Research Article

Impact of *EGFR* Genetic Variants on Glioma Risk and Patient Outcome

Bruno Marques Costa[1,2], Marta Viana-Pereira[1,2], Ricardo Fernandes[1,2], Sandra Costa[1,2], Paulo Linhares[3], Rui Vaz[3], Céline Pinheiro[1,2], Jorge Lima[4], Paula Soares[4,5], Ana Silva[6], Fernando Pardal[6], Júlia Amorim[7], Rui Nabiço[7], Rui Almeida[8], Carlos Alegria[8], Manuel Melo Pires[9], Célia Pinheiro[10], Ernesto Carvalho[10], Pedro Oliveira[11], José M. Lopes[4,5], and Rui M. Reis[1,2,12]

Abstract

Background: The epidermal growth factor receptor (EGFR) regulates important cellular processes and is frequently implicated in human tumors. Three *EGFR* polymorphisms have been described as having a transcriptional regulatory function: two single-nucleotide polymorphisms in the essential promoter region, -216G/T and -191C/A, and a polymorphic $(CA)_n$ microsatellite sequence in intron 1. We aimed to elucidate the roles of these *EGFR* polymorphisms in glioma susceptibility and prognosis.

Methods: We conducted a case–control study with 196 patients with glioma and 168 cancer-free controls. Unconditional multivariate logistic regression models were used to calculate ORs and 95% confidence intervals. A Cox regression model was used to evaluate associations with patient survival. False-positive report probabilities were also assessed.

Results: None of the *EGFR* -216G/T variants was significantly associated with glioma risk. The -191C/A genotype was associated with higher risk for glioma when the $(CA)_n$ alleles were classified as short for ≤ 16 or ≤ 17 repeats. Independently of the $(CA)_n$ repeat cutoff point used, shorter $(CA)_n$ repeat variants were significantly associated with increased risk for glioma, particularly glioblastoma and oligodendroglioma. In all tested models with different $(CA)_n$ cutoff points, only -191C/A genotype was consistently associated with improved survival of patients with glioblastoma.

Conclusions: Our findings implicate *EGFR* -191C/A and the $(CA)_n$ repeat polymorphisms as risk factors for gliomas, and suggest -191C/A as a prognostic marker in glioblastoma.

Impact: Our data support a role of these *EGFR* polymorphisms in determining glioma susceptibility, with potential relevance for molecularly based stratification of patients with glioblastoma for individualized therapies. *Cancer Epidemiol Biomarkers Prev; ©2011 AACR.*

Introduction

Glioma is a broad category of tumors divided into histologic subgroups based on the type of glial cell of origin or morphologic similarities between tumor and normal glial cells: astrocytomas (astrocytic lineage), oligodendrogliomas (oligodendroglial lineage), and oligoastrocytomas (mixed lineage) are the major subtypes, whereas ependymomas (ependymal lineage) are less common (1, 2). They are the most common primary central nervous system tumors and account for approximately 80% of those that are malignant (2), for which efficient therapies are not available (3–5). Despite recent advances in the field of neuro-oncology, the prognosis of patients with glioma remains very poor (6); in particular, patients with glioblastoma, the most common and aggressive (WHO grade IV) form of glioma, present median survival time ranging from 12 to 15 months and the majority die within 2 years (5). Their etiology remains largely unknown: so far, only exposure to high-dose therapeutic

Authors' Affiliations: [1]Life and Health Sciences Research Institute (ICVS), School of Health Sciences, University of Minho, Braga, Portugal; [2]ICVS/3B's - PT Government Associate Laboratory, Braga/Guimarães, Portugal; [3]Department of Neurosurgery, Hospital S. João, Porto, Portugal; [4]IPATIMUP – Institute of Molecular Pathology and Immunology, University of Porto, Porto, Portugal; [5]Medical Faculty, Department of Pathology, University of Porto, Porto, Portugal; Departments of [6]Pathology, [7]Oncology, [8]Neurosurgery, Hospital S. Marcos, Braga, Portugal; [9]Unit of Neuropathology, [10]Department of Neurosurgery, Hospital S. António, Porto, Portugal; [11]Department of Population Studies, ICBAS, University of Porto, Porto, Portugal; [12]Molecular Oncology Research Center, Barretos Cancer Hospital, Barretos, São Paulo, Brazil

Corresponding Author: Rui M. Reis, Life and Health Sciences Research Institute (ICVS), School of Health Sciences, University of Minho, Campus de Gualtar, Braga 4710-057, Portugal. Phone: 351253604825; Fax: 351253604820; E-mail: rreis@ecsaude.uminho.pt

radiation has been firmly established as a risk factor, but other plausible causes include genetic syndromes, familial aggregation, and genetic polymorphisms (7, 8). Equally mysterious are truly clinically relevant prognostic factors for patients with glioma; patient age and clinical performance status are clearly associated with patient outcome, but recent evidences suggest that molecular traits of tumor may also be important determinants of prognosis (9–11).

