Leveraging ethnic group incidence variation to investigate genetic susceptibility to glioma: a novel candidate SNP approach

Daniel I. Jacobs1*, Kyle M. Walsh2, Margaret Wrensch3, John Wiencke2,3, Robert Jenkins4, Richard S. Houlston5, Melissa Bondy6, Matthias Simon7, Marc Sanson8, Konstantinos Gousias9, Johannes Schramm10, Marianne Labussière4, Anna Luisa Di Stefano9, H.-Erich Wichmann10,11,12, Martina Müller-Nurasyid11,13,14, Stefan Schreiber15,16, Andre Franke16, Susanne Moebus17, Lewin Eisele18, Andrew T. Dewan1† and Robert Dubrow1†

1 Yale School of Public Health, Yale School of Medicine, New Haven, CT, USA
2 Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA, USA
3 Department of Neurological Surgery, University of California San Francisco, San Francisco, CA, USA
4 Laboratory Medicine and Pathology, Mayo Clinic College of Medicine, Rochester, MN, USA
5 Division of Genetics and Epidemiology, Institute of Cancer Research, Sutton, UK
6 Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, TX, USA
7 Neurochirurgische Universitätsklinik, Universitätskliniken Bonn, Bonn, Germany
8 Centre de Recherche de l’Institut du cerveau et de la moelle épinière, Université Pierre et Marie Curie-Paris VI, Paris, France
9 AP-HP Groupe Hospitalier Pitié-Salpétrière, Service de Neurologie Mazarin, Paris, France
10 Institute of Epidemiology I, Helmholtz Zentrum München – German Research Center for Environmental Health, Neuherberg, Germany
11 Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany
12 Klinikum Grosshadern, Munich, Germany
13 Institute of Genetic Epidemiology, Helmholtz Zentrum München – German Research Center for Environmental Health, Neuherberg, Germany
14 Department of Medicine I, University Hospital Grosshadern, Ludwig-Maximilians-Universität, Munich, Germany
15 First Medical Department, University Clinic Schleswig-Holstein, Kiel, Germany
16 Institute of Clinical Molecular Biology, Christian-Albrechts-University Kiel, Kiel, Germany
17 Institute for Medical Informatics, Biometry and Epidemiology, University Hospital of Essen, University Duisburg-Essen, Essen, Germany
18 Department of Haematology, University Hospital of Essen, University Duisburg-Essen, Essen, Germany

Edited by: Bahram Assani, University of Alabama at Birmingham, USA
Reviewed by: Stella Aslibekyan, University of Alabama at Birmingham, USA
Digna Velez Edwards, Vanderbilt University, USA
*Correspondence: Daniel I. Jacobs, Yale School of Public Health, 60 College Street, P.O. Box 208034, New Haven, CT 06520-8034, USA. e-mail: daniel.jacobs@yale.edu
†Andrew T. Dewan and Robert Dubrow have contributed equally to this work.

Objectives: Using a novel candidate SNP approach, we aimed to identify a possible genetic basis for the higher glioma incidence in Whites relative to East Asians and African-Americans. Methods: We hypothesized that genetic regions containing SNPs with extreme differences in allele frequencies across ethnicities are most likely to harbor susceptibility variants. We used International HapMap Project data to identify 3,961 candidate SNPs with the largest allele frequency differences in Whites compared to East Asians and Africans and tested these SNPs for association with glioma risk in a set of White cases and controls. Top SNPs identified in the discovery dataset were tested for association with glioma in five independent replication datasets. Results: No SNP achieved statistical significance in either the discovery or replication datasets after accounting for multiple testing or conducting meta-analysis. However, the most strongly associated SNP, rs879471, was found to be in linkage disequilibrium with a previously identified risk SNP, rs6010620, in RTEL1. We estimate rs6010620 to account for a glioma incidence rate ratio of 1.34 for Whites relative to East Asians. Conclusion: We explored genetic susceptibility to glioma using a novel candidate SNP method which may be applicable to other diseases with appropriate epidemiologic patterns.

Keywords: glioma, candidate SNP association study, ancestry informative markers, admixture, race, ethnicity, brain cancer
grade IV glioma (glioblastoma, or GBM) and non-GBM tumors (Dubrow and Darefsky, 2011).

