Genome-Wide High-Density SNP Linkage Search for Glioma Susceptibility Loci: Results from the Gliogene Consortium

Sanjay Shete1, Ching C. Lau2, Richard S. Houlston3, Elizabeth B. Claus4,5, Jill Barnholtz-Sloan6, Rose Lai7, Dora Il'yasova8, Joellen Schildkraut9, Siegal Sadetzki10,11, Christoffer Johansen12, Jonine L. Bernstein8, Bridget J. McCarthy17, Eastwood Hon-chiu Leung2, Caleb Davis2, Rita Cheng2, Fay J. Hosking3, Beatrice S. Melin14, and Melissa L. Bondy2

Abstract

Gliomas, which generally have a poor prognosis, are the most common primary malignant brain tumors in adults. Recent genome-wide association studies have shown that inherited susceptibility plays a role in the development of glioma. Although first-degree relatives of patients exhibit a two-fold increased risk of glioma, the search for susceptibility loci in familial forms of the disease has been challenging because the disease is relatively rare, fatal, and heterogeneous, making it difficult to collect sufficient biosamples from families for statistical power. To address this challenge, the Genetic Epidemiology of Glioma International Consortium (Gliogene) was formed to collect DNA samples from families with two or more cases of histologically confirmed glioma. In this study, we present results obtained from 46 U.S. families in which multipoint linkage analyses were undertaken using nonparametric (model-free) methods. After removal of high link disequilibrium single-nucleotide polymorphism, we obtained a maximum nonparametric linkage score (NPL) of 3.39 ($P = 0.0005$) at 17q12-21.32

Authors’ Affiliations: 1Department of Epidemiology, The University of Texas MD Anderson Cancer Center; 2Department of Pediatrics, Division of Hematology-Oncology, Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, Texas; 3Section of Cancer Genetics, Institute of Cancer Research, Sutton, Surrey, United Kingdom; 4Department of Epidemiology and Public Health, Yale University School of Medicine, New Haven, Connecticut; 5Department of Neurosurgery, Brigham and Women’s Hospital, Boston, Massachusetts; 6Case Comprehensive Cancer Center, Case Western Reserve University School of Medicine, Cleveland, Ohio; 7The Neurological Institute of Columbia University; 8Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, New York; 9Control and Prevention Program, Department of Community and Family Medicine, Duke University Medical Center, Durham, North Carolina; 10Cancer and Radiation Epidemiology Unit, Gertner Institute, Chaim Sheba Medical Center, Tel Hashomer, Israel; 11Sackler School of Medicine, Tel-Aviv University, Tel-Aviv, Israel; 12Department of Neurology; 13Institute of Cancer Epidemiology, Danish Cancer Society, Copenhagen, Denmark; 14Mayo Clinic Comprehensive Cancer Center, Mayo Clinic, Rochester, Minnesota; 15Department of Radiation Sciences Oncology, Umea University, Umea, Sweden; 16Evangelist Kellogg Cancer Care Center NorthShore University HealthSystem, Evanston, Illinois; 17Department of Neurological Surgery, University of California, San Francisco, San Francisco, California; and 18Division of Epidemiology and Biostatistics, University of Illinois at Chicago, Chicago, Illinois

The members of the Gliogene Consortium: Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, Texas (Sanjay Shete, Robert K. Yu); Department of Genetics, The University of Texas MD Anderson Cancer Center, Houston, Texas (Christopher Amos); Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, Texas (Kenneth D. Aaldape); Department of Neuro-Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas (Jeffrey Weinberg); Department of Pediatrics, Section of Hematology and Oncology, Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, Texas (Ching C. Lau, Eastwood Hon-chiu Leung, Caleb Davis, Rita Cheng, Chloe Man, Rudy Guerra, Sivashan karappap Gurusiddappa, Michael E. Scheuer, Melissa L. Bondy, Georgina N. Armstrong, Yanhong Liu); Section of Cancer Genetics, Institute of Cancer Research, Sutton, United Kingdom (Richard S. Houlston, Fay J. Hosking, Lindsay Robertson, Eli Papamannulis); Department of Epidemiology and Public Health, Yale University School of Medicine, New Haven, Connecticut (Elizabeth B. Claus); Department of Neurosurgery, Brigham and Women’s Hospital, Boston, Massachusetts (Elizabeth B. Claus); Case Comprehensive Cancer Center, Case Western Reserve University School of Medicine, Cleveland, Ohio (Jill Barnholtz-Sloan, Andrew E. Sloan, Gene Barnett, Karen Devine, Yingli Wolinsky); The Neurological Institute of Columbia University, New York, New York (Rose Lai, Erika Florendo, Delciaivas, Christina Corpuz); Cancer Control and Prevention Program, Department of Community and Family Medicine, Duke University Medical Center, Durham, North Carolina (Dora Il’yasova, Joellen Schildkraut); Cancer and Radiation Epidemiology Unit, Gertner Institute, Chaim Sheba Medical Center, Tel Hashomer, Israel (Siegal Sadetzki, Galit Hirsh Yecheskel, Revital Bar-Sade); Sackler School of Medicine, Tel-Aviv University, Tel-Aviv, Israel (Sigal Sadetzki); Department of Neurology; Institute of Cancer Epidemiology, Danish Cancer Society, Copenhagen, Denmark (Christoffer Johansen, Hanne Bodcher); Neurosurgery Department, Rigshospitalet, University Copenhagen (Michael Kostelanetz); Neurology Department, Rigshospitalet, University Copenhagen (Helle Brøholm); Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, New York (Jonine L. Bernstein, Sara H. Olson, Erica Schubert); Department of Neurology, Memorial Sloan-Kettering Cancer Center, New York, New York (Lisa Defelipe); Mayo Clinic Comprehensive Cancer Center, Mayo Clinic, Rochester, Minnesota (Robert B. Jenkins, Ping Yang, Amanda Rynearson); Department of Radiation Sciences Oncology, Umea University, Umea, Sweden (Britnace S. Melin, Roger Henriksen, Ulrika Andersson); Department of Medical Biosciences, Umea University, Umea, Sweden (Thomas Brannstrom); Evanston Kellogg Cancer Care Center, North Shore University Health System, Evanston, Illinois (Nicholas A. Vick); Departments of Neurological Surgery and Epidemiology and Biostatistics (Margaret Wrensch, John Wieneke, Joe Wiemels, Lucie McCoy) Division of Epidemiology and Biostatistics, University of Illinois at Chicago, Chicago, Illinois (Bridget J. McCarthy, Faith G. Davis).

