y				
<40	383 (29.0)	279 (14.8)	239 (29.9)	68 (8.4)
40–59	587 (44.5)	831 (44.0)	304 (38.0)	281 (34.8)
≥60	350 (26.5)	779 (41.2)	256 (32.0)	459 (56.8)
Sex				
Female	530 (40.2)	755 (40.0)	335 (41.9)	407 (50.4)
Male	790 (59.8)	1134 (60.0)	464 (58.1)	401 (49.6)
Histology				
Astrocytoma	281 (21.3)	NA	145 (19.2)	NA
Glioblastoma	574 (43.5)	NA	359 (47.5)	NA
Oligoastrocytoma	215 (16.3)	NA	106 (14.0)	NA
Oligodendroglioma	250 (18.9)	NA	145 (19.2)	NA
Missing	0	NA	44	NA
Grade				
2	401 (30.4)	NA	182 (24.1)	NA
3	327 (24.8)	NA	214 (28.3)	NA
4	591 (44.8)	NA	359 (47.5)	NA
Missing	1	NA	44	NA
***IDH* mutation status**				
IDH mutant	622 (47.1)	NA	390 (48.8)	NA
IDH wild-type	698 (52.9)	NA	409 (51.2)	NA
Molecular subtype based on *IDH* mutation and 1p/19q codeletion*				
IDH mutant 1p/19q codeleted	245 (19.5)	NA	145 (18.1)	NA
IDH mutant 1p/19q non-codeleted	316 (25.1)	NA	245 (30.7)	NA
IDH wild-type	698 (55.4)	NA	409 (51.2)	NA
Missing	61	NA	0	NA
Molecular subtype based on *IDH* mutation, *TERT* promoter mutation, and 1p/19q codeletion**				
IDH-mutation only	241 (24.7)	NA	214 (38.4)	NA
TERT and *IDH* mutations	39 (4.0)	NA	13 (2.3)	NA
TERT-mutation only	419 (43.0)	NA	159 (28.5)	NA
Triple-negative	87 (8.9)	NA	33 (5.9)	NA
Triple-positive	189 (19.4)	NA	138 (24.8)	NA
Missing	345	NA	242	NA

*Tumors were required to have both *IDH* mutation and 1p/19q results available in order to classify into subtypes.
**Tumors were required to have *IDH* mutation, *TERT* promoter mutation, and 1p/19q results available in order to classify into subtypes.
NA: not applicable to controls.

IDH-mutated non-codeleted GWAS.—Twelve variants passed genome-wide significance for patients with *IDH*-mutated non-codeleted glioma versus controls (Fig. 1B; Supplementary Figure 2B; Supplementary Table 4). Most of the variants were in known genes, including *CCDC26* and *PHLDB1*.[1,5] One variant within *D2HGDH*, rs1106639, reached genome-wide significance (discovery OR = 1.7, meta *P*-value = 4.96×10^{-8}) (Table 2). The variant remained significantly associated with risk after adjustment for the *IDH1* variant rs7572263 on chromosome 2 ($P = 5.0 \times 10^{-6}$; Supplementary Table 5).[3] The Hi-C analyses for rs1106639 showed similar results as rs5839764 (Supplementary Figure 3).

IDH-mutated 1p/19q-codeleted GWAS.—Twenty-eight variants passed genome-wide significance in *IDH*-mutated codeleted glioma versus controls; all variants were in *CCDC26*[1] (Supplementary Figure 2C; Supplementary Figure 4; Supplementary Table 6).

