

Original Paper

DNA Repair Gene Polymorphisms in Relation to Non-Small Cell Lung Cancer Survival

Yuliang Su^a Huan Zhang^a Fangxiu Xu^a Jinyu Kong^a Herbert Yu^b Biyun Qian^{a,c}

^aDepartment of Cancer Epidemiology and Biostatistics, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center of Cancer; Key Laboratory of Cancer Prevention and Therapy, Ministry of Education, Tianjin, China; ^bEpidemiology Program, University of Hawaii Cancer Center, Honolulu, HI 96813, USA; ^cSchool of Public Health, Shanghai Jiao Tong University, Shanghai, China

Key WordsLung cancer • *hOGG1* • *XPA* • Polymorphism • Survival**Abstract**

Background: Single nucleotide polymorphisms (SNPs) in the DNA repair genes are suspected to be related to the survival of lung cancer patients due to their possible influence on DNA repair capacity (DRC). However, the study results are inconsistent. **Methods:** A follow-up study of 610 non-small cell lung cancer (NSCLC) patients was conducted to investigate genetic polymorphisms associated with the DNA repair genes in relation to NSCLC survival; 6 SNPs were genotyped, including *XRCC1* (rs25487 G>A), *hOGG1* (rs1052133 C>G), *MUTYH* (rs3219489 G>C), *XPA* (rs1800975 G>A), *ERCC2* (rs1799793 G>A) and *XRCC3* (rs861539 C>T). Kaplan-Meier survival curve and Cox proportional hazards regression analyses were performed. SNP-SNP interaction was also examined using the survival tree analysis. **Results:** Advanced disease stage and older age at diagnosis were associated with poor prognosis of NSCLC. Patients with the variant 'G' allele of *hOGG1* rs1052133 had poor overall survival compared with those with the homozygous wild 'CC' genotype, especially in female patients, adenocarcinoma histology, early stage, light smokers and without family history of cancer. For never smoking female lung cancer patients, individuals carrying homozygous variant 'AA' genotype of *XPA* had shorter survival time compared to those with wild 'G' alleles. Furthermore, females carrying homozygous variant *XPA* and *hOGG1* genotypes simultaneously had 2.78-fold increased risk for death. Among all 6 polymorphisms, the homozygous variant 'AA' of *XPA* carriers had poor prognosis compared to the carriers of wild 'G' alleles of *XPA* together with other base excision repair (BER) polymorphisms. **Conclusions:** Besides disease stage and age, the study found DNA repair gene polymorphisms were associated with lung cancer survival.

Copyright © 2015 S. Karger AG, Basel

Introduction

Lung cancer is one of the most commonly diagnosed malignancies, and remains to be the leading cause of cancer-related mortality globally. Each year there are 1,095,000 new male cases and 951,000 deaths in men, and 513,600 new female cases and 427,400 deaths in women [1]. Survival from lung cancer has slowly been improved during the past decade, but the prognosis remains poor (5-year survival only 10%) [2]. Tobacco smoking is the most significant cause of lung cancer, and DNA damage induced by tobacco carcinogens is considered an important pathogenic pathway involved in lung carcinogenesis. Recent research indicates that single nucleotide polymorphisms (SNPs) in the DNA repair genes may influence DNA repair capacity (DRC), and therefore affect the risk of lung cancer and survival of patients [3, 4]. Although previous studies focused largely on Asians and less on Caucasians, evidence supporting the genetic involvement of DNA repair genes in lung cancer risk and survival remains inconclusive; more studies with large sample size are still needed. We previously completed a case-control study analyzing the DNA repair gene polymorphisms in relation to lung cancer risk in a Chinese population, and found evidence of associations between some of the SNPs and lung cancer [5]. Based on the hypothesis that DNA repair capacity may also influence lung cancer patient survival, we selected 6 common SNPs which are located either in the coding regions or promoter of the genes, including *XRCC1* Arg399Gln (rs25487 G>A) in 19p13.2, *HOGG1* Ser326Cys (rs1052133 C>G) in 3q26.2, *MUTYH* Gln324His (rs3219489 G>C) in 1p34.1, *XPA* (rs1800975 G>A) in 9q22.3, *ERCC2* Asp312Asn (rs1799793 G>A) in 19q13.3 and *XRCC3* Thr241Met (rs861539 C>T) in 14q32.3. These genes belong to three DNA repair pathways, and the chosen SNPs have been suggested to have impacts on DRC, which may have influence on lung cancer patient survival.

Specifically, *XRCC1*, which is involved in the repair of single strand breaks following the base excision repair (BER) pathway, is a scaffold like protein which helps to remove the DNA adducts. *XRCC1* (rs25487 G>A), the most characterized and extensively studied SNP, has been evaluated in a number of studies with regard to NSCLC survival. Previous studies reported that this SNP has an association with DNA repair activity [6, 7]. *HOGG1* regulates the excision and removal of 8-OH-dG adducts through the BER pathway. The *HOGG1*(rs1052133 C>G) SNP may influence the DNA repair capacity as the mutant genotype has a reduced protein activity [8]. *MUTYH* acts as a counterpart of *HOGG1* in repairing the oxidative DNA damage. Although previous study has reported that a SNP in the gene, *MUTYH* (rs3219489 G>C), has no association with survival in Japanese lung cancer patients [9], we considered this evaluation should be repeated. *XPA* encodes a zinc-finger DNA-binding protein which involves both global genome and transcription-coupled repair pathways and plays central roles in the nucleotide excision repair (NER) pathway. A polymorphism rs1800975 located in the fourth nucleotide upstream from the initiation codon (ATG) was found to affect DNA repair capacity in a host cell reactivation assay [10]. *ERCC2* encodes a helicase which is subunit of the NER machinery, and has been well studied in lung cancer. The SNP (rs1799793 G>A) in *ERCC2* was shown to affect DRC [11]. The protein encoded by *XRCC3* belongs to the double-strand break repair (DSBR) pathway which repairs the DNA damages caused by ionizing irradiation and oxidative free radicals. A polymorphism in codon 241 (rs861539 C>T) of *XRCC3* has been linked to an increased level of DNA adducts in healthy people [7]. Based on the evidence, we selected these SNPs to examine their associations with lung cancer survival in a Chinese population.

