IJC
International Journal of Cancer

An *ERCC4* regulatory variant predicts grade-3 or -4 toxicities in patients with advanced non-small cell lung cancer treated by platinum-based therapy

Ruoxin Zhang[1], Ming Jia[1,2], Yuan Xu[1,2], Danwen Qian[1,2], Mengyun Wang[1], Meiling Zhu[3], Menghong Sun[1,4], Jianhua Chang[1,5] and Qingyi Wei[1,2,6]

[1] Cancer Institute, Collaborative Innovative Center for Cancer Medicine, Fudan University Shanghai Cancer Center, 270 Dong An Road, Xuhui District, Shanghai, 200032, People's Republic of China
[2] Department of Oncology, Shanghai Medical College, Fudan University Shanghai Cancer Center, 270 Dong An Road, Shanghai, 200032, People's Republic of China
[3] Department of Oncology, Xinhua Hospital affiliated to Shanghai Jiaotong University, No. 1665 Kong Jiang Road, Shanghai, 200092, People's Republic of China
[4] Department of Pathology, Fudan University Shanghai Cancer Center, 270 Dong An Road, Shanghai, 200032, People's Republic of China
[5] Department of Medical Oncology, Fudan University Shanghai Cancer Center, 270 Dong An Road, Shanghai, 200032, People's Republic of China
[6] Duke Cancer Institute, Duke University Medical Center, 10 Bryn Searle Dr., Durham, NC 27710, USA

Cancer Epidemiology

Platinum-based chemotherapy (PBC) in combination with the 3rd generation drugs is the first-line treatment for patients with advanced non-small cell lung cancer (NSCLC); however, the efficacy is severely hampered by grade 3–4 toxicities. Nucleotide excision repair (NER) pathway is the main mechanism of removing platinum-induced DNA adducts that contribute to the toxicity and outcome of PBC. We analyzed data from 710 Chinese NSCLC patients treated with PBC and assessed the associations of 25 potentially functional single nucleotide polymorphisms (SNPs) in nine NER core genes with overall, gastrointestinal and hematologic toxicities. Through a two-phase study, we found that *ERCC4* rs1799798 was significantly associated with overall and gastrointestinal toxicities [all patients: GA/AA vs. GG, odds ratio $(OR)_{adj}$=1.61 and 2.35, 95% confidence interval (CI)=1.11–2.33 and 1.25–4.41, and P_{adj}=0.012 and 0.008, respectively]. Our prediction model for the overall toxicity incorporating rs1799798 demonstrated a significant increase in the area under the curve (AUC) value, compared to that for clinical factors only (all patients: AUC = 0.61 vs. 0.59, 95% CI = 0.57–0.65 vs. 0.55–0.63, P = 0.010). Furthermore, the *ERCC4* rs1799798 A allele was associated with lower *ERCC4* mRNA expression levels according to the expression quantitative trait loci (eQTL) analysis. Our study provided some new clue in future development of biomarkers for assessing toxicity and outcomes of platinum drugs in lung cancer treatment.

Introduction

Lung cancer is the leading cause of cancer mortality worldwide, with an estimation of 1.8 million new cases and 1.59 million deaths in 2012,[1] of which non-small cell lung cancer (NSCLC) represents approximately 85% of the total incident cases.[2] Over the past decade, platinum-based chemotherapy

Key words: non-small cell lung cancer, platinum-based chemotherapy, nucleotide excision repair pathway, regulatory single nucleotide polymorphism, grade 3/4 toxicity

Abbreviations: PBC: platinum-based chemotherapy; NSCLC: non-small cell lung cancer; rSNP: regulatory single nucleotide polymorphism; NER: nucleotide excision repair; XP: xeroderma pigmentosum; ERCC: excision repair cross-complementing; ECOG: Eastern cooperative oncology group; BMI: body mass index; GI: gastrointestinal; eQTL: expression quantitative trait loci; OR: odds ratio; CI: confidence interval; ROC: receiver operating characteristic; AUC: area under the curve; UTR: untranslated region

Additional Supporting Information may be found in the online version of this article.

Conflict of interest: The authors declare that they have no conflicts of interest.

Grant sponsor: China's Thousand Talents Program, Fudan University, Shanghai, People's Republic of China

DOI: 10.1002/ijc.31153

History: Received 1 June 2017; Accepted 7 Nov 2017; Online 14 Nov 2017

Correspondence to: Jianhua Chang, Department of Medical Oncology, Fudan University Shanghai Cancer Center, 270 Dong An Road, Xuhui, Shanghai, 200032, People's Republic of China, E-mail: changjianhua@163.com; Tel: 86-21-64175590; Fax: 86-21-64175590; Qingyi Wei, Cancer Institute, Collaborative Innovation Center for Cancer Medicine, Fudan University Shanghai Cancer Center, 270 Dong An Road, Xuhui, Shanghai 200032, People's Republic of China, E-mail: weiqingyi@fudan.edu.cn; Tel: 86-21-64175590; Fax: 86-21-64174774 or Duke Cancer Institute, Duke University Medical Center, Durham, NC 27710, USA, E-mail: qingyi.wei@duke.edu; Tel: 919-660-0652; Fax: 919-660-0708

What's new?

