Received: 20 March 2018 | Revised: 22 June 2018 | Accepted: 19 July 2018
DOI: 10.1002/mc.22877

RESEARCH ARTICLE

WILEY Molecular Carcinogenesis

Genetic variants in nucleotide excision repair pathway predict survival of esophageal squamous cell cancer patients receiving platinum-based chemotherapy

Ruoxin Zhang[1,2] | Fei Zhou[3] | Lei Cheng[1,2] | Alexandria Yu[4] | Meiling Zhu[5] | Mengyun Wang[1,2] | Zhuanxu Zhang[6] | Jiaqing Xiang[1,2,7] | Qingyi Wei[1,8]

[1] Cancer Institute, Fudan University Shanghai Cancer Center, Shanghai, China

[2] Department of Oncology, Shanghai Medical College, Fudan University, Shanghai, China

[3] Department of Oncology, Shanghai Jiao Tong University affiliated Shanghai General Hospital, Shanghai, China

[4] Department of Public Health, Johns Hopkins University, Baltimore, Maryland

[5] Department of Oncology, Xinhua Hospital Affiliated to Shanghai Jiao Tong University, Shanghai, China

[6] Department of Pathology, Fudan University Shanghai Cancer Center, Shanghai, China

[7] Department of Thoracic Surgery, Fudan University Shanghai Cancer Center, Shanghai, China

[8] Duke Cancer Institute, Duke University Medical Center, Durham, North Carolina

Correspondence
Jiaqing Xiang, Department of Thoracic Surgery, Fudan University Shanghai Cancer Center, Shanghai 200032, China.
Email: j.q.xiang@hotmail.com
Qingyi Wei, Cancer Institute, Fudan University Shanghai Cancer Center, Shanghai 200032, China.
Email: qingyi.wei@duke.edu; weiqingyi@yahoo.com

Funding information
China's Thousand Talents Program at Fudan University, Shanghai, China; National Human Genetic Resources Sharing Service Platform, Grant number: 2005DKA21300

The benefits of platinum-based chemotherapy (PBC) on survival of esophageal squamous cell carcinoma (ESCC) patients are inexplicit due to the varied therapeutic effects. Nucleotide excision repair (NER) pathway plays a vital role in removing platinum-DNA adducts in tumor cells and hence may modulate the therapeutic effect and survival outcome. The present study assessed the associations of 26 potentially functional regulatory single nucleotide polymorphisms (rSNPs) in nine core NER genes with disease-free survival (DFS) and overall survival (OS) in 339 ESCC patients. We found that *ERCC2* rs2097215 T and rs3916788 A, *ERCC5* rs3759497 A and *XPC* rs3731054 C alleles were associated with unfavorable DFS. Patients carrying high-risk allele group (HRG, 5-8 risk alleles) had a significantly shorter DFS, compared with those carrying low-risk alleles (LRG, 0-4 risk alleles) [adjusted hazards ratio (HR_{adj}) = 1.64, 95% CI = 1.23-2.19, $P_{adj} < 0.001$]. Three of these SNPs (ie, *ERCC2* rs2097215 T and rs3916788 A and *ERCC5* rs3759497 A) were also significantly associated with a poorer OS (HRG vs LRG: $HR_{adj} = 1.75$, 95%CI = 1.23-2.47, $P_{adj} = 0.002$). The expression quantitative trait loci (eQTL) analysis revealed significant genotype-expression correlations for *ERCC5* rs3759497 and *ERCC2* 2097215 and rs3916788, which suggest regulatory roles of these SNPs. It appears that these NER variants may independently or jointly exert an impact on survival outcome of Chinese ESCC patients undergoing adjuvant platinum-based therapy. Large studies are warranted to validate these findings.

KEYWORDS
disease-free survival, overall survival, regulatory single nucleotide polymorphism

Abbreviations: CI, confidence interval; DFS, disease-free survival; eQTL, expression quantitative trait loci; ESCC, esophageal squamous cell carcinoma; HR, hazards ratio; HRG, high-risk allele group; LRG, low-risk allele group; NER, nucleotide excision repair; OS, overall survival; PBC, platinum-based chemotherapy; rSNP, regulatory single nucleotide polymorphism; SNP, single nucleotide polymorphism.

Ruoxin Zhang and Fei Zhou contributed equally to this work and should be considered as co-first authors.

1 | INTRODUCTION

Esophageal cancer (EC) is the eighth leading cause of cancer and the sixth most common cancer-related death, and worldwide, there are approximately 450 000 new cases and approximately 400 000 EC related deaths each year.[1] China, South Africa, and North Central Asia have the highest incidence, with approximately 50% of all the cases occurring in China.[1] Histologic types of EC showed varied geographic distribution, with up to 90% of esophageal squamous-cell carcinoma (ESCC) cases occurring in China (or Eastern Asian countries),[2] whereas adenocarcinoma is more common in western counties. Most of patients have been diagnosed with an advanced stage disease, followed by poor survival. The 5-year survival rates range from 10 to 24% for surgery-alone patients and less than 5% for stage IV patients.[3] To improve prognostic outcome of ESCC patients, most surgeries are followed by adjuvant chemotherapies. However, investigations of chemotherapy (consisting of cisplatin and fluorouracil) effects have yielded conflicting results in various clinical trials, in which varied response rates and improvement of survival outcome are seen in different patient groups.[4–7] Apart from conventional prognostic factors including TNM stage, weight loss, dysphagia, advanced age and lymphatic micro-metastases,[3,4] genetic variations may also contribute to different prognostic outcomes of ESCC patients.

Platinum-based chemotherapy (PBC) drugs (eg, cisplatin) function by crosslinking to DNA, forming intra- and inter-strand adducts that consequently distorting the DNA helix, inhibiting DNA replication, and driving cell apoptosis.[8] However, the effectiveness of this treatment varies from patient to patient and has an overall response rate of 6-32%, which is possibly due to varied individual inherited abilities to repair damaged DNA.[9,10] Evidence has shown that common genetic variation, such as single nucleotide polymorphisms (SNPs) and levels of gene expression are key factors influencing DNA repair capacity, in which nucleotide excision repair (NER) pathway is a versatile DNA repair system that repairs DNA damage induced by PBC.[10,11]

The NER pathway is one of the five major DNA repair mechanisms that include four primary steps. The initial DNA lesion recognition step is mediated by xeroderma pigmentosum, complex group C (XPC), and E (XPE).[12] Second, the damaged DNA is unwound and remodeled by helicases XPB/ERCC3 and XPD/ERCC2, and bound to XPA and replication protein A (RPA).[13] The third step DNA incision involves nucleases XPG/ERCC5 and XPF-ERCC1, which excise the DNA from either side of the identified lesion. In the last step, a patch is resynthesized by proliferating cell nuclear antigen, replicative polymerases δ, ε, or κ and a ligase.[14,15] DNA repair safeguards genome integrity and stability, with mounting evidence suggesting that inefficient DNA repair can lead to diseases such as xeroderma pigmentosum or Cockayne syndrome;[13] dysregulated expression of NER genes may lead to cisplatin resistance in ovarian, glioma, bladder, and lung cancer cells which may in turn affect recipients' survival outcome.[16,17]

Previously, our group investigated the associations between six SNPs in the NER genes and the survival outcome of ESCC patients receiving PBC adjuvant chemotherapy.[18] In the present study we expanded the scale by studying the impact of all the potentially functional regulatory SNPs (rSNPs) in the nine core genes of the NER pathway (*ERCC1*, *ERCC2*, *ERCC3*, *ERCC4*, *ERCC5*, *RAD23B*, *XPA*, *XPC*, and *XPE*) on survival of ESCC patients who had received PBC treatment.