Materials and Methods

Study subjects

A patient cohort of lung cancer was established between January 2006 and July 2010, and the project was approved by the ethical review committee at the Tianjin Medical University Cancer Hospital. During the study, blood samples were collected consecutively from the patients who underwent surgery for lung cancer

at the Tianjin Medical University Cancer Hospital. Clinical and pathological information was obtained from medical records and pathology reports of the patients, including age of diagnosis, histological type, tumor size, disease stage, lymph node metastasis, chemotherapy types and other demographic information. All patients had newly diagnosed, and histologically confirmed primary non-small cell lung cancer. Furthermore, patients with previous a medical history of cancer, radiotherapy or chemotherapy were excluded. Face-to-face or telephone interviews to patients or their family members were conducted to ascertain information on tobacco use and survival status. Surrogate persons we obtained patients' information from must live with the patients or take care of the patients after cancer was diagnosed. Information on tobacco use includes smoking status, age at first use, years of smoking, number of cigarettes smoked per day, and the status and age of quitting smoking. Individuals who smoked at least one cigarette per day for half a year were considered smokers. Individuals who smoked less than 100 cigarettes in their whole life were considered never smoker.

Genotyping

Peripheral blood leukocytes obtained from the buffy coat of 1mL of whole blood sample was used for DNA extraction. Genomic DNA was isolated using the DNA Blood Mini kit (Qiagen, Valencia, CA) according to the manufacturer's instruction. The SNPs of *XRCC1* (rs25487 G>A), *hOGG* (rs1052133 C>G), and *MUTYH* (rs3219489 G>C) from the BER pathway; *XPA* (rs1800975 G>A) and *ERCC2* (rs1799793 G>A) from the NER pathway and *XRCC3* (rs861539 C>T) from the DSB pathway were genotyped with a polymerase chain reaction (PCR)-based on fluorescence 5' nuclease assay (TaqMan) described previously [5].

In the genotype analysis, about 5% of the samples were randomly selected for repeat, and the results were in agreement with the original results. The genotyping results were also validated by direct sequencing of the PCR products.

Statistical analysis

The associations between overall survival and individual SNPs were evaluated using the Kaplan-Meier method and the Log-rank test. Hazard ratios (HR) and their 95% confidence interval (CI) were calculated from the Cox proportional hazards regression model. Both univariate and multivariate analyses were performed. The multivariate model included age at diagnosis, gender, histology type, disease stage, treatment and type of surgery. Since patients with different stage of disease may receive different treatment after surgery, we also included treatment and the type of surgery in the multivariate survival analysis for adjustment. All statistical analyses were two-sides and *P* values < 0.05 were considered statistically significant. Statistical software SAS 9.1 (SAS, Cary, NC) was used for data analysis. In order to evaluate the high order gene-gene interactions and to obtain subgroups with distinct death risk, survival tree analysis was performed. In the survival tree analysis, we examined all the covariates and chose the split nodes that had the best separation of homogeneous groups. Tree building and tree pruning were carried out with the R (version 2.13.0) software for statistical computing, using the part packages.

Results

Patient characteristics and clinical features

The median follow-up time of the patients was 28 months ranging from 1 to 56 months. Of the 610 patients, there were 222 deaths and 25 lost to follow-up. Table 1 summarized the clinical features of the patients. The mean age of patients at diagnosis was 60.30±9.14 years, ranging from 23 to 83 years. The 2-year overall survival rate was 71.67%. As anticipated, younger patients had longer survival time compared with the older ones. The results of Cox regression analyses indicated that the risk of death was nearly 3-fold higher in the patients with stage III-IV diseases compared with those with stage I-II (HR=2.91, 95%CI:2.20-3.86) after adjusting for age, gender, histological type, treatment and type of surgery. Furthermore, we divided the treatment into surgery, surgery + chemotherapy, surgery+ radiotherapy, surgery +chemotherapy +radiotherapy and found that patients who took surgery+ chemotherapy and surgery+ radiotherapy had lower risk of death than those who took only surgery as treatment (HR=0.62 and 0.57, respectively). No significant differences in survival

Table 1. Patient characteristics and clinical features. a: adjusted by gender, treatment and type of surgery. b: adjusted by age, treatment and type of surgery. c: SCC (Squamous cell carcinomas); ACC (Adenocarcinomas). d: TP(Taxol+Platinum); NP (Vinorelbine+Platinum); GP (Gemcitabine+Platinum); DP (docetaxel+platinum). e: adjusted by age, gender, histological type, stage, treatment and type of surgery. f: adjusted by age, gender, histological type, stage and type of surgery. treatment refer to surgery, surgery+chemotherapy, surgery+radiotherapy, surgery+chemotherapy+radiotherapy. g: data is partly missing