Note: Sanjay Shete, Ching C. Lau, Richard S. Houlston, Beatrise S. Melin, and Melissa L. Bondy contributed equally to this work.

Corresponding Authors: Sanjay Shete, and Melissa Bondy, Department of Epidemiology, The University of Texas MD Anderson Cancer Center, 1155 Pressler Street, Unit 1340, Houston, TX 77030, Phone: 713-745-3020, Fax: 713-792-8261, E-mail: sshete@mdanderson.org and mbondy@bcm.edu

doi: 10.1158/0008-5472.CAN-11-0013

© 2011 American Association for Cancer Research.
and the Z-score of 4.20 \( (P = 0.000007) \). To replicate our findings, we genotyped 29 independent U.S. families and obtained a maximum NPL score of 1.26 \( (P = 0.008) \) and the Z-score of 1.47 \( (P = 0.035) \). Accounting for the genetic heterogeneity using the ordered subset analysis approach, the combined analyses of 75 families resulted in a maximum NPL score of 3.81 \( (P = 0.00001) \). The genomic regions we have implicated in this study may offer novel insights into glioma susceptibility, focusing future work to identify genes that cause familial glioma. Cancer Res 71(24): 7568–75. ©2011 AACR.

Introduction

Gliomas account for approximately 40% of all primary malignant brain tumors (PBT) and are responsible for approximately 13,000 cancer-related deaths in the United States each year. Irrespective of treatment, most gliomas are associated with a poor prognosis with the most common type of glioma, glioblastoma (GBM), having a median overall survival of only 10 to 15 months (1).

Evidence strongly suggests that inherited susceptibility plays a role in the development of glioma, as first-degree relatives of patients with glioma have a 2-fold increased risk of glioma (2–4). Hereditary genetic disorders such as neurofibromatosis type I and II, and Li-Fraumeni and Turcot’s syndromes are known to predispose to glioma (5, 6). These syndromes are, however, rare and collectively make only a minor contribution to the familial risk of glioma.

Families segregating glioma outside the context of these syndromes provide a strong rationale for seeking to identify moderate-high risk susceptibility loci for glioma through genome-wide linkage scans. To date only one genome-wide linkage scan of glioma has been conducted, which was based on the analysis of 4 Finnish families. Evidence of linkage of glioma to 15q23–26 was provided; however, no genetic mutation was identified in this region (7).

To facilitate the collection of glioma families informative for linkage analysis, we established the “Genetic Epidemiology of Glioma International Consortium” (Gliogene-Linkage) in 2006 (8). Gliogene now includes 15 institutions in the United States, United Kingdom, Sweden, Denmark, and Israel. We have undertaken a genome-wide scan of 75 glioma families in the United States ascertained through Gliogene. This search was conducted using high-density single-nucleotide polymorphism (SNP) arrays, thereby allowing us to maximize the power to identify a disease-causing locus. Here, we report evidence for Mendelian predisposition to glioma and strong evidence for a disease locus at 17q12–21.32.