The toxicity of platinum-based chemotherapy (PBC) causes unpredictable adverse effects among non-small cell lung cancer (NSCLC) patients. DNA repair capacity, particularly the nucleotide excision repair (NER), has been associated with PBC efficacy and toxicity. After assessing 25 potentially functional regulatory single nucleotide polymorphisms (SNPs) in nine core NER genes in 710 Chinese NSCLC patients, the authors found a novel *ERCC4* variant that was significantly associated with overall and gastrointestinal toxicities induced by PBC. This study provides clues that may help to unravel the molecular mechanisms underlying PBC-induced toxicity and develop biomarkers for predicting toxicity in patients with advanced NSCLC.

(PBC) has continued to be the standard care for patients with advanced-stage NSCLC,[3] with a modest response rate of 30–50% [4,5] and a 5-year survival of <15%.[6]

The anti-neoplastic effect of PBC is through the formation of inter- or intra-strand DNA adducts that result in distortion and lesion of the DNA,[7] which further activate apoptotic pathways that lead to cell death.[8] The cytotoxic profile of PBC may cause unpredictable and occasionally severe adverse effects that impair the therapy adherence and quality of life of the patients. Clinical studies have shown marked differences in the incidence and severity of toxicities among patients treated with PBC.[9,10] Therefore, it is urgent to identify predictive and prognostic markers for patients who undergo PBC. Besides the established clinical-related risk factors for the toxicity, such as elderly age, poor performance status and histologic stage, single nucleotide polymorphisms (SNPs) in DNA repair pathways may also contribute to the observed variation in response to PBC.[11,12]

An individual's DNA repair capacity plays a vital role in maintaining genomic stability by three major mechanisms: nucleotide excision repair (NER), base excision repair and mismatch repair. NER is mainly responsible for the repair of platinum-DNA adducts, including all three types of DNA intrastrand crosslinks (1,2-d(ApG), 1,2-d(GpG), and 1,3-d(GpNpG)).[11] NER uses a cut-and-paste method in a sequential order involving the collaborative functions of 20–30 proteins.[13] The binding of the XPC/HHR23B complex to the entire repair protein apparatus signifies the initial step of damage recognition, followed by opening the DNA helix around the lesion, being orchestrated by XPA[14] and employing helicases of the transcription factor II H (TFIIH) complex (including ERCC3 and ERCC2 helicases).[15] Subsequently, the dual incision of the lesion takes place with the facilitation of the ERCC5 and ERCC1/ERCC4 complex. Finally, the gap-filling repair synthesis is collaboratively performed with replication protein A, replication factor C, proliferating cell nuclear antigen and DNA polymerase δ and ε.[16] Theoretically, suboptimal DNA repair capacity may either increase individual's positive response to chemotherapy by promoting cancer cell death (i.e., the treatment effect) or lead to an increased toxicity because of inefficient removal of DNA lesions in normal bystander cells exposed to the treatment, which eventually results in a high rate of cell death in normal tissue as well (i.e., the side-effects). Therefore, it is conceivable that genetic variation in NER pathway genes may serve as a prognostic factor for PBC-induced toxicity in NSCLC patients.

In the present study, we assessed predictive and prognostic effects of the selected common genetic variants on PBC-induced toxicity by analyzing 25 potentially regulatory SNPs (rSNPs) in nine core NER genes (*ERCC1, ERCC2, ERCC3, ERCC4, ERCC5, RAD23B, XPA, XPC* and *XPE*) in 710 patients with advanced NSCLC. All these 25 rSNPs are predicted by bioinformatics tools to be potentially functional in regulating gene expression (Supporting Information, Table S1).

Materials and Methods

Study population

The present study included patients diagnosed with histologically advanced NSCLC from Fudan University Shanghai Cancer Center (FUSCC). The discovery group included 355 patients recruited between February 2009 and February 2012, whereas the replication group consisted of 355 patients recruited from March 2012 to November 2013. Further details about the recruitment have been described elsewhere.[17] In brief, the recruited patients were unrelated Han Chinese with inoperable TNM stages III to IV tumors of NSCLC without prior history of cancer other than *in situ* carcinoma; received PBC as the first-line treatment; the Eastern Cooperative Oncology Group performance (ECOG) status 0 to 2 with normal blood and uronoscopy laboratory test results; no active infection, serious medical or psychological conditions that might affect adverse events; and at least two chemotherapy cycles except for those who developed severe toxicity after one cycle. Clinical data, including age at treatment, sex, smoking status, ECOG performance, TNM stage, histological type and grade, chemotherapy regimen options and radiotherapy treatment, were collected from patients' medical records. A written informed consent was obtained from each participant, and their blood samples were collected and stored in the FUSCC Tissue Bank in the hospital. This research protocol was approved by the Institutional Review Board of FUSCC, and all the methods were implemented in accordance with relevant guidelines and regulations.

Toxicity data

Patients' toxicity profiles were evaluated through examination of their laboratory test results, when toxicity events occurred.

