Genetic Variants in the 9p21.3 Locus Associated with Glioma Risk in Children, Adolescents, and Young Adults: A Case–Control Study

Anna M. Dahlin, Carl Wibom, Ulrika Andersson, David M. Hougaard, Jonas Bybjerg-Grauholm, Isabelle Deltour, Christina M. Hultman, Anna K. Kåhler, Robert Karlsson, Ulf Hjalmar, Beatrice Melin

Abstract

Background: Genome-wide association studies have identified germline genetic variants in 25 genetic loci that increase the risk of developing glioma in adulthood. It is not known if these variants increase the risk of developing glioma in children and adolescents and young adults (AYA). To date, no studies have performed genome-wide analyses to find novel genetic variants associated with glioma risk in children and AYA.

Methods: We investigated the association between 8,831,628 genetic variants and risk of glioma in 854 patients diagnosed up to the age of 29 years and 3,689 controls from Sweden and Denmark. Recruitment of patients and controls was population based. Genotyping was performed using Illumina BeadChips, and untyped variants were imputed with IMPUTE2. We selected 41 established adult glioma risk variants for detailed investigation.

Results: Three adult glioma risk variants, rs634537, rs2157719, and rs145929329, all mapping to the 9p21.3 (CDKN2B-AS1) locus, were associated with glioma risk in children and AYA. The strongest association was seen for rs634537 (odds ratio: 1.21; 95% confidence interval = 1.09–1.35; P = 5.8 × 10⁻⁴). In genome-wide analysis, an association with risk was suggested for 129 genetic variants (P < 1 × 10⁻⁵).

Conclusions: Carriers of risk alleles in the 9p21.3 locus have an increased risk of glioma throughout life. The results from genome-wide association analyses require validation in independent cohorts.

Impact: Our findings line up with existing evidence that some, although not all, established adult glioma risk variants are associated with risk of glioma in children and AYA. Validation of results from genome-wide analyses may reveal novel susceptibility loci for glioma in children and AYA.

Introduction

To date, six genome-wide association studies (GWAS) and additional fine-mapping efforts have identified genetic variants in 25 different genetic loci that influence the risk of developing glioma (1–10). These studies have included increasingly large sample sizes, enabling identification of glioma subtype-specific risk loci. None of these studies have, however, included pediatric patients. Pediatric patients are typically diagnosed with glioma subtypes different than adult patients. For example, pilocytic astrocytoma is the most common form of glioma in pediatric patients, but it is rarely found in adults. Conversely, glioblastoma is the most common form of glioma in adults, but it is rarely found in children. It is yet to be established if the glioma risk loci found in the adult setting also increase the risk of developing glioma in childhood, adolescence, and early adulthood.

The international CEFALO study, which includes pediatric brain tumors from four European countries, has reported the following shared risk loci for adult and pediatric brain tumors: CDKN2B-AS1 (9p21.3), RETL1 (2q13.33), TERT (5p15.33), CCDC26 (8q24.21), and EGRF (7p11.2) (11, 12). However, this study preceded two of the most extensive GWAS to date and did not investigate the 19 novel risk loci reported (1, 2). In addition, because the number of samples included in the CEFALO study was limited, further investigation of reported findings was warranted. In this study, which includes a large population-based dataset with young glioma patients and controls, we investigated the association between risk of glioma in children and adolescents and young adults (AYA) and genetic variants mapping to the 25 loci that are known to increase glioma risk in adults. We have also performed genome-wide association analyses to find novel genetic variants that are associated with glioma risk in children and AYA.