2 | MATERIALS AND METHODS

2.1 | Study population

The present study was conducted on 344 patients diagnosed with ESCC in Fudan University Shanghai Cancer Center (FUSCC) between March 2009 and December 2010. The recruited patients must fulfill all the following criteria: 1) unrelated Han Chinese with TNM stage II-III tumors of ESCC without a prior history of cancer other than in situ carcinoma; 2) not received neoajduvant chemoradiotherapy and underwent lymphadenectomy; 3) received four cycles of PBC as the adjuvant chemotherapy and without postoperative chemoradiotherapy; and 4) without recent myocardial infarction, cardiac arrhythmia, active congestive heart failure or cerebral apoplexy, crankiness, or depression. Clinical data including age at treatment, sex, smoking and drinking status, TNM stage, histological grade, vessel, and neural invasion were collected from patients' medical records. Tumor staging was conducted according to the 7th edition of the American Joint Committee on Cancer Staging system. As a result, the final analysis included 344 patients who met the above-mentioned criteria and completed four cycles of adjuvant chemotherapy (ie, oxaliplatin 135 mg/m^2 d1 or cisplatin 40 mg/m^2 d1-3 plus 5-Fu 750 mg/m^2 d1-5).

2.1.1 | Survival data

Survival data were collected from patient's next of kin through a telephone follow-up and outpatient medical records of the follow-up. Overall survival (OS) time was defined as the time period between the surgical resection and the date of the last follow-up or the death. Disease free survival (DFS) was calculated from the date of the surgical resection until the date of the first local recurrence or metastases in the last follow-up. Patients without progression were censored at the date of the last contact. All participants signed an informed consent form upon the start of this study. The Institutional Review Board of Fudan University Shanghai Cancer Center (FUSCC) approved this study, with all methods performed in accordance with the guidelines and regulations of FUSCC.

2.1.2 | SNP selection

Twenty-nine rSNPs in eight out of the nine (none found in *DDB1/XPE*) core NER genes were selected by using SNP/GeneView in dbSNP database (http://www.ncbi.nlm.nih.gov/snp/). The selection criteria were set as follows: minor allele frequency (MAF) ≥5% in Han Chinese, in the regulatory region (5′ near gene, 5′UTR, intron, 3'near gene, or 3′ UTR), have predicted functions (e.g., transcription factor binding site, splicing, miRNA binding site, or significant expression quantitative trait loci [eQTL]) by SNPinfo (http://snpinfo.niehs.nih.gov/snpfunc.htm)

and GTEx portal (http://www.gtexportal.org/home/). The detailed characteristics of all investigated genes and rSNPs are shown in Supplemental Table S1.

2.1.3 | SNPseq genotyping

Genomic DNA was extracted from the blood samples of all study participants by using the Qiagen Blood DNA Mini kit (Qiagen, Valencia, CA). All the rSNPs were genotyped using FastTaget, a next generation sequencing-based method using Illumina Miseq2000 Platform (2 × 250 bp, Illumina, CA). Prior to sequencing, 5% of the samples were randomly selected and subjected to 1% agarose gel electrophoresis for quality control. Genomic regions containing investigating rSNPs were amplified using the FastTarget™ technology (Genesky Biotechnologies Inc, Shanghai, China). The primers used for genotyping are listed in Supplemental Table S3. DNA fragments containing the rSNPs were ligated with the adaptor by using Q5 DNA polymerase Kit (New England Biolabs, MA), and further purified by Agencourt AMPure XP (Beckman Coulter, CA). Next-generation sequencing of the amplification products was then carried out by MiSeq 2000 Sequencer following the manufacturer's standard protocols. Sequencing depth of more than 30× was achieved for over 90% of the samples. Output sequence data were trimmed and then compared with fragment reference sequences (hg19) using the Blat program.[28] Burrows-Wheeler Aligner (BWA, V 0.7.5a) was used to map the reads,[19] followed by Sequence Alignment/Map (SAM)-to-BAM conversion, sorting, and removal of duplicates using SAM tools (v0.1.19).[20] Combined rSNP calling was performed on the resulting BAM files using Genome Analysis Toolkit (GATK, https://software.broadinstitute.org/gatk/best-practices/) and VarScan programs.[21] Finally SNP annotation was done by the Annovar program.[22]

2.1.4 | eQTL analysis

GTEx portal was used to assess the correlation between survival-associated NER genetic variants and NER core genes mRNA expression levels (http://www.gtexportal.org/home/). The latest version (Oct. 2015 release) includes genotype data from approximately 450 donors and 9600 RNA-seq samples across 51 tissues and 2 cell lines. The associations of genetic variants and gene expression are investigated in relevant tissue types.

2.1.5 | Statistical analysis

The association between the clinical or genetic variable and DFS/OS was evaluated using Cox proportional hazards regression model, calculated as HRs with their corresponding 95% CIs. Kaplan-Meier test was used to assess each genetic variant on the cumulative probability of DFS and OS.[23] Log-rank test was used to examine the difference in survival between groups. In order to correct for multiple tests, false-positive report probability (FPRP) analysis was performed with a prior probability of 0.01, 0.1, and 0.2, respectively, being assigned to detect an HR of 1.5 (for a risk allele) or 0.67 (for a protective allele) for an association with genotypes or alleles of each SNP. Only significant result with an FPRP value less than 0.20 was considered a noteworthy association.[24] Stratification analysis was also carried out using Cox model by patients' clinical characteristics. For survival prediction model construction, independent predictors including selected clinical variables and genetic variants were selected by the backward-stepwise Cox proportional hazards regression analysis. The receiver operating characteristic (ROC) analysis with the area under the curve (AUC) value was used to compare sensitivity and specificity of the OS and DFS prediction by the included parameters. Predictive values of selected variables were evaluated by I/D AUC of the ROC curves for censored data and C index for comparison of survival models. The I/D ROC and I/D AUC were calculated and plotted by RisksetROC package of R software (The R Project for Statistical Computing, Austria, version 3.2.3).[25] All statistical analyses were performed by SAS software (SAS Institute, NC, version 9.4) and R software. Unless stated otherwise, all *P* values were two-sided with a significance level of $P < 0.05$.

3 | RESULTS

3.1 | Characteristics of the study population

The present study consisted of 339 ESCC patients who received esophagectomy and PBC, with complete data on demographic, clinical characteristics, DFS, and OS. Of all the patients, 291 were males and 48 were females, with a median age of 59 (a range of 37-77) years, and 219 (over 329, because 70 subjects had missing data on smoking; 66.6%) were smokers. All subjects had an intermediate TNM stage (II or III) esophageal squamous cancer, with 276 (81.4%) and 63 (18.6%) underwent two fields and three fields lymphadenectomy, respectively. The median follow-up time was 38.8 months, during which 193 (56.9%) patients were deceased. In the univariate analysis, TNM stage (HR = 1.81, 95%CI = 1.34-2.45, $P < 0.001$), neural invasion (HR = 1.60, 95%CI = 1.21-2.12, $P = 0.001$) and lymphadenectomy (HR = 1.62, 95% CI = 1.17-2.23, $P = 0.003$) were significantly associated with DFS. Additionally, smoking was also significantly associated with OS of the ESCC patients (HR = 1.46, 95%CI = 1.05-2.03, $P = 0.023$) (Table 1). Age, sex, and smoking, but not vessel invasion, were still included in the further multivariate Cox regression analysis, due to the evidence on their effects on survival.[3]

3.2 | NER rSNPs and survival outcomes

Out of the 29 rSNPs that are located in the regulatory region of the eight NER genes, three SNPs with a call rate of <95% (rs2607735, rs1007616, and rs7507745) and five samples with poor genotyping outcome were excluded from further analysis, which amounted to a total of 339 patients and 26 rSNPs. In both univariate analysis and multivariate analysis with adjustment for clinical variables, we found that DFS of the patients was significantly associated with *ERCC5* rs3759497 (GA/AA vs GG, $HR_{adj} = 1.53$, 95%CI = 1.07-2.19, $P_{adj} = 0.020$), *XPC* rs3731054 (GG vs CC/CG, $HR_{adj} = 0.50$, 95%CI = 0.26-0.97, $P_{adj} = 0.040$), *ERCC2* rs2097215 (CT/TT vs CC, $HR_{adj} = 1.62$, 95%CI = 1.12-2.33, $P_{adj} = 0.010$), and