The chemotherapy-related toxicities were assessed twice a week until 1 week after the completion of chemotherapy, including overall, gastrointestinal (GI) (including nausea, vomiting and diarrhea) and hematologic toxicity (including leukocytopenia, agranulocytosis, anemia and thrombocytopenia). The grade of toxicity was recorded according to the National Cancer Institute Common Toxicity Criteria, version 3.0.[18] However, due to the limited information in the medical reports, details about the grade 3 or 4 classified GI toxicities and the occurring time of the toxicity events were not available.

Chemotherapy regime

All patients included in the present study were having inoperable late-stage tumors and thus had received first-line platinum-based chemotherapy. The drugs paired with platinum formed 10 combinations: cisplatin (75 mg/m^2) or carboplatin (area under the curve [AUC] 6 mg/ml·min), administered with paclitaxel (175 mg/m^2) on day one every 3 weeks, docetaxel (75 mg/m$^{2)}$ on day one every 3 weeks, gemcitabine (1250 mg/m^2) on day one and eight every 3 weeks, pemetrexed (500 mg/m^2) on day one every 3 weeks or vinorelbine (25 mg/m^2) on day one and eight every 3 weeks, and cisplatin (100 mg/m^2) or carboplatin (AUC 6 mg/ml·min) administered on day one every 4 weeks, in combination with etoposide (100 mg/m^2) on days one to three every 4 weeks. All chemotherapeutic drugs were administered intravenously.

SNP selection

To specifically explore the association between rSNPs in the NER core genes and toxicity outcome of NSCLC patients subjected to PBC, all the rSNPs were queried from the NER gene regions under the study by using SNP/GeneView in dbSNP database (http://www.ncbi.nlm.nih.gov/snp/) with the GRCh38 reference build of the human genome. A total of 25 rSNPs in eight (out of nine, and none found in *DDB1/XPE*) NER core genes were chosen, with detailed characteristics of all investigated genes and rSNPs shown in Supporting Information, Table S1. The selection was based on the following criteria: minor allele frequency (MAF) $\geq$5% in Han Chinese, in the regulatory region (5' near gene, 5'UTR, intron, 3'near gene or 3'UTR), in low linkage disequilibrium (LD) with each other (r^2<0.8), and having predicted functions [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/). A full list of the NER genes analyzed in our study, their region coordinates, their start and stop sites, and the characteristics of genotyped variants are also summarized in Supporting Information, Table S1.

SNPseq genotyping

Genomic DNA was extracted from the whole blood of all study subjects by using DNA Blood Mini Kit (Qiagen, Valencia, CA). The purity [optical density $(OD)_{260/280}$ at 1.7–2.0] and concentration (>20 ng/μl) met the sequencing requirements. Genotyping of all rSNPs was conducted by FastTarget, a next generation sequencing-based method using Illumina Miseq2000 Platform (2 × 250 bp, Illumina, CA, USA). 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). After multiple PCR reactions, DNA fragments were ligated with the adaptor by using Q5 DNA polymerase Kit (New England Biolabs, MA, USA), and further purified by Agencourt AMPure XP (Beckman Coulter, CA, USA).

Next-generation sequencing of the amplification products was carried out by MiSeq 2000 Sequencer (Illumina, Inc., San Diego, CA, USA), following the manufacturer's standard protocols. Output sequence data were trimmed and then compared to fragment reference sequences (hg19) using the Blat program. 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 were done 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]

The eQTL analysis

GTEx portal was used to assess the correlation between toxicity-related genetic variants and NER gene 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, with sufficient power to detect eQTL in 44 tissues. The associations of genetic variant and gene expression are investigated in various tissue types.

Statistical analysis

We used the χ^2 test for comparisons between individuals with severe toxicity (NCI-CTC grades 3–4) and those with mild or no toxicity (NCI-CTC grades 0–2) as well as for adverse events that were dichotomized by presence or absence of (a) any grade 3–4 toxicity, and (b) any grade 3–4 GI or hematologic toxicities. The association between each genetic variant and severe chemotherapy toxicity was assessed by odds ratios (ORs) and their 95% confidence intervals (CIs) estimated by unconditional logistic regression models with or without adjustment for age at treatment, sex, smoking status, TNM stage, BMI, histological type, histologic grade, ECOG performance status, chemotherapy regimens and palliative radiotherapy. For toxicity prediction model construction, independent predictors including selected clinical variables and genetic variant were selected by the backward-stepwise logistic regression process. The receiver

Table 1. Clinical characteristics and their associations with overall, gastrointestinal and hematologic toxicities in Chinese NSCLC patients

