Human Molecular Genetics, 2008, Vol. 17, No. 6 ***800–805***
doi:10.1093/hmg/ddm351
Advance Access published on November 29, 2007

Comprehensive analysis of the role of DNA repair gene polymorphisms on risk of glioma

Lara Bethke[1], Emily Webb[1], Anne Murray[1], Minouk Schoemaker[2], Christoffer Johansen[3], Helle Collatz Christensen[3], Kenneth Muir[4], Patricia McKinney[5], Sarah Hepworth[5], Polyxeni Dimitropoulou[4], Artitaya Lophatananon[4], Maria Feychting[6], Stefan Lönn[7], Anders Ahlbom[6], Beatrice Malmer[8], Roger Henriksson[8], Anssi Auvinen[9,10], Anne Kiuru[10], Tiina Salminen[9,10], Anthony Swerdlow[2] and Richard Houlston[1,*]

[1]Section of Cancer Genetics, [2]Section of Epidemiology, Institute of Cancer Research, 15 Cotswold Rd, Sutton, Surrey SM2 5NG, UK, [3]Institute of Cancer Epidemiology, Danish Cancer Society, Copenhagen, Denmark, [4]Division of Epidemiology and Public Health, University of Nottingham Medical School Queen's Medical Centre, Nottingham, UK, [5]Centre for Epidemiology and Biostatistics, University of Leeds, UK, [6]Division of Epidemiology, Institute of Environmental Medicine, [7]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden, [8]Department of Radiation Sciences, Oncology Umeå University, Sweden, [9]Department of Epidemiology, Tampere School of Public Health, University of Tampere, Finland and [10]Department of Research and Environmental Surveillance, Radiation and Nuclear Safety Authority, STUK, Helsinki, Finland

Received October 16, 2007; Revised and Accepted November 28, 2007

Much of the variation in inherited risk of glioma is likely to be explained by combinations of common low risk variants. The established relationship between glioma risk and exposure to ionizing radiation led us to examine whether variants in the DNA repair genes contribute to disease susceptibility. We evaluated 1127 haplotype-tagging single-nucleotide polymorphisms (SNPs) supplemented with 388 putative functional SNPs to capture most of the common variation in 136 DNA repair genes, in five unique case–control series from four different countries (1013 cases, 1016 controls). We identified 16 SNPs associated with glioma risk at the 1% significance level. The highest association observed across the five independent case–control data-sets involved rs243356, which maps to intron 3 of *CHAF1A* (trend odds ratio, 1.32; 95% confidence interval 1.14–1.54; $P = 0.0002$; false-positive report probability = 0.055, based on a prior probability of 0.01). Our results provide additional support for the hypothesis that low penetrance variants contribute to the risk of developing glioma and suggest that a genetic variant located in or around the *CHAF1A* gene contributes to disease risk.

INTRODUCTION

Gliomas account for ~80% of primary malignant brain tumours and ~21 000 cases are diagnosed each year in the USA (1). Evidence for an inherited predisposition to risk of developing glioma is provided by the single gene disorders neurofibromatosis type 1 and 2, tuberous sclerosis, retinoblastoma, Li-Fraumeni syndrome and Turcot's syndrome (1,2). While these syndromes confer substantial risk of developing glioma, they are rare and collectively only make a minor contribution to the 2–3-fold increased risk seen in first-degree relatives of patients (1,3). To date, few environmental risk factors have unambiguously been shown to influence glioma risk, and it is likely that the residual familial risk is mediated by as yet unidentified genetic factors including low risk variants, some of which may be common and therefore detectable through association analyses.

Inherited dysfunction of DNA repair is the underlying basis of genetic susceptibility to many cancers. This, coupled with the established relationship between glioma risk and exposure

*To whom correspondence should be addressed. Tel: +44 2087224175; Fax: +44 2087224059; Email: richard.houlston@icr.ac.uk

to ionizing radiation (4), provides a strong rationale to explore whether variation in the DNA repair genes contributes to disease susceptibility.

Here we report an analysis of the relationship between common variation in 136 DNA repair genes and the risk of developing glioma. We based our analysis on five case–control series established in association with the Interphone Study (5). In addition to assaying 1127 tagging single-nucleotide polymorphisms (SNPs) selected to capture most of the common variation in the DNA repair genes, we also genotyped 388 putative functional SNPs, including 69 non-synonymous coding SNPs, that have the potential to impact directly on the function of expressed proteins.

