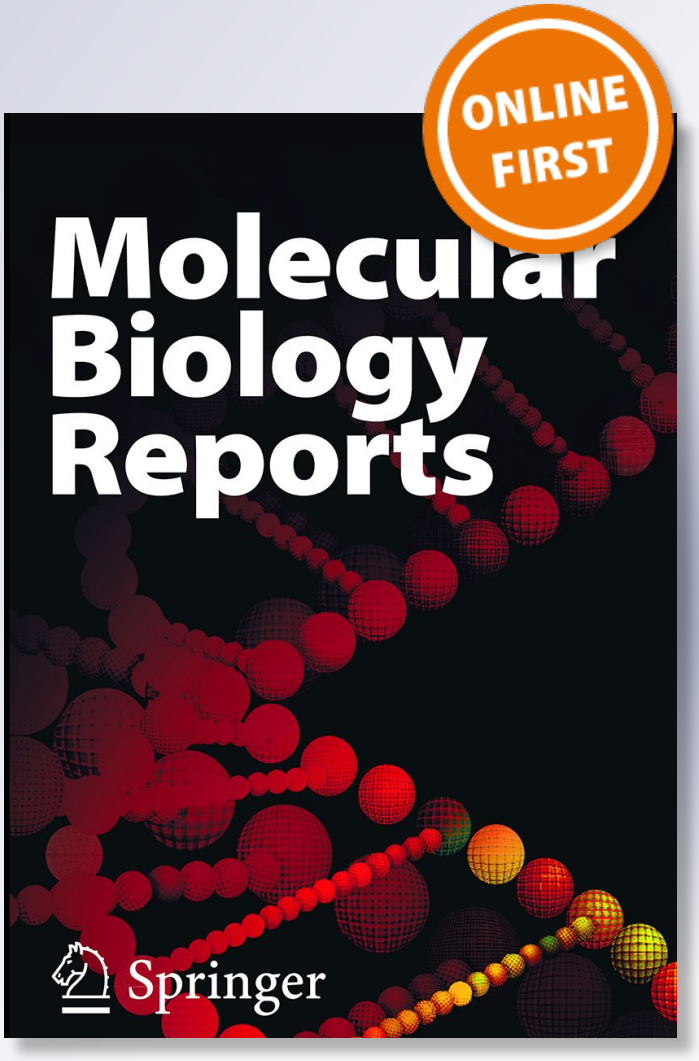
EGF+61 A>G polymorphism is not associated with lung cancer risk in the Brazilian population

Ana Carolina Laus, Flavia Escremim de Paula, Marcos Alves de Lima, Carolina Dias Carlos, Izabela Natalia Faria Gomes, et al.

Molecular Biology Reports
An International Journal on Molecular and Cellular Biology

ISSN 0301-4851

Mol Biol Rep
DOI 10.1007/s11033-019-04702-0



Springer

Molecular Biology Reports
https://doi.org/10.1007/s11033-019-04702-0

ORIGINAL ARTICLE

EGF+61 A>G polymorphism is not associated with lung cancer risk in the Brazilian population

Ana Carolina Laus[1] · Flavia Escremim de Paula[1] · Marcos Alves de Lima[2] · Carolina Dias Carlos[1] · Izabela Natalia Faria Gomes[1] · Pedro de Marchi[3] · Jenna Kadja Neves Valente[3] · Ana Beatriz Maringolo Pioltini[3] · José Elias Miziara[4] · Carlos Maciel da Silva[4] · Luciano de Souza Viana[3] · Cristovam Scapulatempo-Neto[5] · Rui Manuel Reis[1,6,7]

Received: 27 October 2018 / Accepted: 12 February 2019

Abstract

Epidermal growth factor (EGF) and its receptor (EGFR) play an important role in lung carcinogenesis. A functional single nucleotide polymorphism (SNP) in *EGF* promoter region (*EGF*+61 A>G—rs4444903) has been associated with cancer susceptibility. Yet, in lung cancer, the *EGF*+61 A>G role is unclear. The aim of this study was to evaluate the risk of lung cancer associated with *EGF*+61 A>G SNP in the Brazilian population. For that, 669 lung cancer patients and 1104 controls were analyzed. *EGF*+61 A>G genotype was assessed by PCR-RFLP and TaqMan genotyping assay. Both patients and controls were in Hardy–Weinberg equilibrium. As expected, uni- and multivariate analyses showed that tobacco consumption and age were significant risk factors for lung cancer. The genotype frequencies in lung cancer patients were 27.3% of AA, 47.4% of AG and 25.3% of GG, and for controls were 25.3% of AA, 51.6% of AG and 23.1% of GG. The allele frequencies were 51.1% of A and 48.9% of G for both cases and controls. No significant differences for the three genotypes (AA, AG and GG—codominant model) were observed between cases and controls. We then grouped AG and GG (recessive model) genotypes, as well as AA and AG (dominant model), and again, no significant differences were also found. This is the largest study to explore *EGF*+61 A>G polymorphism association with lung cancer risk and suggests that this SNP is not a risk factor for lung cancer in the Brazilian population.

Keywords Lung cancer · SNP · Risk factor · *EGF*+61 A>G polymorphism

Introduction

Lung cancer is a malignancy with high incidence and mortality [1]. Globocan 2018 estimated more than 2 million new cases/year worldwide and more than 1.7 million deaths [1]. In Brazil, the National Cancer Institute estimated 31,000 new cases of lung and trachea cancer in 2018, ranking lung cancer as the second most common cancer in men and the fourth in women [2]. The Brazilian Ministry of Health calculated that 15,514 men and 10,978 women died due to lung cancer in 2015 [3]. Depending on the region of the country, the estimative can also change, since lung cancer is more frequent in South and Southwest

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11033-019-04702-0) contains supplementary material, which is available to authorized users.