	Overall toxicity			Hematologic toxicity			Gastrointestinal toxicity		
Patient characteristics	Events/No.[3]	OR (95% CI)	p^2	Events/No.	OR (95% CI)	p^2	Events/No.	OR (95% CI)	p^2
Median age range									
≤58	120/385	ref.		100/385	ref.		27/385	ref.	
>58	125/325	1.38 (1.01–1.88)	**0.042**	110/325	1.46 (1.06–2.02)	**0.022**	23/325	1.01 (0.57–1.80)	0.974
Sex									
Male	170/508	ref.		150/508	ref.		31/508	ref.	
Female	75/202	1.17 (0.84–1.65)	0.354	60/202	1.01 (0.71–1.44)	0.963	19/202	1.60 (0.88–2.90)	0.123
Performance status									
0–1	150/407	ref.		131/407	ref.		30/407	ref.	
2	95/303	0.78 (0.57–1.07)	0.128	79/303	0.74 (0.53–1.03)	0.078	20/303	0.89 (0.49–1.60)	0.692
BMI									
≤22	97/300	ref.		88/300	ref.		15/300	ref.	
>22	139/384	1.19 (0.86–1.63)	0.292	115/384	1.04 (0.75–1.44)	0.815	33/384	1.71 (0.93–3.13)	0.083
Missing	9/26			7/26			2/26		
Smoking status									
Nonsmokers	119/334	ref.		95/334	ref.		28/334	ref.	
Former-smokers	13/41	0.84 (0.42–1.68)	0.620	13/41	1.17 (0.58–2.35)	0.663	0/41	N/A	0.974
Current-smokers	113/335	0.92 (0.67–1.27)	0.606	102/335	1.10 (0.79–1.54)	0.570	22/335	0.77 (0.43–1.37)	0.373
TNM stage									
III	95/288	ref.		89/288	ref.		10/288	ref.	
IV	150/422	1.12 (0.82–1.54)	0.481	121/422	0.90 (0.65–1.25)	0.523	40/422	2.91 (1.43–5.92)	**0.003**
Histological type									
Squamous cell carcinoma	56/138	ref.		52/138	ref.		6/138	ref.	
Adenocarcinoma	153/478	0.69 (0.47–1.02)	0.062	126/478	0.59 (0.40–0.88)	**0.010**	38/478	1.90 (0.79–4.59)	0.154
Other[1]	27/65	1.04 (0.57–1.89)	0.897	25/65	1.03 (0.56–1.90)	0.915	3/65	1.07 (0.26–4.40)	0.931
NSCLC-NOS	9/29	0.66 (0.28–1.55)	0.340	7/29	0.53 (0.21–1.32)	0.170	3/29	2.54 (0.60–10.80)	0.207
Histologic grade									
Poorly differentiated	151/415	ref.		133/415	ref.		26/415	ref.	
Well-moderately differentiated	94/295	0.82 (0.60–1.12)	0.212	77/295	0.75 (0.54–1.04)	0.087	24/295	1.33 (0.75–2.36)	0.338
Chemotherapy regimen 1									
Docetaxel/paclitaxel	97/237	ref.		94/237	ref.		7/237	ref.	
Etoposide	2/7	0.58 (0.11–3.04)	0.517	2/7	0.61 (0.12–3.20)	0.558	1/7	5.48 (0.58–51.78)	0.138
Gemcitabine	54/132	1.00 (0.65–1.54)	0.997	40/132	0.66 (0.42–1.04)	0.074	17/132	4.86 (1.96–12.05)	**<0.001**

Table 1. Clinical characteristics and their associations with overall, gastrointestinal and hematologic toxicities in Chinese NSCLC patients (Continued)

	Overall toxicity			Hematologic toxicity			Gastrointestinal toxicity		
Patient characteristics	**Events/No.[3]**	**OR (95% CI)**	**p^2**	**Events/No.**	**OR (95% CI)**	**p^2**	**Events/No.**	**OR (95% CI)**	**p^2**
Vinorelbine	11/34	0.69 (0.32–1.48)	0.342	10/34	0.63 (0.29–1.39)	0.253	1/34	1.00 (0.12–8.35)	0.997
Pemetrexed	81/300	0.53 (0.37–0.77)	**0.001**	64/300	0.41 (0.28–0.60)	**<0.001**	24/300	2.86 (1.21–6.75)	**0.017**
Chemotherapy regimen 2									
Cisplatin-combinations	179/573	ref.		150/573	ref.		41/573	ref.	
Carboplatin-combinations	66/137	2.05 (1.40–2.99)	**<0.001**	60/137	2.20 (1.50–3.23)	**<0.001**	9/137	0.91 (0.43–1.93)	0.810
Radiotherapy									
No	170/491	ref.		141/491	ref.		38/491	ref.	
Yes	75/219	0.98 (0.70–1.38)	0.922	69/219	1.14 (0.81–1.61)	0.452	12/219	0.69 (0.35–1.35)	0.280

[1]Other carcinomas include adenosquamous carcinoma, mixed cell, neuroendocrine carcinoma, and undifferentiated carcinoma.
[2]p value was calculated by χ^2 test under unconditional logistic regression analysis.
[3]Events refers to the number of toxicity incidents in each subgroup, whereas No. means the total number of patients in each subgroup; The results are in **bold** if $p < 0.05$. Abbreviations: OR-odds ratio; BMI-body mass index; TNM-tumor, lymph nodes, metastasis; NSCLC-NOS-non-small cell lung cancer not otherwise specified; chemotherapy regimen 1- types of non-platinum drug; chemotherapy regimen 2- types of platinum drug.

operating characteristic curve (ROC) analysis with the area under the curve (AUC) value was used to compare sensitivity and specificity of the toxicity prediction by the selected parameters. All statistical analyses were performed using the SAS software, version 9.4 (SAS Institute, Inc., Cary, NC). Unless stated otherwise, all p values were two-sided with a significance level of $p < 0.05$.