The epidermal growth factor receptor (EGFR) pathway plays prominent roles in regulating cell growth, apoptosis, and differentiation, a process that is tightly regulated in normal epithelial cells (12). Deregulation of this pathway can occur, for example, via somatic *EGFR* mutations (e.g., *EGFRvIII*), gene amplifications, and ligand-independent constitutive stimulation signaling loops, and has been implicated in several human tumors, including gliomas (13). Aberrant EGFR signaling may ultimately affect major hallmarks of cancer including tumor growth, invasion, malignancy, and prognosis (13–16). In addition to these somatic molecular alterations, germ line polymorphic variants in the EGF/EGFR pathway have also been implicated in tumor risk, therapy response, and patient outcome (17–20). We have previously shown that a functional *EGF* polymorphism (*EGF* +61) is associated with glioma risk (19), which adds to the growing hypothesis that polymorphisms in the EGFR pathway may be relevant in the context of glioma.

Two single-nucleotide polymorphisms (SNP) were recently found in the essential promoter region of the *EGFR* gene ($-216G/T$ and $-191C/A$), and the T variant in the -216 SNP was shown to increase the promoter activity in an Sp1-dependent manner, resulting in higher gene expression both *in vitro* and *in vivo* (21). A study by Carpentier and colleagues (22) implicated the $-216T$ variant as a risk factor for glioblastoma. Likewise, a highly polymorphic microsatellite sequence $(CA)_n$ repeat was identified in the first intron of *EGFR* (23). It has been proposed that transcriptional activity of *EGFR* declines with increasing number of

$(CA)_n$ repeats, suggesting that this polymorphism may play a role in cancer susceptibility (24).

In the context of the widely acknowledged role of EGFR signaling in gliomas and the functional relevance of particular *EGFR* polymorphisms, we investigated the relevance of *EGFR* $-216G/T$, $-191C/A$, and $(CA)_n$ repeat polymorphisms in glioma susceptibility and patient prognosis. To do so, we conducted a population-based case–control study of 196 patients with glioma and 168 cancer-free controls from Portugal.

Materials and Methods

Study population

We studied a total of 196 patients with glioma recruited from Hospital S. Marcos, Braga, Hospital S. João, and Hospital S. António. Tumors were classified by experienced neuropathologists according to WHO standards (1). The control group is composed of 168 cancer-free individuals randomly selected from blood donors' bank. All patients and controls were from Northwest Portugal and of Caucasian ethnicity. Table 1 summarizes the clinicopathologic features of patients and controls. This study was conducted in accordance with institutional ethical committees. Blood samples were obtained following informed signed consent. All the samples enrolled were unlinked and unidentified from their donors.

Genotyping

Genomic DNA from patients with glioma was obtained from blood ($n = 70$) or formalin-fixed, paraffin-embedded tumor samples ($n = 126$) as previously reported (19). *EGFR* $-216G/T$ and $-191C/A$ genotyping was carried out by PCR amplification of the promoter region from nucleotides -352 to -155 (Fw primer: 5′-CTCCTCC-TCCTCTGCTCCTC-3′; Rv primer: 5′-GGGGCTAGC-TCGGGACTC-3′), followed by RFLP with *Bse*RI and *Sac*II respectively, as previously described (21). To evaluate

Table 1. Clinicopathologic features of gliomas and controls

Groups (WHO grade)	Number of cases	Age, y (mean ± SD)	Male/female ratio
Controls	168	44.2 ± 15.6	1.43
Gliomas (II–IV)	196	52.0 ± 14.1	1.23
Astrocytoma (II-IV)	136	54.3 ± 13.5	1.19
Diffuse astrocytoma (II)	15	40.1 ± 13.5	0.36
Anaplastic astrocytoma (III)	5	37.4 ± 11.6	All males
Glioblastoma (IV)	113	56.8 ± 12.1	1.31
Gliosarcoma (IV)	3	59.7 ± 5.69	0.50
Oligodendroglioma (II–III)	58	47.0 ± 14.3	1.32
Oligodendroglioma (II)	30	42.2 ± 16.1	1.31
Anaplastic oligodendroglioma (III)	28	52.0 ± 10.0	1.33
Oligoastrocytoma	2	42.0 ± 8.49	1.00
Anaplastic oligoastrocytoma (III)	2	42.0 ± 8.49	1.00

EGFR $(CA)_n$ repeats in intron 1, a 5-carboxyfluorescein (5-FAM)-labeled PCR reaction was carried out (Fw primer: 5-FAM–labeled 5′-GGCTCACAGCAAACTTCTCC-3′; Rv primer: 5′-AAGCCAGACTCGCTCATGTT-3′) and the products subsequently analyzed by single capillary genetic analysis, as previously described (25). Briefly, PCR cycles were 95°C for 7 minutes; followed by 35 cycles of 95°C for 45 seconds, 60°C for 30 seconds, and 72°C for 30 seconds; and a final step at 72°C for 7 minutes. PCR fragments ranging from 194 to 212 bp containing the polymorphic region were obtained. One microliter of PCR product and 0.6 μL of GeneScan 500 TAMRA size standard (Applied Biosystems) were mixed with 14.4 μL of formamide. PCR products were then separated using an ABI Prism 310 (Applied Biosystems), and the fragment lengths were determined with GeneScan Analysis software version 3.7 (Applied Biosystems).