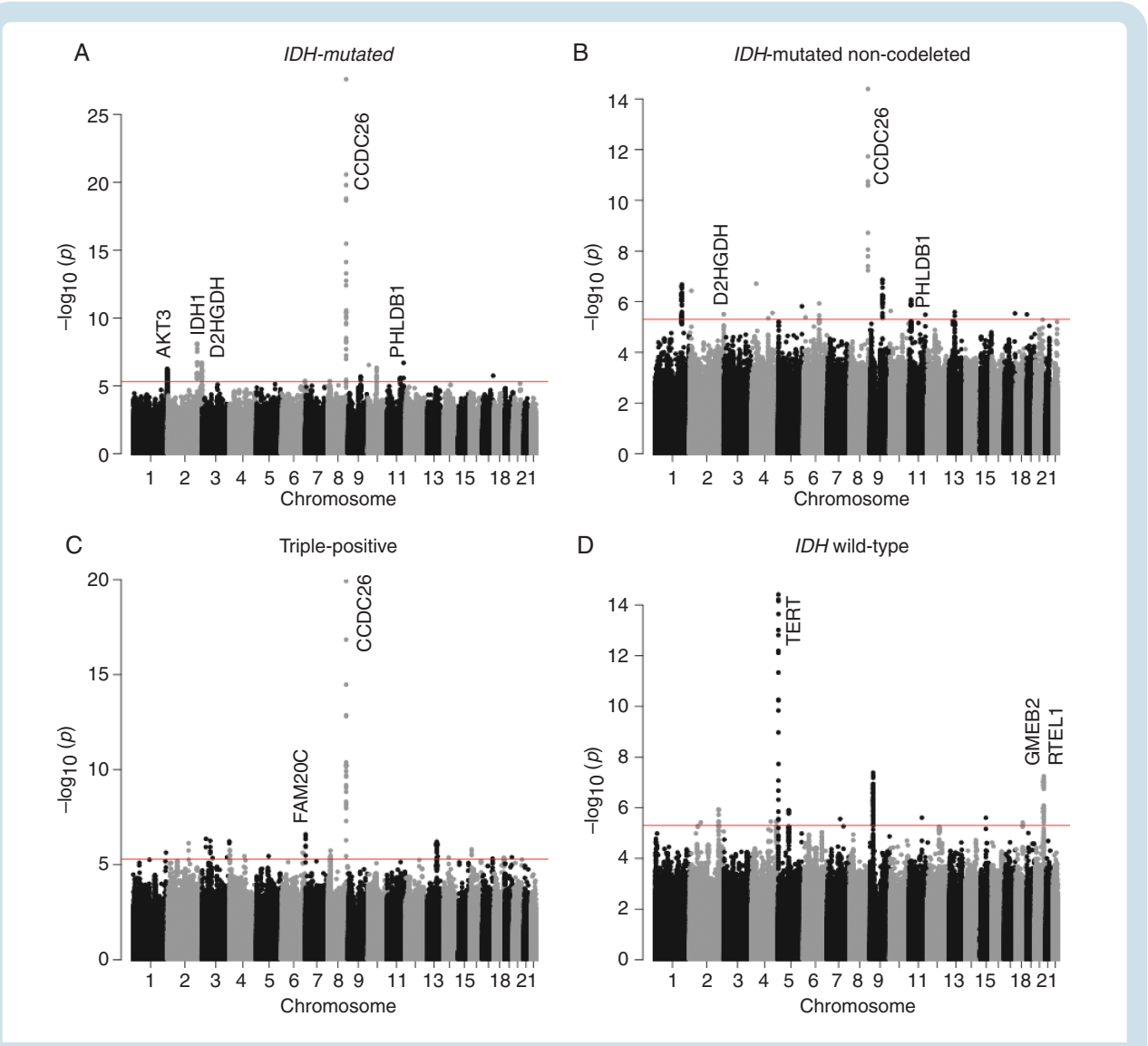

Fig. 1 Manhattan plot for (A) *IDH*-mutated glioma, (B) *IDH* mutated non-codeleted glioma, (C) Triple-positive glioma, and (D) *IDH* wild-type glioma. Results from the discovery set are shown and genes that passed genome-wide significance ($P < 5 \times 10^{-8}$) in the meta analysis are annotated.

IDH wild-type GWAS.—Eighty-five variants were associated with *IDH* wild-type glioma versus controls at genome-wide significance (Fig. 1D; Supplementary Figure 2D; Supplementary Table 7). Most of the variants were in known genes, including *TERT* and regulator of telomere elongation helicase 1 (*RTEL1*).[5,8] Thirty-six variants in or near glucocorticoid modulatory element binding protein 2 (*GMEB2*) on chromosome region 20q13 were genome-wide significant (Table 2; Fig. 3). The most significant variant in the *GMEB2* region was rs4809313 (discovery OR = 0.66, meta *P*-value = 2.60×10^{-10}), and this variant remained significant after adjustment for the known *RTEL1* glioma risk variant (rs2297440) that is nearby on chromosome 20 ($P = 0.029$; Supplementary Table 8).[3,5] There were no significant eQTL in cerebellar hemisphere tissue. Hi-C interactions were observed between rs4809313 and nearby genes, including *RTEL1* (Fig. 3). Since we observed Hi-C interactions between the *GMEB2* variant and the *RTEL1* region, we also evaluated Hi-C interactions with the *RTEL1* glioma risk variant rs2297440. Hi-C interactions were observed between the *RTEL1* variant rs2297440 and the *GMEB2* region (Supplementary Figure 5).

GWAS by Glioma Subtypes: Subtypes Stratified by *TERT* Promoter Mutation

Further stratifying *IDH*-mutated tumors by *TERT* promoter mutation resulted in the following numbers of patients in each of the meta-analyses: 455 IDH-mutated only (*IDH*-mutated, *TERT* wild-type, and 1p/19q non-codeleted), 52 *IDH*- and *TERT*-mutated (*IDH*-mutated, *TERT*-mutated, and 1p/19q non-codeleted), and 327 triple-positive (*IDH*-mutated, *TERT*-mutated, and 1p/19q codeleted) glioma (Table 1). The meta-analyses for *IDH* wild-type tumors were further stratified by *TERT* promoter mutation into