Results

Characteristics of the study population

The analysis included 710 NSCLC patients treated with PBC as the first-line chemotherapy. The demographic, clinical characteristics and main toxicity events are summarized in Table 1. Of all the patients, 508 were males and 202 were females, with a median age of 58 (a range of 23–83) years. The numbers of never, former and current smokers were 334 (47%), 41 (5.8%) and 335 (47.2%), respectively. All patients had an advanced TNM stage (III or IV) cancer, with 478 (67.3%) being adenocarcinoma and 138 (19.4%) being squamous cell carcinoma. Among different chemotherapy combinations, 237 (33.4%) patients received platinum-docetaxel/paclitaxel, while 300 (42.3%) received platinum-pemetrexed treatment. Of all the patients, 219 (30.8%) of the patients received palliative radiotherapy.

Clinical characteristics of NSCLC patients and chemotherapy-induced toxicities

To assess whether clinical variables have contributed to chemotherapy-induced toxicities (i.e. overall, hematologic and GI toxicities), we conducted univariate logistic regression analysis of the clinical characteristics against those toxicity outcomes (Table 1). We found that elderly people (>58 years) had a higher risk of overall and hematologic toxicities (OR = 1.38 and 1.46, 95% CI = 1.01–1.88 and 1.06–2.02, $p = 0.042$ and 0.022, respectively), compared to patients <58. The chemotherapy regimen containing carboplatin was associated with a significantly higher risk of overall and hematologic toxicities (OR = 2.05 and 2.20, 95% CI = 1.40–2.99 and 1.50–3.23; both $p < 0.001$, respectively), compared to the cisplatin-based chemotherapy. Compared to patients having a TNM stage III disease, patients with a stage IV disease exhibited a higher risk of GI toxicity (OR = 2.91, 95% CI = 1.43–5.92, and $p = 0.003$). Platinum-pemetrexed was the predominant chemotherapy used to treat patients with adenocarcinoma. Despite its opposing toxic effects on GI and hematologic systems, an overall protective effect of this treatment was observed (OR = 0.53, 95% CI = 0.37–0.77, $p = 0.001$), compared to the docetaxel/paclitaxel regimes. Compared to patients with squamous cell carcinoma, adenocarcinoma patients tended to have a lower risk of hematologic toxicity (OR = 0.59, 95% CI = 0.40–0.88, $p = 0.010$). Patients with or without radiotherapy treatment has shown no significant difference in any type of toxicity outcome. However, given the established toxic effect of radiotherapy in cancer patients, it was still included as a covariate in further

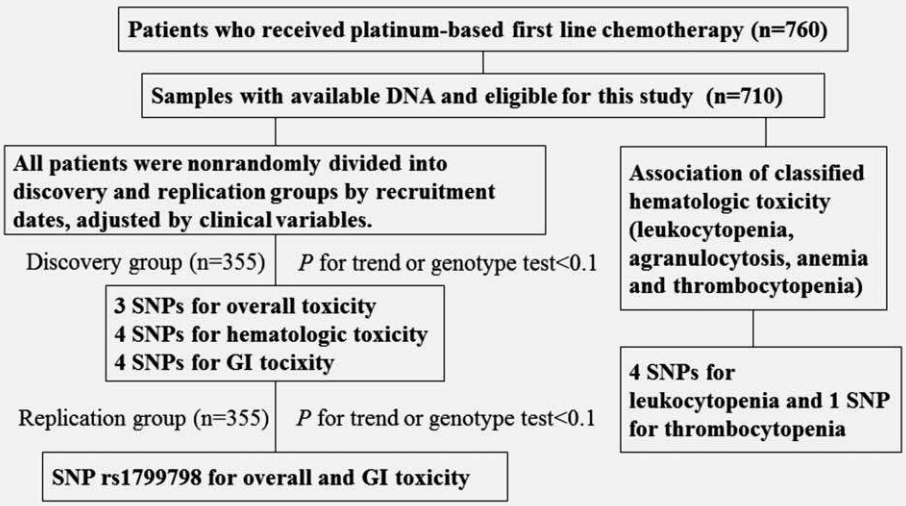

Figure 1. The two-phase screening of single nucleotide polymorphisms (SNPs) associated with grade 3/4 toxicities of platinum-based chemotherapy.

multivariate logistic regression analysis. To control for any possible confounding factors on the main effects associated with the rSNPs on the PBC toxicity, all the clinical variables, including age at treatment, sex, ECOG status, BMI, smoking status, TNM stage, histological type, histologic grade, chemotherapy regimens and palliative radiotherapy, were further adjusted for in the subsequent multivariate logistic regression analysis.