The only established environmental risk factor for glioma is exposure to high-dose ionizing radiation (Bondy et al., 2008; Ostrom and Barnholtz-Sloan, 2011), which accounts for a small number of cases; furthermore, studies have demonstrated a consistent inverse association with history of allergy (Schoemaker et al., 2010; Lachance et al., 2011) as well as evidence of interaction effects between history of allergy and several established glioma risk alleles (Schoemaker et al., 2010). However, epidemiologic studies have provided no conclusive evidence for diagnostic radiation (Davis et al., 2011), electromagnetic field exposure from residential power lines (Wrensch et al., 1999), smoking (Mandelzweig et al., 2009), alcohol consumption (Efrid et al., 2004), nutritional factors (Bondy et al., 2008), or cell phone use (Cardis et al., 2010) as risk factors. Collectively, these observations suggest that ethnic group associated genetic variants, rather than environmental factors, underscore variation in glioma incidence among ethnic groups. Such an assertion is supported by a number of studies suggesting genetic pathways to glioma may differ across ethnicities (Mochizuki et al., 1999; Chen et al., 2001; Das et al., 2002; Wiencke et al., 2005). Following on from this it is possible that the frequencies of haplotypes associated with glioma susceptibility will differ between Whites and East Asians/Africans, such that haplotypes harboring alleles associated with an increased glioma risk would be more prevalent among Whites and conversely haplotypes associated with decreased glioma risk would be more prevalent among East Asians and Africans. Identification of these haplotypes offers the prospect of gaining valuable insight into genes influencing glioma risk.

Here we employed a candidate SNP approach to identify previously unknown genetic variants associated with glioma risk through the identification of SNPs that may tag glioma-related haplotypes. Our primary hypothesis is based on the premise that the same alleles confer protection against glioma in both East Asians and Africans. Consequently, we propose that these alleles are carried at a greater frequency by both East Asians and Africans than by Whites and that genetic regions (i.e., haplotypes) containing SNPs with the greatest allele frequency differences between Whites and both East Asians and Africans (with the same direction of difference) are particularly likely to harbor these alleles. To take into account the possibility that alleles that confer protection in East Asians differ from alleles that confer protection in Africans, we also propose a secondary hypothesis that genetic regions containing SNPs with the greatest allele frequency differences between Whites and either East Asians or Africans, but not both, are likely to harbor protective alleles which are distinct from those identified under the primary hypothesis.

Given that ethnic incidence differences are broadly similar for GBM and non-GBM glioma (Dubrow and Darefsky, 2011), we postulate that polymorphisms driving these incidence differences are common across these glioma subtypes, and therefore consider all gliomas combined without stratification. Since large differences in allele frequency are needed to account for even a relatively small portion of the White/East Asian or White/African incidence rate ratio (Figure 1), we restrict our analyses to SNPs showing the largest frequency differences.

![FIGURE 1 | White/East Asian incidence rate ratios for varying allele distributions and genotypic relative risks.](image-url)

Plots were generated by calculating incidence rate ratios (IRRs) according to varying genotypic relative risks (GRR) and ethnic group allele frequencies. For example, suppose the GRR for glioma for persons with one B allele is 2.00, and the GRR for persons with two B alleles is 3.00 (relative to those homozygous for the A allele). If the frequency of allele A in Whites is 0.20 (p = 0.2), the proportions of AA (p^2), AB (2pq), and BB (q^2) genotypes are 0.04, 0.32, and 0.64, respectively, assuming Hardy-Weinberg equilibrium. To calculate a normalized incidence rate, the genotype proportion is multiplied by the associated GRR risk: 0.04 (1.00) + 0.32 (2.00) + 0.64 (3.00) = 2.60. Given an East Asian allele A frequency of 0.80, the East Asian normalized incidence rate is 0.64 (1.00) + 0.32 (2.00) + 0.04 (3.00) = 1.40. The White/East Asian IRR is 1.86 (2.60/1.40) in this scenario. The same calculations apply for White/African IRRs.
MATERIALS AND METHODS

SELECTION OF CANDIDATE SNPs

To select candidate SNPs we used allele frequency data on unrelated individuals from six populations included in the International HapMap Project Phase III (Altshuler et al., 2010): 113 Utah residents with Northern and Western European ethnicity (CEU); 102 Toscans from Italy (TSI); 137 Han Chinese from Beijing, China (CHB); 113 Japanese from Tokyo, Japan (JPT); 147 Yoruba from Ibadan, Nigeria (YRI); and 110 Luhya from Webuye, Kenya (LWK). We grouped CEU and TSI together as Whites, CHB and JPT as East Asians, and YRI and LWK as Africans. Data from Phase III, release 28 were downloaded from the HapMap Project File Transfer Protocol1. In this release, frequency of genotype missingness per SNP was required to be <0.05 per population, and SNPs were excluded with Hardy Weinberg P < 10⁻⁵. We used in-house Perl scripts to calculate allele frequencies and call rates, according to ethnic group, for SNPs genotyped previously for our discovery genome-wide association study (GWAS) of high-grade adult glioma (Wrensch et al., 2009). SNPs with call rates <95% among Whites, East Asians, or Africans, respectively, were excluded. For each SNP, differences in reference allele frequencies were calculated for Whites vs. East Asians and Whites vs. Africans, as well as the average difference if differences were in the same direction (e.g., the frequency of allele A is low in Whites but high in both East Asians and Africans).

To test our primary hypothesis, we selected SNPs with the highest average allele frequency differences (provided equivalent directionality) as defined by three categories: “Highest” (≥0.70), “High” (0.60 to <0.70), and “Moderate” (0.40 to <0.60). To test our secondary hypothesis, we selected SNPs for which the allele frequency difference was Highest (≥0.70) in one population comparison, but Low (<0.40) in the other (“Highest/Low”).