Data analyses and statistical methods

EGFR $-216G/T$ and $-191C/A$ groups were categorized as previously described (21). Because *EGFR* $(CA)_n$ repeats vary between $n = 14$ and $n = 23$ and considering different approaches used in other studies (26–30), we evaluated *EGFR* $(CA)_n$ repeat length on the basis of 2 different methods. First, by the genotype-based method (26, 30), the genotypes of *EGFR* $(CA)_n$ repeat were categorized into 3 groups considering a cutoff point of 17 (CA) repeats: homozygous short (SS) contained 2 short alleles (both alleles ≤ 17 CA repeats), homozygous long (LL) contained 2 long alleles (both alleles >17 CA repeats), and heterozygous short/long (SL) contained one short (≤ 17) and one long (>17) allele. Because the cutoff value to consider short and long alleles has been somewhat controversial in the literature (30–33), we also tested 2 additional cutoff values to define short alleles with ≤ 16 and ≤ 18 CA repeats, and long alleles with >16 and >18 CA repeats, respectively (33). Second, we investigated the *EGFR* $(CA)_n$ repeat polymorphism by the summed allele lengths method (28), where individuals were categorized into short (sum of alleles length <36) or long (sum of alleles length ≥ 36).

The χ^2 test was used to assess whether the observed allele distributions of all polymorphisms were in Hardy–Weinberg equilibrium in the control group. ORs and 95% confidence intervals (95% CI) for the effect of *EGFR* variants on the risk for each glioma type were estimated by multivariate logistic regression analyses, adjusted for *EGF* $+61A/G$ genotype, which was previously reported to affect glioma susceptibility (19), and patient age (as a continuous variable) and sex. The false-positive report probability (FPRP) was calculated for significant associations observed in multivariate tests according to the study of Wacholder and colleagues (34). Associations between *EGFR* variants [$-216G/T$, $-191C/A$, and $(CA)_n$ repeat] and patient survival were assessed using a multivariate Cox regression model adjusted for *EGF* $+61A/G$, patient age, and sex (35).

Results

EGFR $-216G/T$, $-191C/A$, and $(CA)_n$ repeat polymorphisms and risk of glioma

We studied a total of 196 patients with glioma and 168 cancer-free control individuals. Successful determination of *EGFR* genotypes was achieved for all samples. The distributions of *EGF* $+61A/G$, *EGFR* $-216G/T$, and $(CA)_n$ allele frequencies in the control group were in Hardy–Weinberg equilibrium ($P = 0.121$, $P = 0.899$, and $P = 0.150$, respectively); however, *EGFR* $-191C/A$ alleles in controls were not in equilibrium ($P = 0.034$). A summary of clinicopathologic features of the controls and cases is presented in Table 1.

The frequencies of each *EGFR* genotype in controls and cases are presented in Table 2. The *EGFR* allele frequencies in our cancer-free control population are similar to those found in other control populations (21, 31, 33). The most frequent genotypes for each polymorphism in the control group (homozygous $-216G/G$, homozygous $-191C/C$, and heterozygous $(CA)_n$ repeat SL) were considered as references. To assess associations between each variant of the studied *EGFR* polymorphisms and risk for glioma, we used an unconditional multivariate logistic regression analysis adjusted for potential confounding variables (*EGF* $+61A/G$ genotype, patient age, and sex). The control group was compared with all glioma cases (WHO grades II–IV) and also with glioblastoma (WHO grade IV) and oligodendroglioma (WHO grades II and III; Table 2), as these were the most frequent subtypes in our series (Table 1).

No statistically significant associations were found between *EGFR* $-216G/T$ genotype variants and risk for glioma, glioblastoma, or oligodendroglioma (Table 2; all values of $P > 0.05$). In contrast, when compared with the $-191C/C$ reference genotype, the heterozygous $-191C/A$ genotype was significantly associated with increased risk for glioma (OR = 1.81; 95% CI, 1.01–3.24; Table 2). No significant associations between the $-191A/A$ genotype and risk for glioma, glioblastoma, or oligodendroglioma were found (Table 2; all values of $P > 0.05$). In addition, the SS genotype of $(CA)_n$ repeat (considering short alleles ≤ 17 CA repeats) was significantly associated with increased risk for glioma (OR = 2.38; 95% CI, 1.42–3.98), glioblastoma (OR = 2.25; 95% CI, 1.20–4.25), and oligodendroglioma (OR = 2.45; 95% CI, 1.17–5.12), as well as the LL genotype was significantly associated with increased risk for glioma (OR = 1.95; 95% CI, 1.02–3.73). As expected (19), *EGF* $+61A/G$ and G/G genotypes were significantly associated with increased risk for glioma (OR = 1.75; 95% CI, 1.03–2.96 and OR = 1.90; 95% CI, 1.03–3.52, respectively; Table 2), particularly oligodendroglioma (OR = 2.80; 95% CI, 1.18–6.67 and OR = 2.75; 95% CI 1.06–7.09, respectively; Table 2) and nearly statistically significant for glioblastoma (OR = 1.72; 95% CI, 0.91–3.26 and OR = 1.92; 95% CI, 0.91–4.03, respectively; Table 2).