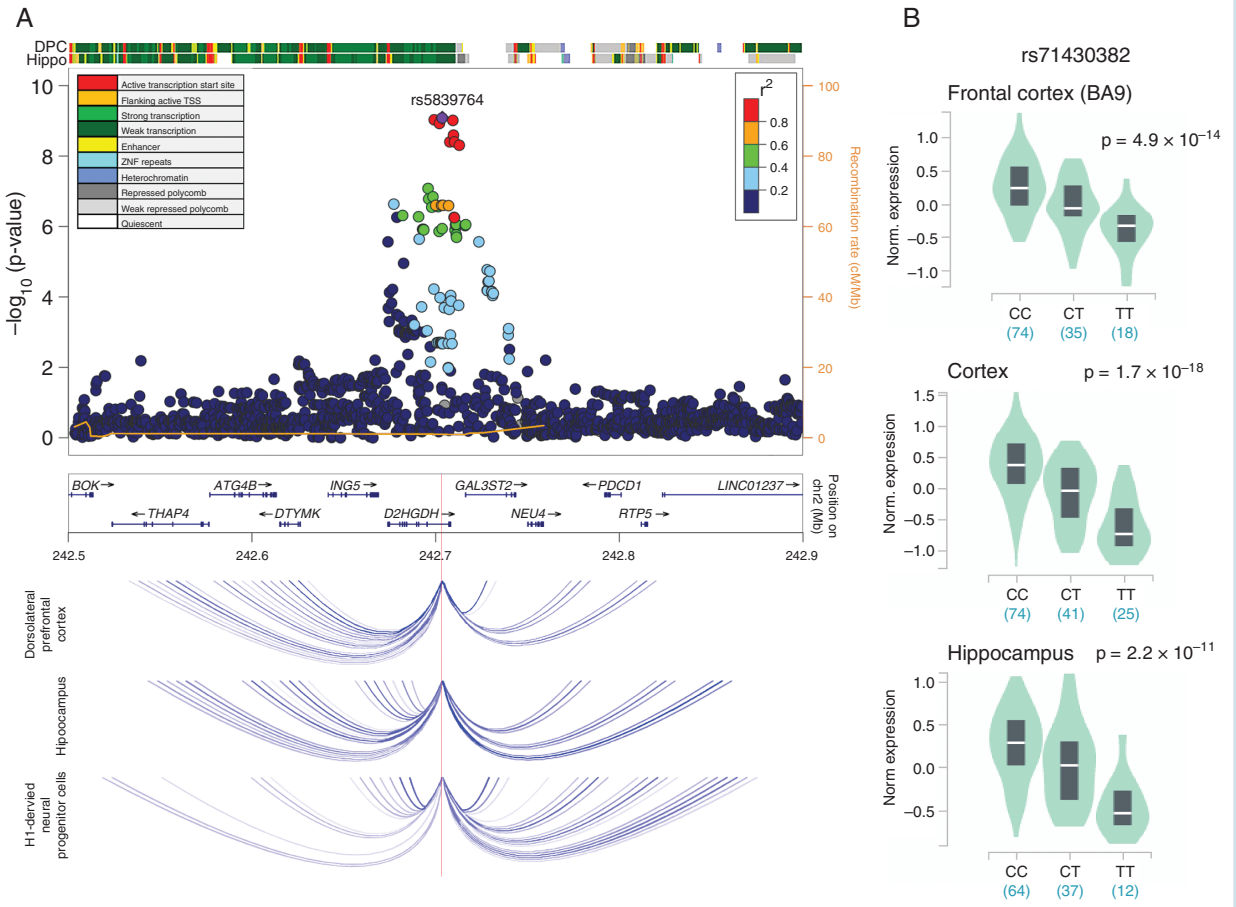

Fig. 2 *D2HGDH* functional analyses. (A) Genomic region for 200 kb surrounding rs5839764 is shown. The heatmap shows DNA states predicted by the ChromHMM algorithm in the dorsolateral prefrontal cortex (DPC) and hippocampus (Hippo) in brains from normal donors. The locus zoom plot shows the *P*-values from the meta-analysis for all variants in the region and the recombination rate. DNA-DNA interactions were measured by Hi-C analysis in the dorsolateral prefrontal cortex, hippocampus, and H1-derived neural progenitor cells; each blue line denotes an interaction. (B) eQTL was examined using GTEx for rs71430382, the second most significant variant (Supplementary Table 2), as no data were available in GTEx for rs5839764. Significant eQTL were observed in frontal cortex (top), cortex (middle), and hippocampus (bottom).

578 *TERT*-mutated only (*IDH* wild-type, *TERT*-mutated, and 1p/19q non-codeleted) and 120 triple-negative (*IDH* wild-type, *TERT* wild-type, and 1p/19q non-codeleted) glioma (Table 1).

Triple-positive GWAS.—Twenty-nine variants were associated with triple-positive glioma versus controls at genome-wide significance (Fig. 1C; Supplementary Figure 6a; Supplementary Table 9). Most variants were in *CCDC26*[1]; however, rs111976262 was in a novel region on chromosome 7 near family with sequence similarity 20, member C (*FAM20C*) (discovery OR = 3.52, meta *P*-value = 9.56×10^{-9}) (Table 2). There were no significant eQTL in normal brain tissue. Hi-C interactions were observed between rs111976262 and nearby regions (Fig. 4).

IDH-mutated only GWAS.—Seven variants were associated with *IDH*-mutated only glioma versus controls at genome-wide significance; all variants were in *CCDC26*[1] and previously known (Supplementary Figure 6b; Supplementary Figure 7a; Supplementary Table 10).

IDH- and TERT-mutated GWAS.—No variants were observed to be associated at genome-wide significance with *IDH*- and *TERT*-mutated glioma versus controls (Supplementary Figure 6c; Supplementary Figure 7b).

TERT-mutated only GWAS.—Thirty-eight variants were associated with *TERT*-mutated only glioma at genome-wide significance (Supplementary Figure 8a; Supplementary Figure 9a; Supplementary Table 11). All variants were in or near previously established glioma risk regions including *TERT* and *RTEL1*.[5]

Triple-negative GWAS.—No variants were genome-wide significant in triple-negative glioma versus controls (Supplementary Figure 8b; Supplementary Figure 9b).

Table 2. Variants associated with particular glioma molecular subtypes

SNP	Gene	Chr	Position (hg19)	Ref Allele	Alt Allele	Discovery Set				Validation Set		Meta
						Case AAF	Control AAF	OR	*P*-value	OR	*P*-value	*P*-value
IDH mutated glioma												
rs5839764	*D2HGDH*	2	242703618	C	G	0.458	0.372	1.512	3.62E-07	1.564	0.0001264	2.82E-10
IDH mutated non-codeleted glioma												
rs1106639	*D2HGDH*	2	242690675	G	A	0.337	0.260	1.705	3.20E-06	1.718	0.002258	4.96E-08
Triple-positive glioma (IDH mutated, TERT mutated, 1p19q codeleted)												
rs111976262	*FAM20C*	7	188634	C	A	0.074	0.031	3.516	1.02E-06	3.051	0.001252	9.56E-09
IDH wild-type glioma												
rs4809313	*GMEB2*	20	62238086	G	A	0.173	0.232	0.663	1.10E-06	0.627	4.79E-05	2.60E-10

Abbreviations: Alt = alternate; AAF = alternate allele frequency; Chr = chromosome; OR = odds ratio; Ref = reference; SNP = single nucleotide polymorphism.