✉ Rui Manuel Reis
ruireis.hcb@gmail.com

[1] Molecular Oncology Research Center, Barretos Cancer Hospital, Antenor Duarte Villela St, 1331, Barretos, SP 14784-400, Brazil

[2] Epidemiology and Biostatistics Department, Barretos Cancer Hospital, Barretos, Brazil

[3] Medical Oncology Department, Barretos Cancer Hospital, Barretos, Brazil

[4] Surgery Department, Barretos Cancer Hospital, Barretos, Brazil

[5] Pathology Department, Barretos Cancer Hospital, Barretos, Brazil

[6] Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, Braga, Portugal

[7] ICVS/3B's – PT Government Associate Laboratory, Braga/Guimarães, Portugal

Published online: 19 February 2019

regions than Northwest and North, both for men and women [2].

Lung cancer can be histologically classified according to cell type in small-cell lung cancer and non-small-cell lung cancer, which can be categorized as adenocarcinoma, squamous cell carcinoma and large cell carcinoma [4]. The main risk factor associated is tobacco consumption, but the incidence of lung cancer in non-smokers has been increasing in the last years, suggesting that other risk factors, both exogenous, such as air pollution, environmental exposure, and genetic, like hereditary and single nucleotide polymorphisms (SNP) can be associated with lung cancer [4].

In the context of molecular pathways involved in lung cancer, activation by the epidermal growth factor receptor (EGFR) and epidermal growth factor (EGF) binding are extremely relevant, since EGF/EGFR ligation activate signaling pathways (RAS-RAF-MEK-ERK MAPK, PI3K-AKT-mTOR and PLC-1-PKC, JNK, and JAK-STAT pathways), which lead to uncontrolled proliferation, angiogenesis, inhibition of apoptosis, invasion, metastasis, and immortalization [5–7]. EGFR plays an important role in lung carcinogenesis, and present significant rates of mutations, amplification and overexpression [8–10]. SNPs located in the promoter region of *EGFR* have also been associated with increased risk of lung cancer or altered response to drug therapy [11].

In 2002, Shahbazi et al. screened the promoter region of *EGF* and identified an A to G nucleotide substitution at position 61, called *EGF*+61 A>G (rs4444903), that was further genotyped in 135 melanoma patients and 99 healthy controls. A significant association between homozygosity for the *EGF* 61*G allele and the development of malignant melanoma was found [12]. They also showed that cells from EGF 61*A homozygous individuals produced significantly less EGF than cells from GG or AG individuals [12]. In 2007, Costa et al. analyzed *EGF*+61 A>G in 197 glioma patients and 570 cancer-free individuals and showed that the *EGF* 61*G allele conferred higher risk for gliomas and *in vitro* assays demonstrated a significant higher transcriptional activity of EGF+61*G allele, when compared with EGF+61*A allele [13].

Following these studies, many other groups evaluated the association between this SNP and other tumor types, with distinct results, leading consequently to several meta-analyzes with thousands of participants. Significant associations were found in gastric cancer [14], hepatocellular carcinoma [15], and gliomas [16].

In the lung cancer context, Lim et al. [17] found the G allele as a risk factor in the Korean population, whereas Kang et al. [18], also evaluating Korean patients, did not find any association. In a Portuguese population, AG, GG and AG + GG were risk factors for lung cancer [19], whereas Masroor et al. observed that Indian patients harboring AG, GG and AG + GG presented higher risk for lung cancer [20].

Therefore, the aim of this study was to assess the risk of lung cancer associated with the SNP *EGF*+61 A>G in a large Brazilian study population.

Methods

Study population

This is a Brazilian case-control study, which included patients with lung cancer, diagnosed between 2001 and 2015 at Barretos Cancer Hospital, and a control group consisting of cancer-free individuals recruited between 2012 and 2014 also at Barretos Cancer Hospital. Demographic and clinical pathological data from cases and controls are summarized in Table 1.

The patient group consisted of 669 cases (403 males and 266 females). The demographic and clinical-pathological data were obtained by medical records review, and included information about diagnosis, staging, treatment and follow up. The median age was 66.3 years (SD ± 11.29), with 512 individuals reporting being of white skin color, 322 were current and 182 were former smokers. Moreover, 70.3% of the cases presented adenocarcinomas, 24.2% squamous cell carcinomas, and 5.5% other histology (Table 1).

The control group was composed of 1218 individuals (597 males, 500 females and for 121 gender was not reported). The median age was 56.6 years (SD ± 12.5), 813 reporting being of white skin color, and 164 were current and 313 were former smokers (Table 1).

DNA isolation

For all 1218 controls, DNA was isolated from blood using the QIAsymphony automated system and QIAsymphony DNA Mini Kit (Qiagen), according to manufacture instructions.

In the patient group, from a subset of 232 individuals, the DNA was isolated from blood using the QIAsymphony automated system and QIAsymphony DNA Mini Kit (Qiagen). From 437 patients, since blood was not available, formalin-fixed paraffin-embedded (FFPE) tumor tissue was the DNA source and was isolated using QIAamp DNA FFPE Tissue Kit (Qiagen), also according to manufacture instructions.