NER rSNPs and grade-3 or −4 toxicities

We performed a two-stage screening to evaluate the association of rSNPs in the eight NER core genes with grade-3 or −4 toxicities. In the discovery group, we performed multivariate logistic regression analysis to assess the associations of these rSNPs with toxicities. We found that three SNPs were associated with overall toxicity, four SNPs were associated with hematologic toxicity, and four SNPs associated with GI toxicity ($p < 0.1$ by the trend test or genotype test, Figure 1, Supporting Information, Table S2–4). The positive associations in the discovery group were further validated in the replication group using the same significance threshold criteria. Out of all the SNPs that passed the significance threshold in the discovery phase, only *ERCC4* rs1799798 SNP remained to be significantly associated with overall and GI toxicity. Patients with *ERCC5* rs751402 AG and AG/AA genotypes exhibited a significantly higher risk of hematologic toxicity in the discovery group, when compared to the GG genotype, but not in the validation set (Fig. 1, Table 2, Supporting Information, Table S2–4). When we combined the two datasets and increased the statistical threshold in multivariate logistic regression analysis of all 710 patients with adjustment for clinical variables, both SNPs (*ERCC4* rs1799798 and *ERCC5* rs751402) exhibited significant associations (Fig. 1, Table 2, Supporting Information, Table S2–4). *ERCC4* rs1799798 GA and GA/AA genotypes were associated with an increased risk of overall toxicities in all patients (GA and GA/AA vs. GG, OR_{adj}= 1.65 and 1.61, 95% CI= 1.13–2.43 and 1.11–2.33, P_{adj}= 0.010 and 0.012, respectively), suggesting a dominant effect of the A allele, which was also evident in the effect on severe GI toxicity (GA and GA/AA vs. GG, OR_{adj}=2.45 and 2.35, 95% CI= 1.29–4.68 and 1.25–4.41, P_{adj}= 0.006 and 0.008, respectively). It appears that based on the sample size of the present study, only *ERCC4* rs1799798 was consistently associated with GI and overall toxicity in all the datasets.

NER variant-based prediction model for PBC-induced toxicity in NSCLC patients

The cumulative effects of toxicity-associated rSNPs and respective clinical risk factors were further examined by comparing the area under the curve (AUC) of the two. To assess the discriminative accuracy of the prediction models, we calculated AUC of the ROC curves. All the variables were selected using a backward-stepwise logistic regression process for inclusion in the prediction model. The selected variables were carboplatin-based therapy, pemetrexed and *ERCC4* rs1799798 GA/AA for overall toxicity; and TNM, BMI, gemcitabine, pemetrexed and *ERCC4* rs1799798 GA/AA for GI toxicity (Supporting Information, Table S5). In the prediction model of overall toxicity, the addition of *ERCC4* rs1799798 GA/AA significantly increased the AUC to 0.61 (95% CI= 0.57–0.65) in all patients, compared to the model that only included the clinical variables (AUC= 0.59; 95% CI= 0.55–0.63; P_{ChiSq}= 0.010) (Fig. 2*a*, Supporting Information, Table S6). However, the ROC prediction model incorporating the risk genotype for overall toxicity was not significantly different from that incorporating only clinical factor in the discovery set. For GI toxicity prediction, the AUC for the model with the addition of *ERCC4* rs1799798 GA/AA increased slightly, compared to that of the model with clinical variables

Table 2. Association of NER SNPs with grade 3/4 overall, gastrointestinal and hematologic toxicity (leukocytopenia and thrombocytopenia) in NSCLC patients treated with platinum-based chemotherapy in discovery, replication and combined datasets

		Discovery group (n = 355)			Replication group (n = 355)			Combined (n = 710)		
Grade 3/4 Toxicity	Gene/SNPs	Events/No.	Adjusted OR (95% CI)	p^1	Events/No.	Adjusted OR (95% CI)	p^1	Events/No.	Adjusted OR (95% CI)	P^1
Overall toxicity	*ERCC4* rs1799798			0.107[2]			**0.048**[2]			**0.025**[2]
	GG	97/274	1.00 (ref.)		78/273	1.00 (ref.)		175/547	1.00 (ref.)	
	GA	36/76	1.91 (1.11–3.30)	**0.020**	29/73	1.69 (0.94–3.01)	**0.078**	65/149	1.65 (1.13–2.43)	**0.010**
	AA	1/5	0.42 (0.04–3.92)	0.445	4/9	2.26 (0.51–9.92)	0.280	5/14	1.19 (0.38–3.71)	0.759
	Dominant	37/81	1.75 (1.03–2.98)	**0.040**	33/82	1.74 (1.00–3.03)	**0.050**	70/163	1.61 (1.11–2.33)	**0.012**
	Recessive	133/350	0.36 (0.04–3.37)	0.371	107/346	2.03 (0.46–8.91)	0.348	240/696	1.07 (0.35–3.32)	0.906
Gastrointestinal	*ERCC4* rs1799798			**0.030**[2]			**0.082**[2]			**0.019**[2]
toxicity	GG	20/274	1.00 (ref.)		11/273	1.00 (ref.)		31/547	1.00 (ref.)	
	GA	12/76	2.96 (1.26–6.92)	**0.013**	6/73	2.52 (0.82–7.78)	0.107	18/149	2.45 (1.29–4.68)	**0.006**
	AA	0/5	—	0.989	1/9	3.2 (0.33–31.32)	0.318	1/14	1.31 (0.15–11.03)	0.807
	Dominant	12/81	2.82 (1.21–6.58)	**0.017**	7/82	2.61 (0.89–7.61)	**0.080**	19/163	2.35 (1.25–4.41)	**0.008**
	Recessive	32/350	—	0.989	17/346	2.46 (0.26–23.24)	0.431	49/696	1.02 (0.12–8.47)	0.988
Hematologic	*ERCC5/XPG* rs751402			0.184[2]			0.463[2]			0.132[2]
toxicity	GG	39/149	1.00 (ref.)		39/149	1.00 (ref.)		73/286	1.00 (ref.)	
	AG	60/156	1.90 (1.12–3.20)	**0.017**	60/156	1.25 (0.73–2.14)	0.418	114/340	1.52 (1.06–2.17)	**0.024**
	AA	14/50	1.23 (0.58–2.62)	0.590	14/50	1.26 (0.50–3.20)	0.623	23/83	1.23 (0.70–2.18)	0.477
	Dominant	74/206	1.71 (1.04–2.81)	**0.034**	74/206	1.24 (0.74–2.10)	0.419	137/423	1.45 (1.02–2.05)	**0.036**
	Recessive	99/305	0.88 (0.44–1.77)	0.720	99/305	1.10 (0.46–2.63)	0.828	187/626	0.97 (0.57–1.64)	0.904