RESULTS

We genotyped 1515 SNPs in 1013 glioma cases and 1016 controls. Of the 2029 DNA samples submitted for genotyping, 2024 samples were successfully processed, generating in excess of three million genotypes. Genotypes were obtained for 1010 of 1013 cases (99.7%) and 1014 of 1016 controls (99.8%). Of the 2024 samples successfully genotyped, 2 were excluded from analysis due to non-European ethnicity, and 6 were excluded due to unclear identity, leaving 2016 samples (1005 cases and 1011 controls). Mean SNP call rates per sample were 99.8 and 99.9% in cases and controls, respectively. Of the 1515 SNPs submitted for analysis, 1468 SNPs were satisfactorily genotyped (96.9%). Of these, 1 was not polymorphic in control samples, 1 violated Hardy–Weinberg equilibrium (HWE) in the control population from all five study centres, and 7 had a genotype failure rate >5%, leaving 1459 SNPs for which genotype data were informative.

The observed number of SNPs with significant deviation from HWE ($P < 0.05$) was not significantly more than the expected number in the control population from any of the five study centres. The estimate of the stratification parameter of the genomic control method was close to unity in all of five case–control datasets with $\hat{\lambda}$ values of: UK-North, 1.06 [95% confidence interval (CI): 0.93–1.20]; UK-Southeast, 0.75 (95% CI: 0.68–0.86); Sweden, 1.12 (95% CI: 1.00–1.22); Finland, 0.93 (95% CI: 0.82–1.05), Denmark, 1.10 (95% CI: 0.97–1.22). These data provide no evidence of population stratification as a cause of false-positive results. Furthermore, no evidence was found for differences in genotype frequencies of SNPs between male and female controls as a source of potential confounding in subsequent analyses. There was slight evidence of over-inflation of the test statistics as assessed by a quantile-quantile (Q-Q) plot of the Cochran–Armitage test statistics adjusted by logistic regression for study centre, age and sex (Fig. 1), but this was more pronounced among the smaller test statistics and does not substantially change the result of this study.

Seventy-eight of the 1458 SNPs were significantly associated with the risk of glioma at the 5% level (Supplementary Materials, Table S1) and 16 at the 1% level (Table 1). Adjusting for age and sex did not substantially affect study findings (data not shown).

The most highly associated SNP, rs243356, which maps to intron 3 of *CHAF1A*, was significant at the 0.1% level ($P = 0.0002$). This association was consistent across the five cases–control series with little evidence of heterogeneity [$Q = 8.4$, 4 degrees of freedom (df); $P = 0.08$]. The associated false-positive report probability (FPRP) was 0.055, well below the 0.200 threshold for false-positivity (6). Risks associated with hetero- and homozygosity for the minor allele of rs243356 were 1.33 (95% CI: 1.10–1.60) and 1.74 (95% CI: 1.16–2.61) respectively, thereby being entirely compatible with a multiplicative model of action. There was no evidence that risks of glioma associated with rs243356 differed among subtypes—oligodendroglioma (106 cases), odds ratio $(OR)_{trend} = 1.31$ (95% CI: 0.94–1.84); astrocytoma (329 cases), $OR_{trend} = 1.29$ (95% CI: 1.05–1.59) and glioblastoma (447 cases), $OR_{trend} = 1.26$ (95% CI: 1.04–1.52).

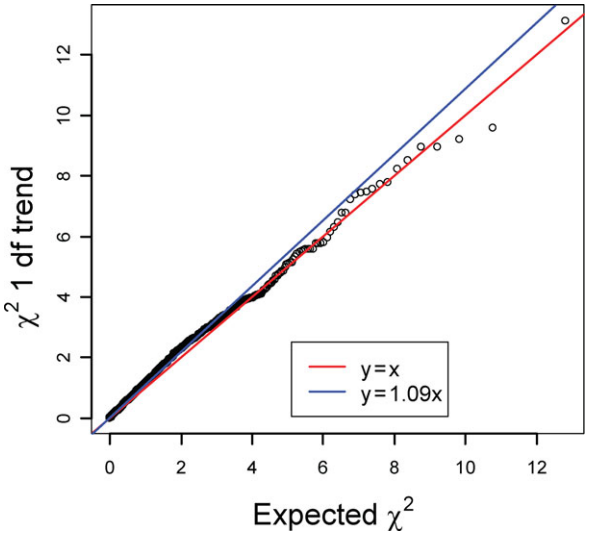

Figure 1. Quantile-quantile plot of test statistics (1 df χ^2 trend test). Black dots represent the test statistics, adjusted by logistic regression for age, sex and study centre, for the 1459 SNPs tested in glioma cases and controls. The blue line $y = 1.09\times$ is fitted to the lower 90% of the distribution of observed test statistics. The red line represents the global null hypothesis of no association at any locus.