EGF+61 A>G genotyping

Genotyping was performed using two methodologies in accordance to the type of tissue used for DNA isolation. Quantitative real time PCR was the first choice since it is the fastest and easier methodology, however, since it showed

Table 1 Description and univariate comparison of lung cancer patients and controls according to demographic and clinical-pathological features

Characteristics	Cases	Controls	p
No	669	1104	
Gender			
Male	403 (60.2%)	597 (54.4%)	**0.017**
Female	266 (39.8%)	500 (45.6%)	
NA-data unavailable	–	7	
Age			
Median (SD*)	66.3 (± 11.29)	56.6 (± 12.5)	< **0.001**
≤65 years	323 (48.3%)	851 (77.7%)	
>65 years	346 (51.7%)	244 (22.3%)	
NA-data unavailable	–	9	
Self-reported skin color			
White	512 (79.3%)	813 (75%)	**0.043**
Other	134 (20.7%)	271 (25%)	
NA-data unavailable	23	20	
Smoking status			
Never	154 (23.4%)	616 (56.4%)	< **0.001**
Current	322 (48.9%)	164 (15%)	
Former	182 (27.7%)	313 (28.6%)	
NA-data unavailable	11	11	
Alcohol consumption			
Never	386 (60.6%)	513 (48%)	< **0.001**
Current	197 (30.9%)	420 (39.3%)	
Former	54 (8.5%)	135 (12.7%)	
NA-data unavailable	32	36	
Histology			
Adenocarcinoma	462 (70.3%)	–	
Squamous cell carcinoma	159 (24.2%)		
Others	36 (5.5%)		
NA-data unavailable	12		
ECOG PS			
0–1	449 (68.7%)	–	
2–3	178 (27.2%)		
4	27 (4.1%)		
NA-data unavailable	15		
TNM stage			
0/I/II	94 (14.4%)	–	
III/IV	559 (85.6%)		
NA-data unavailable	16		
Metastasis at diagnosis			
No metastasis	118 (22.5%)	–	
Metastasis in one organ	162 (30.8%)		
Metastasis in multiple organs	245 (46.7%)		
NA-data unavailable	144		

Bold value indicates $p \leq 0.005$

*SD** standard deviation

low efficiency in DNA isolated from FFPE tissues, RFLP was applied in those samples. Briefly, for DNAs isolated from FFPE tumor samples, genotyping was performed by polymerase chain reaction–restriction fragment length polymorphism (PCR-RFLP), as previously reported [13]. 50 ng of DNA was amplified by PCR using with forward- CAG GTAATGGAGCGAAGCTTTCAT and reverse- GAGTTA AATGCCTACACTGTATCT primers, producing a 242-bp fragment. The product was then digested with AluI (10U/uL) restriction enzyme (Invitrogen) at 37 °C for 16 h, followed by inactivation at 65 °C for 20 min. Digested PCR products were then visualized in agarose gel to distinguish+61 alleles: G alleles produce two fragments of 91 and 102 bp, while A allele produces only one fragment of 193 bp [13].

DNAs isolated from blood samples were genotyped by quantitative Real Time PCR, using a commercially available TaqMan Genotyping Assay (ThermoFisher, USA) [C_27031637_10 (*EGF*+61 A>G)] in the QuantStudio™ 6 Flex Real-Time PCR System (ThermoFisher, USA) and under standard cycling.

Statistical analysis

Hardy–Weinberg equilibrium was calculated in each group. To verify whether there were differences in genotype frequencies depending on the type of sample evaluated (blood or FFPE), agreement Kappa test was used, as well a χ^2 test.

χ^2 test was also used to compare cases and controls according to the frequencies distribution of gender, age, skin color, smoking status, alcohol consumption, genotypes and alleles. *EGF*+61 A>G genotype was analyzed using three different combinations: codominant model—AA versus AG versus GG, recessive model—AA vs AG + GG, and dominant model—AA + AG versus GG. Multivariate analysis was used to access the association between demographic information and risk of lung cancer. Also, *EGF*+61 A>G genotype and risk of lung cancer was estimated by computing Odds Ratio (ORs) with 95% confidence interval (CI), without adjustment and adjusted by age, smoking status and alcohol consumption.

Considering only the lung cancer group, χ^2 test was performed to correlate genotype and clinical pathological features.

All statistical analysis were performed using IBM® SPSS Statistics, version 20, considering a significant p-value < 0.05.

Results

Due to pre-analytic blood collecting issues, which led to poor DNA quality, 114 controls were excluded, totalizing 1104 controls and 669 cases genotyped. Both groups

were in Hardy–Weinberg equilibrium (cases—p = 0.17; controls—p = 0.27).

Since we performed the *EGF*+61 A>G genotyping in DNA derived from both blood and FFPE tissue in the cases group, we initially tested the concordance between both sources in a subset of 40 cases from which both blood and FFPE samples were available, and the Kappa test revealed a high agreement (Kappa value = 0.912). Moreover, in the patient group, the genotypic and allelic frequencies of DNA derived from blood and FFPE were compared, and no differences were observed (Supplemental Table S1). Therefore, in the further statistical analysis, we considered both blood and FFPE samples from patients as a single group.