Variable	Number(%)	Failed/ Censored	2 year survival ratio	Log-rank P	HR (95%CI)	HR ^e (95%CI)
Age^a						
<60	279(45.74%)	86/193	77.11%	0.008	1.00	1.00
>=60	331(54.26%)	136/195	67.09%		1.44(1.10-1.88)	1.56(1.19-2.05)
Gender^b						
Male	411(67.38%)	154/257	68.92%	0.339	1.00	1.00
Female	199(32.62%)	68/131	77.44%		0.87(0.65-1.16)	0.87(0.65-1.17)
Histology^c						
SCC	272(44.59%)	272/100	70.56%	0.132	1.00	1.00
ACC	250(40.98%)	250/84	75.85%		0.93(0.70-1.25)	0.91(0.66-1.27)
Others	88(14.43%)	88/38	63.26%		1.37(0.94-1.99)	1.43(0.97-2.09)
Stage						
Earlier	381(62.46%)	93/288	81.07%	<.001	1.00	1.00
Later	229(37.54%)	129/100	55.60%		3.06(2.34-4.00)	2.91(2.20-3.86)
Family History						
No	503(82.46%)	189/314	70.94%	0.282	1.00	1.00
Yes	107(17.54%)	33/74	75.11%		0.82(0.57-1.18)	0.76(0.52-1.11)
Smoking Status						
Never smoker	177(29.02%)	64/113	74.14%	0.005	1.00	1.00
Former Smoker	142(23.28%)	36/106	79.97%		0.64(0.43-0.96)	0.67(0.43-1.05)
Current Smoker	291(47.70%)	122/169	66.12%		1.18(0.87-1.59)	1.16(0.82-1.65)
Chemotherapy^{d,g}						
TP	165(56.51%)	74/ 91	63.48%	0.532	1.00	1.00
NP	38(13.01%)	15/ 23	76.04%		0.99(0.59-1.68)	0.80(0.46-1.38)
GP	42(14.38%)	14/ 28	70.77%		0.89(0.51-1.53)	0.77(0.44-1.35)
DP	35(11.99%)	11/ 24	73.98%		0.83(0.45-1.52)	0.60(0.32-1.11)
Surgery Types^e						
Lobectomy	114(61.62)	46/68	69.77	0.014	1.00	1.00
wedge resection	56(30.27)	26/30	67.09		3.18(1.73-5.85)	2.26(1.20-4.24)
pneumectomy	15(8.11)	11/4	26.67		1.32(0.87-1.98)	1.00(0.65-1.53)
Treatment^f						
Surgery	299(49.02%)	96/203	76.33%	0.0006	1.00	1.00
Surgery+chemotherapy	237(38.85%)	83/154	71.28%		1.07(0.79-1.43)	0.62(0.42-0.93)
Surgery+radiotherapy	19(3.11%)	6/13	68.02%		0.99(0.44-2.27)	0.57(0.39-0.85)
Surgery+chemotherapy+radiotherapy	55(9.02%)	37/18	49.72%		2.13(1.46-3.11)	0.73(0.31-1.74)

were observed among patients with different gender, histology, family history of cancer, smoking status or chemotherapy types after adjustment.

DNA repair gene polymorphisms

The genotype frequencies of the investigated SNPs listed in Table 2 were all in Hardy-Weinberg equilibrium. Of the 6 SNPs genotyped in our study, we only found that the variant 'CG+GG' genotype of *hOGG1* (rs1052133 C>G) was associated with survival of lung cancer when comparing to the wild 'CC' genotype (Log-rank *P* value=0.026). However, this significance disappeared after adjusting for age at diagnosis, gender, disease stage, histological type and treatment. We further stratified the association of the *hOGG1* (rs1052133 C>G) SNP and lung cancer survival by several covariates. Table 3 shows that the following patients who carried *hOGG1* variant 'G' allele had poor prognosis for lung cancer, early stage (HR=2.22, 95%CI: 1.11-4.44), age (<60 years) (HR=2.02, 95%CI: 1.04-3.92), adenocarcinoma (HR=2.55, 95%CI: 1.11-5.86) and without family history of cancer (HR=1.60, 95%CI: 1.04-2.45).

Among the smokers, patients who carried *hOGG1* variant 'G' allele had a shorter survival time compared to those with wild homozygote, but the difference was not significant (Log-rank *p* value= 0.072, data not shown). Furthermore, when further dividing the smokers into light and heavy ones according to the smoking amount (pack-year), we found that light smokers with less than 40 pack-year who carried variant 'G' allele had nearly 2.6 times higher risk of death compared to those carrying the wild 'CC' genotype (HR=2.63, 95%CI: 1.05-6.59). However, we found no significant association among the heavy smokers.

Table 2. Polymorphisms in DNA repair pathway genes and survival of patients. a: adjusted by age, gender, histological type, stage, treatment and type of surgery. b: adjusted by age, histological type, stage, treatment and type of surgery

Pathway	Gene (rs)	Genotypes	Number (%)		log-rank p value		HR (95% CI)		Adjusted HR (95% CI)	
			All n=610	Female n=199	All	Female	All	Female	All ^a	Female ^b
BER	XRCC1 rs25487	GG	341(55.90%)	111(55.78%)	0.182	0.246	1.00	1.00	1.00	1.00
		GA	232(38.03%)	73(36.68%)						
		AA	37(6.07%)	15(7.54%)						
	hOGG1 rs1052133	GG	341(55.90%)	184(92.46%)	0.182	0.660	1.00	1.00	1.00	1.00
		GA+AA	269(44.10%)	15(7.54%)						
		CC	98(16.07%)	31(15.58%)						
	MUTYH rs3219489	CG	305(50.00%)	108(54.27%)	0.084	0.040	1.00	1.00	1.00	1.00
		GG	207(33.93%)	60(30.15%)						
		CC	98(16.07%)	139(69.85%)						
	XPA rs1800975	CG+GG	512(83.93%)	60(30.15%)	0.549	0.691	1.00	1.00	1.00	1.00
		GG	229(37.54%)	68(34.17%)						
		GA	284(46.56%)	100(50.25%)						
	ERCC2 rs1799793	CC	97(15.90%)	31(15.58%)	0.204	0.828	1.00	1.00	1.00	1.00
		GG	229(37.54%)	168(84.42%)						
		GA+CC	381(62.46%)	31(15.58%)						
	XRCC3 rs861539	GG	156(25.57%)	56(28.14%)	0.656	0.032	1.00	1.00	1.00	1.00
		GA	290(47.54%)	100(50.25%)						
		AA	156(25.57%)	43(21.61%)						
	ERCC2 rs1799793	GG	156(25.57%)	156(78.39%)	0.656	0.038	1.00	1.00	1.00	1.00
		GA+AA	454(74.43%)	43(21.61%)						
		GG	515(84.43%)	167(83.92%)						
	XRCC3 rs861539	GA	92(15.08%)	32(16.08%)	0.204	0.615	1.00	1.00	1.00	1.00
		AA	3(0.49%)	0(0%)						
		GG	515(84.43%)	199(100.00%)						
	XRCC3 rs861539	GA+AA	95(15.57%)	0(0%)	0.151	0.828	1.00	1.00	1.00	1.00
		GG	550(90.16%)	182(91.46%)						
		GA	60(9.84%)	17(8.54%)						
	XRCC3 rs861539	AA	0(0%)	0(0%)	0.151	0.151	1.00	1.00	1.00	1.00
		GG	550(90.16%)	199(100.00%)						
		GA+AA	60(9.84%)	0(0%)						