DISCOVERY DATASET

Descriptive characteristics of discovery set cases and controls are presented in Table 1. Subjects providing genotype data for the discovery phase included 692 high-grade glioma cases and 3,992 controls originally assembled for the 2009 GWAS of glioma by Wrensch et al. (2009). Briefly, cases included 622 individuals of European ethnicity from the San Francisco Adult Glioma Study (AGS) and 70 from The Cancer Genome Atlas (TCGA; McDon- don et al., 2008), aged 20 or older with incident histologically confirmed anaplastic astrocytoma (n = 97) or GBM (n = 595; International Classification of Diseases for Oncology, morphology codes 9380–9481). Controls included 602 subjects from AGS identified using random digit dialing and frequency matched to cases on age, sex, and ethnicity, as well as 3,390 subjects from the Illumina iControl Database2. All subjects were confirmed to be unrelated and of European ethnicity by multidimensional scaling analysis.

DISCOVERY DATASET GENOTYPING

Details of sample preparation and genotyping have been provided previously (Wrensch et al., 2009). Briefly, DNA from all AGS cases and controls was isolated from whole blood using Qiagen's Gentra Puregene DNA isolation kit, and genotyping was conducted using Illumina's HumanCNV370-Duo BeadChip. AGS samples were required to have a call rate of at least 98%. SNPs deviating from Hardy Weinberg equilibrium in AGS or Illumina controls (P < 10⁻⁵) were excluded from further analysis, as were those with greater than 5% missing data in any of the four subject groups (AGS cases or controls, TCGA cases, Illumina controls).

REPLICATION DATASETS

We investigated our top candidate SNPs from the discovery dataset in five independent sets of cases and controls. Detailed procedures of subject selection and genotyping have been described previously (Shete et al., 2009; Wrensch et al., 2009; Sanson et al., 2011). Mayo Clinic cases (n = 176), 65% with GBM and 35% with grade III glioma, were diagnosed in Rochester, Minnesota between 2005 and 2008. Controls (n = 174) were identified from among individuals who had a general medical exam at the Mayo Clinic, and were matched to cases on sex, age, race, and residence. All cases

Table 1 | Discovery set subject characteristics*

| Characteristic | Adult Glioma Study (n = 622) | The Cancer Genome Atlas (n = 70) | Adult Glioma Study (n = 602) | Illumina iControls (n = 3,390) |
|---------------|------------------------------|---------------------------------|----------------------------|-----------------------------|
| Age           | 55 ± 0.5                     | 54 ± 1.7                        | 56 ± 0.6                   | 29 ± 0.4                    |
| GENDER        |                              |                                 |                            |                             |
| Male          | 398 (64%)                    | 40 (57%)                        | 319 (53%)                  | 1,254 (37%)                 |
| Female        | 224 (36%)                    | 30 (43%)                        | 283 (47%)                  | 2,136 (63%)                 |
| TUMOR SUBTYPE |                              |                                 |                            |                             |
| Grade III     | 97 (16%)                     | 0 (0%)                          | –                          | –                           |
| Grade IV      | 525 (84%)                    | 70 (100%)                       | –                          | –                           |

*From Wrensch et al. (2009).

Table values are mean ± standard error for continuous variables and n (column%) for categorical variables.
and controls were genotyped using Illumina Human 610Quad arrays.

The four other replication datasets (UK, US, French, and German) were previously included in a pooled GWAS of glioma (Shete et al., 2009; Sanson et al., 2011). Briefly, the UK GWAS comprised 631 cases ascertained through the INTERPHONE study (Cardis et al., 2010) and 2,699 controls from the 1958 Birth Cohort (Power and Elliott, 2006). The US GWAS comprised 1,247 cases recruited through MD Anderson Cancer Center in Houston, Texas and 2,236 controls from the Cancer Genetic Markers of Susceptibility study (Hunter et al., 2007). The French GWAS comprised 1,423 cases from the Service de Neurologie Mazarin, Groupe Hospitalier Pitié-Salpêtrière Paris, and 1,190 controls from the S.U.V.I.MAX study (Herceberg et al., 2004). The German GWAS comprised 846 cases recruited from the University of Bonn Medical Center, and 1,310 controls from the KORA (Holle et al., 2005; Wichmann et al., 2005), POPGEN (Krawczak et al., 2006), and Heinz Nixdorf RECALL studies (Schmermund et al., 2002). Cases in the UK and US GWAS were genotyped using Illumina Human 610Quad arrays, and cases from the French and German GWAS were genotyped using Illumina HumanHap660 arrays. Controls in the UK GWAS were genotyped using Illumina Human 1M Duo arrays; the US controls on Illumina HumanHap240, 300, and 500 arrays; the French controls on Illumina HumanHap660 arrays; and the German controls using Illumina HumanHap550 arrays.