Materials and Methods

Ascertainment and collection of families

Forty-six glioma families with at least 2 biologically related family members diagnosed with a histologically confirmed glioma [International Classification of Disease—OncoLOGY codes: low grade glioma (WHO grade I, II); juvenile pilocytic astrocytoma (9421/3), fibrillary astrocytoma (9420/3), protoplasmic astrocytoma (9410/3), gemistocytic astrocytoma (9411/3), diffuse astrocytoma (9400/3), oligodendroglioma (9450/3), oligoastrocytoma (9382/3); ependymoma (9391/3), high grade glioma (WHO grade III and IV); anaplastic astrocytoma (9401/3), anaplastic oligodendroglioma (9451/3), anaplastic oligoastrocytoma (9382/3), anaplastic ependymoma (9392/3), gliosarcoma (9442/3), gliomatosis cerebri (9381/3), and glioblastoma multiforme (9440/3)] were ascertained through the Gliogene consortium. We excluded all families with reported or confirmed diagnosis of neurofibromatosis 1, neurofibromatosis 2, Turcot’s syndrome or tuberous sclerosis. The scheme for recruitment, and data collection of families has been previously described (8). These 49 families were ascertained and genotyped in years 2004 to 2009, and are referred as the first stage families. An additional 29 families were collected using the same sampling criterion in 2010 and genotyped in 2011. These 29 families are referred to in this manuscript as the second stage or the replication families. All members of the included families were self-reported to be non-Hispanic white. Blood or saliva samples were obtained from both the offspring and spouse of deceased affected family members whenever possible to facilitate the genotype reconstruction of deceased family members. DNA was extracted from EDTA-venous blood samples and saliva samples that were collected using the Oragene kits (DNA Genotek). Biosamples and clinicopathologic information from patients and family members was collected with informed consent according to protocols approved by each center’s Institutional Review Board in accordance with the tenets of the Declaration of Helsinki.

Genotyping

Before genotyping all DNA samples were quantified by PicoGreen (Invitrogen). Genotyping was conducted using Illumina Hap370K BeadChip arrays according to manufacturer’s protocol. DNA with GenCall scores less than 0.25 at any locus were considered ‘no-calls.’ A SNP was considered to have failed if less than 95% of DNA samples generated a genotype at the locus. Cluster plots were manually inspected to resolve any ambiguities. We excluded SNPs from the analyses if the Hardy–Weinberg proportion \( P \) value was < 0.0001. We also removed SNPs for which the minor allele frequency was less than 5%.

Data processing and error checking

The pedigree relationship testing program Pedcheck software, version 1.1 (November 24, 1998) (9) was implemented to check for pedigree errors. Non-Mendelian error checking of genotypes and generation of linkage format files from Illumina array files was done using in-house generated scripts. The genotyping calls were processed using the BeadStudio Genotyping (BSGT) software, version 3.2.32. The map order and
distances between SNP markers was based on the NCBI GenomeBuild version 36.

Linkage analysis

We conducted multipoint nonparametric linkage analysis using Allegro software, version 2.0 (10). We used an equal family weighting scheme in the exponential model with a scoring function $S_{\text{pair}}$ (11), which is optimal over a range of disease models (12).

Multipoint linkage assumes that markers are in linkage equilibrium. However, for closely spaced markers, this is not always the case thus increasing false-positive evidence for linkage (13). Pairwise LD (as measured by $r^2$ value) between SNPs was calculated using Plink (v1.06; ref. 14). Then, we obtained LD blocks based on the LD value criteria of $r^2 \geq 0.01$ and $r^2 \geq 0.004$. For each block, we retained the SNP from each set with the highest information content, defined by entropy information measure $I_E$ (15). Then, we conducted linkage analysis using 3 different criteria: retaining all SNPs, excluding all SNPs with $r^2 > 0.01$, and excluding all SNPs with $r^2 > 0.004$. For some pedigrees, individual members not informative for linkage were removed to accommodate Allegro pedigree size limitations.

We conducted haplotype analyses using SimWalk2 (version 2.83; refs. 16–18). SimWalk2 uses Markov chain Monte Carlo methods and simulated annealing algorithms to conduct multipoint haplotype analysis, which estimates the most likely set of fully typed maternal and paternal haplotypes of the marker loci for each individual in the pedigree. The input files to SimWalk2 were automatically generated from the linkage analysis files using a utility program Mega2 (Manipulation Engine for Genetic Analysis), version 4.3.1 (19). The options for running Mega2 were set with 'SimWalk2' for analysis option and 'Haplotype analysis' for analysis suboption.

Genetic heterogeneity is an important feature of complex disease etiology such as glioma. We conducted the ordered subset analysis (OSA) using trait-related covariates that provide maximal evidence for linkage (20, 21). In the OSA method, the family members’ covariate information is used to assign a covariate score to each family. Subsequently, families are ordered by family-specific covariate score and linkage analyses are done on all subsets ranked by covariate scores. We used a computer program, FLOSS (version 1.4.1 for Windows), that implemented the OSA method (21). We used age of onset (minimum, maximum, mean, and range) and number of affected individuals per family as covariates in the FLOSS analyses.