The univariate analysis of the demographic and clinical pathological information from patients and controls showed a statistically significant difference between groups according to gender (p = 0.017), age (p < 0.001), skin color (p = 0.043), smoking status (p < 0.001) and alcohol consumption (p < 0.001) (Table 1). In non-adjusted multivariate analysis, the following demographic characteristics were statistically associated with a higher risk of lung cancer: gender (male—OR 1.26, 95% CI 1.04–1.54, p = 0.017), age (> 65 years—OR 3.73, 95% CI 3.03–4.6, p < 0.001), skin color (other than white—OR 0.78, 95% CI 0.62–0.99, p = 0.043) and tobacco consumption (former smoker—OR 2.32, 95% CI 1.80–2.99, p < 0.001; current smoker—OR 7.85, 95% CI 6.06–10.16, p < 0.001). Alcohol consumption revealed as a protective factor (former consumer—OR 0.53, 95% CI 0.37–0.74, p < 0.001; current consumer—OR 0.62, 95% CI 0.50–0.77, p < 0.001) (Table 2). In the analysis adjusted by gender, age, skin color, tobacco and alcohol consumption, only age, smoking status and alcohol consumption remained statistically significant (> 65 years—OR 3.67, 95% CI 2.89–4.68, p < 0.001; former smoker—OR 3.26, 95% CI 2.41–4.42, p < 0.001; current smoker—OR 11.51, 95% CI 8.39–15.79, p < 0.001; former consumer—OR 0.32, 95% CI 0.21–0.48, p < 0.001; current consumer—OR 0.35, 95% CI 0.26–0.47, p < 0.001) (Table 2).

The *EGF*+61 A>G genotypic frequencies in cases were 27.3% of AA, 47.4% of AG and 25.3% of GG, and for controls, they were 25.3% of AA, 51.6% of AG and 23.1% of GG, and the allelic distribution in patients was 51.1% of A and 48.9% of G, and the same frequencies were observed in control group (Table 3). The univariate analysis revealed no differences between cases and controls according to genotype in any model tested (codominant model—p = 0.222; recessive model—p = 0.333, and dominant model—p = 0.301) and also allelic frequencies (p > 0.999) (Table 3).

The multivariate analysis was further performed to access the risk of lung cancer associated with the SNP, both without adjustment and adjusting for age, smoking status and alcohol consumption. In none of the analysis, the polymorphism was associated with increased risk for lung cancer (Table 4). Namely, without any adjustment, codominant model presented for AG genotype an OR of 0.84, 95% CI 0.67–1.06, p = 0.163 and GG genotype OR 1.01, 95% CI 0.77–1.32, p = 0.94; recessive model presented for AG + GG an OR of 0.89, 95% CI 0.72–1.11, p = 0.333, and the dominant model presented for GG genotype an OR of 1.12, 95% CI 0.90–1.40, p = 0.301

Table 2 Multivariate analysis of demographic features, without and with adjustment for gender, age, skin color, tobacco and alcohol consumption

	Cases	Controls	Without adjustment		With adjustment	
			OR (95% CI)	p	OR (95% CI)	p
Gender						
Female	266 (39.8%)	500 (45.6%)	1		1	
Male	403 (60.2%)	597 (54.4%)	1.26 (1.04–1.54)	**0.017**	1.06 (0.82–1.39)	0.615
Skin color						
White	512 (79.3%)	813 (75%)	1		1	
Other	134 (20.7%)	271 (25%)	0.78 (0.62–0.99)	**0.043**	0.86 (0.65–1.14)	0.305
Age						
≤ 65 years	323 (48.3%)	851 (77.7%)	1		1	
> 65 years	346 (51.7%)	244 (22.3%)	3.73 (3.03–4.60)	< **0.001**	3.67 (2.89–4.68)	< **0.001**
Smoking status						
Never	154 (23.4%)	616 (56.4%)	1		1	
Current	322 (48.9%)	164 (15%)	7.85 (6.06–10.16)	< **0.001**	11.51 (8.39–15.79)	< **0.001**
Former	182 (27.7%)	313 (28.6%)	2.32 (1.80–2.99)	< **0.001**	3.26 (2.41–4.42)	< **0.001**
Alcohol consumption						
Never	386 (60.6%)	513 (48%)	1		1	
Current	197 (30.9%)	420 (39.3%)	0.62 (0.50–0.77)	< **0.001**	0.35 (0.26–0.47)	< **0.001**
Former	54 (8.5%)	135 (12.7%)	0.53 (0.37–0.74)	< **0.001**	0.32 (0.21–0.48)	< **0.001**

Bold value indicates $p \leq 0.005$

Table 3 Genotypic and allelic frequencies of lung cancer patients and controls, and univariate analysis by χ^2 test comparing the two groups

Characteristics	Cases	Controls	p
Genotype (codominant model)			
AA	183 (27.3%)	279 (25.3%)	0.222
AG	317 (47.4%)	570 (51.6%)	
GG	169 (25.3%)	255 (23.1%)	
Genotype (recessive model)			
AA	183 (27.4%)	279 (25.3%)	0.333
AG + GG	486 (72.6%)	825 (74.7%)	
Genotype (dominant model)			
AA + AG	500 (74.7%)	849 (76.9%)	0.301
GG	169 (25.3%)	255 (23.1%)	
Allele frequency			
A	51.1%	51.1%	> 0.999
G	48.9%	48.9%	

(Table 4). For adjusted analysis, in codominant model, AG presented OR 0.80, 95% CI 0.61–1.05, p = 0.116, and GG presented OR 1.03, 95% CI 0.75–1.42, p = 0.834. In recessive model, AG + GG presented OR 0.87, 95% CI 0.67–1.12 (p = 0.3), and in dominant model, GG presented OR 1.19, 95% CI 0.91–1.55 (p = 0.197) (Table 4).