TABLE 1 Association between clinical characteristics and disease-free survival (DFS) and overall survival (OS) of patients with ESCC

	Disease-free survival (DFS)					Overall survival (OS)				
	MST (mo)	Event/No.	P^a	HR (95% CI)	P^b	MST (mo)	Event/No.	P^a	HR (95%CI)	P^b
Age										
≤59	22.8	110/179	0.405	1.00 (ref.)	0.404	43.6	95/179	0.194	1.00 (ref.)	0.195
>59	21.2	107/160		1.12 (0.86-1.46)		33.2	98/160		1.21 (0.91-1.60)	
Sex										
Male	21.5	192/291	0.071	1.00 (ref.)	0.336	36.7	171/291	0.165	1.00 (ref.)	0.168
Female	35.9	25/48		0.77 (0.44-1.32)		-	22/48		0.73 (0.47-1.14)	
Smoking										
Never	25.8	59/110	0.075	1.00 (ref.)	0.076	79.5	49/102	**0.022**	1.00 (ref.)	**0.023**
Yes	20.6	147/219		1.31 (0.97-1.78)		32.1	134/219		1.46 (1.05-2.03)	
Missing		70								
Drinking										
No	22.8	75/125	0.110	1.00 (ref.)	0.111	42.2	68/125	0.245	1.00 (ref.)	0.246
Yes	20.9	120/173		1.27 (0.95-1.69)		34.5	107/173		1.20 (0.88-1.62)	
Missing		101								
Vessel invasion										
No	23.7	174/276	0.194	1.00 (ref.)	0.195	41.9	155/276	0.442	1.00 (ref.)	0.442
Yes	16.1	43/63		1.25 (0.89-1.74)		29.1	38/63		1.15 (0.81-1.64)	
Neural invasion										
No	26.3	141/240	**<0.001**	1.00 (ref.)	**0.001**	52.4	121/240	**<0.001**	1.00 (ref.)	**<0.001**
Yes	15.9	76/99		1.60 (1.21-2.12)		26.4	72/99		1.70 (1.27-2.28)	
Lymphadenectomy										
Two fields	23.7	169/276	**0.003**	1.00 (ref.)	**0.003**	44.4	150/276	**0.006**	1.00 (ref.)	**0.006**
Three fields	13.3	48/63		1.62 (1.17-2.23)		25.4	43/63		1.61 (1.15-2.26)	
TNM stage										
II	59.8	57/114	**<0.001**	1.00 (ref.)	**<0.001**	-	51/114	**<0.001**	1.00 (ref.)	**<0.001**
III	19.4	160/225		1.81 (1.34-2.45)		28.1	142/225		1.78 (1.29-2.46)	

CI, confidence interval; DFS, disease-free survival; HR, hazards ratio; mo, month; MST, median survival time; OS, overall survival. $P < 0.05$ are indicated in bold.

[a]*P*-value for Log-rank test.

[b]*P*-value for univariate Cox hazards regression analysis.

rs3916788 (AA vs CC, HR_{adj} = 1.60, 95%CI = 1.07-2.40, P_{adj} = 0.022) (Table 2), in which the effects of three rSNPs (rs3759497, rs2097215, and rs3916788) remained noteworthy with a prior FPRP of 0.1 (Table 3). These three SNPs were also found to be significantly associated with OS of these patients (rs3759497 GA/AA vs GG, HR_{adj} = 1.57, 95%CI = 1.05-2.36, P_{adj} = 0.027; rs2097215 CT/TT vs CC, HR_{adj} = 1.98, 95%CI = 1.32-2.99, P_{adj} = 0.001; rs3916788 AC/AA vs. CC, HR_{adj} = 1.86, 95%CI = 1.24-2.79, P_{adj} = 0.003). When we combined all risk alleles into the number of risk alleles (NRA) to assess their joint effect on DFS and OS, the frequencies of patients with an NRA score ranges of 0-4, 5-6, and 7-8 were 169, 115, and 48 for DFS, whereas the NRA score ranges of 0-2, 3-4, and 5-6 were 119, 141, and 73 for OS. A dose-dependent trend was observed for patients with higher NRA scores, and patients at the top tier of NRA (7-8 risk alleles for DFS) exhibited the highest risk for disease progression (MST: 18 vs 37.6 months, $P_{log\text{-}rank}$ = 0.0003, HR_{adj} = 1.57, 95%CI = 1.05-2.34, P_{adj} = 0.029), whereas those carrying the highest score of NRA (5-6 risk alleles for OS) also showed the highest risk of death ($P_{log\text{-}rank}$ < 0.0001, HR_{adj} = 1.87, 95%CI = 1.31-2.68, P_{adj} < 0.001) (Table 2, Figures 1A and 1B). We further dichotomized all the patients into high-risk group (HRG) and low-risk group (LRG), and patients in the HRG showed a significantly increased risk of disease progression and death when compared with patients in the LRG (DFS: $P_{log\text{-}rank}$ < 0.0001; HR_{adj} = 1.64, 95%CI = 1.23-2.19, P_{adj}<0.001; OS: $P_{log\text{-}rank}$ < 0.0001, HR_{adj} = 1.75, 95% CI = 1.23-2.47, P_{adj} = 0.002) (Table 2, Figures 1C and 1 D).

3.3 | Stratified analysis between the risk genotypes and ESCC survival

Stratified analysis was performed to assess the differential effects of clinical variables on survival outcome associated with the two risk allele groups (HRG or LRG). Overall, HRG carriers tended to have a substantial shorter time to disease progression in most of the subgroups, except for the female patients and the subgroup with lymphadenectomy covering three fields (Table 4). Similarly, compared with LRG patients, a significantly increased risk of death was also observed in HRG patients in most subgroups, except for female patients, those receiving three fields lymphadenectomy, and those who had tumors with vessel invasion (Table 4).

3.4 | In silico analysis of the regulatory SNP linkage disequilibrium block

An extensive in silico analysis was further conducted for the ESCC survival-associated rSNPs. Each rSNP in *ERCC2*, *ERCC5*, or *XPC* was located in different LD blocks. The *ERCC2* rs2097215 block contains a total of 45 SNPs (r^2>0.8), six of which were predicted to have biological functions (Supplemental Table S4). Most of these SNPs have been reported to be associated with risk of different cancers, with rs238406 being associated with ESCC survival by a previous study.[18] Our finding further strengthened the association of this LD block with ESCC survival, suggesting that this block may have one or more potentially functional SNPs. On the other hand, no other variants have been found to be in high LD with *ERCC2* rs3916788 according to the 1000 Genomes Project data.

The *ERCC5* rs3759497 LD block consists of 17 SNPs, among which two were predicted to be functional (Supplemental Table S4). While rs1323697 has not been associated with any type of cancer, rs2094258 has been previously reported to be associated with risk of different cancers and survival of esophageal cancer.[18,26]

Out of the 16 SNPs within the *XPC* rs3731054 LD block (r^2>0.80), four SNPs were predicted to be functional (Supplemental Table S4). The presence of a non-synonymous SNP rs2228000 (Ala499Val) has been widely associated with an increased overall risk of cancers,[27–32] particularly bladder cancer.[33] The 3′UTR variant rs2229090 has been reported to be associated with NSCLC survival,[34,35] but it was not found to be associated with ESCC survival in the present study.