Results

Details of the 75 families: 46 from the first stage of analyses and 29 from the replication study are summarized in Table 1. For the first stage families, the average family size was 24.6 individuals, with the smallest family having 8 individuals and the largest having 44 individuals. A total of 127 individuals (11.2%) were affected with a diagnosis of glioma, 835 (73.8%) were unaffected, and 169 (15.0%) were coded as 'unknown affection status.' Of the 127 affected individuals, 61 (48%) were male compared with 50.1% of the total individuals. We genotyped 415 individuals (36.7%), with a range of 4 to 20 individuals genotyped per family. Twenty-eight pedigrees had 2 affected relatives per pedigree, 12 pedigrees had 3 affected relatives per pedigree, 1 pedigree had 4 affected relatives, 3 pedigrees had 5 affected relatives per pedigree, and 2 pedigrees had 6 affected relatives per pedigree. The average age of diagnosis for the probands was 46.9 years (SD = 13.4) and the average age of diagnosis for the other affected members in the pedigrees (excluding proband) was 46.9 years (SD = 19.6; Table 1).

Data quality

A number of parameters were used throughout the study to determine data quality. The average SNP call rate per array was more than 96%. Three pedigrees from the stage 1 samples had excessive Mendelian inheritance errors, which were likely due to misspecified relationships or sample switching. These 3 pedigrees were removed from the analyses. On average, Mendelian inconsistency was found in less than 0.4% of the SNPs. These error SNPs were randomly distributed across the genome and were removed from further analyses.

Linkage analysis

The linkage analyses based on the SNPs retained using the 3 LD criterions (see the methods section) resulted in LOD score peaks at identical locations (data not shown). The first 2 criterions based on retaining all the SNPs and excluding all SNPs with $r^2 > 0.01$ gave higher LOD scores than the criterion based on excluding all SNPs with $r^2 > 0.004$. Conservatively, we report linkage analysis results based on the last criterion (i.e., excluding all SNPs with $r^2 > 0.004$) which is based on a total of 23,476 SNPs across the genome.

For the initial 46 families, the most significant genome-wide nonparametric linkage (NPL) score of 3.39 was obtained at physical location 42,504,408 on chromosome 17q12–21.32 ($P = 0.0005$; Table 2 and Fig. 1). The corresponding Z-score in this region was 4.20 ($P = 0.00007$). In addition, 3 other chromosomal regions had maximum NPL scores exceeding 2.0. These regions were chromosome 6 (max NPL = 2.10, $P = 0.02$ at physical location 22,393,991, 6p22.3), chromosome 12 (max NPL = 2.07, $P = 0.02$ at physical location 24,347,720, 12p13.33–12.1), and chromosome 18 (max NPL = 2.45, $P = 0.008$ at physical location 75,238,190, 18q23).

We further investigated the contribution of each family to the LOD scores on chromosome 17. We investigated the age of onset of the affected members of these families linked to the region on chromosome 17. Eleven families showed maximal evidence of linkage at chromosome 17. Each of these families had individual NPL score greater than 0.59 (equivalent to a point-wise significance of 0.05). The age of onset for all individuals affected with glioma in the families linked to the chromosome 17 region was lower than the age of onset for all individuals affected with glioma coming from unlinked families (36.8 vs. 43.9 years); however, the difference was not statistically significant ($P = 0.17$). Most of the families linked to chromosome 17 have affected siblings with glioma (8 of 11 families). No significant glioma histologic differences were noted between families.
To better understand the 17q region and to identify risk haplotypes shared by affected individuals from the 11 families linked to this region, we genotyped 5 highly polymorphic microsatellite markers in 17q region (D17S932, D17S950, D17S791, D17S943, and D17S1869). The average information content value for all the markers was 0.56 to 0.70. Haplotype analyses revealed that the risk haplotype segregated among affected relatives. However, we also observed that a number of unaffected individuals were carriers of the disease haplotype implying decreased penetrance. The highest NPL score of 4.02 \( (P = 0.000009) \) was obtained at the marker D17S950 using microsatellite-based multipoint linkage analyses.

To replicate the 17q12–21.32 region, we genotyped 29 independent glioma families ascertained using the same criteria as families in the first stage. The average family size was 19 individuals, with the smallest family having 9 individuals and the largest had 32 individuals. A total of 70 individuals (12.6%) were affected with a confirmed diagnosis of glioma. Of the 70 affected individuals, 41 (58.6%) were male which is a higher percentage than the percentage of affected males in the first stage families; however, the difference is not statistically significant. We genotyped a total 309 individuals (56.0%) which is a higher percentage than the number of individuals genotyped in the first stage families. We were able to get DNA samples on more individuals in the second stage which could be due to the fact that stage 2 families were more recently ascertained and recruitment staff was more experienced. The majority of the families had only 2 affected relatives \( (n = 21) \), and the other 8 families had 3 to 6 affected relatives. The average age of diagnosis for the probands was 47.6 years (SD = 14.8), and the average age of diagnosis for the affected members (excluding proband) was 46.6 years (SD = 20.5; Table 1) which was very similar to the average age of diagnosis for the affected individuals in the first stage families.