Moreover, we evaluated whether *EGF*+61 A>G was associated with lung cancer clinical pathological features. The χ^2 test revealed, in codominant and dominant models, an association between patients with GG genotypes and T3/T4 tumor status (codominant model—p = 0.05 and dominant model—p = 0.035) (Table 5). No other association was found between genotype and patients' clinical pathological characteristics (Table 5).

Discussion

The main risk factors associated with lung cancer are well known and include tobacco consumption, air pollution and environmental exposure, together with genetic variations in genes associated with carcinogeneis [4, 21]. In this study, we analyzed genotypic and allelic frequencies of *EGF*+61 A>G polymorphism in a Brazilian population composed of 669 lung cancer patients and 1104 controls. To our knowledge, this is the largest study in lung cancer that included the highest number of participants, and our results suggest no association between this SNP and lung cancer risk.

The *EGF*+61 A>G polymorphism was previously evaluated as risk factor for several cancers, such as melanoma [12], esophageal carcinoma [22], prostate [23] and colorectal cancer [24]. Many of these studies revealed an association between the SNP and susceptibility to cancer, and indeed, this association is supported by studies that indicated increased levels of EGF in cells carrying *EGF*+61*G allele [12, 13]. More recently, meta-analysis revealed an association between *EGF*+61*G allele and higher risk to hepatocellular carcinoma [15], gastric cancer [14] and gliomas [16], but did not find an association with breast cancer, cervical cancer, and melanoma [25].

In lung cancer, the results are limited and conflicting [17–20]. So far, only four studies analyzed *EGF*+61 A>G SNP and lung cancer risk, being three in an Asiatic population and one in Causasians (Portuguese). Lim et al. evaluated a Korean population, composed by 122 patients and 132 controls, and found an association between risk and the SNP (OR 2.32, 95% CI 1.60–3.36), while Kang et al. that also studied a Korean population (432 patients and 432 controls), did not find the same association (AA – OR 0.81, 95% CI 0.51–1.29; AG–OR 1.02, 95% CI 0.77–1.37; and AA + AG–OR 0.98, 95% CI 0.74–1.29) [17, 18]. In 2012, de Mello et al. evaluated 112 lung cancer patients and 126 controls from Portugal and found that AG and GG genotypes

Table 4 Multivariate analysis by logistic regression for *EGF*+61 A>G genotypes, without and with adjustment for age, smoking status and alcohol consumption

	Cases	Controls	Without adjustment		With adjustment	
			OR (95% CI)	p	OR (95% CI)	p
Codominant model						
AA	183 (27.3%)	279 (25.3%)	1		1	
AG	317 (47.4%)	570 (51.6%)	0.84 (0.67–1.06)	0.163	0.80 (0.61–1.05)	0.116
GG	169 (25.3%)	255 (23.1%)	1.01 (0.77–1.32)	0.94	1.03 (0.75–1.42)	0.834
Recessive model						
AA	183 (27.4%)	279 (25.3%)	1		1	
AG + GG	486 (72.6%)	825 (74.7%)	0.89 (0.72–1.11)	0.333	0.87 (0.67–1.12)	0.300
Dominant model						
AA + AG	500 (74.7%)	849 (76.9%)	1		1	
GG	169 (25.3%)	255 (23.1%)	1.12 (0.90–1.40)	0.301	1.19 (0.91–1.55)	0.197

Table 5 Univariate analysis by χ^2 test in case group, associating clinical pathological characteristics and *EGF*+61 A>G genotype, considering the three models

Characteristics	Codominant model				Recessive model			Dominant model		
	AA	AG	GG	p	AA	AG + GG	p	AA + AG	GG	p
Histology										
Adenocarcinoma	127 (71.3)	212 (68.4)	123 (72.7)		127 (71.3)	335 (69.9)		339 (69.5)	123 (72.7)	
Squamous cell carcinoma	42 (23.6)	77 (24.8)	40 (23.7)	0.629	42 (23.6)	117 (24.4)	0.926	119 (24.4)	40 (23.7)	0.413
Other	9 (5.1)	21 (6.8)	6 (3.6)		9 (5.1)	27 (5.7)		30 (6.1)	6 (3.6)	
T										
T0/T1/T2	47 (26.7)	99 (31.9)	36 (21.6)	**0.050**	47 (26.7)	135 (28.3)	0.686	146 (30.0)	36 (21.6)	**0.035**
T3/T4	129 (73.3)	211 (68.1)	131 (78.4)		129 (73.3)	342 (71.7)		340 (70.0)	131 (78.4)	
N										
N0	43 (24.2)	66 (21.4)	30 (18.1)	0.387	43 (24.2)	96 (20.3)	0.278	109 (22.4)	30 (18.1)	0.237
N1/N2/N3	135 (75.8)	242 (78.6)	136 (81.9)		135 (75.8)	378 (79.7)		377 (77.6)	136 (81.9)	
M										
M0	77 (43.0)	119 (38.8)	54 (32.3)	0.121	77 (43.0)	173 (36.5)	0.126	196 (40.3)	54 (32.3)	0.067
M1	102 (57.0)	188 (61.2)	113 (67.7)		102 (57.0)	301 (63.5)		290 (59.7)	113 (67.7)	
TNM										
0/I/II	29 (16.2)	47 (15.3)	18 (10.8)	0.293	29 (16.2)	65 (13.7)	0.419	76 (15.6)	18 (10.8)	0.123
III/IV	150 (83.8)	260 (84.7)	149 (89.2)		150 (83.8)	409 (86.3)		410 (84.4)	149 (89.2)	
Metastasis at diagnosis										
No metastasis	39 (27.3)	52 (21.7)	27 (19.0)		39 (27.3)	79 (20.7)		91 (23.8)	27 (19.0)	
Metastasis in one organ	39 (27.3)	77 (32.1)	46 (32.4)	0.512	39 (27.3)	123 (32.2)	0.236	116 (30.2)	46 (32.4)	0.511
Metastasis in multiple organs	65 (45.4)	111 (46.2)	69 (48.6)		65 (45.4)	180 (47.1)		176 (46.0)	69 (48.6)	