Of the seven SNPs we genotyped that map to *CHAF1A*, three in addition to rs243356 were significant at the 1% level: rs243341 and rs105038, which are in tight linkage disequilibrium (LD) and map to introns 1 and 3, and rs2992, which maps to the 3′-UTR and a putative exonic sequence enhancer (Fig. 2). The risk allele for rs243356 is found in only one of the six haplotypes predicted in the LD block defined by these seven SNPs, and none of the haplotypes defined by these SNPs is more strongly associated with risk of glioma than the SNP rs243356 on its own.

Other associations identified included two SNPs mapping to *ERCC1*, three SNPs mapping to *RPA3* and two SNPs mapping to *DCLRE1B*. Single SNP associations were identified for *NEIL3, TP53, POLD1, ATR* and *MSH5* (Table 1).

Interactions between the 16 SNPs that showed some association with risk ($P < 0.01$) were examined by fitting full logistic regression models for each pair, generating 120 models, and comparing these with the main effects model. Five pairs

Table 1. SNPs associated with risk of glioma at the 1% significance level based on a Cochran–Armitage trend test, and their heterozygote and homozygote odds ratios and false-positive report probabilities

SNP	Gene	$MAF_{controls}$	MAF_{cases}	P_{trend}-value	FPRP	OR_{trend}	OR_{het}	OR_{hom}	Putatitive SNP function
rs243356	*CHAF1A*	0.21	0.26	0.0002	0.055	1.32 (1.14–1.54)	1.33 (1.10–1.60)	1.74 (1.16–2.61)	tag
rs12645561	*NEIL3*	0.13	0.16	0.0026	0.363	1.31 (1.10–1.56)	1.29 (1.05–1.59)	1.81 (1.00–3.28)	tag
rs243341	*CHAF1A*	0.26	0.30	0.0028	0.268	1.23 (1.07–1.42)	1.25 (1.04–1.50)	1.50 (1.07–2.09)	tag
rs105038	*CHAF1A*	0.26	0.30	0.0028	0.268	1.23 (1.07–1.42)	1.25 (1.04–1.50)	1.50 (1.07–2.09)	tag
rs3212986	*ERCC1*	0.27	0.23	0.0031	0.298	0.81 (0.70–0.93)	0.76 (0.63–0.92)	0.73 (0.51–1.04)	tag (non-syn on opp. strand)
rs2160138	*RPA3*	0.44	0.48	0.0038	0.302	1.20 (1.06–1.36)	1.11 (0.90–1.36)	1.46 (1.14–1.88)	tag
rs3761936	*DCLRE1B*	0.19	0.15	0.0042	0.419	0.79 (0.67–0.93)	0.89 (0.73–1.08)	0.36 (0.20–0.65)	tag, PHAST, ESE
rs8079544	*TP53*	0.06	0.09	0.0042	0.660	1.43 (1.12–1.82)	1.34 (1.04–1.72)	NA	TFBS
rs11920625	*ATR*	0.08	0.10	0.0044	0.597	1.38 (1.10–1.72)	1.40 (1.11–1.77)	1.49 (0.47–4.73)	tag
rs2992	*CHAF1A*	0.26	0.30	0.0044	0.353	1.22 (1.06–1.40)	1.23 (1.03–1.48)	1.47 (1.05–2.04)	ESE, 3′-UTR
rs12022378	*DCLRE1B*	0.19	0.15	0.0048	0.443	0.79 (0.67–0.93)	0.89 (0.73–1.08)	0.36 (0.20–0.65)	tag, non-syn (H to Y), ESE
rs3212955	*ERCC1*	0.26	0.22	0.0049	0.388	0.81 (0.71–0.94)	0.79 (0.66–0.96)	0.70 (0.48–1.00)	tag
rs6947203	*RPA3*	0.33	0.37	0.0064	0.417	1.20 (1.05–1.36)	1.16 (0.96–1.40)	1.47 (1.11–1.94)	tag
rs4140805	*RPA3*	0.40	0.45	0.0065	0.416	1.19 (1.05–1.35)	1.14 (0.94–1.39)	1.43 (1.11–1.85)	tag
rs1673041	*POLD1*	0.26	0.22	0.0078	0.495	0.82 (0.71–0.95)	0.90 (0.75–1.08)	0.53 (0.35–0.79)	tag
rs707938	*MSH5*	0.36	0.32	0.0096	0.512	0.84 (0.74–0.96)	0.90 (0.75–1.09)	0.67 (0.50–0.89)	synon (Q), PHAST, ESE