Table 3. hOGG1 gene Polymorphism and survival of lung carcinomas in different categories. a: adjusted by age, histological type, stage, treatment and type of surgery. b: adjusted by age, gender, histological type, stage, treatment and type of surgery. c: adjusted by age, gender, histological type, treatment and type of surgery. d: adjusted by gender, histological type, stage, treatment and type of surgery. e: adjusted by age, gender, stage, treatment and type of surgery. f: SCC (Squamous cell carcinomas); ACC (Adenocarcinomas)

		HRs(95%CI)					
		CC (1)	CG (2)	GG (3)	GG vs. CC (3 vs. 1)	GG+CG/CC (2+3 vs. 1)	GG/CC+CG (3 vs. 1+2)
Gender ^a	Male	67	197	147	1.22(0.73-2.04)	1.35(0.85-2.14)	0.93(0.66-1.30)
	Female	31	108	60	2.68(1.09-6.60)	2.10(0.90-4.88)	1.64(0.99-2.69)
Smoking status ^b	Smoker	72	205	156	1.55(0.93-2.58)	1.59(0.99-2.54)	1.07(0.77-1.49)
	<40 Packyear	28	101	74	2.24(0.85-5.88)	2.63(1.05-6.59)	0.91(0.56-1.46)
	≥40 Packyear	44	104	82	1.26(0.68-2.36)	1.10(0.62-1.94)	1.29(0.82-2.03)
	Never smoker	26	100	51	1.63(0.66-4.01)	1.52(0.68-3.39)	1.18(0.67-2.06)
Stage ^c	I + II	64	193	124	1.88(0.89-3.99)	2.22(1.11-4.44)	0.93(0.60-1.44)
	III+IV	34	112	83	1.32(0.76-2.28)	1.22(0.74-2.02)	1.18(0.82-1.71)
Age ^d	<60	46	147	86	2.24(1.08-4.68)	2.02(1.04-3.92)	1.36(0.85-2.17)
	≥60	52	158	121	1.37(0.78-2.39)	1.42(0.85-2.39)	1.02(0.72-1.45)
Histology ^{e,f}	SCC	52	130	90	1.13(0.62-2.05)	1.32(0.78-2.24)	0.86(0.55-1.32)
	ACC	34	133	83	2.34(0.99-5.53)	2.55(1.11-5.86)	1.43(0.91-2.23)
Family History ^b	No	85	248	170	1.55(0.97-2.49)	1.60(1.04-2.45)	1.07(0.79-1.45)
	Yes	13	57	37	1.17(0.32-4.30)	1.15(0.34-3.89)	1.05(0.50-2.24)

Since female lung cancer is believed to be a distinct disease from male lung cancer based on epidemiological and biological evidence, we further evaluated the association of DNA repair gene polymorphisms and lung cancer survival only among female patients (Table 2).

Our results indicated that female patients carrying variant 'AA' genotype of the XPA gene had a 2-fold increase in risk for death compared to those with the wild 'G' allele (HR=1.72, 95%CI:1.02-2.90). The association for hOGG1 was borderline significant. Kaplan-Meier

Fig. 1. Kaplan-Meier curves for female nonsmokers by *XPA* rs1800975 genotypes. The patients who carried the *XPA* rs1800975 homozygous variant 'AA' genotype have poor survival than those who carried 'GG' or 'GA' genotype (Log-rank $P=0.021$).

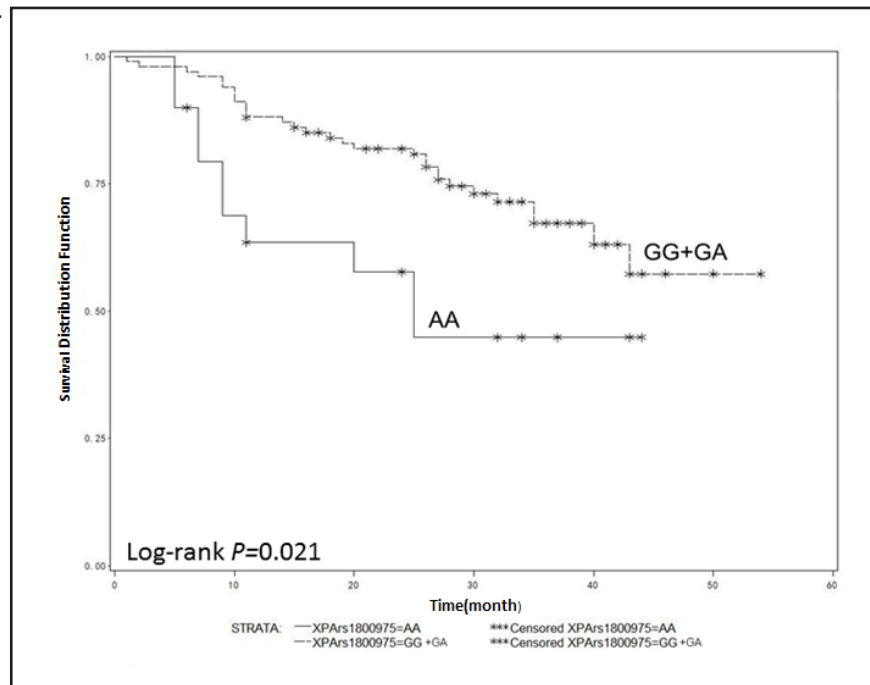


Table 4. Association between *hOGG1* and *XPA* polymorphism on female lung cancer patients survival. a: adjusted by age, histological type, stage, treatment and type of surgery