Association of D2HGDH, Allergy, and Glioma

Alternate allele A of germline variant rs34290285 within *D2HGDH* has been shown to be protective of allergy and asthma,[25–27] and allergy has been shown to be protective of adult diffuse glioma.[28,29] We evaluated the association of the *D2HGDH* variant rs34290285 with glioma molecular subtypes and observed significant association within *IDH*-mutated glioma (discovery OR = 1.54, meta *P*-value = 6.40×10^{-8}), but not *IDH* wild-type glioma (meta *P*-value = 0.47) (Supplementary Table 12). Further stratifying *IDH*-mutated glioma by 1p/19q codeletion status, rs34290285 was associated with both *IDH*-mutated codeleted (discovery OR = 1.52, meta *P*-value = 0.0019) and *IDH*-mutated non-codeleted glioma (discovery OR = 1.64, meta *P*-value = 4.98×10^{-8}) (Supplementary Table 12).

Discussion

Two novel glioma regions were identified to be associated with risk of specific glioma molecular subtypes in these first GWAS to be conducted by glioma molecular subtype. Variants within *D2HGDH* were associated with *IDH*-mutated glioma, and a variant near *FAM20C* was associated with gliomas that are *IDH*-mutated, *TERT*-mutated, and 1p/19q codeleted. We also identified a possible independent region near *RTEL1*; variants in *GMEB2* were associated with *IDH* wild-type glioma. Interestingly, all three regions are located near telomeres.

We demonstrated genome-wide significance of *D2HGDH* variants with *IDH*-mutated glioma, and the more homogeneous subset of *IDH*-mutated non-codeleted glioma. One of these variants, rs1106639, was previously shown to have a candidate association with non-GBM (OR = 1.29, $P = 1.11 \times 10^{-5}$); however, it was not previously reported to be genome-wide significant.[30] In the meta-analyses reported herein, rs5839764 was observed to be more significant than rs1106639 in the analysis of *IDH*-mutated glioma. And we observed a highly significant correlation between rs71430382 (surrogate for rs5839764) genotype and *D2HGDH* gene expression. *D2HGDH* is a ubiquitously expressed enzyme found in mitochondria where it converts low levels of naturally produced D2HG to alpha-ketoglutarate (αKG).[31] DNA interactions between rs5839764 and the 5′ region of *D2HGDH* in the hippocampus and cultured neural progenitor cells provide support that rs5839764 (or another variant in linkage disequilibrium) regulates expression of *D2HGDH* through a long-range interaction. This is further supported by the ChromHMM data that assigned promoter DNA in areas where DNA interactions were mapping to the 5′ region of *D2HGDH*.

A variant in *D2HGDH* (rs34290285) has been reported to be associated with asthma and allergic disease.[25–27] Though there have been some discrepant results in the literature, a history of allergies seems to be protective of adult diffuse glioma[28,29] and confers a better prognosis for patients who develop glioma.[32] Together, these results suggest a link between the immune system and *IDH* mutant gliomas, mediated by *D2HGDH*, which may drive the glioma–allergy