**STATISTICAL ANALYSES**

Odds ratios and 95% confidence intervals for the association of candidate SNPs with glioma in the discovery and replication sets were calculated using unconditional logistic regression under an additive model (0, 1, or 2 copies of the minor allele). Potential population stratification was adjusted for using principal components derived by the EIGENSTRAT method and included in the logistic regression model (Price et al., 2006).

Discovery set results were evaluated in comparison to Bonferroni-adjusted significance thresholds based on a study-wide significance threshold of 0.05, calculated separately for each of the four subgroups of candidate SNPs [Highest allele frequency difference (≥0.70), High (0.60 to <0.70), Moderate (0.40 to <0.60), and Highest/Low (≥0.70 in one population, <0.40 in the other population)] such that the significance thresholds accounted for the prior probability of association with glioma according to our hypotheses (i.e., $P = 0.0125$ per subgroup). With 38 SNPs in the Highest allele frequency difference category, the significance threshold for this category was $0.0125/38 = 3.29 \times 10^{-4}$. Statistical thresholds for the High, Moderate, and Highest/Low subgroups were $4.70 \times 10^{-5}$, $4.34 \times 10^{-6}$, and $1.61 \times 10^{-5}$, respectively (Table 2). For replication set analyses we used a nominal significance level of 0.05. All $P$-values reported (discovery and replication) are one-sided because of the directionality inherent in the hypothesis being tested. The generic inverse variance method was used (assuming a fixed effects model) to obtain meta-analysis results for combined discovery and replication set data.

**RESULTS**

**SELECTED CANDIDATE SNPs**

Of 275,895 SNPs for which genotype data on discovery glioma cases and controls were previously available, HapMap data were not available or were of insufficient quality for 1,188 (0.43%). Pre-determined allele frequency difference criteria were met for 3,961 of the 274,707 remaining SNPs (Figure 2). We identified 2,883 SNPs in the Moderate (0.40 to <0.60) allele frequency difference category, 266 in the High (0.60 to <0.70) category, 38 in the Highest (≥0.70) category, and 774 in the Highest/Low category (≥0.70 in one population comparison, but <0.40 in the other; Table 2).

**DISCOVERY SET**

A Manhattan plot of the 3,961 SNP-glioma associations is shown in Figure 4. The most strongly associated SNP, rs879471 in STM3N on chromosome 20q13 ($P = 1.72 \times 10^{-10}$), maps 39.9 kb from rs6010620, a SNP intronic to RETL1 that was previously identified as a top hit in the GWAS conducted by Wrensch et al. (2009). Conditioning rs879471 on rs6010620 did not, however, provide evidence of a separate signal ($P = 0.22$), so rs879471 was excluded from further analysis.

While no SNP association attained our predetermined Bonferroni-adjusted significance levels (Table 2), 10 genes (SMARCA2, BRE, SLCO3A1, MORN5, C10orf11, RBM27, PTPRJ, SMARCA2, C10orf11, PTPRJ, and NMNAT1) were identified containing at least one SNP with $P < 0.01$. In order to investigate SNPs that may be markers of glioma risk but were excluded by our strict allele frequency criteria, we tested 260 additional SNPs in these 10 genes and within 5 kb upstream and downstream for association with glioma risk (regardless of ethnic allele frequency differences). We excluded four genes (SLCO3A1, C10orf11, PTPRJ, and NMNAT1) from further analysis because the direction of

**Table 2 | Number of selected SNPs by allele frequency difference category with corresponding Bonferroni-adjusted significance thresholds.**

| Category               | Number of SNPs | Significance threshold |
|------------------------|----------------|-----------------------|
| Primary hypothesis     |                |                       |
| Highest mean allele frequency difference (≥0.70) | 38             | $3.29 \times 10^{-4}$  |
| High mean allele frequency difference (0.60 to <0.70) | 266            | $4.70 \times 10^{-5}$  |
| Moderate mean allele frequency difference (0.40 to <0.60) | 2,883          | $4.34 \times 10^{-6}$  |
| Secondary hypothesis   |                |                       |
| Highest/low allele frequency differences (≥0.70 in one comparison and <0.40 in the other) | 774             | $1.61 \times 10^{-5}$  |
FIGURE 2 | Flow diagram of candidate SNP selection from HapMap.