SNP, single-nucleotide polymorphism; MAF, minor allele frequency; FPRP, false-positive report probability based on OR_{trend} (assuming prior probability of 0.01 and power to detect an OR of 1.3); OR_{het}, odds ratio of heterozygote versus common homozygote genotypes; OR_{hom}, odds ratio of rare homozygote versus common homozygote genotypes; tag, haplotype tagging SNP; PHAST, log odds conservation score based on the multiple alignment of human, chimp, mouse, rat, dog, chicken, fugu and zebrafish >15; ESE, exonic splicing enhancer; TFBS, transcription factor binding site; UTR, untranslated region; non-syn, non-synonymous sequence variation; synon, synonymous sequence variation; NA, not applicable.

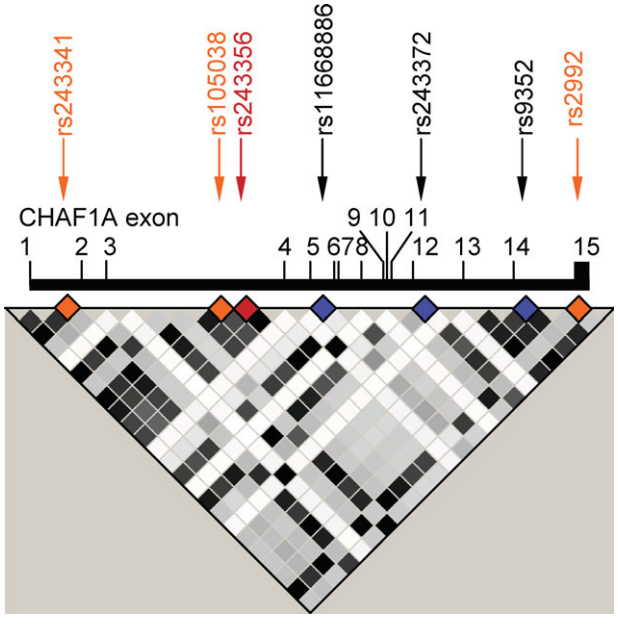

Figure 2. SNPs tested across the *CHAF1A* locus. The exon positions of *CHAF1A* are shown relative to 26 HapMap CEU SNPs with MAF >0.1% represented in the LD heatmap (exon map is not to physical scale). The heatmap shows estimates of the square of the correlation coefficient (r^2) in greyscale (black, 1; white, 0), calculated for each pair-wise comparison of SNPs derived by the Haploview (v.3.2) programme. The SNPs tested in this analysis are as labelled, with the SNP most significantly associated with glioma shown in red ($P_{trend} = 0.002$), and the three other SNPs associated at the 1% significance level shown in orange.

of SNPs showed nominally significant interaction at the 5% level, but the strongest interaction, between rs6947203 and rs4140805 ($P = 0.02$), was non-significant after correction for multiple testing.

DISCUSSION

Our study has systematically evaluated the association of DNA repair genes with glioma risk, on the basis of a priori evidence implicating such variation in cancer risk. While gliomas are the most frequent primary brain tumour, they are relatively rare and most previous association studies have been conducted on small numbers of cases and controls, thereby having modest power to identify associations. Moreover, most studies to date have only evaluated a very restricted set of loci. A major strength of our study is that it has been based on several case–control series collected in the same consistent fashion to generate a uniquely large dataset. Rather than rely on the standard *P*-value criterion of 0.05 to define statistical significance we have computed associated FPRP values for associations identified. This metric depends not only on the observed *P*-value, but also on the prior probability (i.e. strength of belief) that the association between the variant and the disease is real, and on the statistical power of the study. While this approach is in essence Bayesian, it retains the principle of distinguishing those observations that are noteworthy from those that are likely to be simply a consequence of chance (type I errors). Use of the FPRP therefore avoids correcting for multiple testing using a simple Bonferroni procedure which is inherently conservative especially when SNPs are correlated, as is the case in our study when many polymorphisms tag the same region. Furthermore, the approach does not require specification of a prior probability distribution for the OR, which can be burdensome. One limitation to the FPRP approach, however, is assigning a prior probability. On the basis of the biological plausibility of candidate genes selected for evaluation, we have imposed a low prior in keeping with the work of others and to be relatively conservative (7).