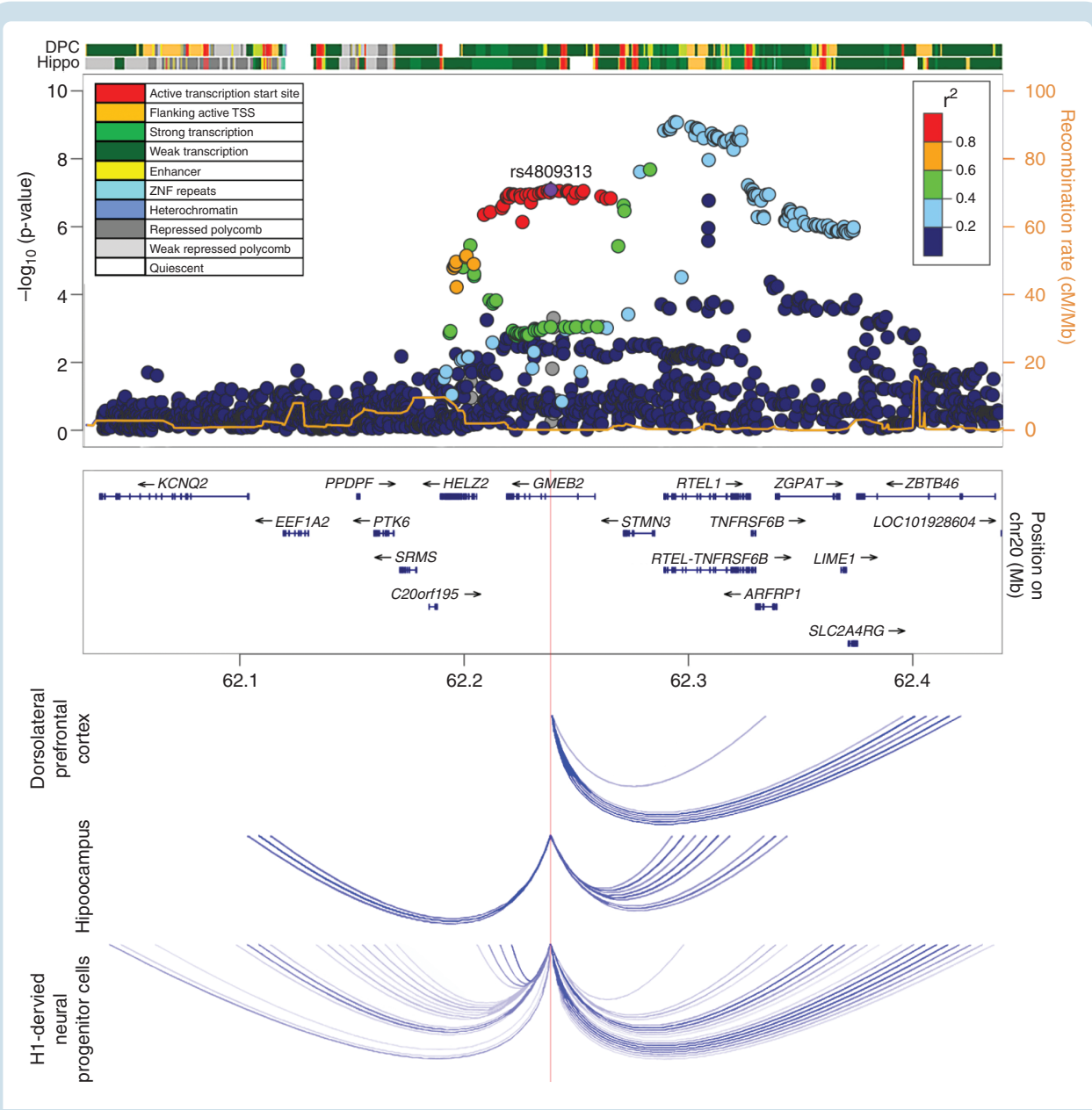

Fig. 3 *GMEB2* functional analyses. Genomic region for 200 kb surrounding rs4809313 is shown. The heatmap shows DNA states predicted by the ChromHMM algorithm in the dorsolateral prefrontal cortex (DPC) and hippocampus (Hippo) in brains from normal donors. The locus zoom plot shows the *P*-values from the meta-analysis for all variants in the region and the recombination rate. DNA-DNA interactions were measured by Hi-C analysis in the dorsolateral prefrontal cortex, hippocampus, and H1-derived neural progenitor cells; each blue line denotes an interaction.

association. The oncometabolite 2-hydroxyglutarate is likely present at increased levels due to *D2HGDH* dysfunction and *IDH* mutation. D2HG not only promotes glioma cytosine-phosphate-guanine island methylator phenotype (G-CIMP), but also specifically plays a role in the epigenetic regulation and promotion of T-cell differentiation. Future studies are important to understand how this genetic relationship leads to altered immunobiology.[33]

We observed an association with a variant near *FAM20C* (also known as *DMP-4* and *G-CK*) and gliomas that have *IDH* mutation, *TERT* promoter mutation, and 1p/19q codeletion. Located near the 7p telomere, *FAM20C* is a promiscuous serine kinase that localizes to the lumen of the Golgi apparatus. A recent investigation of *FAM20C* in differentiation of human dental pulp cells found that ten-eleven translocation methylcytosine dioxygenase 1 (*TET1*) binds to the promoter region of *FAM20C*, leading to increased expression of the gene due to conversion of 5-methylcytosine (5-mc) to 5-hydroxymethylcytosine (5-hmc).[34] *FAM20C* primarily phosphorylates proteins with Ser-x-Glu/pSer motifs, including approximately 80% of secreted phosphoproteins.[35] Known substrates of *FAM20C* that are implicated in various cancers include apolipoproteins,[36] insulin-like growth factor binding proteins,[37] and Serpins (serine