Patients and Methods

Study subjects

Patient subjects. Glioma diagnosed in children and AYA (13) up to the age of 29 were identified from the national cancer registries in
Sweden and Denmark. Dried blood spot samples were collected from the Swedish Phenylketonuria Screening Registry and the Danish Newborn Screening Biobank, which are national biobanks containing dried blood spot samples from newborns (14–18). All patients were diagnosed between 1976 and 2008. Clinical data for the Swedish patients were retrieved from the Swedish Cancer Registry (SCR); since 1958, all cancer diagnoses in Sweden have been recorded in the SCR. The SCR classifies a cancer as glioma after careful consideration of the entries made by the pathology department and the responsible clinician. The SCR has high coverage and keeps data of high quality (19, 20). Tumor grade is registered as low (grade I and II) or high (grade III and IV) according to the pathology codes that have been used since 1958 in conjunction with ICD-7 coding. For some patients, histopathological diagnosis was available from the Swedish Pediatric Brain Tumor Quality Registry, which started 1982. Because this registry is not mandatory, information on detailed histopathological diagnosis is missing for 11.1% of all Swedish cases (Supplementary Table S1). No clinical data were available for the Danish cases.

Control subjects. For each Danish patient, a control subject matched by date of birth and sex was identified among samples that were physically located close to the patient sample in the biobank. Control subjects from a previous GWAS of schizophrenia were included as a source of Swedish controls (‘Sw5’ and ‘Sw6’ in Ripke and colleagues; ref. 21). These controls were identified using Statistics Sweden’s population registers.

In Sweden, the study was approved by the Data Inspection Board and the Regional Ethical Review Board in Umeå. The Regional Ethical Review Board approved the use of samples from deceased Swedish patients without informed consent from close relatives. Swedish control subjects were initially recruited to a study of schizophrenia, which was approved by the Regional Ethical Review Board in Stockholm. We validated our findings from interaction analyses in data from the Swedish arm of the Glioma International Case-Control Study (GICC; ref. 2), which was approved by the Regional Ethical Review Board in Umeå. All Swedish subjects provided written informed consent. In Denmark, the study was approved by the Research Ethics Committee of the Capital Region (Copenhagen), the Danish Data Protection Agency, and the Danish Newborn Screening Biobank Steering Committee. According to Danish law, the regional Ethics Committee can grant exemption from obtaining informed consent for research projects using biobank samples under certain circumstances (15). For this study, such an exemption was granted. This study was conducted in accordance with the Declaration of Helsinki.

Genotyping and imputation

Extraction and whole-genome amplification of DNA from dried blood spot samples (Swedish and Danish patients and controls) and extraction of DNA from peripheral blood samples (Swedish controls) have been previously described (16–18, 21, 22). Genotyping was performed using the following Illumina BeadChips: HumanOmni2.5Exome (Swedish patients), HumanOmniExpressExome (Danish patients and controls), and HumanOmniExpress (Swedish controls). Pre-imputation quality control was performed using PLINK (version 1.07, http://zzz.bwh.harvard.edu/plink/; ref. 23). We excluded individuals with poor call-rate (<95%–98% in the different datasets) or inconsistencies between reported sex and sex estimated by genotype. We also excluded one of each pair of individuals with spurious relations (PI-HAT > 0.2) and individuals identified as outliers in principle component analyses—that is, individuals exceeding six standard deviations along principal component 1 to 10 calculated using EIGENSOFT version 4.2/6.1.4 (refs. 24, 25; Supplementary Fig. S1). In total, we excluded 49 patients (23 Swedish and 26 Danish) and 128 controls (100 Swedish and 28 Danish) from the study (Supplementary Fig. S2).

Before imputation, we excluded all genetic variants (single nucleotide polymorphisms, SNP) with poor call-rate (<95%), P value from the Hardy–Weinberg test < 1 × 10−6, minor allele frequency <0.05, and all A/T and C/G SNPs (Supplementary Fig. S2). Imputation was based on 472,141 SNPs that passed quality control in all datasets and was performed using IMPUTE2 and SHAPEIT2 software and data from the 1000 Genomes Project as reference (26–29). Imputed SNP with a minor allele frequency <0.01 or imputation info score <0.8 were excluded from all subsequent analyses.