Using Allegro, we obtained NPL score of 1.26 \( (P = 0.008) \) at 17q12–21.32 region for the stage 2 families. The corresponding Z-score in this region was 1.47 \( (P = 0.035) \).

To account for hypothesized genetic heterogeneity, using the FLOSS program, we obtained an ordered subset analyses-based NPL score of 2.16 \( (P = 0.0008) \) at 17q12–21.32 region for the stage 2 families. The corresponding Z-score in this region was 1.47 \( (P = 0.035) \).

### Table 1. Descriptive statistics and characteristics of glioma pedigrees from United States used for linkage analyses

<table>
<thead>
<tr>
<th>No. of affected individuals in a pedigree</th>
<th>No. of affected individuals in generation(s)</th>
<th>First stage ( (n = 46) )</th>
<th>Second stage ( (n = 29) )</th>
<th>Combined ( (n = 75) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>14</td>
<td>8</td>
<td>22</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>13</td>
<td>10</td>
<td>23</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>7</td>
<td>–</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>3</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>1</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>–</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1,2</td>
<td></td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>3</td>
<td>–</td>
<td>3</td>
</tr>
<tr>
<td>1,2</td>
<td></td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>No. of affected male (%)</td>
<td>61 (48.0)</td>
<td>41 (58.6)</td>
<td>102 (51.8)</td>
<td></td>
</tr>
<tr>
<td>No. of affected female (%)</td>
<td>66 (52.0)</td>
<td>29 (41.4)</td>
<td>95 (48.2)</td>
<td></td>
</tr>
<tr>
<td>Avg. age (y) at dx, proband (SD)</td>
<td>46.9 (13.4)</td>
<td>47.6 (14.8)</td>
<td>45.0 (14.4)</td>
<td></td>
</tr>
<tr>
<td>Age range at dx, proband</td>
<td>6–69</td>
<td>22–75</td>
<td>6–75</td>
<td></td>
</tr>
<tr>
<td>Avg. age (y) at dx, family (SD)</td>
<td>46.9 (19.6)</td>
<td>46.6 (20.5)</td>
<td>45.3 (19.1)</td>
<td></td>
</tr>
<tr>
<td>Age range at dx, all affected individuals</td>
<td>1–86</td>
<td>1–85</td>
<td>1–86</td>
<td></td>
</tr>
<tr>
<td>Total no. individuals</td>
<td>1,131</td>
<td>552</td>
<td>1,683</td>
<td></td>
</tr>
<tr>
<td>Avg. family size</td>
<td>24.6</td>
<td>19.0</td>
<td>22.4</td>
<td></td>
</tr>
<tr>
<td>No. genotyped (%)</td>
<td>415 (36.7)</td>
<td>309 (56.0)</td>
<td>724 (43.0)</td>
<td></td>
</tr>
</tbody>
</table>

To better understand the 17q region and to identify risk haplotypes shared by affected individuals from the 11 families linked to this region, we genotyped 5 highly polymorphic microsatellite markers in 17q region (D17S932, D17S950, D17S791, D17S943, and D17S1869). The average information content value for all the markers was 0.56 to 0.70. Haplotype analyses revealed that the risk haplotype segregated among affected relatives. However, we also observed that a number of unaffected individuals were carriers of the disease haplotype implying decreased penetrance. The highest NPL score of 4.02 \( (P = 0.000009) \) was obtained at the marker D17S950 using microsatellite-based multipoint linkage analyses.

### Replication and joint analyses

To replicate the 17q12–21.32 region, we genotyped 29 independent glioma families ascertained using the same criteria as families in the first stage. The average family size was 19 individuals, with the smallest family having 9 individuals and the largest had 32 individuals. A total of 70 individuals (12.6%) were affected with a confirmed diagnosis of glioma. Of the 70 affected individuals, 41 (58.6%) were male which is a higher percentage than the percentage of affected males in the first stage families; however, the difference is not statistically significant. We genotyped a total 309 individuals (56.0%) which is a higher percentage than the number of individuals genotyped in the first stage families. We were able to get DNA samples on more individuals in the second stage which could be due to the fact that stage 2 families were more recently ascertained and recruitment staff was more experienced. The majority of the families had only 2 affected relatives \( (n = 21) \), and the other 8 families had 3 to 6 affected relatives. The average age of diagnosis for the probands was 47.6 years (SD = 14.8), and the average age of diagnosis for the affected members (excluding proband) was 46.6 years (SD = 20.5; Table 1) which was very similar to the average age of diagnosis for the affected individuals in the first stage families.

Using Allegro, we obtained NPL score of 1.26 \( (P = 0.008) \) at 17q12–21.32 region for the stage 2 families. The corresponding Z-score in this region was 1.47 \( (P = 0.035) \).