When other cutoff values were used to classify short and long alleles for the $(CA)_n$ repeat polymorphism,

Table 2. Multivariate logistic regression analysis of associations between *EGFR/EGF* polymorphisms and risk for glioma groups

Polymorphism	Control	Glioma (grades II–IV)	OR (95% CI)	Glioblastoma (grade IV)	OR (95% CI)	Oligodendroglioma (grades II–III)	OR (95% CI)
EGFR −216G/T							
G/G	77	91	—	55	—	28	—
G/T	74	85	1.12 (0.68–1.84)	45	1.09 (0.58–2.03)	25	1.01 (0.49–2.07)
T/T	17	20	1.15 (0.52–2.59)	13	1.30 (0.50–3.40)	5	1.02 (0.31–3.40)
EGFR −191C/A							
C/C	130	140	—	77	—	41	—
C/A	32	49	**1.81 (1.01–3.24)**	30	1.97 (0.96–4.03)	16	2.08 (0.93–4.63)
A/A	6	7	0.68 (0.19–2.45)	6	0.88 (0.21–3.70)	1	0.47 (0.05–4.84)
EGFR (CA)$_n$ repeat[a]							
SL	92	75	—	42	—	23	—
LL	28	40	**1.95 (1.02–3.73)**	25	2.14 (0.96–4.76)	10	1.83 (0.68–4.94)
SS	48	81	**2.38 (1.42–3.98)**	46	**2.25 (1.20–4.25)**	25	**2.45 (1.17–5.12)**
EGF +61A/G							
A/A	57	47	—	30	—	9	—
A/G	73	96	**1.75 (1.03–2.96)**	54	1.72 (0.91–3.26)	32	**2.80 (1.18–6.67)**
G/G	38	53	**1.90 (1.03–3.51)**	29	1.92 (0.91–4.03)	17	**2.75 (1.06–7.09)**
Age			**1.04 (1.02–1.06)**		**1.07 (1.04–1.09)**		1.02 (1.00-1.04)
Sex							
Female	69	88	—	49	—	25	—
Male	99	108	0.95 (0.60–1.49)	64	1.02 (0.59–1.78)	33	0.98 (0.51–1.88)

NOTE: Bold-faced values indicate significant difference at 5% level.
[a](CA)$_n$ repeat considered short ≤ 17 and long > 17.

specifically $n = 16$ or $n = 18$ (i.e., S ≤ 16 and L > 16, or S ≤ 18 and L > 18, respectively), similar results were obtained; particularly, the (CA) repeat homozygous SS genotype was again significantly associated with increased risks for glioma, glioblastoma, and oligodendroglioma both for a cutoff value of S ≤ 16 repeats (Supplementary Table S1) and S ≤ 18 repeats (Supplementary Table S2). The LL genotype was also associated with increased risk for glioma and glioblastoma for the cutoff value of S ≤ 16 (Supplementary Table S1) and for glioblastoma only for the cutoff value of S ≤ 18 (Supplementary Table S2). Of note, the heterozygous −191C/A genotype was also associated with increased risk for glioma when the cutoff value for (CA)$_n$ short and long repeats was 16 (Supplementary Table S1) but not in the case of a cutoff value of 18 repeats (Supplementary Table S2). We also analyzed the associations between *EGFR* variants and risk considering the sum of alleles for the (CA)$_n$ repeat polymorphism (28), which in our series varied from 29 to 41 repeats. Because the median value for the sum of alleles in the control group was 36, we categorized the sum of alleles as short <36 and long ≥ 36 but none of the *EGFR* variants were significantly associated with risk in this analysis (Supplementary Table S3).

The calculation of FPRP showed that all of the abovementioned *EGFR* associations with risk remained noteworthy (FPRP ≤ 0.5) when a prior probability of associ-

ation of 10% or greater was considered (Table 3). This was also the case for *EGF* +61A/G associations, except in the case of the G/G genotype in oligodendroglioma risk (FPRP = 0.562; Table 3). For a prior probability of 5% or more, only the associations between (CA)$_n$ SS genotype and risk for glioma and glioblastoma remained noteworthy, which remained significant even for a prior probability of 1% in the case of glioma risk (Table 3).

Taken together, these data strongly suggest that shorter variants of the *EGFR* intron 1 (CA)$_n$ repeat polymorphism increase the risk for gliomas, particularly glioblastoma and oligodendroglioma.

EGFR −216G/T, −191C/A, and (CA)$_n$ repeat polymorphisms and survival of patients with glioblastoma

In a subset of patients with glioblastoma, we also had available follow-up data ($n = 63$). Thus, we investigated the associations between each *EGFR* variant (−216G/T, −191C/A, and (CA)$_n$ repeat) and overall survival by a multivariate Cox proportional hazard model, adjusted for *EGF* +61A/G, patient age, and sex.

None of the *EGFR* −216G/T variants was associated with survival of patients with glioblastoma (Table 4; Supplementary Tables S4–S6 when the (CA)$_n$ repeat variants were classified considering a cutoff value of S ≤ 17, S ≤ 16, S ≤ 18, or sum of alleles < 36, respectively).