[1]p value of χ^2 test, adjusted by age at diagnose, sex, smoking status, BMI, TNM status, ECOG performance status, chemotherapy regimens, histological type, histologic grade and palliative radiotherapy.
[2]Ptrend: p values for trend tests. In the discovery and replication groups, the results are in **bold** if $p < 0.1$; in the combined group, the results are in **bold** if $p < 0.05$.
Abbreviations: CI, confidence interval; OR, odds ratio.

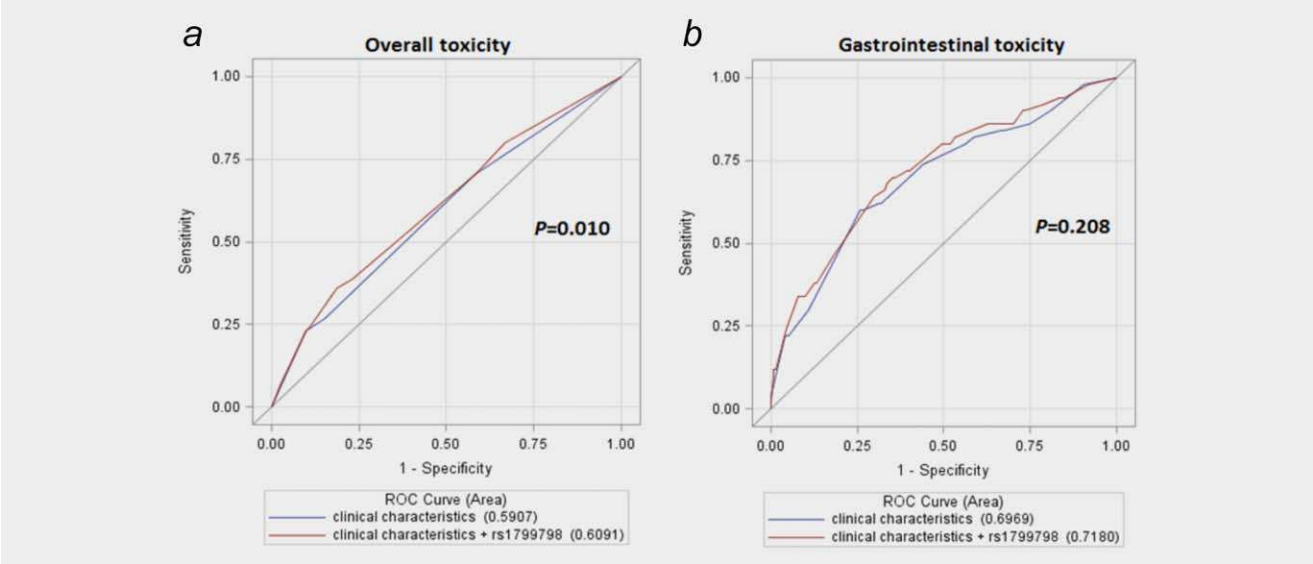

Figure 2. Receiver operating characteristics (ROC) curves and area under the curve (AUC) plots with or without risk SNPs for the prediction of overall, gastrointestinal and hematologic toxicities. (a) ROC and AUC plots for overall toxicity prediction based on risk genotypes (ERCC4 rs1799798 GA/AA) and clinical risk factors (platinum-pemetrexed and carboplatin combinations) or clinical risk factors alone; (b) ROC and AUC plots for gastrointestinal toxicity prediction based on risk genotypes (ERCC4 rs1799798 GA/AA) and clinical risk factors (BMI, TNM stage, platinum-gemcitabine and platinum-pemetrexed) or clinical risk factors alone.

only (AUC= 0.72 vs. 0.70; 95% CI= 0.64–0.80 vs. 0.62–0.78; P_{ChiSq}= 0.208) (Fig. 2*b*). The AUC values of ROC curves for the overall and GI toxicity in all datasets are listed in Supporting Information, Table S6. Given the modest AUC values in general and the limited sample size, these findings indicated that *ERCC4* rs1799798 might independently contribute to the partial prediction power of overall toxicity in PBC-treated NSCLC patients.