	<i>XPA</i> -GG+GA		<i>XPA</i> -AA	
	Number(%)	HR ^a (95%CI)	Number(%)	HR ^a (95%CI)
<i>XRCC1</i> (rs25487)				
GG+GA	145(72.86%)	1.00	39(19.60%)	2.16(1.26-3.71)
AA	11(5.53%)	1.24(0.44-3.48)	4(2.01%)	-
<i>hOGG1</i> (rs1052133)				
CC+CG	110(55.28%)	1.00	29(14.57%)	2.18(1.15-4.27)
GG	46(23.12%)	1.82(1.00-3.28)	14(7.04%)	2.78(1.19-6.54)
<i>MUTYH</i> (rs3219489)				
GG+GC	131(65.83%)	1.00	37(18.59%)	2.15(1.22-3.79)
CC	25(12.56%)	0.91(0.40-2.05)	6(3.02%)	0.87(0.21-3.65)

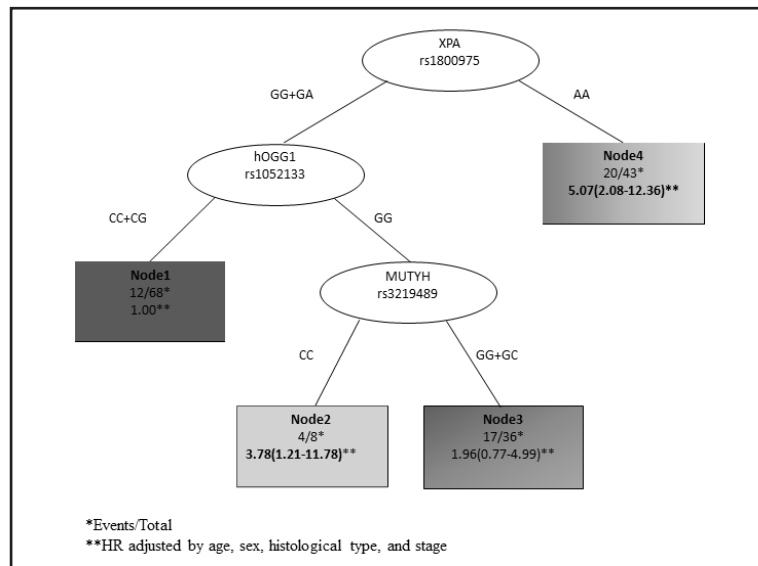
analysis showed that patients with different *XPA* genotypes had quite different survival curves (Log-rank test, $P=0.021$, Fig. 1). In contrast, no significant differences in survival were observed among male patients (data not shown).

Gene-gene interactions

We further analyzed the SNP-SNP interaction in association with lung cancer survival. When we conducted stratified analysis by the *XPA* genotype, we found that the variant *XPA* carriers had higher risk for lung cancer death if these subjects also had the homozygous variant 'GG' genotype of *hOGG1* compared to those who had wild alleles (HR=2.78; 95%CI: 1.19-6.54) (Table 4). We didn't find any significant associations when including either *XRCC1* or *MUTYH* polymorphisms in the analysis.

Survival tree analysis was introduced by Segal in 1988, and then was widely used in identifying subgroups of individuals at higher risk to certain diseases especially cancer [12, 13]. The analysis is based on recursive partitioning, and its graphical output makes it possible to visualize the prognostic subgroups of gene-gene interactions. The analysis is also known as an extension of recursive partitioning for censored survival data. The process is to split the covariate space into regions by a full likelihood estimation and to split the dataset

Fig. 2. Survival tree analysis of all investigated genes in female never smoking lung cancer patients. Figure 2 displays the resultant tree among female never smoking patients after a variety of multivariate survival tree analysis of all the DNA repair pathway genes were computed. Node 1 was chosen as reference, Node 3 has 2.18 fold risk of death when compared with node 1, while Node 4 had 2.37 fold death risk.



into progressively homogeneous groups, until it breaks the stopping rule then stopping the tree growth.

Figure 2 displays the resultant tree among never smoking female patients after a variety of multivariate survival tree analysis of all the DNA repair pathway genes. It was not surprising to find that the top two splits were *XPA* and *hOGG1*, which confirmed the results above. Furthermore, the next best cut point to separate is *MUTYH* polymorphism. Figure 2 also shows the HRs and 95% CI for the four terminal nodes of the survival tree. Compared with node 1 (reference group, consisting of wild alleles of *XPA* and *hOGG1*), patients carrying with *XPA* and *MUTYH* wild types and *hOGG1* variant polymorphism had 2.18-fold increases in risk for death (HR for node 3=2.18; 95%CI:1.18-4.04). Moreover, the subjects with the variant genotype of *hOGG1* had even higher risk for death when carrying only one variant genotype of *XPA* (HRs for node 4=2.37; 95%CI: 1.32-4.27).

Discussion

Consistent with the literature, we found advanced disease stage and older ages at diagnosis were independently associated with NSCLC survival in our study. We also detected a possible association between DNA repair gene polymorphisms and overall survival of lung cancer among Chinese patients. The variant 'G' allele of *hOGG1* rs1052133 was associated with increased risk of death, especially in female, early stage disease, adenocarcinoma, light smokers, or patients without a family history of cancer. Moreover, for female patients, the genetic polymorphism of *XPA* rs1800975 was shown some influences on lung cancer survival besides the BER gene polymorphism.

A significant association between the *hOGG1* (rs1052133 C>G) polymorphism and increased lung cancer susceptibility was demonstrated in a meta-analysis [14]. Genetic polymorphisms have emerged as a risk factor not only for cancer development, but also for cancer progression and prognosis [15]. A study reported by Shen et al. showed that patients with the variant 'CG/GG' genotypes of *hOGG1* had a shorter survival time compared to those with the wild 'CC' genotype, although the difference was not significant (Log-rank $p=0.098$) [16]. *In vitro* experiments demonstrated that the variant 'G' allele of *hOGG1* was associated with higher level of 8-OH-dG protein, but lower enzymatic activity of 8-OH-dG compared to the wildtype 'C' allele [17, 18]. In addition, the *hOGG1* protein encoded by the 'G' allele could promote the accumulation of oxidative adducts, such as 8-oxo-G, which might increase the mutation rates [19]. Based on these findings, it is not surprised that lung cancer patients

who carried the *hOGG1* 'G' allele might have poor survival outcomes. Our study did show that patients carrying the *hOGG1* variant 'G' allele had poor survival, which was especially evident in female patients, early stage of cancer, adenocarcinoma, light smokers, and those without a family history of cancer. Our results were basically consistent with those observed by Shen et al., except for a few small differences which were probably due to different sample sizes, ethnic variations, and less prevalence of the 'G' allele in the Caucasian population.