FIGURE 3 | White vs. African and White vs. East Asian allele frequency differences and selected candidate SNPs. Each SNP is plotted by its allele frequency difference between Whites and Africans vs. its allele frequency difference between Whites and East Asians. Green SNPs in the upper-right and lower-left quadrants represent those with mean allele frequency differences of 0.40 to <0.60 (Moderate), red SNPs represent those with mean allele frequency differences of 0.60 to <0.70 (High), and black SNPs represent those with mean allele frequency differences ≥0.70 (Highest). Orange SNPs around the perimeter represent those in which the allele frequency difference was at least 0.70 in one comparison, but less than 0.40 in the other (Highest/Low).
There were some limitations of this study which may have prevented the detection of glioma-associated SNPs. While our allele frequency difference criteria for candidate SNPs were designed to be inclusive of SNPs that could be responsible for a meaningful risk difference across ethnic groups, it is possible that these criteria excluded SNPs that are, in fact, associated with glioma risk but did not meet our criteria. Additionally, given 692 cases and 3,992 controls in our discovery set, our power calculations demonstrate that for a moderately common putative risk allele (0.20 allele frequency), we had 80% power to detect an odds ratio as low as 1.46. Yet for a relatively rare risk allele (0.05 allele frequency), we had 80% power to detect an odds ratio no lower than 1.88. Thus, it is plausible that our set of candidate SNPs includes one or more variants with low to moderate association with glioma risk, but that we were underpowered to detect such an association. Furthermore, it should be noted that although we postulated that polymorphisms driving ethnic group incidence differences are common across glioma subtypes, it is possible that differences in the glioma subtype distribution in the discovery and replication sets impacted the replicability of our findings. The ability of our approach to detect SNPs that tag glioma-related haplotypes may also have been degraded by heterogeneity across ethnic groups in the haplotype that a given tagging SNP represents. Finally, we note that our study was unable to assess potential interaction effects between risk loci, gene-environment interactions, or the role of rare variants.

Although this study did not lead to the discovery of novel glioma-associated SNPs, it is noteworthy that our most strongly associated candidate SNP, rs879471, was in strong linkage disequilibrium with rs6010620 ($D' = 0.78$ in HapMap CEU + TSI, data from Haploview version 4.2), a top hit from the Wrensch et al. (2009) GWAS. This suggests the successful identification of a haplotype that differs in frequency across ethnic groups and
Table 3 | Discovery and replication set results.

| rsID    | Gene     | Discovery set | UCSF discovery set (n = 692) |
|---------|----------|---------------|------------------------------|
|         |          | P-valuea      |                             |
| rs1127125 | BRE      | 5.75E-03      | 0.83 (0.71, 0.96)           |
|         |          |               | 1.25 (0.87, 1.81)           |
|         |          |               | 0.89 (0.77, 1.04)           |
| rs4464229 | BRE      | 5.05E-04      | 0.77 (0.65, 0.90)           |
|         |          |               | 1.37 (0.93, 2.01)           |
| rs1666020 | BRE      | 2.90E-04      | 0.76 (0.64, 0.89)           |
|         |          |               | 1.37 (0.93, 2.02)           |
| rs4666022 | BRE      | 3.37E-03      | 0.82 (0.70, 0.95)           |
| rs10175508 | BRE     | 7.40E-04      | 0.79 (0.68, 0.91)           |
|         |          |               | 1.21 (0.84, 1.74)           |
| rs10173528 | BRE     | 9.00E-04      | 0.86 (0.77, 0.97)           |
|         |          |               | 1.11 (0.82, 1.50)           |
| rs10173426 | BRE     | 4.09E-03      | 0.85 (0.75, 0.96)           |
|         |          |               | 1.10 (0.80, 1.50)           |
| rs13402525 | BRE     | 5.40E-03      | 0.85 (0.75, 0.96)           |
|         |          |               | 1.08 (0.79, 1.48)           |
| rs3963   | BRE      | 9.10E-03      | 1.15 (1.02, 1.30)           |
| rs11131194 | RPUSD3  | 1.98E-03      | 0.84 (0.75, 0.95)           |
|         |          |               | 0.73 (0.53, 1.02)           |
| rs2279696 | RBM27    | 6.95E-03      | 1.17 (1.03, 1.33)           |
|         |          |               | 0.96 (0.89, 1.34)           |
| rs11952506 | RBM27   | 7.75E-04      | 1.23 (1.06, 1.40)           |
|         |          |               | 0.86 (0.62, 1.21)           |
| rs19930900 | RBM27   | 1.12E-03      | 1.22 (1.07, 1.39)           |
|         |          |               | 0.87 (0.62, 1.22)           |
| rs2304032 | RBM27    | 1.05E-03      | 1.22 (1.06, 1.39)           |
|         |          |               | 0.87 (0.62, 1.22)           |
| rs13551519 | RBM27   | 1.17E-03      | 1.22 (1.07, 1.39)           |
|         |          |               | 0.87 (0.62, 1.22)           |
| rs2839070 | SMARCA2  | 2.69E-03      | 0.76 (0.63, 0.92)           |
|         |          |               | 0.72 (0.42, 1.24)           |
| rs2838007 | SMARCA2  | 2.17E-04      | 0.71 (0.58, 0.96)           |
|         |          |               | 1.00 (0.61, 1.66)           |
| rs10946907 | SMARCA2 | 2.19E-03      | 0.74 (0.61, 0.91)           |
|         |          |               | 0.88 (0.51, 1.52)           |
| rs870272 | MORN5    | 1.02E-03      | 0.83 (0.73, 0.93)           |
|         |          |               | 1.04 (0.75, 1.44)           |
| rs7454515 | SPIB     | 1.60E-03      | 0.81 (0.70, 0.93)           |
|         |          |               | 0.78 (0.54, 1.14)           |