Bold value indicates $p \leq 0.005$

were associated with increased lung cancer susceptibility and tumor aggressiveness (AG–OR 2.14, 95% CI 1.17–3.92; GG–OR 2.39, 95% CI 1.15–4.96; and AG + GG–OR 6.86, 95% CI 2.47–19.08) [19]. Finally, Masroor et al. showed an association between AG and GG genotypes and reduced overall survival and higher risk to develop lung cancer in Indian population (AG–OR 2.61, 95% CI 1.31–5.18; GG–OR 3.25, 95% CI 1.31–8.06; and AG + GG–OR 2.74, 95% CI 1.41–5.32) [20] (Table 6).

The discrepant results observed in the literature may be explained by the fact that different populations were evaluated in these studies. Indeed, according to 1000Genomes Project, the frequency of *EGF*+61 A>G alleles are different among human populations. Asian and African populations present a G frequency of around 70%, while 40% of Europeans carry the G allele [21]. Obviously, these differences may influence the association studies. In Brazil, the allele frequencies were unknown. Our results showed 51.1% of A and 48.9% of G, and since the Brazilian population is highly admixed, with contribution from European, African, Asian and Native American ancestries, it could be speculated that the frequencies of the two alleles were similar [26].

Moreover, some studies included a low number of individuals, which could impair their statistical power. In fact, the publications that suggested the association between *EGF*+61 A>G polymorphism and lung cancer susceptibility included around 100 in each group [17, 19, 20], while Kang et al., who evaluated more participants (432 cases and 432 controls) [18], as well as our study, which included 669 cases and 1104 controls, did not find the same association. On the other hand, the meta-analysis associating the previous studies from Lim et al. [17] and Kang et al. [18], included 554 cases and 564 controls, and indicated the association between the SNP and risk [27]. Other potential confounding effect among studies could be the controls used [28]. One limitation of our study was the absence of match by demographic parameters, such as gender, age, skin color, tobacco and alcohol consumption. However, following an adjusted multivariate analysis we showed that *EGF*+61 A>G polymorphism continued not affecting lung cancer risk.

Finally, is important to consider that the previous studies differ from the present one when histology is taken into account. Lim et al. did not even describe the clinical and pathological information from their patients [17], whereas Masroor et al. included only adenocarcinoma patients [20]. Kang et al. [18] and de Mello et al. [19] included 48.6% of adenocarcinomas and 32.6% of squamous cell carcinomas, and 59.8% of adenocarcinomas and 25% of squamous cell carcinomas, respectively. Our study included 70.3% of adenocarcinomas and 24.2% of squamous cell carcinomas.

Table 6 Comparison between the present study and other reports from literature which evaluated the *EGF*+61 A>G polymorphism in lung cancer patients versus controls

First author	Year	Country	Cases	Controls	Cases			Controls			p value	OR (95% CI)
					AA (%)	AG (%)	GG (%)	AA (%)	AG (%)	GG (%)		
Lim et al. [17]*	2005	Korea	122	132	10 (8.2)	48 (39.3)	64 (52.4)	35 (26.5)	55 (41.6)	42 (31.8)	**<0.0001** (AA vs. AG vs. GG)	A-ref G – 2.32 (1.60–3.36)
Kang et al. [18]	2007	Korea	432	432	44 (10.1)	191 (44.2)	197 (45.6)	49 (11.3)	185 (42.8)	198 (45.8)	–	GG-ref AG – 1.02 (0.77–1.37) AA – 0.81 (0.51–1.29) AA + AG – 0.98 (0.74–1.29)
Zhang et al. [27]* Meta-analysis	2010	Korea	554	564	54 (9.7)	239 (43.1)	261 (47.2)	84 (14.9)	240 (42.55)	240 (42.55)	–	AA-ref AG + GG – 2.05 (0.59–7.18)
De Mello et al. [19]*	2012	Portugal	112	126	26 (23.2)	58 (51.8)	28 (25.0)	51 (40.5)	52 (41.3)	23 (18.3)	–	AA-ref AG – 2.14 (1.17–3.92) GG – 2.39 (1.15–4.96) AG + GG – 6.86 (2.47–19.08)
Masroor et al. [20]*	2015	India	100	100	17 (17)	63 (63)	20 (20)	36 (36)	51 (51)	13 (13)	**0.008** **(AA vs. AG vs. GG)**	AA-ref AG – 2.61 (1.31–5.18) GG – 3.25 (1.31–8.06) AG + GG – 2.74 (1.41–5.32)
Present study	2017	Brazil	669	1104	183 (27.4)	317 (47.4)	169 (25.3)	279 (25.3)	570 (51.6)	255 (23.1)	0.22 (AA vs. AG vs. GG)	AA-ref AG – 0.80 (0.61–1.05) GG – 1.03 (0.75–1.42) AG + GG – 0.87 (0.67–1.12) GG-ref AA + AG – 1.19 (0.91–1.55)