Our strongest finding was the association of risk with rs243356, which maps to the gene *CHAF1A*. Although we cannot exclude the possibility that the association was attributable to chance, the association had a low FPRP, providing some measure of robustness to our observation.

CHAF1A is a part of the chromatin assembly factor-1 (CAF-1) complex, which is thought to mediate chromatin assembly in DNA replication and repair by assembling histones into nucleosomes. Although the biological function of CAF-1 has yet to be fully elucidated, several studies have already linked this complex to cancer biology. Expression of CAF-1 has been associated with cell proliferation in breast cancer (8), and CAF-1-mediated deregulation of cell proliferation and DNA repair has been linked to the aggressive behaviour of squamous carcinoma (9). CAF-1 also interacts with the gene product associated with the recessively inherited Bloom's syndrome (MIM 210900), which is characterized by genomic instability and cancer susceptibility (10). Therefore, while the functional role of CHAF1A in primary brain tumours is currently unsubstantiated, a biological relevance in development of glioma is highly plausible.

Haplotype analysis was not informative, as it did not suggest that glioma risk was substantially elevated among individuals with specific haplotypes, compared with individuals with single SNP risk alleles. It is however unlikely that rs243356 is a causal factor itself, and more probable that the association identified is a consequence of LD with a causal variant. On the basis that we failed to detect an association with any of the common nsSNPs of *CHAF1A*, it is likely that the functional effects are mediated through LD with sequence changes within the promoter or enhancer elements of the gene influencing differential expression of *CHAF1A*. Alternatively, there is the possibility that the rs243356 tags a causative variant located outside of the *CHAF1A* locus.

Variants in a number of the other genes we identified represent plausible candidates for determining glioma risk on the basis of their biology. These include *ATR* (MIM 601215) and *RPA3* (MIM 179837), both of which participate in cell cycle checkpoint signalling pathways required for cellular response to DNA damage and for genome stability.

Our study is the first to evaluate the risk of glioma associated with a comprehensive set of DNA repair genes. Our analysis provides evidence that common low risk variants contribute to risk of developing glioma and implicates variants of *CHAF1A* as risk factors. As with all association studies, however, it is highly desirable that our findings are replicated in independent case–control series. Moreover, for some datasets it may prove highly advantageous to conduct analyses of the interaction of genotype with environmental risk factors, notably radiation exposure, where such information has robustly been collected.

MATERIALS AND METHODS

Study subjects

The study was based on five case–control studies that contributed to the Interphone Study (5) and which have been previously used for some candidate gene analyses (6). Briefly, the Interphone Study was an international multi-centre case–control study of primary brain tumours coordinated by the International Agency for Research on Cancer (IARC), with material collected between September 2000 and February 2004. Five population-based case–control studies of primary adult brain tumours were assembled in the Thames regions of Southeast England; the Northern UK including central Scotland, the West Midlands, West Yorkshire and the Trent area; the Stockholm, Lund, Göteborg and Umeå regions of Sweden; throughout Denmark; and in all regions of Finland except Northern Lapland and Åland.

Adult brain tumour cases were identified through neurosurgery, neuropathology, oncology and neurology centres and cancer registries. Eligible cases were patients with glioma [International Classification of Diseases (ICD), 10th revision, code C71; ICD for oncology (ICD-O), 2nd edn, codes 9380–9384, 9390–9411, 9420–9451 and 9505] aged 20–69 years in the Nordic countries, 18–69 in the Northern UK and 18–59 years in Southeast England and resident in the study region at diagnosis. Cases with previous brain tumours were excluded.

Controls in the Nordic centres were randomly selected from population registers within each country and frequency-matched to cases on age, sex and region. In the UK, controls were selected from general practitioners' practice lists, individually matched to cases by age and sex in the Northern UK, and frequency-matched on sex, age and region in the Southeastern UK. Controls had no history of a cerebral tumour.