3.5 | Correlations between survival risk genotypes and mRNA expression levels

To examine the genotypic effect of the survival-associated rSNPs on gene expression, the eQTL analysis of the four NER rSNPs was further conducted by using GTEx database. In normal muscle skeletal and esophagus tissues, the *ERCC5* rs3759497 A allele was associated with a significantly higher mRNA levels of *ERCC5* and *BIVM*, respectively. Similarly, *ERCC2* rs2097215 C and rs3916788 C alleles were associated with a significantly decreased level of *ERCC2* mRNA expression in esophageal tissues ($P = 0.32 \times 10^{-4}$ for rs2097215 and P = 8.6e-8 for rs3916788, Figure 2). Therefore, it is biologically plausible that these variants may collectively influence ESCC survival via regulating the expressions of NER genes.

3.6 | NER variant-based prediction model for ESCC patients' survival

The cumulative effects of NER gene rSNPs on ESCC patients' DFS and OS were further examined by the receiver operation characteristics (ROC) analysis. To assess the discriminative power of the prediction model, we also calculated incident/dynamic (I/D) area under the curve (AUC) of the time-dependent ROC curves and estimated the C index. We performed the backward-stepwise Cox regression analysis to further select optimal predictors of survival in ESCC patients, using covariates listed in Table 1 and HRG (5-8 risk alleles for DFS and 3-6 risk alleles for OS). The final selected variables for DFS prediction model include disease stage, lymphadenectomy methods, and HRG, whereas smoking status was added in addition to the variables listed above in the prediction model for OS (Supplemental Table S2). The addition of HRG showed some effects on improving the discriminative power of prediction models, yet this effect was not statistically significant (with and without HRG: I/C AUC: 0.61 vs 0.58, 95%CI: 0.57-0.65 vs 0.55-0.63, P = 0.232 for DFS; I/C AUC: 0.63 vs. 0.59, 95%CI: 0.58-0.66 vs 0.55–0.63, P = 0.236 for OS) (Supplemental Figure S1). It is likely that the study sample size was not large enough for the needed study power to detect such a genetic effect in the multivariable model.

TABLE 2 Associations of NER rSNPs with DFS and OS in Chinese ESCC patients

	DFS					OS				
Gene/SNP	Event/No.	MST (mo)	P^a	Adjusted HR^b (95%CI)	P^b	Event/No.	MST (mo)	P^a	Adjusted HR^b (95%CI)	P^b
ERCC5										
rs3759497										
GG	46/90	50.8	**0.004**	1.00 (ref)		36/90	-	**<0.001**	1.00 (ref)	
GA	123/185	21.0		1.50 (1.03-2.17)	**0.034**	111/185	36.4		1.46 (0.96-2.24)	0.081
AA	48/64	16.7		1.63 (1.05-2.54)	**0.033**	46/64	25.1		1.91 (1.15-3.16)	**0.012**
Trend					**0.026**					**0.010**
GG	46/90	50.8	**0.003**	1.00 (ref)		36/90	-	**<0.001**	1.00 (ref)	
GA/AA	171/249	19.5		1.53 (1.07-2.19)	**0.020**	157/249	30.2		1.57 (1.05-2.36)	**0.027**
GG/GA	169/275	23.7	**0.020**	1.00 (ref)	0.244	147/275	47.2	**0.003**	1.00 (ref)	
AA	48/64	16.7		1.23 (0.87-1.74)	0.242	46/64	25.1		1.29 (0.78-2.12)	0.326
XPC										
rs3731054										
CC	100/155	19.5	0.243	1.00 (ref)		93/155	31.4	0.194	1.00 (ref)	
CG	103/158	21.8		0.86 (0.64-1.16)	0.301	88/158	41.4		0.94 (0.68-1.31)	0.716
GG	13/25	55.9		0.46 (0.24-0.90)	**0.024**	11/25	-		0.51 (0.24-1.08)	0.076
Trend					**0.032**					0.175
CC	100/155	19.5	0.308	1.00 (ref)		93/155	31.4	0.179	1.00 (ref)	
CG/GG	116/183	23.7		0.80 (0.60-1.06)	0.122	99/183	44.4		0.88 (0.64-1.22)	0.455
CC/CG	203/313	21.5	0.118	1.00 (ref)		181/313	35.9	0.124	1.00 (ref)	
GG	13/25	55.9		0.50 (0.26-0.97)	**0.040**	11/25	-		0.52 (0.25-1.10)	0.086
ERCC2/XPD										
rs2097215										
CC	46/86	39.7	**0.007**	1.00 (ref)		38/86	-	**0.001**	1.00 (ref)	
CT	101/158	21.5		1.48 (1.01-2.17)	**0.045**	88/158	36.4		1.73 (1.13-2.66)	**0.012**
TT	69/94	16.5		1.77 (1.17-2.68)	**0.007**	66/94	26.9		2.26 (1.43-3.58)	**<0.001**
Trend					**0.005**					**<0.001**
CC	46/86	39.7	**0.012**	1.00 (ref)		38/86	-	**0.004**	1.00 (ref)	
CT/TT	170/252	19.5		1.62 (1.12-2.33)	**0.010**	154/252	30.1		1.98 (1.32-2.99)	**0.001**
CC/CT	147/244	26.3	**0.007**	1.00 (ref)		126/244	47.5	**0.002**		
TT	69/94	16.5		1.37 (1.00-1.87)	0.050	66/94	26.9		1.58 (1.12-2.22)	**0.008**

(Continues)

TABLE 2 (Continued)

	DFS					OS				
Gene/SNP	**Event/No.**	**MST (mo)**	***P*[a]**	**Adjusted HR[b] (95%CI)**	***P*[b]**	**Event/No.**	**MST (mo)**	***P*[a]**	**Adjusted HR[b] (95%CI)**	***P*[b]**
rs3916788										
CC	47/82	35.32	**0.019**	1.00 (ref)		39/82	-	**0.002**	1.00 (ref)	
AC	94/152	21.7		1.29 (0.89-1.89)	0.180	81/152	41.4		1.55 (1.01-2.36)	**0.043**
AA	73/99	16.5		1.60 (1.07-2.40)	**0.022**	70/99	26.8		2.12 (1.36-3.29)	**<0.001**
Trend					**0.021**					**<0.001**
CC	47/82	35.3	0.081	1.00 (ref)		39/82	-	**0.021**	1.00 (ref)	
AC/AA	167/251	19.5		1.42 (0.99-2.04)	0.057	151/251	30.6		1.86 (1.24-2.79)	**0.003**
CC/AC	141/234	26.3	**0.007**	1.00 (ref)		120/234	47.7	**0.001**	1.00 (ref)	
AA	73/99	16.5		1.35 (0.99-1.84)	0.057	70/99	26.8		1.62 (1.16-2.27)	**0.005**
Number of risk allele (NRA)[c]						**Number of risk allele (NRA)[d]**				
0-4	94/169	37.6	**<0.001**	1.00 (ref)		52/119	-	**<0.0001**	1.00 (ref)	
5-6	82/115	16.4		1.37 (1.02-1.85)	**0.037**	82/141	29.1		1.07 (0.78-1.49)	0.667
7-8	37/48	18.0		1.57 (1.05-2.34)	**0.029**	56/73	25.7		1.87 (1.31-2.68)	**<0.001**
Trend					**<0.001**					**<0.001**
LRG[e]	94/169	37.9	**<0.001**	1.00 (ref)		52/119	-	**<0.001**	1.00 (ref)	
HRG[f]	119/163	16.5		1.64 (1.23-2.19)	**<0.001**	138/214	27.6		1.75 (1.23-2.47)	**0.002**

DFS, disease-free survival; OS, overall survival; MST, median survival time; mo, month; HR, hazards ratio; CI, confidence interval; LRG, low-risk allele group; HRG, high-risk allele group. $P < 0.05$ are indicated in bold.