*Studies that found a significant association between genotype and risk to lung cancer (bold)

Since recent results from genome international consortiums showed that genetic alterations in EGFR pathway seems more relevant in adenocarcinomas, than in squamous cell carcinomas [8–10], the comparison between heterogeneous studies according to histology subtype seems impaired. In this context, the presence of specific SNPs seems to be associated with some histological types of lung cancer. In fact, the polymorphism −191C>A located in the *EGFR* gene present different frequencies in adenocarcinomas when compared to other histology and was considered a risk factor for patients younger than 64 years [29]. Also, the *EGFR* rs2072454 SNP revealed as a risk factor only for lung adenocarcinomas in Jordanian population [30].

In the present study, 232 patients were genotyped using DNA isolated from blood samples and qPCR technique, while 437 patients provided tumor tissue and genotype was accessed by the RFLP-PCR method. The statistic analysis comparing genotype frequencies between the two subgroups of patients revealed no differences in these frequencies, suggesting that, even being a germline feature, it is possible to evaluate it in tumor cells.

Conclusions

In conclusion, this is the largest study that evaluated the association of the *EGF*+61 A>G polymorphism with lung cancer risk. Our findings indicate that *EGF*+61 A>G polymorphism is not associated with lung cancer risk in the Brazilian population.

Acknowledgements The authors would like to thank Barretos Cancer Hospital Research Support Department (NAP) for sample collection and Barretos Cancer Hospital Biobank for sample processing.

Funding This study was partially supported by FINEP - CT-INFRA (02/2010) and Barretos Cancer Hospital Research Fund (PAIP).

Data Availability All data generated or analyzed during this study are included in this published article (and its supplementary information files).

Compliance with ethical standards

Conflict of interest None of the authors have any financial or non-financial conflict of interests.

Ethical approval All procedures performed in this study were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki declaration or its later amendments or comparable ethical standards. The study was approved by Barretos Cancer Hospital Ethics Committee (#596/2012).

Informed consent Signed informed consents were obtained from all control individuals and cases included in the study that provided blood samples. From the FFPE cases, due to the retrospective nature of the study, the Barretos Cancer Hospital Ethics Committee exempted informed consent.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. https://doi.org/10.3322/caac.21492
2. INCA (2018) Estimativas 2018. In. INCA. http://www.inca.gov.br/estimativa/2018/. Accessed 04 Oct 2018
3. DATASUS (2018) Sistema de informações sobre mortalidade. In: Ministério da Saúde. http://www.datasus.gov.br. Accessed 14 Jan 2019
4. Akhtar N, Bansal JG (2017) Risk factors of Lung Cancer in non-smoker. Curr Probl Cancer. https://doi.org/10.1016/j.currproblcancer.2017.07.002
5. Wee P, Wang Z (2017) Epidermal Growth Factor Receptor Cell Proliferation Signaling Pathways. Cancers (Basel). https://doi.org/10.3390/cancers9050052
6. Dibble CC, Cantley LC (2015) Regulation of mTORC1 by PI3K signaling. Trends Cell Biol 25(9):545–555. https://doi.org/10.1016/j.tcb.2015.06.002
7. Normanno N, Bianco C, De Luca A, Salomon DS (2001) The role of EGF-related peptides in tumor growth. Front Biosci J Virtual Libr 6:D685–D707
8. TCGA The Cancer Genome Atlas. https://tcga-data.nci.nih.gov/docs/publications/tcga/?. Accessed 28 Sept 2018
9. ICGC International Cancer Genome Consortium. https://dcc.icgc.org/
10. Jordan EJ, Kim HR, Arcila ME, Barron D, Chakravarty D, Gao J, Chang MT, Ni A, Kundra R, Jonsson P, Jayakumaran G, Gao SP, Johnsen HC, Hanrahan AJ, Zehir A, Rekhtman N, Ginsberg MS, Li BT, Yu HA, Paik PK, Drilon A, Hellmann MD, Reales DN, Benayed R, Rusch VW, Kris MG, Chaft JE, Baselga J, Taylor BS, Schultz N, Rudin CM, Hyman DM, Berger MF, Solit DB, Ladanyi M, Riely GJ (2017) Prospective comprehensive molecular characterization of lung adenocarcinomas for efficient patient matching to approved and emerging therapies. Cancer Discov 7(6):596–609. https://doi.org/10.1158/2159-8290.CD-16-1337
11. Jurisic V, Obradovic J, Pavlovic S, Djordjevic N (2018) Epidermal Growth Factor Receptor Gene in Non-Small-Cell Lung Cancer: The Importance of Promoter Polymorphism Investigation. Anal Cell Pathol 2018:6192187. https://doi.org/10.1155/2018/6192187
12. Shahbazi M, Pravica V, Nasreen N, Fakhoury H, Fryer AA, Strange RC, Hutchinson PE, Osborne JE, Lear JT, Smith AG, Hutchinson IV (2002) Association between functional polymorphism in EGF gene and malignant melanoma. Lancet 359(9304):397–401. https://doi.org/10.1016/S0140-6736(02)07600-6
13. Costa BM, Ferreira P, Costa S, Canedo P, Oliveira P, Silva A, Pardal F, Suriano G, Machado JC, Lopes JM, Reis RM (2007) Association between functional EGF + 61 polymorphism and glioma risk. Clin Cancer Res Off J Am Assoc Cancer Res 13(9):2621–2626. https://doi.org/10.1158/1078-0432.CCR-06-2606
14. Wu SJ, Jiang SY, Wu J, Xiong GL (2015) Association between EGF+61 A>G polymorphism and gastric cancer risk: A meta-analysis. J Huazhong Univ Sci Technol Med Sci 35(3):327–332. https://doi.org/10.1007/s11596-015-1432-3
15. Sun S, Jin GJ, Zhao Y, Kang H (2015) Association between the epidermal growth factor 61*A/G polymorphism and hepatocellular carcinoma risk: a meta-analysis. Asian Pac J Cancer Prev APJCP 16(7):3009–3014