Selection of SNPs

In addition to genome-wide analyses, we focused on genetic variants in the 25 genetic loci that have been previously associated with risk of glioma in GWAS and additional fine-mapping efforts (Supplementary Table S2; refs. 1–10). In each locus, we selected the variants that were presented as the lead SNP (i.e., the SNP with the lowest P value) by Melin and colleagues and Kinnersley and colleagues, two recent GWAS (1, 2). Both studies used imputation to create a more densely tagged map of the genome compared with earlier studies that relied on array-based genotyping. In this present study, the reported lead SNPs in the TERT locus (rs10069690 and rs72709458) had low imputation info scores. Therefore, we included rs2736100 instead, which was among the first reported glioma risk variants from GWAS (5). In total, we investigated 41 SNPs, including the lead SNPs reported by Melin and colleagues and Kinnersley and colleagues in each of the 25 loci and the separate lead SNPs for glioblastoma, non-glioblastoma, and all glioma reported by Kinnersley and colleagues (1, 2). Of the 41 adult glioma susceptibility variants, three were directly genotyped and 38 were imputed (Supplementary Table S2).

Statistical analysis

The association between genetic variants and glioma risk in children and AYA was calculated with a frequentist test under an additive model using SNPTEST v2.5.2 and the score method to account for genotype uncertainty of imputed variants (30). For the 41 adult glioma risk variants, we also calculated association with the assumption of a dominant and recessive model. For the adult glioma risk variants, a statistical test with P < 0.0019 was considered statistically significant, corresponding to Bonferroni correction for testing 26 independent genetic regions (as indicated by linkage disequilibrium (LD) structure between SNPs in the same loci; Supplementary Table S2). In this article, P-values are not adjusted for multiple testing. Association analyses on subgroups of patients based on age, tumor grade, and histopathological diagnosis were restricted to the Swedish patients for whom subgroup information was available, although these analyses included all Swedish controls. Interaction analysis was carried out using unconditional logistic regression in R version 3.4.1 (31). For the purpose of interaction analyses and calculation of LD, genotypes were called on the basis of their imputed genotype.
probabilities. A genotype call was set to "missing" in subjects when all three genotype probabilities for a variant were < 0.9. Subjects with missing genotype calls were excluded from interaction analyses and calculation of LD. All tests of association were adjusted for sex and the top five principal components. This was done to adjust for the associations between principal components and case-control status that were found up to principal component five (P < 0.05; Supplementary Fig. S1). We tested the possibility of false-positive findings due to genotyping errors by excluding each genotyped SNP with P < 1 × 10⁻⁵ and repeated pre-phasing, imputation, and association testing of the SNP and imputed variants within 500 kb. On the basis of these analyses, we suspected genotyping errors in three SNPs with P values that changed from 4 × 10⁻⁵⁷, 1 × 10⁻⁶⁷, and 3 × 10⁻⁶⁶ to 0.55, 0.10, and 0.33, respectively. These false-positive SNPs were excluded from all further analyses. We also excluded imputed SNPs in one 16 kb region due to a dramatic increase in P values after re-imitation without the false positive SNP in this region.