It is well known that smoking is the predominant risk factor of lung cancer. However, not all lung cancer patients have a history of smoking, especially for female patients. Females are more likely than males to have non-smoking lung cancer, although the etiology of lung cancer for non-smoking females is unclear. Based on different epidemiological and biological features between male and female lung cancer patients, female lung cancer has been suggested to be treated differently as a distinct disease [1, 20-22]. That notion is further confirmed by molecular and genetic analysis which show unique features of methylation index and mutations in *P53*, *EGFR* and *KRAS* [23]. Our results suggested that *XPA* (rs1800975 G>A) polymorphism had a possible influence on lung cancer survival among female patients in addition to the polymorphism of *hOGG1* gene. Moreover, survival tree analysis which showed good discriminative survival profiles from different gene variation further indicated the combined effect of *XPA* and *hOGG1* on survival outcomes.

Our previous study showed that the variant 'AA' genotype of *XPA* was significantly associated with increased risk of NSCLC [5], and this association was also supported by a large meta-analysis [24]. Similar to our results, some studies suggested the variant 'A' allele of *XPA* rs1800975 was a risk factor for survival in lung cancer [25-28]. One study focused on younger patients and ever smokers with less than 36 pack-year [25]. Other studies indicated that the 'A' allele-associated unfavorable overall survival of NSCLC may be related to the poor response of patients to treatment [26-28]. *XPA* polymorphisms had been reported to affect patient response to platinum-based chemotherapy [29].

The SNP at *XPA* (rs1800975 G>A), is particularly interesting because it is located in the region near the transcription initiation codon at the 5'UTR, and it may affect the transcription of the *XPA* gene which plays a critical role in the NER pathway. In addition, this substitution may remove a CpG site, which can interfere with methylation of the *XPA* promoter and proper regulation of *XPA* expression [30]. The *XPA* allele may also be in linkage disequilibrium with some alleles from the adjacent genes which are susceptible for survival outcomes [26, 31]. Finally, a recent study found that the *XPA* (rs1800975 G>A) SNP may alter the occurrence of *TP53* mutations in lung cancer, which can lead to a more aggressive phenotype, resulting in poor survival [32]. However, functional evaluations and population studies are still needed to explain the finding of our study in female non-smoking patients.

It has been reported that *XRCC1* rs25487 is associated with the survival of non-smoking female patients with lung adenocarcinoma [33]. However, we did not find a similar association in our results. Different findings between studies may be due to patient differences with regard to their treatment -46% of patients received chemotherapy in our study, while 88% of patients received chemotherapy or radiotherapy in the other study. Treatment may influence the course of tumor progression. For the polymorphisms of *MUTYH*, *ERCC2*, and *XRCC3*, we found no significant associations with lung cancer survival. These results were similar to those of previous studies [34, 35].

Poor survival in NSCLC is partly attributed to drug resistance. It has been reported that genetic Variants and protein expressions might affect the platinum-based chemotherapy response and prognosis of advanced NSCLC patients [36, 37]. Platinum-based chemotherapy as a main treatment for lung cancer exerts its anti-cancer effect largely through the formation of bulky platinum-DNA adducts, which result in destabilization of the DNA double helix that otherwise blocks replication and inhibits transcription. Since adduct-induced DNA damage is repaired mainly by NER and BER, these DNA repair functions play an important role in determining the sensitivity of tumor cells to platinum treatment [38]. Polymorphisms in the DNA repair pathway genes, such as *ERCC1*, *XRCC1*, *ERCC2* and *XRCC3*, may affect the activities of DNA repair and therefore influence patient's response to platinum-based chemotherapy,

ultimately impacting on disease outcomes and patient survival [27, 38-44]. The complexity of the DNA repair mechanisms along with the heterogeneity of tumor cells, however, may complicate our effort to understand the effect of genetic polymorphisms on DNA repair functions and substantially on patient's response to chemotherapy which is based the DNA damage mechanisms. Despite the complication, we did find evidence to support the possibility that genetic polymorphisms involved in DNA repair capacity may affect treatment influencing patient survival.

Our study, however, was still relatively small and had limited follow-up time. The observations were retrospective without sufficient control for confounding. To confirm our findings, prospective studies with larger sample size and longer follow-up time are needed with adjustment for more confounding variables. The detailed mechanisms of chemotherapy-induced DNA damage and repair remain to be elucidated.

In conclusion, for non-smoking female NSCLC, genetic polymorphisms in the NER pathway may play a role in survival. SNPs both in NER and BER may have a joint effect on NSCLC survival.

Abbreviations

SNPs (Single Nucleotide Polymorphisms); DRC (DNA Repair Capacity); NSCLC (Non-Small Cell Lung Cancer); BER (Base Excision Repair); PCR (Polymerase Chain Reaction); HR (Hazard Ratios); CI (Confidence Interval).

Acknowledgements

This work was supported by National Natural Science Foundation of China (No. 81372229) and Natural Science Foundation of Shanghai (No. 15ZR1424400).

Disclosure Statement

All authors have read and approved the paper for submission. We declare no conflict of interest.