16. Chen X, Yang G, Zhang D, Zhang W, Zou H, Zhao H, Zhang X, Zhao S (2014) Association between the epidermal growth factor + 61 G/A polymorphism and glioma risk: a meta-analysis. PLoS ONE 9(4):e95139. https://doi.org/10.1371/journal.pone.0095139
17. Lim YJ, Kim JW, Song JY, Hong MS, Jin SY, Yoon SH, Park HJ, Choe BK, Lee JJ, Yim SV, Hong SI, Baik HH, Ha E, Park YH (2005) Epidermal growth factor gene polymorphism is different between schizophrenia and lung cancer patients in Korean population. Neurosci Lett 374(3):157–160. https://doi.org/10.1016/j.neulet.2004.10.055
18. Kang HG, Choi JE, Lee WK, Kam S, Cha SI, Kim CH, Jung TH, Park JY (2007) +61 A>G polymorphism in the EGF gene does not increase the risk of lung cancer. Respirology 12:902–905. https://doi.org/10.1111/j.1440-1843.2007.01152.x
19. de Mello RA, Ferreira M, Costa S, Costa BM, Pires FS, Neves I, Almeida MI, Cunha J, Oliveira P, Hespanhol V, Reis RM (2012) Association between EGF + 61 genetic polymorphisms and non-small cell lung cancer increased risk in a Portuguese population: a case-control study. Tumour Biol 33(5):1341–1348. https://doi.org/10.1007/s13277-012-0382-7
20. Masroor M, Amit J, Javid J, Mir R, Prasant Y, Imtiyaz A, Mariyam Z, Mohan A, Ray PC, Saxena A (2015) Clinical implication of EGF A61G polymorphism in the risk of non small cell lung adenocarcinoma patients: a case control study. Asian Pac J Cancer Prev 16(17):7529–7534
21. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR (2015) A global reference for human genetic variation. Nature 526(7571):68–74. https://doi.org/10.1038/nature15393
22. Cui L, Pan XM, Ma CF, Shang-Guan J, Yu HB, Chen GX, Wang J (2010) Association between epidermal growth factor polymorphism and esophageal squamous cell carcinoma susceptibility. Dig Dis Sci 55(1):40–45. https://doi.org/10.1007/s10620-008-0700-5
23. Teixeira AL, Ribeiro R, Cardoso D, Pinto D, Lobo F, Fraga A, Pina F, Calais-da-Silva F, Medeiros R (2008) Genetic polymorphism in EGF is associated with prostate cancer aggressiveness and progression-free interval in androgen blockade-treated patients. Clin Cancer Res 14(11):3367–3371. https://doi.org/10.1158/1078-0432.CCR-07-5119
24. Wu GY, Hasenberg T, Magdeburg R, Bonninghoff R, Sturm JW, Keese M (2009) Association between EGF, TGF-beta1, VEGF gene polymorphism and colorectal cancer. World J Surg 33(1):124–129. https://doi.org/10.1007/s00268-008-9784-5
25. Li TF, Ren KW, Liu PF (2012) Meta-analysis of epidermal growth factor polymorphisms and cancer risk: involving 9779 cases and 15,932 controls. DNA Cell Biol 31(4):568–574. https://doi.org/10.1089/dna.2011.1394
26. Callegari-Jacques SM, Grattapaglia D, Salzano FM, Salamoni SP, Crossetti SG, Ferreira ME, Hutz MH (2003) Historical genetics: spatiotemporal analysis of the formation of the Brazilian population. Am J Hum Biol 15(6):824–834. https://doi.org/10.1002/ajhb.10217
27. Zhang YM, Cao C, Liang K (2010) Genetic polymorphism of epidermal growth factor 61A>G and cancer risk: a meta-analysis. Cancer Epidemiol 34(2):150–156. https://doi.org/10.1016/j.canep.2010.02.004
28. Rose S, Laan MJ (2009) Why match? Investigating matched case-control study designs with causal effect estimation. Int J Biostat 5(1):Article 1. https://doi.org/10.2202/1557-4679.1127
29. Elez-Burnjakovic N, Ugrin M, Obradovic J, Miletic N, Racic M, Kulic M, Pavlovic S, Jurisic V (2018) Distribution of EGFR SNPs - 191C/A and 181946G/A in patients with lung cancer depending on smoking status in the Republic of Srpska, Bosnia and Herzegovina. J BUON Off J Balk Union Oncol 23(2):384–390
30. Bashir NA, Ragab ES, Khabour OF, Khassawneh BY, Alfaqih MA, Momani JA (2018) The Association between Epidermal Growth Factor Receptor (EGFR) Gene Polymorphisms and Lung Cancer Risk. Biomolecules. https://doi.org/10.3390/biom8030053