To account for hypothesized genetic heterogeneity, using the FLOSS program, we obtained an ordered subset analyses-based NPL score of 2.16 from the 29 replication families \( (P = 0.0008) \) with the number of affected individuals per family as significant family-specific covariate. Even though, the NPL score based on stage 2 families did not reach the LOD score threshold of 3.0, it should be noted that we are not conducting genome-wide analyses with these replication families. Therefore, the point-wise \( P \) value of 0.0008 should be considered as a replication of the genetic region identified using the first stage families.
We further conducted joint OSA on 75 families (46 from stage 1 and 29 from stage 2). Using the FLOSS, the combined analyses gave us an NPL score of 3.81 \( (P = 0.0001) \) with the number of affected individuals per family as significant family-specific covariate. When we reanalysed the data excluding those families deemed unlinked from the FLOSS analysis, the LOD score increased to 4.24 \( (P = 0.000005) \), further supporting linkage to this region.

We also conducted joint OSA on 75 families at the 3 chromosomal regions 6p22.3, 12p13.33–12.1, and 18q23 that were suggestive of linkage based on the first stage families. Using the FLOSS, the combined analyses gave respective NPL scores 2.34, 1.75, and 2.04 with the number of affected individuals per family as significant family-specific covariate.

### Discussion

Here, we provide evidence for a moderate–high penetrance susceptibility to glioma mapping to 17q12–21.32. We searched the National Center for Biotechnology Information (NCBI) Hapmap database and identified potentially important target genes related to cancers in this area. \( \text{WNT9B} \) (17q21), mapping at 44,941,702 bps, belongs to the \( \text{WNT} \) gene family, is associated with basal cell carcinoma. The \( \text{WNT} \) gene family consists of several structurally related genes that encode signalling proteins that have been implicated in oncogenesis and in several developmental processes, including the regulation of cell fate and patterning during embryogenesis (22, 23). Another gene that resides in this region is \( \text{CDC27} \), at location 17q21.32, which is involved in brain ischemia (27).

We also identified 3 additional regions, on chromosomes 6p, 12p, and 18q, which are suggestive for linkage using data from first stage families. The associated allele-sharing LOD scores also exceeded thresholds for suggestive evidence of linkage. The exact \( P \) values for NPL scores in these regions were between

### Table 2. The NPL scores and exact \( P \) values for United States glioma pedigrees after the removal of all SNPs with \( r^2 > 0.004 \) and minor allele frequency less than 0.05

<table>
<thead>
<tr>
<th>Chr</th>
<th>Physical location</th>
<th>SNP</th>
<th>Max NPL</th>
<th>NPL exact ( P ) value</th>
<th>Left physical location</th>
<th>Right physical location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>158,235,632</td>
<td>rs6676862</td>
<td>1.6120</td>
<td>0.0556</td>
<td>157,007,286</td>
<td>158,871,097</td>
</tr>
<tr>
<td>2</td>
<td>1,657,741</td>
<td>rs7560004</td>
<td>1.9369</td>
<td>0.0283</td>
<td>1,443,268</td>
<td>4,151,604</td>
</tr>
<tr>
<td>3</td>
<td>178,984,865</td>
<td>rs6764952</td>
<td>1.9355</td>
<td>0.0284</td>
<td>175,719,120</td>
<td>180,045,647</td>
</tr>
<tr>
<td>4</td>
<td>7,375,903</td>
<td>rs4689705</td>
<td>1.1218</td>
<td>0.1320</td>
<td>6,691,863</td>
<td>7,852,481</td>
</tr>
<tr>
<td>5</td>
<td>103,137,679</td>
<td>rs320654</td>
<td>1.8552</td>
<td>0.0338</td>
<td>89,093,506</td>
<td>124,487,446</td>
</tr>
<tr>
<td>6</td>
<td>22,393,991</td>
<td>rs1205961</td>
<td>2.0998</td>
<td>0.0195</td>
<td>22,382,745</td>
<td>23,359,338</td>
</tr>
<tr>
<td>7</td>
<td>100,639,893</td>
<td>rs1468358</td>
<td>1.4495</td>
<td>0.0755</td>
<td>100,151,356</td>
<td>101,310,585</td>
</tr>
<tr>
<td>8</td>
<td>72,212,959</td>
<td>rs4689705</td>
<td>1.1218</td>
<td>0.1320</td>
<td>6,691,863</td>
<td>7,852,481</td>
</tr>
<tr>
<td>9</td>
<td>1,901,291</td>
<td>rs10811102</td>
<td>1.6935</td>
<td>0.0473</td>
<td>1,381,887</td>
<td>2,049,341</td>
</tr>
<tr>
<td>10</td>
<td>37,456,382</td>
<td>rs1878249</td>
<td>1.2281</td>
<td>0.1111</td>
<td>1,050,218</td>
<td>33,053,017</td>
</tr>
<tr>
<td>11</td>
<td>122,434,697</td>
<td>rs11551598</td>
<td>1.7056</td>
<td>0.0461</td>
<td>122,356,773</td>
<td>123,153,543</td>
</tr>
<tr>
<td>12</td>
<td>24,344,720</td>
<td>rs565934</td>
<td>2.0733</td>
<td>0.0208</td>
<td>2,152,566</td>
<td>25,405,864</td>
</tr>
<tr>
<td>13</td>
<td>105,045,125</td>
<td>rs10508158</td>
<td>1.5536</td>
<td>0.0622</td>
<td>101,097,159</td>
<td>114,108,121</td>
</tr>
<tr>
<td>14</td>
<td>19,946,093</td>
<td>rs1760897</td>
<td>1.5150</td>
<td>0.0669</td>
<td>19,536,664</td>
<td>93,349,018</td>
</tr>
<tr>
<td>15</td>
<td>41,144,698</td>
<td>rs12050604</td>
<td>1.5787</td>
<td>0.0593</td>
<td>37,598,478</td>
<td>41,154,373</td>
</tr>
<tr>
<td>16</td>
<td>86,488,782</td>
<td>rs8058389</td>
<td>1.4061</td>
<td>0.0817</td>
<td>84,489,852</td>
<td>87,500,555</td>
</tr>
<tr>
<td>17</td>
<td>42,504,408</td>
<td>rs4074249</td>
<td>3.3913</td>
<td>0.0005</td>
<td>33,196,340</td>
<td>45,510,424</td>
</tr>
<tr>
<td>18</td>
<td>75,238,190</td>
<td>rs1110784</td>
<td>2.4501</td>
<td>0.0082</td>
<td>73,651,628</td>
<td>76,085,595</td>
</tr>
<tr>
<td>19</td>
<td>2,561,822</td>
<td>rs8104096</td>
<td>1.0822</td>
<td>0.1404</td>
<td>5,495,085</td>
<td>54,286,819</td>
</tr>
<tr>
<td>20</td>
<td>2,051,264</td>
<td>rs6075668</td>
<td>1.3404</td>
<td>0.0918</td>
<td>16,749</td>
<td>3,398,635</td>
</tr>
<tr>
<td>21</td>
<td>45,595,912</td>
<td>rs2838891</td>
<td>0.5032</td>
<td>0.3048</td>
<td>42,200,186</td>
<td>46,852,308</td>
</tr>
<tr>
<td>22</td>
<td>26,003,211</td>
<td>rs134806</td>
<td>1.1607</td>
<td>0.1240</td>
<td>15,873,708</td>
<td>33,553,460</td>
</tr>
</tbody>
</table>