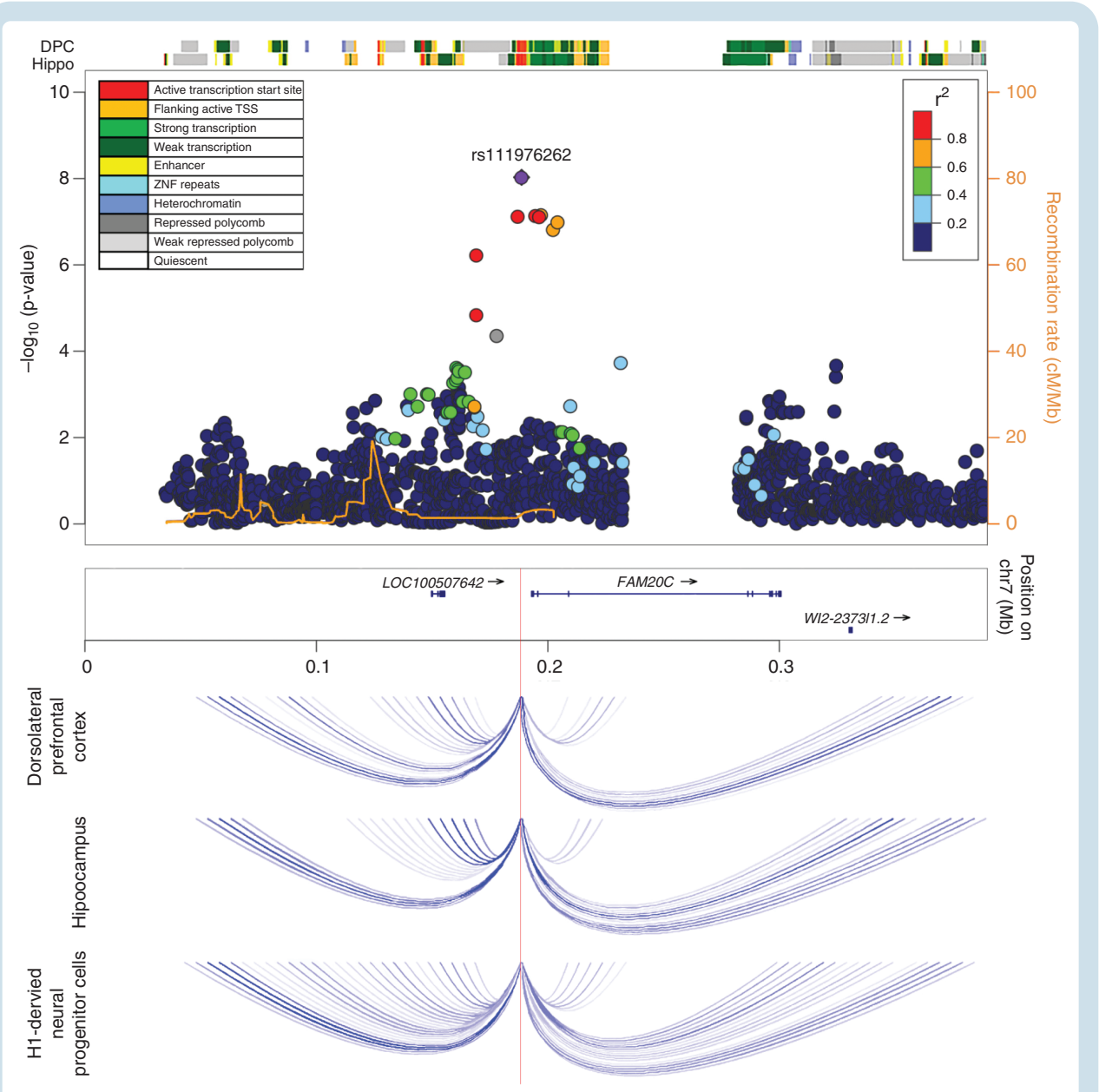

Fig. 4 *FAM20C* functional analyses. Genomic region for 200 kb surrounding rs111976262 is shown. The heatmap shows DNA states predicted by the ChromHMM algorithm in the dorsolateral prefrontal cortex (DPC) and hippocampus (Hippo) in brains from normal donors. The locus zoom plot shows the *P*-values from the meta-analysis for all variants in the region and the recombination rate. DNA-DNA interactions were measured by Hi-C analysis in the dorsolateral prefrontal cortex, hippocampus, and H1-derived neural progenitor cells; each blue line denotes an interaction.

protease inhibitors).[38] Changes in phosphorylation of one or more of these substrates due to altered expression of *FAM20C* may be responsible for the increased risk of *IDH*-mutated 1p/19q codeleted glioma in individuals carrying the rs111976262 variant.

We observed an association with variants in *GMEB2* and *IDH* wild-type glioma. *GMEB2* modulates glucocorticoid-mediated gene expression by binding to the glucocorticoid receptor.[39] The glucocorticoid receptor regulates gene expression via both transactivation of anti-inflammatory genes and transrepression of pro-inflammatory genes by binding to other transcription factors, including nuclear factor-kappaB (NF-kB).[40] Aberrant activation of the NF-kB pathway in glioma is common in *IDH* wild-type glioma, particularly in those classified as mesenchymal.[41] *GMEB2* is located on chromosome arm 20q13.33, which is a region near *RTEL1* that is known to be associated with adult diffuse glioma.[8] Hi-C interactions were observed between the *RTEL1* and *GMEB2* regions, and each of the 2 regions appears to have DNA interactions with similar loci. This suggests that both variants may be involved in DNA-DNA interactions in cells of the brain, and that they may cooperate in the disease process.

There are some limitations with this study. The small sample sizes in some of the molecular groups limited power

to detect variants with small effect sizes for these GWAS (eg, triple-negative gliomas and gliomas that are non-codeleted and have *TERT* and *IDH* mutations). Because of the limited sample size, the discovery *P*-value threshold was relaxed to 5×10^{-6}; however, the meta-analysis *P*-value threshold was set at the genome-wide level of 5×10^{-8}. We reported the findings of 9 GWAS studies, where each GWAS represented a particular molecular subtype. Because many of these GWAS were highly correlated, we utilized the accepted genome-wide significance threshold of 5×10^{-8}. Overall, the results presented herein demonstrate that novel germline variants were detected when a GWAS is performed within more homogeneous subsets of glioma. Larger collaborative efforts across institutions will be required in order to identify variants with smaller effect sizes. It is also important to acknowledge that the experimental design utilized readily available GWAS data for both discovery and validation. The discovery set included cases and controls that were all run on the OncoArray platform, and thus imputation was performed across cases and controls. The validation set included cases that were obtained from TCGA and genotyped on the Affymetrix 6.0 genotyping array, and controls that were obtained by the Mayo Clinic Biobank and genotyped on the Illumina Omni Express genotyping array. The overlap between the Affymetrix 6.0 and Illumina Omni Express arrays was ~209 000 variants, and thus in order to achieve adequate imputation results, imputation and quality control were performed within the cases and controls separately. As an additional quality control step, the alternative allele frequencies from the discovery set for the newly identified significant variants were compared with the 1000 Genome frequencies.

The discovery of glioma germline variants has helped us to understand how gliomas arise, and has opened new avenues for etiologic research. Using 25 germline variants, patient age, and sex, glioma risk models were developed to estimate relative and lifetime absolute risks of adult diffuse glioma and subtype models to predict glioma subtypes (eg, *IDH* mutated vs *IDH* wild-type).[42] The identification of additional novel germline variants will help to unravel the etiology of adult diffuse glioma, as well as improve the accuracy of these genetic-based risk models.