| Maya Clinic (n = 176) | UK series (n = 631) | US series (n = 1,247) | French series (n = 1,423) | German series (n = 846) | Meta-analysis (n = 5,015)d | P-value for heterogeneityb |
|-----------------------|---------------------|-----------------------|---------------------------|-------------------------|-----------------------------|-----------------------------|
| 0.89 (0.77, 1.04)    | 1.17 (1.03, 1.31)*  | 1.06 (0.93, 1.21)    | 1.00 (0.86, 1.17)         | 0.99 (0.93, 1.05)       | 0.61                         |
| 0.77 (0.65, 0.90)    | NA                  | NA                   | NA                        | NA                      | 0.91 (0.81, 1.02)           | 0.28                         |
| 0.91 (0.77, 1.07)    | 1.16 (1.02, 1.31)*  | 1.04 (0.90, 1.20)    | 1.02 (0.87, 1.19)         | 0.98 (0.92, 1.04)       | 0.41                         |
| 0.82 (0.70, 0.95)    | NA                  | NA                   | NA                        | NA                      | NA                          | 0.77                         |
| 0.89 (0.77, 1.04)    | 1.13 (1.01, 1.27)*  | 1.06 (0.93, 1.21)    | 1.00 (0.86, 1.16)         | 0.98 (0.93, 1.04)       | 0.46                         |
| 0.92 (0.81, 1.05)    | 1.02 (0.92, 1.13)   | 1.00 (0.90, 1.12)    | 1.10 (0.97, 1.24)         | 0.98 (0.94, 1.03)       | 0.37                         |
| 0.89 (0.76, 1.01)*   | 1.00 (0.90, 1.11)   | 0.98 (0.87, 1.09)    | 1.08 (0.86, 1.19)         | 0.98 (0.93, 1.03)       | 0.41                         |
| 0.92 (0.80, 1.50)    | NA                  | NA                   | NA                        | NA                      | NA                          | 0.77                         |
| 0.89 (0.78, 1.01)*   | 1.00 (0.90, 1.11)   | 0.98 (0.87, 1.09)    | 1.08 (0.86, 1.23)         | 0.98 (0.93, 1.03)       | 0.81                         |

Odds ratio (two-sided 95% confidence interval)

- P-values are one-sided.
- n = number of cases.
- NA, data not available.
- Heterogeneity assessed using Cochran's Q statistic (Cochran, 1954).
- *Statistical significance at P < 0.05.
is related to glioma risk. In this respect the A allele of rs6010620, which is protective against glioma (OR = 0.68, 95% CI: 0.58–0.79), is considerably more common in East Asians than Whites (frequency of 0.697 vs. 0.228, respectively), but did not meet our strict allele frequency difference criteria because the frequency in Africans was 0.019. On the basis of a calculation similar to that presented in Figure 1, we would conclude that this SNP is sufficient to account for an incidence rate ratio of 1.34 for Whites relative to East Asians. Notably, another top hit from the same GWAS, rs1412829, has a risk allele C (OR = 1.39, 95% CI: 1.24–1.57) that is more common in Whites than East Asians or Africans but also did not meet our allele frequency criteria (frequency of 0.402, 0.104, and 0.009, respectively). Based on our calculations, this SNP can account for a White to East Asian incidence rate ratio of 1.22, and a White to African incidence rate ratio of 1.30. Thus, these two SNPs alone may account for a meaningful proportion of the observed inter-ethnic incidence rate ratios. None of the other five established glioma susceptibility loci contribute to the inter-ethnic incidence rate differences (Table 4).

Table 4 | Contribution to ethnic incidence rate ratios by established glioma susceptibility loci.

| SNP       | Reference allele frequency | Odds ratio (95% confidence interval)* | White-East Asian incidence rate ratio | White-African incidence rate ratio |
|-----------|----------------------------|--------------------------------------|---------------------------------------|------------------------------------|
| rs6010620 | A 0.23 0.70 0.02           | 0.68 (0.58, 0.79)                     | 1.34                                  | 0.90                               |
| rs1412829 | C 0.40 0.10 0.01           | 1.39 (1.24, 1.57)                     | 1.22                                  | 1.30                               |
| rs2736100 | G 0.54 0.41 0.42           | 1.27 (1.19, 1.37)                     | 1.05                                  | 1.05                               |
| rs4299527 | G 0.15 0.24 0.16           | 1.36 (1.29, 1.43)                     | 0.95                                  | 0.99                               |
| rs4977756 | G 0.37 0.23 0.34           | 1.24 (1.19, 1.30)                     | 1.06                                  | 1.01                               |
| rs4988772 | T 0.28 0.24 0.08           | 1.18 (1.13, 1.24)                     | 1.01                                  | 1.07                               |
| rs2252586 | T 0.28 0.02 0.32           | 1.18 (1.11, 1.25)                     | 1.09                                  | 0.99                               |

*Odds ratios were calculated according to an additive model (0, 1, or 2 copies of the reference allele).