[a]*P*-value from Log-rank tests.

[b]Data were calculated by using Cox hazards regression analysis with adjustment for age at treatment, sex, smoking status, drinking status, stage, operation, neuron invasion.

[c]NRA for DFS include rs3759497 A, rs3731054 C, rs2097215 T, rs3916788 A alleles.

[d]NRA for OS include rs3759497 A, rs2097215 T, rs3916788 A alleles. Each line corresponds to 0-2, 3-4, 5-6 alleles.

[e]LRG stands for low-risk group. LRG for DFS includes 0-4 alleles; LRG for OS includes 0-2 alleles.

[f]HRG stands for high-risk group. HRG for DFS includes 5-8 alleles; HRG for OS includes 3-6 alleles.

TABLE 3 False-positive report probability values for associations between ESCC survival outcomes and genotypes of NER gene rSNPs

	DFS						OS					
				Prior probability						Prior probability		
Genotypes	HR[a] (95% CI)	P[b]	Statistical power[c]	0.2	0.1	0.01	HR[a] (95% CI)	P[b]	Statistical power[c]	0.2	0.1	0.01
ERCC5 rs3759497												
GA vs. GG	1.50 (1.03-2.17)	**0.034**	0.650	0.173	0.320	0.838						
AA vs. GG	1.63 (1.05-2.54)	**0.033**	0.437	0.232	0.405	0.882	1.91 (1.15-3.16)	**0.012**	0.394	0.109	0.215	0.751
Trend		**0.026**	0.965	0.097	**0.195**	0.727		**0.010**	0.945	0.041	**0.087**	0.512
GA/AA vs. GG	1.53 (1.07-2.19)	**0.020**	0.702	0.102	0.204	0.738	1.57 (1.05-2.36)	**0.027**	0.653	0.142	0.271	0.804
ERCC2 rs2097215												
CT vs. CC	1.48 (1.01-2.17)	**0.045**	0.614	0.227	0.398	0.879	1.73 (1.13-2.66)	**0.012**	0.557	0.079	**0.162**	0.681
TT vs. CC	1.77 (1.17-2.68)	**0.007**	0.508	0.052	**0.110**	0.577	2.26 (1.43-3.58)	**<0.001**	0.470	0.004	**0.009**	0.095
Trend		**0.005**	0.983	0.020	**0.044**	0.335		**<0.001**	0.970	0.001	**0.003**	0.030
CT/TT vs. CC	1.62 (1.12-2.33)	**0.010**	0.683	0.055	**0.116**	0.592	1.98 (1.32-2.99)	**0.001**	0.629	0.006	**0.014**	0.136
TT vs. CT/CC							1.58 (1.12-2.22)	**0.008**	0.046	0.046	**0.097**	0.543
ERCC2 rs3916788												
AC vs. CC	1.29 (0.89-1.89)	0.180					1.55 (1.01-2.36)	**0.043**	0.548	0.239	0.414	0.886
AA vs. CC	1.60 (1.07-2.40)	**0.022**	0.510	0.147	0.280	0.810	2.12 (1.36-3.29)	**<0.001**	0.468	0.008	0.017	0.160
Trend		**0.021**	0.984	0.078	**0.161**	0.679		**<0.001**	0.971	0.002	**0.005**	0.048
AC/AA vs. CC							1.86 (1.24-2.79)	**0.003**	0.619	0.019	**0.042**	0.324
AA VS. CC/AC							1.62 (1.16-2.27)	**0.005**	0.679	0.029	**0.062**	0.422
XPC rs3731054												
GG vs. CC	0.46 (0.24-0.90)	**0.024**	0.274	0.260	0.441	0.897						
Trend		**0.032**	0.934	0.120	0.236	0.772						
GG vs. CC/CG	0.50 (0.26-0.97)	**0.040**	0.308	0.342	0.539	0.928						

CI, confidence interval; DFS, disease free survival; HR, hazards ratio; OS, overall survival. $P < 0.05$ are indicated in bold.

[a]Data were calculated by using Cox hazards regression analysis with adjustment for age at treatment, sex, smoking status, drinking status, stage, operation, neuron invasion.

[b]*P*-value were calculated by chi-square test.

[c]Statistical power was calculated using the number of observations in the subgroup and the HR and *P*-values in this table.

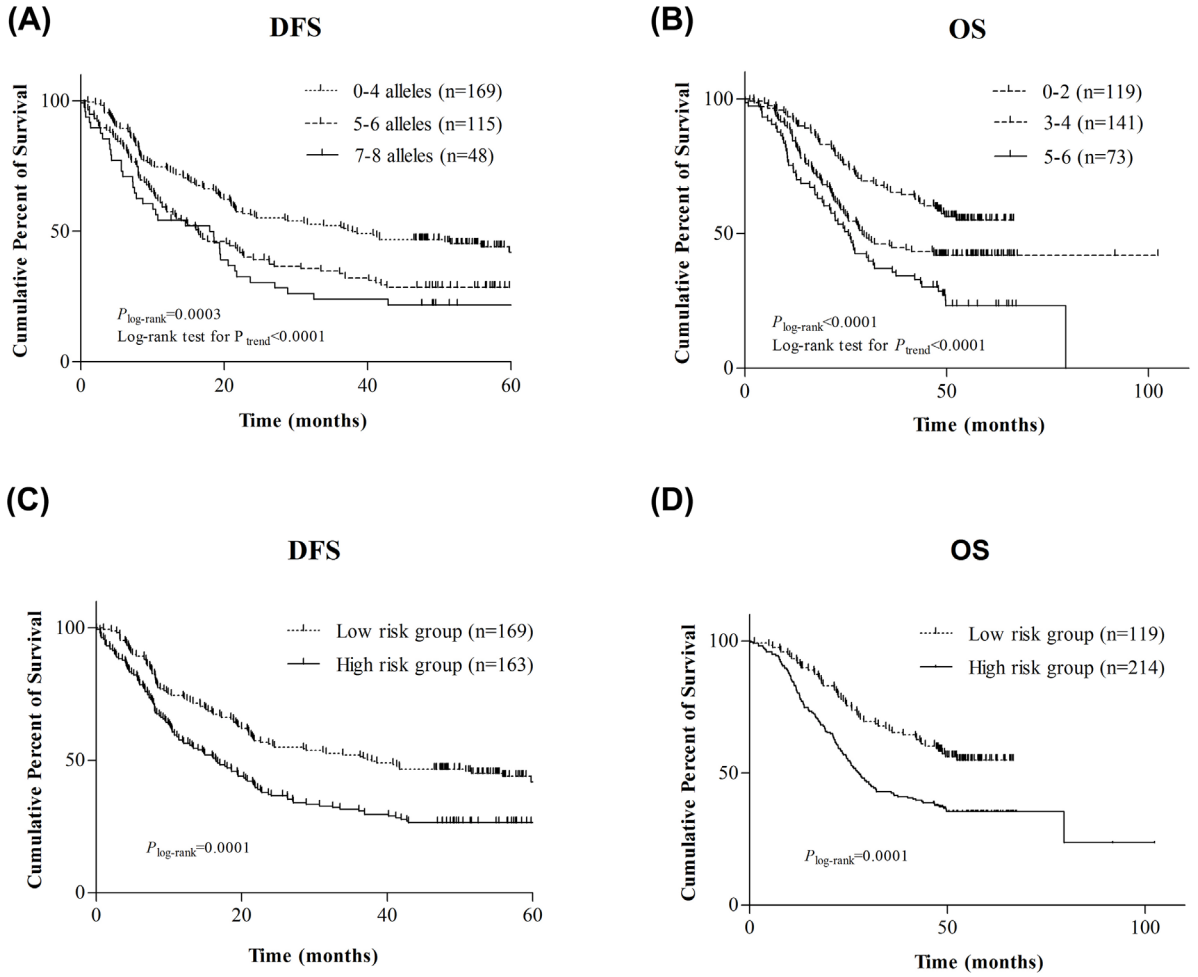

FIGURE 1 Kaplan-Meier survival curves for ESCC patients by number of risk alleles (NRA) and high/low-risk groups.