**Results**

We investigated the association between germline genetic variants and risk of glioma in 854 pediatric and AYA patients and 3,689 control subjects who passed quality control filtering (Supplementary Fig. S2). Figure 1 presents the P values for 8,831,628 genetic variants that were directly genotyped or imputed with high quality (imputation info score ≥ 0.8). Q-Q plots and inflation factor lambda indicated no inflation of P values in analyses adjusted for sex and five principal components (λ_adjusted = 1.00; Supplementary Fig. S1). Of the 41 adult glioma susceptibility SNPs, 38 were directly genotyped or imputed with high quality and thus included in the association analyses (Supplementary Table S2). Three variants in the 9p21.3 (CDKN2B-AS1) locus—that is, rs634537 (odds ratio [OR]_C, 1.21; 95% confidence interval [CI], 1.09–1.35; P = 0.0006), rs2157719 (OR_C, 1.21; 95% CI, 1.09–1.35; P = 0.0006), and rs145929329 (OR_ATT, 1.19; 95% CI, 1.07–1.33; P = 0.0017)—were statistically significantly associated with glioma risk in children and AYA after correcting for multiple testing of 26 genomic loci (Fig. 2; Supplementary Table S3). The effects of these associations were in the same direction as previously published (1, 2) and evident when restricting analyses to patients with a low grade tumor, with an astrocytic tumor, or ages 20 years or younger at diagnosis (Supplementary Table S4). The three 9p21.3 (CDKN2B-AS1) variants (rs634537, rs2157719, and rs145929329) were all in strong LD with each other (Supplementary Table S1). When investigating a region ±500 kbp from these three variants, the lowest P value was found for rs1063192 (OR_C, 0.81; 95% CI, 0.73–0.91; P = 0.0002; Supplementary Fig. S3). None of the other investigated adult glioma risk variants were statistically significantly associated with glioma risk in children and AYA after correcting for multiple comparisons (Supplementary Table S3), and SNPs with lower P values than the established glioma risk-SNPs were found in all candidate loci (Supplementary Fig. S3). Results were not altered in analyses restricted to patients ages 20 years or younger at diagnosis (Supplementary Table S4). In a previous report by Adel Fahmideh and colleagues (11), the association between rs2736100 (TERT) and childhood brain tumor risk was stronger in a model assuming a dominant effect. Also, in the same study, an association between SNPs in the RETI loci (rs6089953, rs6010620, rs2297440, rs4809324) and childhood brain tumor risk was found only when restricting analyses to astrocytoma (11). Similarly, in the present study, the association between rs2736100 (TERT) and risk of glioma in children and AYA was stronger in a model assuming a dominant effect (OR_C, 1.20; 95% CI, 1.00–1.44; P = 0.046; Supplementary Table S3), and the association between rs2297440 (RTEL) and risk of glioma in children and AYA was strongest in analyses restricted to pilocytic astrocytoma (OR_C, 1.37; 95% CI, 1.01–1.86; P = 0.042; Supplementary Table S4).
although these findings were not statistically significant after correcting for multiple comparisons.

One of the genes in the 9p21.3 locus is **CDKN2A**, which codes for two different proteins with known tumor-suppressor function: P16INK4A (cyclin-dependent kinase inhibitor 2A), a cyclin-dependent kinase inhibitor with cell-cycle regulating function in the Rb pathway, and P14ARF (tumor suppressor ARF), which promotes P53 (Cellular tumor antigen p53) function by inactivation of MDM2 (E3 ubiquitin-protein ligase Mdm2), a negative regulator of P53. Another protein that can negatively regulate P53 activity is MDM4. Interestingly, a variant in the **MDM4** gene (rs4252707) was also associated with pediatric glioma risk, with effects in the same direction as previously published (2), although this finding was not statistically significant after correction for multiple comparisons (OR, 1.24; 95% CI, 1.07–1.42; P = 0.0034; Supplementary Table S3). Of the 41 adult glioma risk variants, a low frequency variant located in the **TP53** gene, rs78378222, had the largest effect size with direction consistent with previous reports (2, 7), although the association was not statistically significant (OR, 1.32; 95% CI, 0.87–2.01; P = 0.190;
Discussion

Here, we report on three genetic variants in the 9p21.3 locus that were associated with risk of glioma in children and AYA. Our findings confirm the results of the previous CEFALO study of pediatric brain tumors (11). Variants in the 9p21.3 locus have also been reported to modify the risks of cancer of different types, including adult glioma, as well as the risks of other diseases (32). The strongest association in the region was found for rs1063192. This variant is located within the 3′ untranslated region of CDKN2B and has previously been associated with glioma in adults and pediatric brain tumors (5, 11). Another obvious candidate gene within this locus is CDKN2A, which codes for tumor-suppressor proteins P16INK4A and P14ARF. In this locus, somatic homozygous deletion is a common event in both adult and pediatric high-grade glioma (33, 34), but it is not as frequent in lower-grade adult glioma (35) and was not observed in any of the pilocytic astrocytomas investigated in a recent pediatric pancyt cancer study (33). The exact mechanism through which germline variants in this locus increase the risk of cancer is not known.