While we did not identify novel glioma susceptibility variants in this analysis, we conclude that the additional risk in White populations conferred by rs6010620 and rs1412829 lends support to our initial hypothesis, and provides an impetus for a larger discovery set and/or pursuing admixture mapping. Given the rarity of glioma among African-Americans and the resultant difficulties inherent in collecting enough African-American cases to perform an admixture mapping study, further application of our method may be the preferred approach.

ACKNOWLEDGMENTS

This study was supported by the Nancy Hildreth Memorial Summer Fellowship in Chronic Disease Epidemiology, Yale School of Public Health. Work at University of California, San Francisco was supported by the National Institutes of Health (grant numbers R01CA52689, P50CA097257, R25CA126831, and CA112355), as well as the National Brain Tumor Foundation, the UCSF Lewis Chair in Brain Tumor Research and by donations from families and friends of John Berardi, Helen Glaser, Elvera Olsen, Raymond E. Cooper, and William Martinusen. Assembly of the French cohort was supported by Ligue Nationale contre le Cancer and the Institut National du Cancer (INCA; PL 046). Computational work was supported by the Yale University Biomedical High Performance Computing Center and the National Institutes of Health (RR19895). We thank Drs. Josephine Hoh, Yuri Galachyants, and Alan Wu for helpful early conceptual discussions.
REFERENCES

Altshuler, D. M., Gibbs, R. A., Peltonen, L., Dermitzakis, E., Schaffner, S. F., Yu, F., et al. (2010). Integrating common and rare genetic variation in diverse human populations. Nature 467, 52–58.

Barrett, J. C., Fry, B., Maller, J., and Daly, M.J. (2005). Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21, 263–265.

Blankenberg, D., Küster, G. V., Coraor, N., Anzanda, G., Lazarus, R., Man- gan, M., et al. (2011). Galaxy: a web-based genome analysis tool for experimentalists. Curr. Protoc. Mol. Biol. 89, 1–21.

Bondy, M. L., Scheurer, M. E., Malfer, B., Barnholtz-Sloan, J. S., Davis, F. G., Il'Yasova, D., et al. (2008). Brain tumor epidemiology: consensus from the Brain Tumor Epidemiology Consortium. Cancer 113, 1953–1968.

Cardis, E., Deltoeur, L., Vrijheid, M., Combã, F., Moissonnier, M., Tardy, H., et al. (2010). Brain tumour risk in relation to mobile telephone use: results of the INTERPHONE international case–control study. Int. J. Epidemiol. 39, 675–684.

Chen, P., Aldape, K., Wierc, J. K., Kelsey, K. T., Miki, R., Davis, R. L., et al. (2001). Ethnicity delineates different genetic pathways in malignant glioma. Cancer Res. 61, 3949–3954.

Chrochon, W. G. (1954). The combination of estimates from different experiments. Biometrics 10, 101–129.

Darefsky, A. S., and Dubrow, R. (2009). International variation in the incidence of adult malignant neoplasms of the brain and central nervous system. Cancer Causes Control 20, 1593–1604.

Das, A., Tan, W.-L., Teo, J., and Smith, D. R. (2002). Glioblastoma multiforme in an Asian population: evidence for a distinct genetic pathway. J. Neurooncol. 60, 117–125.

Davis, F., Il'Yasova, D., Rankin, K., McCarthy, B., and Bigner, D. D. (2011). Medical diagnostic radiation exposure and risks of glioma. Radiat. Res. 175, 790–796.

Dubrow, R., and Darefsky, A. S. (2011). Demographic variation in incidence of adult glioma by subtype, United States, 1992–2007. Appl. Radiat. Isot. 69, 325–325. doi:10.1016/j.apradiso.2011.11.006

Efrid, I., Friedman, G., Sidney, S., Klatasky, A., Habel, L., Udaltsova, N., et al. (2004). The risk for malignant primary-onset glioma in a large, multietnic, managed-care cohort: cigarette smoking and other lifestyle behaviors. J. Neurooncol. 68, 57–69.

Herceg, S., Galan, P., Preziosi, P., Bertrán, S., Mennén, L., MaIpy, D., et al. (2004). The SU.VI.MAX study: a randomized, placebo-controlled trial of the health effects of antioxidants and vitamins and minerals. Arch. Intern. Med. 164, 2335–2342.

Holle, R., Happich, M., Löwel, H., and Wichmann, H. E. (2005). KORA – a research platform for population-based health research. Gesundheitswesen 67, 19–25.

Hunt, D. J., Kraft, P., Jacobs, K. B., Cox, D. G., Yeager, M., Hankinson, S. E., et al. (2007). A genome-wide association study identifies alleles in FGF2 associated with risk of sporadic postmenopausal breast cancer. Nat. Genet. 39, 870–874.