In Sweden, the Southeast region of the UK (UK-Southeast), the Northern UK (UK-North) and Denmark, a blood sample was drawn from cases and controls who participated in the interview and agreed to provide a blood sample. Finnish investigators did not attempt to collect blood from all interviewed participants; rather, they drew blood from a convenience sample, of predetermined size, from patients in three hospitals and controls living in the area served by these hospitals.

Samples and clinico-pathological information from participants were obtained with informed consent and ethical review board approval in accordance with the tenets of the Declaration of Helsinki.

To minimize population stratification, cases and controls with ethnic origin other than the country of recruitment were excluded from the present study. All remaining cases with sufficient DNA quantity and quality were included. Controls genotyped met the same requirements for DNA quantity and quality as cases, and were randomly frequency-matched by gender to eligible cases for each study centre. The number, sex and age of cases and controls in each of the five studies analyzed in the current study were: UK-North: 370 cases (230 male, 140 female; mean age at diagnosis 49 years; SD $\pm$ 12) and 369 controls (231 male, 138 female; mean age 51 years; SD $\pm$ 11); UK-Southeast: 211 cases (140 male, 71 female; mean age at diagnosis 42 years; SD $\pm$ 11) and 214 controls (142 male, 72 female; mean age 47 years; SD $\pm$ 9); Sweden: 197 cases (121 male, 76 female; mean age at diagnosis 50 years; SD $\pm$ 13) and 197 controls (121 male, 76 female; mean age 52 years; SD $\pm$ 12); Denmark: 128 cases (71 male, 57 female; mean age at diagnosis 48 years; SD $\pm$ 12) and 131 controls (74 male, 57 female; mean age 51 years; SD $\pm$ 12); Finland: 99 cases (56 male, 43 female; mean age at diagnosis 48 years; SD $\pm$ 12) and 100 controls (37 male, 63 female; mean age 53 years; SD $\pm$ 12).

Selection of DNA repair genes and SNPs

A list of 141 candidate DNA repair genes was initially compiled from the inventory of human DNA repair genes published by Wood *et al*. (11). The gene symbols, chromosome position and accession numbers for these genes were updated to the current genome build (RefSeq 36.1) using the UCSC Genome Browser (12,13, http://genome.ucsc.edu/) and the National Center for Biotechnology Information (NCBI) Entrez Gene database (http://www.ncbi.nlm.nih.gov). SNPs mapping to these genes were extracted from dbSNP version 126 (http://www.ncbi.nlm.nih.gov/projects/SNP/).

Of over 35 000 SNPs identified in these 141 autosomal genes, 2765 common tag SNPs were found in 136 genes [MAF >0.1, $r^2 \geq 0.8$ in HapMap CEPH (CEU) Utah residents with ancestry from northern and western Europe], using an Illumina in-house algorithm (Illumina Inc., San Diego, CA, USA) based on LD select by Carlson *et al*. (14) and data provided by the International HapMap Project (15, http://www.hapmap.org). After eliminating SNPs unsuitable for genotyping on the Illumina GoldenGate platform, one or two tag SNPs were selected per bin, prioritizing those SNPs most likely to perform well on the genotyping platform and putative functional SNPs. In total, 1127 tag SNPs were selected across 136 genes, tagging a total of 863 bins, with between 1 and 59 bins tagged per gene (average = 6 bins per gene, standard deviation ± 7.6, median = 5). In addition to these 1127 tag SNPs, 388 putative functional SNPs with CEU MAF ≥ 0.05 were added to the array, including 69 non-synonymous coding SNPs, 53 synonymous coding SNPs, 79 SNPs in UTR regions, 50 SNPs at or near putative splice sites, 124 SNPs with Phast conservation score >15 [log odds conservation score based on the multiple alignment of human, chimp, mouse, rat, dog, chicken, fugu and zebrafish (16)] and 13 putative transcription factor binding SNPs. Putative functional SNPs were selected using the Illumina in-house algorithm for SNP selection and PupaSuite web-based software (17, http://pupasuite.bioinfo.cipf.es/).