NOTE: Chr, chromosome; physical location, the physical location at which maximum NPL score was observed; SNP, rs number of the SNP at which maximum NPL score was observed; max NPL, maximum NPL score obtained; NPL exact \( P \) value, the exact \( P \) value associated with maximum NPL score from Allegro; left physical location and right physical location, boundaries based on 1-LOD drop (approximating 95% CI for the location of peak).
0.01 and 0.02. However, using the joint OSA on 75 families, only 6p22.3 and 18q23 regions remained suggestive for linkage. The contribution of individual families to the genetic region implicated in this study varied, suggesting that gliomas have strong genetic heterogeneity. This provides justification for the further study of larger number of families at high risk for this relatively rare cancer.

Recently, our group conducted 3 independent genome-wide association studies (28–30) which identified 7 risk loci for glioma at 5p15.33 (TERT rs2736100), 7p11.2 (EGFR rs11979158 and rs2252586), 8q24.21 (CCDC26 rs4295627), 9p21.3 (CDKN2A-CDKN2B rs977756), 11q23.3 (PHLDB1 rs98872), and 20q13.33 (RTEL1 rs6010620). Collectively these loci accounted for only 7% to 14% of the excess familial risk of glioma (28). The linkage peak for the GWAS hits was not significant in the families analyzed herein, underscoring that the genes associated with sporadic glioma have little contribution to familial glioma.

Thus, additional studies are needed to discover additional risk loci for familial glioma.

In summary, our analysis provides evidence for disease locus for glioma susceptibility at 17q12–21.3. The next step for this study is to sequence this region in the affected and unaffected members of families that showed high linkage and determine if there is a pathogenic mutation in the germline. We will look for deleterious mutations (frame shift truncating or nonsense mutations, missense variants, insertions, or deletions), that could be cancer promoting genes. If we identify specific genes from the sequencing analysis, we will then screen for these mutations in the linking relatives. To date, this is the largest genetic linkage study of glioma. The regions implicated in this study underscore glioma heterogeneity especially because they are different from those identified in previous GWAS. Further studies could potentially provide novel insights into the biological mechanisms and may ultimately lead to the identification of a causal gene or genes in these regions and to guidelines for genetic counseling.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were not disclosed.