4 | DISCUSSION

The cispatin and 5-fluorouracil (FU)-based therapy is a standard chemotherapy for ESCC in China, in addition to esophagectomy with lymph node dissection.[36] However the survival outcome from combinatory approach was far less than satisfactory, partially due to chemotherapy resistance and attenuated therapeutic efficacy. Platinum drugs form bulky DNA adducts that lead to apoptosis of cancer cells—a process in which the adducts are majorly repaired by the NER mechanism. DNA repair capacity has generally been inversely associated with survival in NSCLC patients under PBC treatment.[11] The NER pathway primarily involves in the repair of DNA adducts induced by PBC drugs. Therefore, suboptimal NER capacity may confer sensitivity to such ESCC treatments and thus improve the response rate of these drugs. To date, the reported SNPs in NER genes associated with survival of ESCC are largely focused on the coding region,[37,38] with very few reports on SNPs in the regulatory regions for their associations with ESCC outcomes.

In the present study, we employed a hypothesis-based approach with a main focus on common regulatory variants with predicted biological functions. Three rSNPs, *ERCC5* rs3759497 and *ERCC2* rs2097215 and *ERCC2* rs3916788, were significantly associated with both DFS and OS in ESCC patients treated with adjuvant PBC, the effects of which remained significant with a prior FPRP of 0.1. Additionally, *XPC* rs3731054 was associated with DFS alone. To date, none of these SNPs have been reported to be associated with ESCC survival or PBC efficacy, although two SNPs (*ERCC2* rs238406 and rs2094258) within the same LD block of the significant rSNPs were reported to be associated with ESCC survival.[18,26] Our finding has further strengthened that these LD blocks may contribute to survival of Chinese ESCC patients. Collectively, patients carrying higher NRA of the rSNPs have exhibited an elevated risk for disease progression and death, compared with the low NRA carriers (Table 2, Figure 1), which suggests that these variants may collectively interact with platinum drugs. Based on the eQTL results, both *ERCC2* rs2097215 T and rs3916788 A minor alleles were shown to up-regulate the mRNA expression of *ERCC2* in esophageal tissues, which were then linked to unfavorable survival outcomes. The *ERCC5* rs3759497 A allele was associated with a higher mRNA expression of *ERCC5* in muscle skeletal tissues and a higher expression of *BIVM* in various tissues.

ERCC2/XPD is generally recognized as the major repair gene in response to DNA damage induced by chemotherapy drugs such as cisplatin.[39] It encodes a helicase that is involved in DNA unwinding

TABLE 4 Stratified multivariate analysis of clinical variables for association between HRG and LRG in ESCC patients treated with platinum-based chemotherapy

Patient characteristics	Disease free survival (DFS)				Overall survival (OS)			
	No. of patients (LRG/HRG[a])	Progression no. (%) (LRG/HRG[a])	HR (95% CI)	P[c]	No. of patients (LRG/HRG[b])	Death no. (%) (LRG/HRG[b])	HR (95% CI)	P[c]
All subjects								
Age								
<60	103/90	55 (46.6)/66 (26.67)	1.67 (1.16-2.42)	0.007	74/119	30 (59.5)/75 (37.0)	1.88 (1.22-2.90)	0.004
≥60	73/73	43 (41.1)/53 (27.4)	1.69 (1.12-2.55)	0.012	51/95	25 (51.1)/63 (33.7)	1.76 (1.09-2.83)	0.020
Sex								
Male	151/140	88 (41.7)/104 (25.7)	1.63 (1.22-2.17)	0.001	107/184	51 (52.3)/120 (34.8)	1.74 (1.25-2.42)	0.001
Female	25/23	10 (60)/15 (34.8)	2.05 (0.87-4.79)	0.100	18/30	4 (77.8)/18 (40.0)	2.49 (0.80-7.78)	0.116
Smoking								
No	57/57	28 (50.9)/39 (31.6)	2.84 (1.56-5.18)	0.001	40/74	12 (70)/44 (40.5)	2.35 (1.21-4.56)	0.012
Yes	118/106	70 (40.7)/80 (24.5)	1.43 (1.01-2.02)	0.042	84/140	43 (48.8)/94 (32.9)	1.65 (1.14-2.38)	0.007
Drinking								
No	85/72	43 (49.4)/49 (31.9)	1.91 (1.26-2.90)	0.003	63/94	24 (61.9)/57 (39.4)	2.07 (1.28-3.36)	0.003
Yes	91/91	55 (39.6)/70 (23.1)	1.65 (1.13-2.39)	0.009	62/120	31 (50.0)/81 (32.5)	1.61 (1.05-2.45)	0.028
TNM stage								
II	58/56	22 (62.1)/35 (37.5)	2.10 (1.21-3.64)	0.008	42/72	10 (76.2)/41 (43.1)	3.28 (1.62-6.63)	0.001
III	118/107	76 (35.6)/84 (21.5)	1.55 (1.13-2.13)	0.007	83/142	45 (45.8)/97 (31.7)	1.54 (1.08-2.21)	0.018
Lymphadenectomy								
Two fields	147/129	78 (46.9)/91 (29.5)	1.67 (1.22-2.28)	0.001	103/173	43 (58.3)/107 (38.2)	1.82 (1.27-2.60)	0.001
Three fields	29/34	20 (31.0)/28 (17.7)	1.65 (0.85-3.20)	0.136	22/41	12 (45.5)/31 (24.4)	1.85 (0.88-3.88)	0.103
Vessel invasion								
No	139/137	76 (45.3)/98 (28.5)	1.67 (1.23-2.26)	0.001	102/174	45 (55.9)/110 (36.8)	1.82 (1.29-2.59)	0.001
Yes	37/26	22 (40.5)/21 (19.2)	2.49 (1.27-4.89)	0.008	23/40	10 (56.5)/28 (30.0)	1.83 (0.87-3.87)	0.112
Neural invasion								
No	133/107	68 (48.9)/73 (31.8)	1.81 (1.29-2.55)	0.001	92/148	35 (62.0)/86 (41.9)	1.85 (1.25-2.75)	0.002
Yes	43/56	30 (30.2)/46 (17.9)	1.65 (1.01-2.70)	0.045	33/66	20 (39.4)/52 (21.2)	1.78 (1.05-3.02)	0.032

CI, confidence interval; DFS, disease free survival; HR, hazards ratio; HRG, high-risk allele group; LRG, low-risk allele group; OS, overall survival.

[a]In DFS analysis, LRG consisted 0-4 risk alleles and HRG consisted of 5-8 risk alleles.

[b]In OS analysis, LRG included 0-2 risk alleles and HRG included 3-6 risk alleles.

[c]Adjusted by all the variables listed with the stratified factor in each stratum excluded.