SNP genotyping and data manipulation

DNA was extracted from samples using conventional methodologies and quantified using PicoGreen (Invitrogen Corp., Carlsbad, CA, USA). Genotyping of samples was performed using customized Illumina GoldenGate Arrays (Illumina Inc.). DNA samples with GenCall scores <0.25 at any locus were considered 'no calls'. A DNA sample was deemed to have failed if it generated genotypes at fewer than 80% of loci. A SNP was deemed to have failed if fewer than 95% of DNA samples generated a genotype at the locus. To ensure quality of genotyping, a series of duplicate samples were genotyped and cases and controls were genotyped in the same batches.

Statistical methods

Statistical analyses were undertaken using R (http://www.r-project.org) and STATA Software (StataCorp LP, College Station, TX, USA). The adequacy of the case–control matching and possibility of differential genotyping of cases and controls were formally evaluated using Q-Q plots of test statistics. Genotypic frequencies in control subjects for each SNP were tested for departure from HWE using an exact test. To detect possible population stratification, we employed the genomic control approach (18) using all SNPs to estimate the stratification parameter and its associated 95% CI. The possibility of sex differences as a source of population substructure was evaluated by a genotype test for each SNP in male and female controls and the number of significant results at the 5% level was compared with the number expected by χ^2 test.

For each SNP, we tested the null hypothesis of no association with glioma using the Cochran–Armitage trend test calculated by logistic regression. As age and sex were not significantly associated with glioma risk within the data, we restricted adjustment to study centre. We assumed an additive co-dominant model by fitting the number of rare alleles carried as an ordinal covariate. The likelihood ratio test for each SNP was therefore a 1 df test comparing a model that included terms for genotype and study with a model including only a term for study. The risk of glioma associated with each SNP was quantified by trend ORs and their 95% CIs adjusted by logistic regression for study centre. To address multiple testing we calculated the FPRP for specific associations (19). We assumed that the study had power to detect an OR of 1.3 and that the prior probability for association with glioma risk was 0.01 (16). FPRP was calculated based on OR_{trend} statistics as per Wacholder *et al*. (19), and we considered an FPRP <0.200 to designate a noteworthy association in accordance with published criteria (7,19). Under certain conditions, a two-stage process incorporating estimates of pair-wise interactions between significant SNPs can yield greater power to detect association (20). To investigate epistatic interactions, each pair of SNPs that displayed a significant association at the 10% level was evaluated by fitting a saturated logistic regression model and the log likelihood ratio statistic for comparison with the main effects model was computed. This was compared against a χ^2 distribution with 1 df.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG Online.

ACKNOWLEDGEMENTS

We are grateful to all the patients and individuals for their participation. We would also like to thank the clinicians and other hospital staff, cancer registries and study staff and the funding bodies who contributed to the blood sample and data collection for this study and who are listed in our previous publications. The views expressed in the publication are those of the authors and not necessarily those of the funding bodies.

Conflict of Interest statement. None declared.

FUNDING

The current study was supported by grants from Cancer Research UK. The Interphone Study was supported by the

European Commission Fifth Framework Program 'Quality of Life and Management of Living Resources' (Contract QLK4-CT-1999-01563) and the International Union against Cancer (UICC; RCA/01/08). The UICC received funds for this study from the Mobile Manufacturers' Forum and the Global System for Mobile Communications Association. Provision of funds to the Interphone study investigators via UICC was governed by agreements that guaranteed Interphone's complete scientific independence. These agreements are publicly available at http://www.iarc.fr/ENG/Units/RCAd.html. Additional support was given to the Danish partner by the Danish Cancer Society. The Finnish partner received further financing from the Emil Aaltonen Foundation and Finnish Cancer Organizations. The Swedish partner received additional grant funding from the Swedish Research Council, The Swedish Cancer Society and the Cancer Foundation of Northern Sweden. The UK-Southeast and Northern centres were also supported by the Mobile Telecommunications and Health Research (MTHR) Program and the UK-North by the Health and Safety Executive, Department of Health and Safety Executive and the UK Network Operators (O_2, Orange, T-Mobile, Vodafone and '3').