Supplementary Material

Supplementary data are available at *Neuro-Oncology* online.

Keywords

allergy | glioblastoma | GWAS glioma | molecular subtype

Funding

Work at Mayo Clinic was supported by National Brain Tumor Society, National Institutes of Health (P50CA108961, P30 CA15083, RC1NS068222Z, and R01 CA230712), Bernie and Edith Waterman Foundation, and Ting Tsung and Wei Fong Chao Family Foundation. Work at University of California San Francisco was supported by the National Institutes of Health (R01CA52689, P50CA097257, R01CA126831, R01CA139020, R01CA163687, and R25CA112355), as well as the Loglio Collective, the National Brain Tumor Society, 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. Work at the University of Illinois was supported by the National Brain Tumor Society.

Acknowledgments

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. 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 authors acknowledge study participants, clinicians, and research staff at participating medical centers, Katherine Cornelius, Mayo Clinic Comprehensive Cancer Center Biospecimens and Processing and Genotyping Shared Resources, the UCSF Helen Diller Family Comprehensive Cancer Center Genome Analysis Core which is supported by a National Cancer Institute Cancer Center Support Grant (5P30CA082103), the UCSF Neurosurgery Tissue Bank, and The Gliogene Consortium. The results published here include data generated by the Glioma International Case Control Study established by the Gliogene Consortium and funded by NCI (R01CA139020). Information about study can be found at the following publication: The Glioma International Case-Control Study: A Report From the Genetic Epidemiology of Glioma International Consortium. Amirian ES, Armstrong GN, Zhou R, Lau CC, Claus EB, Barnholtz-Sloan JS, Il'yasova D, Schildkraut J, Ali-Osman F, Sadetzki S, Johansen C, Houlston RS, Jenkins RB, Lachance D, Olson SH, Bernstein JL, Merrell RT, Wrensch MR, Davis FG, Lai R, Shete S, Amos CI, Scheurer ME, Aldape K, Alafuzoff I, Brännström T, Broholm H, Collins P, Giannini C, Rosenblum M, Tihan T, Melin BS, Bondy ML. Am J Epidemiol. 2016 Jan 15;183(2):85–91. doi: 10.1093/aje/kwv235. The results published here are in part based upon data generated by The Cancer Genome Atlas managed by the NCI and NHGRI.

Information about TCGA can be found at http://cancergenome.nih.gov.

Conflict of interest statement. There are no conflicts of interest to disclose.

Authorship statement. Experimental design: JEEP, DHL, MW, RBJ. Implementation: TMK, TR, AC, KLD, CP, MP, HMH, LSM, PMB, LC, JW, JKW, TCB, CG. Analysis and interpretation of the data: JEEP, PAD, MLK, AMM, KLD, TMK, AB, AA, TB, JSS, SF, DHL, MW, RBJ. All authors were involved in the writing of the manuscript and have read and approved the final version.