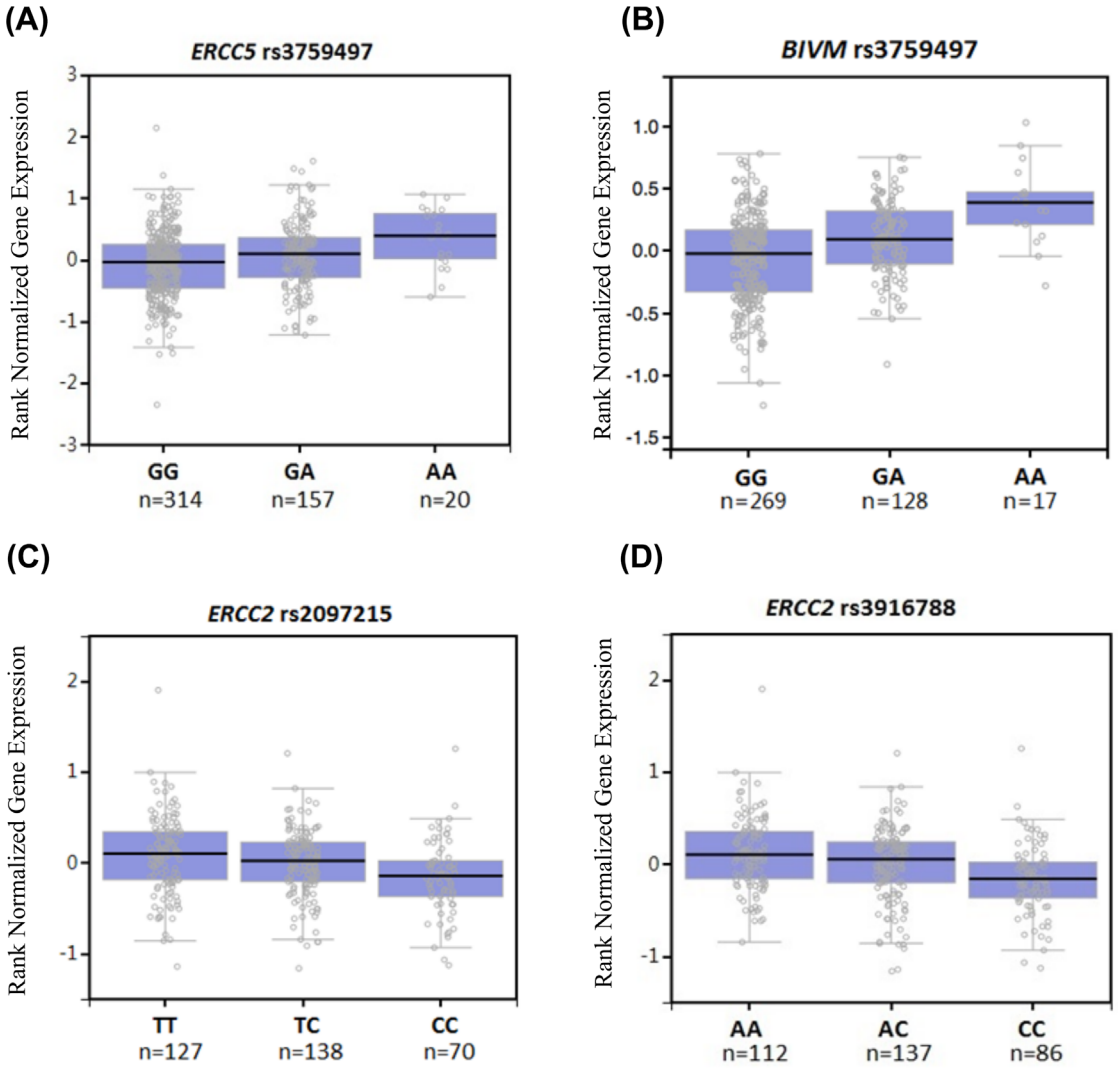

FIGURE 2 eQTL analysis of NER rSNPs from GTEx database.

after damage recognition and is a structural component of the transcription factor IIH (TFIIH) complex. Studies investigating *XPD* SNPs and cisplatin resistance have yielded varied results, among which the most well-studied Lys751Gln SNP has been associated with a decreased risk of death in patients with pancreatic, ovarian, and esophageal cancers, but a longer period of time for disease progression in patients with pancreatic and lung cancer after PBC treatment,[37,40,41] and additional conflicting results have been also reported for NSCLC.[42,43] A number of studies have reported an association of increased *XPD* expression in glioma and colon cancer cell lines with PBC resistance, [44,45] or the correlation of *XPD/ERCC1* expression with cisplatin response in NSCLC cell lines,[39] while others found no associations in other cancer cell lines.[46,47] It is possible that inter-tissue variation exists for *ERCC2* expression, which may be responsible for PBC response and survival outcome.

The polymorphic site of rs3759497 is located in the intron of *ERCC5* and is predicted to be overlapping with a transcription factor binding site (TFBS). SNPs located at TFBS are likely to affect gene expression by augmenting the binding affinity of transcription factors with DNA sequence. In our data, the rs3759497 A allele was associated with a significant up-regulation of *ERCC5* mRNA expression in muscle skeletal tissues and up-regulation of a nearby gene *BIVM* (basic immunoglobin-like variable motif containing) in various tissues (Figure 2). *BIVM* is located upstrean of *ERCC5*, the encoding products of the two can form a fusion protein BIVM-ERCC5.[48] *BIVM* resides in the human chromosome 13q32-q33 region which is linked to bipolar disorder and schizophrenia, the encoding protein is ubiquitously expressed in human tissues.[49] It is possible that rs3759497 has a modest functional effect and may be in LD with other untyped functional SNPs, which in turn may alter the function of *ERCC5* or the nearby genes.

There are several inherent limitations in the present study. First, the recruitment of patients treated in the same hospital may lead to selection bias so that caution is required when referring results to the general population. Second, under the aim of investigating rSNPs, we did not incorporate the coding region SNPs with their known effects, which may have contributed to the weakened power of our survival prediction model. Third, this is a retrospective study with a relatively

small sample size, and therefore study populations from a prospective design or other hospitals are warranted to validate our findings.

In conclusion, the present study suggested that a number of rSNPs in the NER core genes may have an effect on survival outcome of Chinese ESCC patients undergoing platinum-based adjuvant chemotherapy. This exploratory study may serve as the basis for further functional studies to unravel the molecular mechanisms underlying the observed associations, which will allow for further development of biomarkers for patients stratification and personalized management and treatment.

ACKNOWLEDGMENTS

The authors acknowledge the staff members at the tissue bank of the Fudan University Shanghai Cancer Centre for their tremendous support in providing stored blood DNA samples. This research was funded by and China's Thousand Talents Program at Fudan Univerisity, Shanghai, China and National Human Genetic Resources Sharing Service Platform (2005DKA21300).