Table 3. FPRP for significant associations with risk

Polymorphism	OR (95% CI)	Power[a]	Reported P	Prior probability			
				0.1	0.05	0.01	0.001
Glioma risk							
EGFR −191C/A (C/C ref.)	1.81 (1.01–3.24)	0.632	0.046	**0.395**	0.579	0.878	0.980
EGFR (CA)$_n$ LL (SL ref.)	1.95 (1.02–3.73)	0.530	0.044	**0.425**	0.609	0.890	0.988
EGFR (CA)$_n$ SS (SL ref.)	2.38 (1.42–3.98)	0.254	0.001	**0.033**	**0.066**	**0.270**	0.789
EGF +61A/G (A/A ref.)	1.75 (1.03–2.96)	0.691	0.037	**0.325**	0.504	0.841	0.982
EGF +61G/G (A/A ref.)	1.90 (1.03–3.51)	0.565	0.040	**0.391**	0.576	0.876	0.986
Glioblastoma risk							
EGFR (CA)$_n$ SS (SL ref.)	2.25 (1.20–4.25)	0.358	0.012	**0.238**	**0.398**	0.775	0.972
Oligodendroglioma risk							
EGFR (CA)$_n$ SS (SL ref.)	2.45 (1.17–5.12)	0.295	0.017	**0.344**	0.525	0.852	0.998
EGF +61A/G (A/A ref.)	2.80 (1.18–6.67)	0.224	0.020	**0.447**	0.630	0.899	0.989
EGF +61G/G (A/A ref.)	2.75 (1.06–7.09)	0.255	0.036	0.562	0.730	0.934	0.993

Note: Bold-faced values indicate the FPRP ≤ 0.5 for the most likely prior probability.
[a]Estimation of statistical power to detect an OR of 2.0 with an α level equal to the observed P value.

In contrast, the heterozygous −191C/A genotype was significantly associated with improved overall survival of patients with glioblastoma, as compared with homozygous −191C/C, which was consistent across all analyses, that is, when (CA)$_n$ alleles were classified as S $\leq$ 17 (OR = 0.37; 95% CI, 0.16–0.88; Table 4), S $\leq$ 16 (OR = 0.35; 95% CI, 0.15–0.82; Supplementary Table S4), S $\leq$ 18 (OR = 0.37; 95% CI, 0.16–0.82; Supplementary Table S5), or sum of alleles < 36 (OR = 0.39; 95% CI, 0.17–0.86; Supplementary Table S6).

The homozygous SS and LL (CA)$_n$ repeat genotypes were also significantly associated with a longer survival of patients with glioblastoma when the (CA)$_n$ repeat variants were classified as S $\leq$ 17 (OR = 0.33; 95% CI, 0.11–0.95 for LL genotype and OR = 0.41; 95% CI, 0.18–0.93 for SS genotype; Table 4). In the case of the SS genotype, this association with longer survival was maintained when a cutoff value of S $\leq$ 16 was considered (OR = 0.42; 95% CI, 0.18–0.95; Supplementary Table S4).

FPRP calculations showed that the association between −191C/A and (CA)$_n$ repeat SS genotypes and longer survival of patients with glioblastoma remained noteworthy for a prior probability of association of 10% or more (FPRP < 0.5 for all associations), which was not the case for LL genotypes (FPRP > 0.5, data not shown).

Taken together, these data strongly suggest that *EGFR* −191C/A and intron 1 (CA)$_n$ repeat polymorphisms are prognostic markers in patients with glioblastoma, whereas −216G/T variants do not seem to predict the outcome of patients with glioblastoma.

Discussion

Gliomas are the most frequent and malignant primary central nervous system tumors. These result in more years of life lost than do any other tumors (36) and are a significant source of cancer-related death (37). It is generally assumed that genetic and environmental factors contribute to gliomagenesis, but the etiology of gliomas remains very poorly understood. Presently, one of the several lines of brain tumor research focuses on the relevance of germ line genetic polymorphisms in glioma risk, grade, prognosis, and response to specific therapies. We have recently shown that an SNP in the *EGF* gene (which encodes one of the main EGFR ligands), *EGF* +61A/G, has functional consequences and associated the G allele with increased risk for glioma, particularly glioblastoma and oligodendroglioma (19).

The EGFR pathway is commonly altered in gliomas. Approximately 50% of glioblastomas show *EGFR* amplification and overexpression, 40% of which express the mutant form *EGFRvIII*, resulting in constitutive activation of the EGFR pathway (13, 38–40). Previously, Carpentier and colleagues (22) showed an association between the −216T allele and increased risk for glioblastoma. We attempted to replicate these findings, and examined, for the first time, the implication of −216G/T and −191C/A SNPs in other types of glioma. In opposition to the results of the study of Carpentier and colleagues, we have not seen association between *EGFR* −216T allele and increased risk for glioblastomas or any other glioma subtype. This discrepancy may be partially explained by distinct population sampling and the fact that we conducted a logistic regression adjusted for 2 other *EGFR* polymorphisms, together with *EGF* +61A/G, patient age, and sex. Importantly, the allele distribution of the −216G/T polymorphism in the control cancer-free population, we report here (67.9% G, 32.1% T, n = 168) is very similar to that reported by other studies [68.3% G, 31.7% T, n = 60 (ref. 21); 68.2% G, 31.8% T, n = 22 (ref. 41)]. Inversely, the study of

Table 4. Multivariate Cox proportional hazard model analysis of associations between *EGFR/EGF* polymorphisms and survival of patients with glioblastoma (adjusted for patient age and sex)

Polymorphism	Glioblastoma (grade IV)	OR (95% CI)
EGFR −216G/T		
G/G	33	—
G/T	23	0.49 (0.23–1.01)
T/T	7	0.61 (0.21–1.73)
EGFR −191C/A		
C/C	41	—
C/A	18	**0.37 (0.16–0.88)**
A/A	4	0.56 (0.12-2.69)
EGFR (CA)$_n$ repeat[a]		
SL	17	—
LL	17	**0.33 (0.11–0.95)**
SS	29	**0.41 (0.18–0.93)**
EGF +61A/G		
A/A	16	—
A/G	28	1.06 (0.47–2.38)
G/G	19	2.10 (0.89–4.99)
Age		1.00 (0.97–1.03)
Sex		
Female	25	—
Male	38	1.02 (0.54–1.92)

NOTE: Bold-faced values indicate significant difference at 5% level.
[a](CA)$_n$ repeat considered short ≤ 17 and long > 17.