References

- 1 Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D: Global cancer statistics. *CA Cancer J Clin* 2011;61:69-90.
- 2 Sant M, Allemani C, Santaquilani M, Knijn A, Marchesi F, Capocaccia R: EURO-CARE-4. Survival of cancer patients diagnosed in 1995-1999. Results and commentary. *Eur J Cancer* 2009;45:931-991.
- 3 Nakamura H, Saji H, Idiris A, Kawasaki N, Hosaka M, Ogata A, Saijo T, Kato H: Chromosomal instability detected by fluorescence in situ hybridization in surgical specimens of non-small cell lung cancer is associated with poor survival. *Clin Cancer Res* 2003;9:2294-2299.
- 4 Sun Z, Chen J, Aakre J, Marks RS, Garces YY, Jiang R, Idowu O, Cunningham JM, Liu Y, Pankratz VS, Yang P: Genetic variation in glutathione metabolism and DNA repair genes predicts survival of small-cell lung cancer patients. *Ann Oncol* 2010;21:2011-2016.
- 5 Qian B, Zhang H, Zhang L, Zhou X, Yu H, Chen K: Association of genetic polymorphisms in DNA repair pathway genes with non-small cell lung cancer risk. *Lung Cancer* 2011;73:138-146.

- 6 Abdel-Rahman SZ, El-Zein RA: The 399Gln polymorphism in the DNA repair gene XRCC1 modulates the genotoxic response induced in human lymphocytes by the tobacco-specific nitrosamine NNK. *Cancer Lett* 2000;159:63-71.
- 7 Matullo G, Palli D, Peluso M, Guarrera S, Carturan S, Celentano E, Krogh V, Munnia A, Tumino R, Polidoro S, Piazza A, Vineis P: XRCC1, XRCC3, XPD gene polymorphisms, smoking and (32)P-DNA adducts in a sample of healthy subjects. *Carcinogenesis* 2001;22:1437-1445.
- 8 Hill JW, Evans MK: Dimerization and opposite base-dependent catalytic impairment of polymorphic S326C OGG1 glycosylase. *Nucleic Acids Res* 2006;34:1620-1632.
- 9 Osawa K, Nakarai C, Uchino K, Yoshimura M, Tsubota N, Takahashi J, Kido Y: XRCC3 Gene Polymorphism Is Associated with Survival in Japanese Lung Cancer Patients. *Int J Mol Sci* 2012;13:16658-16667.
- 10 Wu X, Zhao H, Wei Q, Amos CI, Zhang K, Guo Z, Qiao Y, Hong WK, Spitz MR: XPA polymorphism associated with reduced lung cancer risk and a modulating effect on nucleotide excision repair capacity. *Carcinogenesis* 2003;24:505-509.
- 11 Affatato AA, Wolfe KJ, Lopez MS, Hallberg C, Ammenheuser MM, Abdel-Rahman SZ: Effect of XPD/ERCC2 polymorphisms on chromosome aberration frequencies in smokers and on sensitivity to the mutagenic tobacco-specific nitrosamine NNK. *Environ Mol Mutagen* 2004;44:65-73.
- 12 Wu X, Lu C, Ye Y, Chang J, Yang H, Lin J, Gu J, Hong WK, Stewart D, Spitz MR: Germline genetic variations in drug action pathways predict clinical outcomes in advanced lung cancer treated with platinum-based chemotherapy. *Pharmacogenet Genomics* 2008;18:955-965.
- 13 Dai J, Gu J, Lu C, Lin J, Stewart D, Chang D, Roth JA, Wu X: Genetic variations in the regulator of G-protein signaling genes are associated with survival in late-stage non-small cell lung cancer. *PLoS One* 2011;6:e21120.
- 14 Zhong D, Li G, Long J, Wu J, Hu Y: The hOGG1Ser326Cys polymorphism and increased lung cancer susceptibility in Caucasians: an updated meta-analysis. *Sci Rep* 2012;2:548.
- 15 Loktionov A: Common gene polymorphisms, cancer progression and prognosis. *Cancer Lett* 2004;208:1-33.
- 16 Shen J, Deininger P, Hunt JD, Zhao H: 8-Hydroxy-2'-deoxyguanosine (8-OH-dG) as a potential survival biomarker in patients with nonsmall-cell lung cancer. *Cancer* 2007;109:574-580.
- 17 Kohno T, Shinmura K, Tosaka M, Tani M, Kim SR, Sugimura H, Nohmi T, Kasai H, Yokota J: Genetic polymorphisms and alternative splicing of the hOGG1 gene, that is involved in the repair of 8-hydroxyguanine in damaged DNA. *Oncogene* 1998;16:3219-3225.
- 18 Bravard A, Vacher M, Moritz E, Vaslin L, Hall J, Epe B, Radicella JP: Oxidation status of human OGG1-S326C polymorphic variant determines cellular DNA repair capacity. *Cancer Res* 2009;69:3642-3649.
- 19 Rihs HP, Marczyński B, Lotz A, Raulf-Heimsoth M, Bruning T: Modulation of oxidative DNA damage by repair enzymes XRCC1 and hOGG1. *J Toxicol Environ Health A* 2012;75:588-596.
- 20 Yang GH, Wang JF, Wan X, Wang LJ, Chen AP: [Quantitative analysis of factors affected mortality trend in Chinese, 2002]. *Zhonghua Liu Xing Bing Xue Za Zhi* 2005;26:934-938.
- 21 Toh CK, Gao F, Lim WT, Leong SS, Fong KW, Yap SP, Hsu AA, Eng P, Koong HN, Thirugnanam A, Tan EH: Never-smokers with lung cancer: epidemiologic evidence of a distinct disease entity. *J Clin Oncol* 2006;24:2245-2251.
- 22 Heigener DF: Non-small cell lung cancer in never-smokers: a new disease entity? *Onkologie* 2011;34:202-207.
- 23 Subramanian J, Govindan R: Molecular genetics of lung cancer in people who have never smoked. *Lancet Oncol* 2008;9:676-682.
- 24 Kiyohara C, Yoshimasu K: Genetic polymorphisms in the nucleotide excision repair pathway and lung cancer risk: a meta-analysis. *Int J Med Sci* 2007;4:59-71.
- 25 Butkiewicz D, Rusin M, Sikora B, Lach A, Chorazy M: An association between DNA repair gene polymorphisms and survival in patients with resected non-small cell lung cancer. *Mol Biol Rep* 2011;38:5231-5241.
- 26 Feng J, Sun X, Sun N, Qin S, Li F, Cheng H, Chen B, Cao Y, Ma J, Cheng L, Lu Z, Ji J, Zhou Y: XPA A23G polymorphism is associated with the elevated response to platinum-based chemotherapy in advanced non-small cell lung cancer. *Acta Biochim Biophys Sin (Shanghai)* 2009;41:429-435.