References

1. Jenkins RB, Xiao Y, Sicotte H, et al. A low-frequency variant at 8q24.21 is strongly associated with risk of oligodendroglial tumors and astrocytomas with IDH1 or IDH2 mutation. *Nat Genet.* 2012;44(10):1122–1125.
2. Kinnersley B, Labussière M, Holroyd A, et al. Genome-wide association study identifies multiple susceptibility loci for glioma. *Nat Commun.* 2015;6:8559.
3. Melin BS, Barnholtz-Sloan JS, Wrensch MR, et al; GliomaScan Consortium. Genome-wide association study of glioma subtypes identifies specific differences in genetic susceptibility to glioblastoma and non-glioblastoma tumors. *Nat Genet.* 2017;49(5):789–794.
4. Rajaraman P, Melin BS, Wang Z, et al. Genome-wide association study of glioma and meta-analysis. *Hum Genet.* 2012;131(12):1877–1888.
5. Shete S, Hosking FJ, Robertson LB, et al. Genome-wide association study identifies five susceptibility loci for glioma. *Nat Genet.* 2009;41(8):899–904.
6. Stacey SN, Sulem P, Jonasdottir A, et al; Swedish Low-risk Colorectal Cancer Study Group. A germline variant in the TP53 polyadenylation signal confers cancer susceptibility. *Nat Genet.* 2011;43(11):1098–1103.
7. Walsh KM, Codd V, Smirnov IV, et al; ENGAGE Consortium Telomere Group. Variants near TERT and TERC influencing telomere length are associated with high-grade glioma risk. *Nat Genet.* 2014;46(7):731–735.
8. Wrensch M, Jenkins RB, Chang JS, et al. Variants in the CDKN2B and RTEL1 regions are associated with high-grade glioma susceptibility. *Nat Genet.* 2009;41(8):905–908.
9. Sanson M, Hosking FJ, Shete S, et al. Chromosome 7p11.2 (EGFR) variation influences glioma risk. *Hum Mol Genet.* 2011;20(14):2897–2904.
10. Louis DN, Perry A, Reifenberger G, et al. The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta Neuropathol.* 2016;131(6):803–820.
11. Eckel-Passow JE, Lachance DH, Molinaro AM, et al. Glioma groups based on 1p/19q, IDH, and TERT promoter mutations in tumors. *N Engl J Med.* 2015;372(26):2499–2508.
12. Killela PJ, Pirozzi CJ, Healy P, et al. Mutations in IDH1, IDH2, and in the TERT promoter define clinically distinct subgroups of adult malignant gliomas. *Oncotarget.* 2014;5(6):1515–1525.
13. Labreche K, Kinnersley B, Berzero G, et al. Diffuse gliomas classified by 1p/19q co-deletion, TERT promoter and IDH mutation status are associated with specific genetic risk loci. *Acta Neuropathol.* 2018;135(5):743–755.
14. Labussière M, Di Stefano AL, Gleize V, et al. TERT promoter mutations in gliomas, genetic associations and clinico-pathological correlations. *Br J Cancer.* 2014;111(10):2024–2032.
15. Jenkins RB, Wrensch MR, Johnson D, et al. Distinct germ line polymorphisms underlie glioma morphologic heterogeneity. *Cancer Genet.* 2011;204(1):13–18.
16. Felini MJ, Olshan AF, Schroeder JC, et al. Reproductive factors and hormone use and risk of adult gliomas. *Cancer Causes Control.* 2009;20(1):87–96.
17. Wrensch M, Lee M, Miike R, et al. Familial and personal medical history of cancer and nervous system conditions among adults with glioma and controls. *Am J Epidemiol.* 1997;145(7):581–593.
18. Wiemels JL, Wiencke JK, Sison JD, Miike R, McMillan A, Wrensch M. History of allergies among adults with glioma and controls. *Int J Cancer.* 2002;98(4):609–615.
19. Ceccarelli M, Barthel FP, Malta TM, et al; TCGA Research Network. Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. *Cell.* 2016;164(3):550–563.
20. Olson JE, Ryu E, Johnson KJ, et al. The Mayo Clinic Biobank: a building block for individualized medicine. *Mayo Clin Proc.* 2013;88(9):952–962.
21. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics.* 2000;155(2):945–959.
22. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet.* 2006;38(8):904–909.
23. Yang D, Jang I, Choi J, et al. 3DIV: A 3D-genome interaction viewer and database. *Nucleic Acids Res.* 2018;46(D1):D52–D57.
24. Kundaje A, Meuleman W, Ernst J, et al; Roadmap Epigenomics Consortium. Integrative analysis of 111 reference human epigenomes. *Nature.* 2015;518(7539):317–330.
25. Shrine N, Portelli MA, John C, et al. Moderate-to-severe asthma in individuals of European ancestry: a genome-wide association study. *Lancet Respir Med.* 2019;7(1):20–34.
26. Ferreira MA, Vonk JM, Baurecht H, et al; 23andMe Research Team; AAGC collaborators; BIOS consortium; LifeLines Cohort Study. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. *Nat Genet.* 2017;49(12):1752–1757.
27. Zhu Z, Lee PH, Chaffin MD, et al. A genome-wide cross-trait analysis from UK Biobank highlights the shared genetic architecture of asthma and allergic diseases. *Nat Genet.* 2018;50(6):857–864.
28. Lachance DH, Yang P, Johnson DR, et al. Associations of high-grade glioma with glioma risk alleles and histories of allergy and smoking. *Am J Epidemiol.* 2011;174(5):574–581.
29. Amirian ES, Zhou R, Wrensch MR, et al. Approaching a scientific consensus on the association between allergies and glioma risk: a report from the glioma international case-control study. *Cancer Epidemiol Biomarkers Prev.* 2016;25(2):282–290.
30. Kinnersley B, Kamatani Y, Labussière M, et al. Search for new loci and low-frequency variants influencing glioma risk by exome-array analysis. *Eur J Hum Genet.* 2016;24(5):717–724.
31. Lin AP, Abbas S, Kim SW, et al. D2HGDH regulates alpha-ketoglutarate levels and dioxygenase function by modulating IDH2. *Nat Commun.* 2015;6:7768.
32. Lehrer S, Rheinstein PH, Rosenzweig KE. Allergy may confer better survival on patients with gliomas. *Clin Neurol Neurosurg.* 2019;177:63–67.
33. Ryan DG, Murphy MP, Frezza C, et al. Coupling Krebs cycle metabolites to signalling in immunity and cancer. *Nat Metab.* 2019;1:16–33.
34. Li Q, Yi B, Feng Z, Meng R, Tian C, Xu Q. FAM20C could be targeted by TET1 to promote odontoblastic differentiation potential of human dental pulp cells. *Cell Prolif.* 2018;51(2):e12426.

35. Tagliabracci VS, Wiley SE, Guo X, et al. A single kinase generates the majority of the secreted phosphoproteome. *Cell.* 2015;161(7):1619–1632.
36. Ren L, Yi J, Li W, et al. Apolipoproteins and cancer. *Cancer Med.* 2019;8(16):7032–7043.
37. Baxter RC. IGF binding proteins in cancer: mechanistic and clinical insights. *Nat Rev Cancer.* 2014;14(5):329–341.
38. Valiente M, Obenauf AC, Jin X, et al. Serpins promote cancer cell survival and vascular co-option in brain metastasis. *Cell.* 2014;156(5):1002–1016.
39. Kawabe K, Lindsay D, Braitch M, Fahey AJ, Showe L, Constantinescu CS. IL-12 inhibits glucocorticoid-induced T cell apoptosis by inducing GMEB1 and activating PI3K/Akt pathway. *Immunobiology.* 2012;217(1):118–123.
40. Caldenhoven E, van Dijk T, Raaijmakers JA, Lammers JW, Koenderman L, De Groot RP. Activation of the STAT3/acute phase response factor transcription factor by interleukin-5. *J Biol Chem.* 1995;270(43):25778–25784.
41. Verhaak RG, Hoadley KA, Purdom E, et al; Cancer Genome Atlas Research Network. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell.* 2010;17(1):98–110.
42. Eckel-Passow JE, Decker PA, Kosel ML, et al. Using germline variants to estimate glioma and subtype risks. *Neuro Oncol.* 2019;21(4):451–461.