Carpentier and colleagues shows a significantly different distribution of −216G/T alleles in the control group (53.2% G, 46.8% T, $n = 176$) when compared with our and others data [$\chi^2 (3) = 19.4, P < 0.001$] (21, 41). Even though the allele frequencies of this SNP vary greatly based on the ethnic background (21), this feature cannot explain the observed differences because all patients in our and others studies (21, 22, 41) were Caucasians. Concerning the −191C/A SNP, we observed that the heterozygous −191C/A genotype was associated with increased risk for glioma. A recent study by Schwartzbaum and colleagues (42) identified 3 *EGFR* SNPs consistently associated with glioblastoma risk across 4 independent data sets. In addition, these SNPs were highly correlated with 4 other *EGFR* SNPs previously found to be significantly associated with risk for glioma (43). Collectively, these and our data seem to support *EGFR* SNPs as potential risk factors for glioma.

Several studies suggest that *EGFR* expression is dependent on the number of the intron 1 (CA)$_n$ repeats (28, 30, 44, 45). In addition, this polymorphism has been associated with risk of breast, lung, and colorectal cancers

(31–33) but was never studied in gliomas. In this study, we provide the first evidence on the relevance of the *EGFR* (CA)$_n$ repeat length polymorphism in glioma risk and patient survival. Because different criteria have been published for the analysis of this polymorphism and no consensual cutoff point exists to distinguish short and long *EGFR* (CA)$_n$ repeat alleles (30–33), we used 3 different cutoff points to cover most of the previously published analysis (considering short alleles ≤16, ≤17, or ≤18 CA repeats) and evaluated the (CA)$_n$ repeat by the genotype, and the sum of the alleles length (considering the cutoff point for the sum of alleles as the median value in our control group). Our data show that the homozygous SS genotype of the (CA)$_n$ repeat polymorphism was associated with increased risk for glioma, glioblastoma, or oligodendroglioma, regardless of the selected cutoff point. The homozygous LL genotype was also found to be associated with increased risk for glioma and glioblastoma but only in some of the tested cutoff values. Thus, caution must be taken in the interpretation of the results and validation in an independent series is required.

Investigating the prognostic value of these 3 *EGFR* polymorphisms, we found a significant association between the heterozygous −191C/A genotype and improved survival of patients with glioblastoma, regardless of the criteria used for the cutoff value of the (CA)$_n$ repeat genotypes in the multivariate Cox model (Table 4, Supplementary Tables S4–S6). The SS and LL genotypes for the (CA)$_n$ repeat polymorphism were also associated with a better survival of patients with glioblastoma but only for specific cutoff values (SS when S ≤ 17 or S ≤ 16; LL when S ≤ 17); thus, the clinical relevance of this polymorphism warrants further confirmation.

In our study, the genotype assays of most patients with glioma were carried out in tumor tissue, raising the possibility that somatic alterations in the *EGFR* locus (7p12) could lead to misgenotyping. However, we believe that this is not the case because of the following reasons: (i) the overall distribution of *EGFR* −216 and −191 and (CA)$_n$ repeat genotypes and alleles was not statistically different for the group of patients whose genotyping was done in DNA from blood or tumor tissue ($P = 0.241$ for *EGFR* −216; $P = 0.176$ for *EGFR* −191; $P = 0.155$ for (CA)$_n$ repeats; data not shown); (ii) genotyping of all 3 polymorphisms was done with 100% concordance in 20 glioma cases from whom DNA was available from both peripheral blood and tumor tissue; (iii) we found no statistically significant associations between *EGFR* polymorphic variants and particular *EGFR* molecular alterations (*EGFR* amplification and *EGFRvIII* mutation; Supplementary Table S7) we had previously analyzed in a subset of tumors (13). Accordingly, in lung cancer, it was previously shown for the (CA)$_n$ repeats polymorphism that regardless the amplification status there was 100% concordance in the genotyping of tumor and nontumor tissues of 450 cases (15).

In summary, our data consistently indicate that *EGFR* intron 1 homozygous $(CA)_n$ repeat short genotypes confer higher susceptibility to develop different histologic entities of glioma, and implicate the heterozygous −191C/A genotype as a predictive marker of worse survival in patients with glioblastoma. Because EGFR is one of the most frequently altered molecules in high-grade glioma, it is natural to think of it as an attractive therapeutic target. Therefore, further studies are warranted to investigate how these *EGFR* polymorphisms may affect response of patients with glioma to EGFR-targeting therapies.

Disclosure of Potential Conflicts of interest

R.M. Reis has commercial research grant from Schering-Plough, Portugal. No potential conflicts of interest were disclosed by other authors.