- 27 Cheng H, Qin Q, Sun X, Li F, Sun N, Cheng L, Lu Z, Chen B: Predictive effect of XPA and XPD polymorphisms on survival of advanced NSCLC patients treated with platinum-based chemotherapy: a three-dimensional (3-D), polyacrylamide gel-based DNA microarray method. *Technol Cancer Res Treat* 2013;12:473-482.
- 28 Butkiewicz D, Drosik A, Suwinski R, Krzesniak M, Rusin M, Kosarewicz A, Rachtan J, Matuszczyk I, Gawkowska-Suwinska M: Influence of DNA repair gene polymorphisms on prognosis in inoperable non-small cell lung cancer patients treated with radiotherapy and platinum-based chemotherapy. *Int J Cancer* 2012;131:E1100-1108.
- 29 Osawa K: Gene Polymorphisms and Chemotherapy in Non-small Cell Lung Cancer. *Zhongguo Fei Ai Za Zhi* 2009;12:837-840.
- 30 Candiloro IL, Dobrovic A: Detection of MGMT promoter methylation in normal individuals is strongly associated with the T allele of the rs16906252 MGMT promoter single nucleotide polymorphism. *Cancer Prev Res (Phila)* 2009;2:862-867.
- 31 Butkiewicz D, Krzesniak M, Vaitiekunaite R, Sikora B, Bowman ED, Harris CC, Rusin M: A functional analysis of G23A polymorphism and the alternative splicing in the expression of the XPA gene. *Cell Mol Biol Lett* 2010;15:611-629.
- 32 Cho S, Kim MJ, Choi YY, Yoo SS, Lee WK, Lee EJ, Jang EJ, Bae EY, Jin G, Jeon HS, Lee SY, Cha SI, Park TI, Kim CH, Park JY: Associations between polymorphisms in DNA repair genes and TP53 mutations in non-small cell lung cancer. *Lung Cancer* 2011;73:25-31.
- 33 Yin Z, Zhou B, He Q, Li M, Guan P, Li X, Cui Z, Xue X, Su M, Ma R, Bai W, Xia S, Jiang Y, Xu S, Lv Y: Association between polymorphisms in DNA repair genes and survival of non-smoking female patients with lung adenocarcinoma. *BMC Cancer* 2009;9:439.
- 34 Kalikaki A, Kanaki M, Vassalou H, Souglakos J, Voutsina A, Georgoulas V, Mavroudis D: DNA repair gene polymorphisms predict favorable clinical outcome in advanced non-small-cell lung cancer. *Clin Lung Cancer* 2009;10:118-123.
- 35 de las Penas R, Sanchez-Ronco M, Alberola V, Taron M, Camps C, Garcia-Carbonero R, Massuti B, Queralt C, Botia M, Garcia-Gomez R, Isla D, Cobo M, Santarpià M, Cecere F, Mendez P, Sanchez JJ, Rosell R: Polymorphisms in DNA repair genes modulate survival in cisplatin/gemcitabine-treated non-small-cell lung cancer patients. *Ann Oncol* 2006;17:668-675.
- 36 Shi C, Qian J, Ma M, Zhang Y, Han B: Notch 3 protein, not its gene polymorphism, is associated with the chemotherapy response and prognosis of advanced NSCLC patients. *Cell Physiol Biochem* 2014;34:743-752.
- 37 Wu F, Zhang J, Liu Y, Zheng Y, Hu N: HIF1alpha genetic variants and protein expressions determine the response to platinum based chemotherapy and clinical outcome in patients with advanced NSCLC. *Cell Physiol Biochem* 2013;32:1566-1576.
- 38 Sullivan I, Salazar J, Majem M, Pallares C, Del Rio E, Paez D, Baiget M, Barnadas A: Pharmacogenetics of the DNA repair pathways in advanced non-small cell lung cancer patients treated with platinum-based chemotherapy. *Cancer Lett* 2014;353:160-166.
- 39 Tiseo M, Bordi P, Bortesi B, Boni L, Boni C, Baldini E, Grossi F, Recchia F, Zanelli F, Fontanini G, Naldi N, Campanini N, Azzoni C, Bordi C, Ardizzoni A: ERCC1/BRCA1 expression and gene polymorphisms as prognostic and predictive factors in advanced NSCLC treated with or without cisplatin. *Br J Cancer* 2013;108:1695-1703.
- 40 Ke HG, Li J, Shen Y, You QS, Yan Y, Dong HX, Liu JH, Shen ZY: Prognostic significance of GSTP1, XRCC1 and XRCC3 polymorphisms in non-small cell lung cancer patients. *Asian Pac J Cancer Prev* 2012;13:4413-4416.
- 41 Liao WY, Shih JY, Chang GC, Cheng YK, Yang JC, Chen YM, Yu CJ: Genetic polymorphism of XRCC1 Arg399Gln is associated with survival in non-small-cell lung cancer patients treated with gemcitabine/platinum. *J Thorac Oncol* 2012;7:973-981.
- 42 Xu TP, Shen H, Liu LX, Shu YQ: Association of ERCC1-C118T and -C8092A polymorphisms with lung cancer risk and survival of advanced-stage non-small cell lung cancer patients receiving platinum-based chemotherapy: a pooled analysis based on 39 reports. *Gene* 2013;526:265-274.
- 43 Wu W, Li H, Wang H, Zhao X, Gao Z, Qiao R, Zhang W, Qian J, Wang J, Chen H, Wei Q, Han B, Lu D: Effect of polymorphisms in XPD on clinical outcomes of platinum-based chemotherapy for Chinese non-small cell lung cancer patients. *PLoS One* 2012;7:e33200.
- 44 Li XD, Han JC, Zhang YJ, Li HB, Wu XY: Common variations of DNA repair genes are associated with response to platinum-based chemotherapy in NSCLCs. *Asian Pac J Cancer Prev* 2013;14:145-148.