None of the other adult glioma risk variants investigated in this study were associated with risk of glioma in children and AYA, indicating that they are susceptibility variants for glioma in adults but not in children and AYA. Our findings highlight a possible difference in etiology between adult and pediatric glioma, although for some of the variants it is possible that statistically significant associations may have been detected if the study size were even larger. Another explanation of the lack of association in glioma in children and AYA may be a difference in the proportion of histological subtypes found in adult and pediatric patients (13). Although GWAS of adult glioma have found several susceptibility loci that are unique to non-glioblastoma or glioblastoma, no GWAS has analyzed specific histological subtypes of non-glioblastoma. Histological subtypes of glioma that are common in children but rare in adults (e.g., pilocytic astrocytoma) may have different susceptibility variants.

In all loci investigated due to their association with glioma risk in adults, other SNPs than the candidate SNP had the lowest P value. Validation of these SNPs in independent cases and controls are required to find out if they are markers of the same association as the candidate SNP or if they are true findings of novel SNPs associated with risk of glioma in the young. In the search for susceptibility variants that are unique to glioma in children and AYA, we also present the results from genome-wide analyses, corresponding to the discovery phase of a GWAS. A GWAS, however, requires large sample sizes for replication. Moreover, for rare diseases, such as glioma in children and AYA, international collaboration is essential.

Of all investigated adult glioma risk variants in this study, TP53 variant rs78378222 (G) and MDM4 variant rs4252707 (A) displayed the strongest risk-altering effects, which were both in the same direction as previously reported (2, 7, 36), but not statistically significant after correction for multiple comparisons. Because the protein products of CDKN2A and MDM4 (P14ARF and MDM4, respectively) are both regulators of P53 activity, it is noteworthy that 9p21.3 (CDKN2A), MDM4, and TP53 are all among the loci in which we found the associations with the lowest P values and/or strongest effects. We found that the variants in these loci had modifying effects on risk when present in the same individual and that the highest risk was found in individuals carrying the TP53 (rs78378222, G) and MDM4 (rs4252707, A), but not the 9p21.3 (CDKN2A, rs634537, G) risk alleles. Because specific combination of risk variants is in turn unreported, but the possibility that the strong effects observed are chance findings due to the small sample size. On the other hand, we found the same trend in an independent dataset, including adult lower-grade glioma (non-glioblastoma) patients and controls. The TP53 variant rs78378222 is located in the 3′ untranslated region of the gene so it can disrupt the polyadenylation signal sequence. This disruption may result in improper termination and polyadenylation of the gene transcript (36). Melin and colleagues (2) reported an expression quantitative trait locus in the MDM4 locus, which suggests that the increased glioma risk associated with rs4252707 may be linked to increased gene expression. The regulation of P53 activity is complex, including but not limited to regulation by MDM4 and P14ARF, where the latter is in turn regulated by a negative feedback loop. It is therefore plausible that genetic variants that shift the balance in this intricate pathway may have modifying effects on risk when acting together rather than having linearly additive effects. Because the sample size of the present study is limited, further investigation of the interaction between variants in the P53 pathway in a larger dataset is warranted.

In addition to variants in the 9p21.3 locus, the CEFALO study also reported variants associated with childhood brain tumors in the EGFR, RETL1, TERT, and CCDC26 loci (11, 12). In the CEFALO report, the association in the RETL1 locus was most evident in analyses restricted to astrocytoma, which is the most common glioma subtype in children. In our study, the association between RETL1 variant rs2297440 and risk of glioma in childhood and AYA was strongest for pilocytic astrocytoma and in the same direction as previously reported (2), but not statistically significant after correction for multiple comparisons. The association between TERT variant rs2736100 and childhood brain tumors reported by CEFALO was stronger in a model assuming a dominant effect, which was also the case in the present study,
Although not statistically significant after correction for multiple comparisons. In contrast to the CEFALO study (12), we did not find an association between variants in CCDC26 or EGFR and risk of glioma in children and AYA. However, in the CEFALO report, variants in CCDC26 were associated with risk only in analyses restricted to non-astrocytic tumors, and the investigated EGFR variants were not the same as in the present study. When comparing results from CEFALO and the present study, it should be noted that these two studies have a potential, likely small, overlap, as both studies recruited Danish glioma patients in the age range 7 to 19 years between 2004 and 2008.