References

1. Louis DN, Ohgaki H, Wiestler OD, Cavenee WK. WHO classification of tumours of the central nervous system. Lyon, France: IARC; 2007.
2. CBTRUS. Statistical report: primary brain tumors in the United States, 1998–2002. Hinsdale (IL): CBTRUS; 2005.
3. Reardon DA, Rich JN, Friedman HS, Bigner DD. Recent advances in the treatment of malignant astrocytoma. J Clin Oncol 2006;24:1253–65.
4. Jaeckle KA, Ballman KV, Rao RD, Jenkins RB, Buckner JC. Current strategies in treatment of oligodendroglioma: evolution of molecular signatures of response. J Clin Oncol 2006;24:1246–52.
5. Clarke J, Butowski N, Chang S. Recent advances in therapy for glioblastoma. Arch Neurol 2010;67:279–83.
6. Butowski NA, Sneed PK, Chang SM. Diagnosis and treatment of recurrent high-grade astrocytoma. J Clin Oncol 2006;24:1273–80.
7. Ohgaki H, Kleihues P. Epidemiology and etiology of gliomas. Acta Neuropathol (Berl) 2005;109:93–108.
8. Bondy ML, Scheurer ME, Malmer B, Barnholtz-Sloan JS, Davis FG, Il'Yasova D, et al. Brain tumor epidemiology: consensus from the Brain Tumor Epidemiology Consortium. Cancer 2008;113:1953–68.
9. Costa BM, Smith JS, Chen Y, Chen J, Phillips HS, Aldape KD, et al. Reversing HOXA9 oncogene activation by PI3K inhibition: epigenetic mechanism and prognostic significance in human glioblastoma. Cancer Res 2010;70:453–62.
10. Hegi ME, Diserens AC, Gorlia T, Hamou MF, de TN, Weller M, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. N Engl J Med 2005;352:997–1003.
11. Phillips HS, Kharbanda S, Chen R, Forrest WF, Soriano RH, Wu TD, et al. Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. Cancer Cell 2006;9:157–73.
12. Yarden Y, Sliwkowski MX. Untangling the ErbB signalling network. Nat Rev Mol Cell Biol 2001;2:127–37.
13. Viana-Pereira M, Lopes JM, Little S, Milanezi F, Basto D, Pardal F, et al. Analysis of EGFR overexpression, EGFR gene amplification and the EGFRvIII mutation in Portuguese high-grade gliomas. Anticancer Res 2008;28:913–20.
14. Zhu Y, Parada LF. The molecular and genetic basis of neurological tumours. Nat Rev Cancer 2002;2:616–26.
15. Nomura M, Shigematsu H, Li L, Suzuki M, Takahashi T, Estess P, et al. Polymorphisms, mutations, and amplification of the EGFR gene in non-small cell lung cancers. PLoS Med 2007;4:e125.
16. Weiss WA, Burns MJ, Hackett C, Aldape K, Hill JR, Kuriyama H, et al. Genetic determinants of malignancy in a mouse model for oligodendroglioma. Cancer Res 2003;63:1589–95.
17. Araujo A, Ribeiro R, Azevedo I, Coelho A, Soares M, Sousa B, et al. Genetic polymorphisms of the epidermal growth factor and related receptor in non-small cell lung cancer–a review of the literature. Oncologist 2007;12:201–10.
18. Zhang YM, Cao C, Liang K. Genetic polymorphism of epidermal growth factor 61A>G and cancer risk: a meta-analysis. Cancer Epidemiol 2010;34:150–6.
19. Costa BM, Ferreira P, Costa S, Canedo P, Oliveira P, Silva A, et al. Association between functional EGF+61 polymorphism and glioma risk. Clin Cancer Res 2007;13:2621–6.
20. Liu W, He L, Ramirez J, Krishnaswamy S, Kanteti R, Wang YC, et al. Functional EGFR germline polymorphisms may confer risk for EGFR somatic mutations in non-small cell lung cancer, with a predominant effect on exon 19 microdeletions. Cancer Res 2011;7:2423–7.
21. Liu W, Innocenti F, Wu MH, Desai AA, Dolan ME, Cook EH Jr, et al. A functional common polymorphism in a Sp1 recognition site of the epidermal growth factor receptor gene promoter. Cancer Res 2005;65:46–53.
22. Carpentier C, Laigle-Donadey F, Marie Y, Auger N, Benouaich-Amiel A, Lejeune J, et al. Polymorphism in Sp1 recognition site of the EGF receptor gene promoter and risk of glioblastoma. Neurology 2006;67:872–4.
23. Chi DD, Hing AV, Helms C, Steinbrueck T, Mishra SK, Donis-Keller H. Two chromosome 7 dinucleotide repeat polymorphisms at gene loci epidermal growth factor receptor (EGFR) and pro alpha 2 (l) collagen (COL1A2). Hum Mol Genet 1992;1:135.
24. Gebhardt F, Zanker KS, Brandt B. Modulation of epidermal growth factor receptor gene transcription by a polymorphic dinucleotide repeat in intron 1. J Biol Chem 1999;274:13176–80.
25. Butler JM, Buel E, Crivellente F, McCord BR. Forensic DNA typing by capillary electrophoresis using the ABI Prism 310 and 3100 genetic analyzers for STR analysis. Electrophoresis 2004;25:1397–412.
26. Kang D, Gridley G, Huang WY, Engel LS, Winn DM, Brown LM, et al. Microsatellite polymorphisms in the epidermal growth factor receptor (EGFR) gene and the transforming growth factor-alpha (TGFA) gene and risk of oral cancer in Puerto Rico. Pharmacogenet Genomics 2005;15:343–7.
27. McKay JA, Murray LJ, Curran S, Ross VG, Clark C, Murray GI, et al. Evaluation of the epidermal growth factor receptor (EGFR) in colorectal tumours and lymph node metastases. Eur J Cancer 2002;38:2258–64.
28. Amador ML, Oppenheimer D, Perea S, Maitra A, Cusatis G, Iacobuzio-Donahue C, et al. An epidermal growth factor receptor intron 1 polymorphism mediates response to epidermal growth factor receptor inhibitors. Cancer Res 2004;64:9139–43.
29. Giovannucci E, Stampfer MJ, Krithivas K, Brown M, Dahl D, Brufsky A, et al. The CAG repeat within the androgen receptor gene and its relationship to prostate cancer. Proc Natl Acad Sci U S A 1997;94:3320–3.