Kohler, B. A., Ward, E., McCarthy, B. J., Schymura, M. J., Ries, L. G., Eberhardt, C., et al. (2011). Annual report to the nation on the status of cancer, 1975–2007, featuring tumors of the brain and other nervous system. J. Natl. Cancer Inst. 103, 1–23.

Krakwacz, M., Nikolaus, S., Von Ebner, H., Croucher, P. J. P., El Moshtake, N. E., and Schreiber, S. (2006). PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. Public Health Genomics 9, 55–61.

Lachance, D. H., Yong, P., Johnson, D. R., Decker, P. A., Rollinney, T. M., McCoy, L. S., et al. (2011). Associations of high-grade glioma with glioma risk alleles and histories of allergy and smoking. Am. J. Epidemiol. 174, 574–581.

Levy, D., Ehret, G. B., Rice, K., Verwoert, G. C., Launer, L. J., Dehghan, A., et al. (2007). A genome-wide association study identifies five susceptibility loci for glioma. Hum. Mol. Genet. 20, 2897–2904.

Schmermund, A., Möhlenkamp, S., Stang, A., Grünermeyer, D., Seibel, R., Hirche, H., et al. (2002). Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: rationale and design of the Heinz Nixdorf RECALL Study. Am. Heart J. 144, 212–218.

Schoemaker, M. J., Robertson, L., Wiertz, A., Jones, M. E., Hosking, F. J., Fuchting, M., et al. (2010). Interaction between 5 genetic variants and allergy in glioma risk. Am. J. Epidemiol. 171, 1165–1173.

Shete, S., Hosking, F. J., Robertson, L. B., Dobkins, S. E., Sanso, M., Malmer, B., et al. (2009). Genome-wide association study identifies five susceptibility loci for glioma. Nat. Genet. 41, 899–904.

Smith, M. W., and O'Brien, S. J. (2005). Mapping by admixture linkage disequilibrium: advances, limitations and guidelines. Nat. Rev. Genet. 6, 623–632.

Wichmann, H. E., Geiger, C., and Illg, T. (2005). KORA-Gen: resource for population genetics, controls and a broad spectrum of disease phenotypes. Gesundheitswesen 67, 425–32.

Wiencke, J. K., Aldepe, K., McMillan, A., Wiersch, M., Moghadassi, M., Miki, R., et al. (2005). Molecular features of adult glioma associated with patient race/ethnicity, age, and a polymorphism in O6-methylguanine-DNA methyltransferase. Cancer Epidemiol. Biomarkers Prev. 14, 1774–1783.

Wrensch, M., Jenkins, R. B., Chang, J. S., Yeh, R.-F., Xiao, Y., Decker, P. A., et al. (2009). Variants in the CDKN2B and RETL1 regions are associated with high-grade glioma susceptibility. Nat. Genet. 41, 905–908.

Wrensch, M., Yost, M., Miki, R., Lee, G., and Touchstone, J. (1999). Adult glioma in relation to residential power frequency electromag- netic field exposures in the San Francisco Bay area. Epidemiology 10, 523–527.

Zhu, X., Young, J. H., Fox, E., Keating, B. J., Franceschini, N., Kang, S., et al. (2011). Combined admixture mapping and association analysis identifies a novel blood pressure genetic locus on 5p13: contributions from the CARe consortium. Hum. Mol. Genet. 20, 2285–2295.

Conflict of Interest Statement: The authors declare that there is no conflict of interest.

Received: 20 August 2012; accepted: 20 September 2012; published online: 12 October 2012.

Citation: Jacobs DI, Walsh KM, Wrensch M, Wiencek J, Jenkins R, Houlston BS, Bondy M, Simon M, Sanov S, Goussia K, Schramm J, Labussière M, Di Stefano AL, Wichmann H-E, Müller-Nurayisy M, Schrander-Schlauss E, Franke T, Ame G, Schramm J, Di Stefano M, Eisele L, Dewan AT and Dubrow R (2012) Leveraging ethnic group inci- dence variation to investigate genetic sus- ceptibility to glioma: a novel candidate SNP approach. Front. Genet. 3:203. doi: 10.3389/fgene.2012.00203

This article was submitted to Frontiers in Applied Genetic Epidemiology, a specialty of Frontiers in Genetics. Copyright © 2012 Jacobs, Walsh, Wrensch, Wiencek, Jenkins, Houlston, Bondy, Simon, Sanov, Goussia, Schramm, Labusssiere, Di Stefano, Wichmann, Müller-Nurayisy, Schrander, Franke, Ame, Schramm J, Eisele L, Dewan AT and Dubrow R (2012) Leveraging ethnic group inci- dence variation to investigate genetic sus- ceptibility to glioma: a novel candidate SNP approach. Front. Gene. 3:203. doi: 10.3389/fgene.2012.00203

This article is an open-access article distributed under the terms of the Creative Com- mons Attribution License, which per- mitts use, distribution and reproduction in other forums, provided the original authors and source are credited and sub- ject to any copyright notices concerning any third-party graphics etc.