The advantages of the present study are the pre-diagnostic blood sampling and nation-wide inclusion of patients diagnosed over almost 30 years in both Sweden and Denmark. This allowed us to collect a large and population-based case sample of a rare disease. With 854 patients, our study is among the largest genetic studies designed for this study. Because controls were not individually matched to patients by geographic location, we used principal component analyses to account for the possible introduction of population stratification with this design. AYA can be defined as individuals up to the age of 39 (13). The present study was limited to patients who had blood spot samples stored in the national neonatal biobanks in Sweden and Denmark, and could include AYA up to the age of 29. We observed only minor changes in the results when analyses were restricted to patients ages ≤20 years at diagnosis.

In summary, our findings line up with the existing evidence that genetic variants in the 9p21.3 locus increase the risk of glioma in children and AYA. We also found indications of an interaction between genetic variants in three genes involved in the P53 pathway, which warrants further investigation. By presenting the results from this study, we hope to find additional collaborators to be able to validate the findings from genome-wide association analyses and discover susceptibility loci that are unique for glioma in children and AYA.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Disclaimer
The authors alone are responsible for the views expressed in this article and they do not necessarily represent the views, decisions, or policies of the institutions with which they are affiliated.

Authors’ Contributions
Conception and design: A.M. Dahlin, I. Deltour, U. Hjalmars, B. Melin
Development of methodology: U. Hjalmars, B. Melin
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): D.M. Hougaard, J. Byberg-Grauholm, C.M. Hultman, A.K. Kähler, R. Karlsson, U. Hjalmars, B. Melin
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): A.M. Dahlin, C. Wibom, U. Andersson, J. Byberg-Grauholm, I. Deltour, R. Karlsson, U. Hjalmars, B. Melin
Writing, review, and/or revision of the manuscript: A.M. Dahlin, C. Wibom, U. Andersson, D.M. Hougaard, J. Byberg-Grauholm, I. Deltour, C.M. Hultman, A.K. Kähler, R. Karlsson, U. Hjalmars, B. Melin
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): C.M. Hultman
Study supervision: U. Hjalmars, B. Melin

Acknowledgments
We acknowledge the contribution of the late Dr. Mads V. Hollegaard to this article. Dr. Hollegaard was one of the initiators and was the head of the laboratory work of the study. We also acknowledge Dr. Bent Nørgaard-Pedersen at the Danish Neonatal Screening Biobank for his support in the initiation of this study and Dr. Ulrika von Dobel at the Swedish phenylketonuria screening registry for assistance with sample collection. We thank the Swedish Pediatric Brain Tumor Quality Registry group for access to clinical data. This research has been conducted using the Danish National Biobank resource, supported by the Novo Nordisk Foundation. This work was supported by the Swedish Childhood Cancer Foundation (to A.M. Dahlin and B. Melin), the Acta Oncologica Foundation through The Royal Swedish Academy of Science (to B. Melin), the Swedish Cancer Foundation (to B. Melin), the Swedish Research Council [to B. Melin], the Cancer Research Foundation in Northern Sweden (to B. Melin), and Umeå University Hospital (Cutting Edge Grant, to B. Melin).

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 September 20, 2018; revised January 7, 2019; accepted April 26, 2019, published first April 30, 2019.

References

www.aacrjournals.org Cancer Epidemiol Biomarkers Prev; 28(7) July 2019 1257
Dahlin et al.


Ploh Genet 2006;2:e190.
Genetic Variants in the 9p21.3 Locus Associated with Glioma Risk in Children, Adolescents, and Young Adults: A Case–Control Study

Anna M. Dahlin, Carl Wibom, Ulrika Andersson, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/1055-9965.EPI-18-1026

Supplementary Material
Access the most recent supplemental material at:
http://cebp.aacrjournals.org/content/suppl/2019/04/30/1055-9965.EPI-18-1026.DC1

Cited articles
This article cites 34 articles, 1 of which you can access for free at:
http://cebp.aacrjournals.org/content/28/7/1252.full#ref-list-1

Citing articles
This article has been cited by 1 HighWire-hosted articles. Access the articles at:
http://cebp.aacrjournals.org/content/28/7/1252.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, use this link
http://cebp.aacrjournals.org/content/28/7/1252.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.