ORCID

Qingyi Wei http://orcid.org/0000-0002-5265-3930

REFERENCES

1. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015;136:E359–E386.
2. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2006. *CA Cancer J Clin*. 2006;56:106–130.
3. Enzinger PC, Mayer RJ. Esophageal cancer. *N Engl J Med*. 2003;349:2241–2252.
4. Gebski V, Burmeister B, Smithers BM, Foo K, Zalcberg J, Simes J. Survival benefits from neoadjuvant chemoradiotherapy or chemotherapy in oesophageal carcinoma: a meta-analysis. *Lancet Oncol*. 2007;8:226–234.
5. Li S, Liu H, Diao C, et al. Prognosis of surgery combined with different adjuvant therapies in esophageal cancer treatment: a network meta-analysis. *Oncotarget*. 2017;8:36339–36353.
6. Saeed NA, Mellon EA, Meredith KL, et al. Adjuvant chemotherapy and outcomes in esophageal carcinoma. *J Gastrointest Oncol* 2017;8:816–824.
7. Kim GJ, Koshy M, Hanlon AL, et al. The benefit of chemotherapy in esophageal cancer patients with residual disease after trimodality therapy. *Am J Clin Oncol*. 2016;39:136–141.
8. Jordan P, Carmo-Fonseca M. Molecular mechanisms involved in cisplatin cytotoxicity. *Cell Mol Life Sci*. 2000;57:1229–1235.
9. Bahl A, Falk S. Meta-analysis of single agents in the chemotherapy of NSCLC: what do we want to know? *Br J Cancer*. 2001;84:1143–1145.
10. Reed E. Platinum-DNA adduct, nucleotide excision repair and platinum based anti-cancer chemotherapy. *Cancer Treat Rev*. 1998;24:331–344.
11. Wang LE, Yin M, Dong Q, et al. DNA repair capacity in peripheral lymphocytes predicts survival of patients with non-small-cell lung cancer treated with first-line platinum-based chemotherapy. *J Clin Oncol*. 2011;29:4121–4128.
12. Cleaver JE. Cancer in xeroderma pigmentosum and related disorders of DNA repair. *Nat Rev Cancer*. 2005;5:564–573.
13. Cleaver JE, Lam ET, Revet I. Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. *Nat Rev Genet*. 2009;10:756–768.
14. Ogi T, Lehmann AR. The Y-family DNA polymerase kappa (pol kappa) functions in mammalian nucleotide-excision repair. *Nat Cell Biol*. 2006;8:640–642.
15. Hoeijmakers JH. Genome maintenance mechanisms for preventing cancer. *Nature*. 2001;411:366–374.
16. Wu X, Fan W, Xu S, Zhou Y. Sensitization to the cytotoxicity of cisplatin by transfection with nucleotide excision repair gene xeroderma pigmentosun group A antisense RNA in human lung adenocarcinoma cells. *Clin Cancer Res*. 2003;9:5874–5879.
17. Bowden NA. Nucleotide excision repair: why is it not used to predict response to platinum-based chemotherapy? *Cancer Lett*. 2014;346:163–171.
18. Zhou F, Zhu M, Wang M, et al. Genetic variants of DNA repair genes predict the survival of patients with esophageal squamous cell cancer receiving platinum-based adjuvant chemotherapy. *J Transl Med*. 2016;14:154.
19. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010;26:589–595.
20. Li H, Handsaker B, Wysoker A, et al. The sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25:2078–2079.
21. Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res*. 2012;22:568–576.
22. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38:e164.
23. Klein JaM M. *Survival Analysis:Techniques for Censored and Truncated Data*. Berlin: Springer; 1997.
24. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J Natl Cancer Inst*. 2004;96:434–442.
25. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. *Biometrics*. 2005;61:92–105.
26. Huang J, Liu X, Tang LL, et al. XPG gene polymorphisms and cancer susceptibility: evidence from 47 studies. *Oncotarget*. 2017;8:37263–37277.
27. Wang C, Nie H, Li Y, et al. The study of the relation of DNA repair pathway genes SNPs and the sensitivity to radiotherapy and chemotherapy of NSCLC. *Sci Rep*. 2016;6:26526.
28. Hua RX, Zhuo ZJ, Shen GP, et al. Polymorphisms in the XPC gene and gastric cancer susceptibility in a Southern Chinese population. *Onco Targets Ther*. 2016;9:5513–5519.
29. Paszkowska-Szczur K, Scott RJ, Gorski B, et al. Polymorphisms in nucleotide excision repair genes and susceptibility to colorectal cancer in the Polish population. *Mol Biol Rep*. 2015;42:755–764.
30. Carrera-Lasfuentes P, Lanas A, Bujanda L, et al. Relevance of DNA repair gene polymorphisms to gastric cancer risk and phenotype. *Oncotarget*. 2017;8:35848–35862.
31. Wu MY, Huang SJ, Yang F, et al. Detection of nasopharyngeal carcinoma susceptibility with single nucleotide polymorphism analysis using next-generation sequencing technology. *Oncotarget*. 2017;8:52708–52723.
32. Liu J, Sun L, Xu Q, et al. Association of nucleotide excision repair pathway gene polymorphisms with gastric cancer and atrophic gastritis risks. *Oncotarget*. 2016;7:6972–6983.
33. Wang Y, Li Z, Liu N, Zhang G. Association between CCND1 and XPC polymorphisms and bladder cancer risk: a meta-analysis based on 15 case-control studies. *Tumour Biol*. 2014;35:3155–3165.
34. Zhang R, Jia M, Xue H, et al. Genetic variants in ERCC1 and XPC predict survival outcome of non-small cell lung cancer patients treated with platinum-based therapy. *Sci Rep*. 2017;7:10702.

35. Shen M, Berndt SI, Rothman N, et al. Polymorphisms in the DNA nucleotide excision repair genes and lung cancer risk in Xuan Wei, China. *Int J Cancer* 2005;116:768–773.
36. Nakajima M, Kato H. Treatment options for esophageal squamous cell carcinoma. *Expert Opin Pharmacother*. 2013;14:1345–1354.
37. Bradbury PA, Kulke MH, Heist RS, et al. Cisplatin pharmacogenetics, DNA repair polymorphisms, and esophageal cancer outcomes. *Pharmacogenet Genomics*. 2009;19:613–625.
38. Sun X, Li F, Sun N, et al. Polymorphisms in XRCC1 and XPG and response to platinum-based chemotherapy in advanced non-small cell lung cancer patients. *Lung Cancer*. 2009;65:230–236.
39. Weaver DA, Crawford EL, Warner KA, Elkhairi F, Khuder SA. Willey JC. ABCC5, ERCC2, XPA and XRCC1 transcript abundance levels correlate with cisplatin chemoresistance in non-small cell lung cancer cell lines. *Mol Cancer*. 2005;4:18.
40. Khrunin A, Ivanova F, Moisseev A, et al. Pharmacogenomics of cisplatin-based chemotherapy in ovarian cancer patients of different ethnic origins. *Pharmacogenomics*. 2012;13:171–178.
41. Avan A, Pacetti P, Reni M, et al. Prognostic factors in gemcitabine-cisplatin polychemotherapy regimens in pancreatic cancer: XPD-Lys751Gln polymorphism strikes back. *Int J Cancer*. 2013;133:1016–1022.
42. Li F, Sun X, Sun N, et al. Association between polymorphisms of ERCC1 and XPD and clinical response to platinum-based chemotherapy in advanced non-small cell lung cancer. *Am J Clin Oncol*. 2010;33:489–494.
43. Tibaldi C, Giovannetti E, Vasile E, et al. Correlation of CDA, ERCC1, and XPD polymorphisms with response and survival in gemcitabine/cisplatin-treated advanced non-small cell lung cancer patients. *Clin Cancer Res*. 2008;14:1797–1803.
44. Aloyz R, Xu ZY, Bello V, et al. Regulation of cisplatin resistance and homologous recombinational repair by the TFIIH subunit XPD. *Cancer Res*. 2002;62:5457–5462.
45. Plasencia C, Martinez-Balibrea E, Martinez-Cardus A, Quinn DI, Abad A, Neamati N. Expression analysis of genes involved in oxaliplatin response and development of oxaliplatin-resistant HT29 colon cancer cells. *Int J Oncol*. 2006;29:225–235.
46. Damia G, Guidi G, D'Incalci M. Expression of genes involved in nucleotide excision repair and sensitivity to cisplatin and melphalan in human cancer cell lines. *Eur J Cancer* 1998;34:1783–1788.
47. Dabholkar M, Bostick-Bruton F, Weber C, Bohr VA, Egwuagu C, Reed E. ERCC1 and ERCC2 expression in malignant tissues from ovarian cancer patients. *J Natl Cancer Inst*. 1992;84:1512–1517.
48. Prakash T, Sharma VK, Adati N, et al. Expression of conjoined genes: another mechanism for gene regulation in eukaryotes. *PLoS ONE*. 2010;5:e13284.
49. Christian SL, McDonough J, Cy Liu, et al. An evaluation of the assembly of an approximately 15-Mb region on human chromosome 13q32-q33 linked to bipolar disorder and schizophrenia. *Genomics*. 2002;79:635–656.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Zhang R, Zhou F, Cheng L, et al. Genetic variants in nucleotide excision repair pathway predict survival of esophageal squamous cell cancer patients receiving platinum-based chemotherapy. *Molecular Carcinogenesis*. 2018;1–13.
https://doi.org/10.1002/mc.22877