REFERENCES

1. Schwartzbaum, J.A., Fisher, J.L., Aldape, K.D. and Wrensch, M. (2006) Epidemiology and molecular pathology of glioma. *Nat. Clin. Pract. Neurol.*, **2**, 494–503.
2. Hodgson, S. and Maher, E. (1999) *A Practical Guide to Human Cancer Genetics*, 2nd edn. Cambridge University Press, Cambridge, pp. 37–38.
3. Hemminki, K. and Li, X. (2003) Familial risks in nervous system tumors. *Cancer Epidemiol. Biomarkers Prev.*, **12**, 1137–1142.
4. Ohgaki, H. and Kleihues, P. (2005) Epidemiology and etiology of gliomas. *Acta Neuropathol. (Berl.)*, **109**, 93–108.
5. Cardis, E., Richardson, L., Deltour, I., Armstrong, B., Feychting, M., Johansen, C., Kilkenny, M., McKinney, P., Modan, B., Sadetzki, S. *et al.* (2007) The INTERPHONE study: design, epidemiological methods, and description of the study population. *Eur. J. Epidemiol.*, **22**, 647–664.
6. Schwartzbaum, J.A., Ahlbom, A., Lonn, S., Warholm, M., Rannug, A., Auvinen, A., Christensen, H.C., Henriksson, R., Johansen, C., Lindholm, C. *et al.* (2007) An international case-control study of glutathione transferase and functionally related polymorphisms and risk of primary adult brain tumors. *Cancer Epidemiol. Biomarkers Prev.*, **16**, 559–565.
7. Engels, E.A., Wu, X., Gu, J., Dong, Q., Liu, J. and Spitz, M.R. (2007) Systematic evaluation of genetic variants in the inflammation pathway and risk of lung cancer. *Cancer Res.*, **67**, 6520–6527.
8. Polo, S.E., Theocharis, S.E., Klijanienko, J., Savignoni, A., Asselain, B., Vielh, P. and Almouzni, G. (2004) Chromatin assembly factor-1, a marker of clinical value to distinguish quiescent from proliferating cells. *Cancer Res.*, **64**, 2371–2381.
9. Staibano, S., Mignogna, C., Lo Muzio, L., Mascolo, M., Salvatore, G., Di Benedetto, M., Califano, L., Rubini, C. and De Rosa, G. (2007) Chromatin assembly factor-1 (CAF-1)-mediated regulation of cell proliferation and DNA repair: a link with the biological behaviour of squamous cell carcinoma of the tongue? *Histopathology*, **50**, 911–919.
10. Jiao, R., Bachrati, C.Z., Pedrazzi, G., Kuster, P., Petkovic, M., Li, J.L., Egli, D., Hickson, I.D. and Stagljar, I. (2004) Physical and functional interaction between the Bloom's syndrome gene product and the largest subunit of chromatin assembly factor 1. *Mol. Cell Biol.*, **24**, 4710–4719.
11. Wood, R.D., Mitchell, M. and Lindahl, T. (2005) Human DNA repair genes, 2005. *Mutat. Res.*, **577**, 275–283.
12. Karolchik, D., Baertsch, R., Diekhans, M., Furey, T.S., Hinrichs, A., Lu, Y.T., Roskin, K.M., Schwartz, M., Sugnet, C.W., Thomas, D.J. *et al.* (2003) The UCSC Genome Browser Database. *Nucleic Acids Res.*, **31**, 51–54.
13. Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M. and Haussler, D. (2002) The human genome browser at UCSC. *Genome Res.*, **12**, 996–1006.
14. Carlson, C.S., Eberle, M.A., Rieder, M.J., Yi, Q., Kruglyak, L. and Nickerson, D.A. (2004) Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am. J. Hum. Genet.*, **74**, 106–120.
15. (2003) The International HapMap Project. *Nature*, **426**, 789–796.
16. Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S. *et al.* (2005) Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.*, **15**, 1034–1050.
17. Conde, L., Vaquerizas, J.M., Dopazo, H., Arbiza, L., Reumers, J., Rousseau, F., Schymkowitz, J. and Dopazo, J. (2006) PupaSuite: finding functional single nucleotide polymorphisms for large-scale genotyping purposes. *Nucleic Acids Res.*, **34**, W621–W625.
18. Devlin, B. and Roeder, K. (1999) Genomic control for association studies. *Biometrics*, **55**, 997–1004.
19. Wacholder, S., Chanock, S., Garcia-Closas, M., El Ghormli, L. and Rothman, N. (2004) Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J. Natl. Cancer Inst.*, **96**, 434–442.
20. Marchini, J., Donnelly, P. and Cardon, L.R. (2005) Genome-wide strategies for detecting multiple loci that influence complex diseases. *Nat. Genet.*, **37**, 413–417.