30. Etienne-Grimaldi MC, Pereira S, Magne N, Formento JL, Francoual M, Fontana X, et al. Analysis of the dinucleotide repeat polymorphism in the epidermal growth factor receptor (EGFR) gene in head and neck cancer patients. Ann Oncol 2005;16:934–41.

31. Zhang W, Weissfeld JL, Romkes M, Land SR, Grandis JR, Siegfried JM. Association of the EGFR intron 1 CA repeat length with lung cancer risk. Mol Carcinog 2007;46:372–80.

32. Zhang W, Park DJ, Lu B, Yang DY, Gordon M, Groshen S, et al. Epidermal growth factor receptor gene polymorphisms predict pelvic recurrence in patients with rectal cancer treated with chemoradiation. Clin Cancer Res 2005;11:600–5.

33. Brandt B, Hermann S, Straif K, Tidow N, Buerger H, Chang-Claude J. Modification of breast cancer risk in young women by a polymorphic sequence in the egfr gene. Cancer Res 2004;64:7–12.

34. Wacholder S, Chanock S, Garcia-Closas M, El GL, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. J Natl Cancer Inst 2004;96:434–42.

35. Costa BM, Caeiro C, Guimaraes I, Martinho O, Jaraquemada T, Augusto I, et al. Prognostic value of MGMT promoter methylation in glioblastoma patients treated with temozolomide-based chemoradiation: a Portuguese multicentre study. Oncol Rep 2010;23:1655–62.

36. Burnet NG, Jefferies SJ, Benson RJ, Hunt DP, Treasure FP. Years of life lost (YLL) from cancer is an important measure of population burden–and should be considered when allocating research funds. Br J Cancer 2005;92:241–5.

37. Jemal A, Siegel R, Ward E, Hao YP, Xu JQ, Thun MJ. Cancer statistics, 2009. CA Cancer J Clin 2009;59:225–49.

38. Humphrey PA, Wong AJ, Vogelstein B, Zalutsky MR, Fuller GN, Archer GE, et al. Anti-synthetic peptide antibody reacting at the fusion junction of deletion-mutant epidermal growth factor receptors in human glioblastoma. Proc Natl Acad Sci U S A 1990;87:4207–11.

39. Wong AJ, Ruppert JM, Bigner SH, Grzeschik CH, Humphrey PA, Bigner DS, et al. Structural alterations of the epidermal growth factor receptor gene in human gliomas. Proc Natl Acad Sci U S A 1992;89:2965–9.

40. Heimberger AB, Suki D, Yang D, Shi W, Aldape K. The natural history of EGFR and EGFRvIII in glioblastoma patients. J Transl Med 2005;3:38.

41. Liu W, Wu X, Zhang W, Montenegro RC, Fackenthal DL, Spitz JA, et al. Relationship of EGFR mutations, expression, amplification, and polymorphisms to epidermal growth factor receptor inhibitors in the NCI60 cell lines. Clin Cancer Res 2007;13:6788–95.

42. Schwartzbaum JA, Xiao Y, Liu Y, Tsavachidis S, Berger MS, Bondy ML, et al. Inherited variation in immune genes and pathways and glioblastoma risk. Carcinogenesis 2010;31:1770–7.

43. Andersson U, Schwartzbaum J, Wiklund F, Sjöström S, Liu Y, Tsavachidis S, et al. A comprehensive study of the association between the EGFR and ERBB2 genes and glioma risk. Acta Oncol 2010;49:767–75.

44. Buerger H, Packeisen J, Boecker A, Tidow N, Kersting C, Bielawski K, et al. Allelic length of a CA dinucleotide repeat in the egfr gene correlates with the frequency of amplifications of this sequence–first results of an inter-ethnic breast cancer study. J Pathol 2004;203:545–50.

45. Buerger H, Gebhardt F, Schmidt H, Beckmann A, Hutmacher K, Simon R, et al. Length and loss of heterozygosity of an intron 1 polymorphic sequence of egfr is related to cytogenetic alterations and epithelial growth factor receptor expression. Cancer Res 2000;60:854–7.