Acknowledgments

The authors thank the contributions of the following individuals to the overall brain tumor research programs—MD Anderson Cancer Center: Phyllis Adatto, Fabian Morice, Sam Payen, Lacey McQuinn, Rebecca McGaha, Sandra Guerra, Leslie Paith, Katherine Roth, Dong Zeng, Hui Zhang, Dr. Alfred Yung, Dr. Howard Colman, Dr. Charles Conrad, Dr. John de Groot, Dr. Arthur

Figure 1. NPL scores for U.S. glioma pedigrees across each chromosome. Each plot shows NPL scores obtained after excluding all SNPs with $r^2 > 0.004$. The y-axis is NPL score and x-axis is physical location.
Forman, Dr. Morris Groves, Dr. Victor Levin, Dr. Monica Loghin, Dr. Vinay Puduvalli, Dr. Raymond Sawaya, Dr. Amy Heimberger, Dr. Frederick Lang, Dr. Nicholas Levine, Lori Tolestino; Brigham and Women's Hospital; Kate Saunders, Donna Delio Iacono; Case Western Reserve University; Dr. Stanton Gerson, Dr. Warren Selman, Dr. Robert Maciunas, Dr. Nicholas Bambakidis, Dr. David Hart, Dr. Jonathan Miller, Dr. Alan Hoffer, Dr. Mark Cohen, Dr. Lisa Rogers, Dr. Charles J Nock, Wendi Barrett, Anita Merriam, Quinn Ostrom, Sarah Robbins, Perica Davitkov, Dr. Michael Vogelbaum, Dr. Robert Weil, Dr. Mannneet Ahuwalia, Dr. David Peerbooms, Dr. Edward Benzel, Dr. Susan Stangattis, Cathy Schliere, Cathy Brewer, Kathy Smolenski, Diane Fabec, Theresa Naska, Jennifer Hornacek-Gaudek; Columbia University Medical Center; Dr. Steven Rosenfeld; Israel; Dr. Zvi Ram, Dr. Deborah T. Blumenthal, Dr. Felix Bokstein (Tel-Aviv Sourasky Medical Center), Dr. Felix Umansky (Hadassah—Hebrew University Medical Center; Henry Ford Hospital). Dr. Menashe Zaaaroor (Rambam—Health Care Campus) Dr. Avi Cohen (Soroka University Medical Center, Chaim Sheba Medical Center), Dr. Tzeela Trusk-Shina (Rambam Medical Center and Faculty of Medicine, Technion-Israel Institute of Technology); Denmark; Dr. Bo Voldby (Aarhus University Hospital), Dr. René Laursen M.D. (Aalborg University Hospital), Dr. Claus Andersen (Odense University Hospital). Dr. Janmark Brunnum (Glostrup University Hospital), Matilde Bille Henriksen (Institute of Cancer Epidemiology, the Danish Cancer Society); Memorial Sloan-Kettering Cancer Center: Maya Marzouk, Mary Elizabeth Davis, Eamon Boland, Marcel Smith, Ogechukwu Eze, Mahalia Way; NorthShore University HealthSystem: Pat Lada, Maya Marzouk, Mary Elizabeth Davis, Eamon Boland, Marcel Smith, Andrew K. Yung, Noreen McGee, Jon Pfeiffer, Callie Wilson, Pagan Morris, Mary Hughes, Martin Britt-Williams, Jessica Foell, Julia Madsen, Csaba Polony: University of Illinois at Chicago; Candice Zahora, Dr. John Villano, Dr. Herbert Engelhard.

The authors also thank the input of the Gliogene External Advisory Committee (Dr. Ake Borg (Department of Oncology, Lund University, Lund, Sweden), Dr. Stephen K Chanock (National Cancer Institute, U.S. NIH), Dr. Peter Collins (University of Cambridge, United Kingdom), Dr. Robert Elston (Department of Epidemiology and Biostatistics, Case Western Reserve University). Dr. Paul Kleihues (Department of Pathology, University Hospital, Zurich, Switzerland), Carol Kruchko (Central Brain Tumor Registry of the United States), Dr. Gloria Petersen (Health Sciences Research, Mayo Clinic). Dr. Sharon Plon (Baylor Genetic Cancer Clinic, Baylor College of Medicine).

The Danish (C. Johansen), Israeli (S. Sadetzki), and Swedish (B. Melin) sites recruit population-based participants nationwide. The authors also thank the patients and their families for participating in this research.

Grant Support

This work was supported by grants from the NIH, Bethesda, Maryland (5R01 CA119215, 5R01 CA070917, R01CA26869, P05097257, R01CA126381, 5P30CA16672). Additional support was provided by the American Brain Tumor Association, The National Brain Tumor Society, and the Tug McGraw Foundation. For more information about the Gliogene Consortium, refer to the following Web site: http://www.gliogene.org.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received January 6, 2011; revised August 24, 2011; accepted September 27, 2011; published OnlineFirst October 28, 2011.

References


Genome-Wide High-Density SNP Linkage Search for Glioma Susceptibility Loci: Results from the Gliogene Consortium


Updated version
Access the most recent version of this article at:
doi:10.1158/0008-5472.CAN-11-0013

Cited articles
This article cites 30 articles, 6 of which you can access for free at:
http://cancerres.aacrjournals.org/content/71/24/7568.full#ref-list-1

Citing articles
This article has been cited by 1 HighWire-hosted articles. Access the articles at:
http://cancerres.aacrjournals.org/content/71/24/7568.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.