Stratified analysis for associations between the risk genotypes and PBC toxicity

Stratified analysis was also performed to assess potential effects of different demographic or clinical variables on the toxicity outcomes associated with risk genotypes of *ERCC4*. Overall, the *ERCC4* rs1799798 GA/AA carriers exhibited an increased risk of overall toxicity in subgroups of elderly patients (>58 years), non-smokers, poorly differentiated tumors, adenocarcinoma, recipients of gemcitabine, pemetrexed, and cisplatin-based chemotherapies, compared to the GG carriers (Table 3). The carriers of the risk A allele had an increased risk of GI toxicity in subgroups of older, non-smokers, ECOG status 2, well-moderately differentiated, adenocarcinoma, no radiotherapy, and recipients of pemetrexed (Table 3).

NER rSNPs and grade-3 or −4 hematologic toxicity

The associations of NER rSNPs with leukocytopenia, agranulocytosis, anemia and thrombocytopenia were further evaluated by univariate and multivariate logistic regression analyses. Under a dominant model, carriers of the *ERCC1* rs3212924 G, rs3212986 A alleles and *RAD23* rs7041137 T allele exhibited a significantly increased risk of leukocytopenia (*ERCC1*: OR_{adj}=1.77 and 1.63, 95% CI = 1.13–2.76 and 1.03–2.57, P_{adj}=0.013 and 0.036, respectively; *RAD23*: OR_{adj}=2.08, 95% CI = 1.34–3.24, P_{adj}=0.001), while carriers of the *ERCC3* rs13385611 C allele exhibited a similar association under a recessive model (OR_{adj}=5.64, 95% CI = 1.68–18.96, P_{adj}=0.005) (Table 4). With a relatively small number of incidents, a borderline protective effect of thrombocytopenia was found in patients carrying *ERCC5* rs873601 GA and GA/GG genotypes (OR_{adj}=0.35 and 0.39, 95% CI = 0.13–0.91 and 0.17–0.91, P_{adj}=0.031 and 0.029, respectively), compared to those carrying the AA genotype. No significant association was found between these genetic variants and anemia or agranulocytosis.

Correlations between *ERCC4* risk genotypes and mRNA expression levels

To examine the genotypic effect of toxicity-associated risk rSNPs on the gene expression, the eQTL analysis of the *ERCC4* risk SNP rs1799798 was further performed by using publically available database. In the GTEx samples of transformed fibroblast cells, we found that the *ERCC4* rs1799798 A allele was significantly associated with a lower mRNA level of *ERCC4* ($p = 0.81 \times 10^{-8}$, effect size = −0.35, Fig. 3*a*). In addition, significant difference in gene expression levels regulated by this SNP was also found in tissues of esophagus (gastroesophageal junction), thyroid, colon and brain ($p = 0.0013$, effect size = −0.40, Fig. 3*b*). Taken together, the observed association between *ERCC4* rs1799798 and overall and GI toxicity may be explained by the difference in gene expression levels regulated by these two alleles. Therefore, it

Table 3. Stratified multivariate analysis of clinical variables for associations between NER risk SNPs and overall and gastrointestinal toxicities in NSCLC patients

	ERCC4 rs1799798 (Event/N)		Overall toxicity		*ERCC4* rs1799798 (Event/N)		Gastrointestinal toxicity	
Variables	**GG**	**GA/AA**	**Adjusted OR (95% CI)**	**p^1**	**GG**	**GA/AA**	**Adjusted OR (95% CI)**	**p^1**
Median age range, yr								
≤58	87/290	33/95	1.12 (0.66–1.89)	0.672	18/290	9/95	1.54 (0.63–3.77)	0.343
>58	88/257	37/68	2.49 (1.40–4.44)	**0.002**	13/257	10/68	4.42 (1.60–12.23)	**0.004**
Sex								
Male	126/398	44/110	1.42 (0.90–2.25)	0.134	20/398	11/110	1.84 (0.79–4.26)	0.157
Female	49/149	26/53	1.78 (0.88–3.58)	0.107	11/149	8/53	2.61 (0.87–7.85)	0.088
Smoking status								
Non-smokers	80/250	39/84	1.99 (1.16–3.41)	**0.012**	17/250	11/84	2.83 (1.15–7.01)	**0.024**
Current smokers	85/261	28/74	1.26 (0.70–2.26)	0.436	14/261	8/74	1.83 (0.65–5.13)	0.251
TNM								
III	65/222	30/66	1.70 (0.92–3.15)	0.093	5/222	5/66	2.67 (0.66–10.75)	0.168
IV	110/325	40/97	1.47 (0.89–2.43)	0.131	26/325	14/97	2.17 (1.00–4.70)	0.051
ECOG performance status								
0–1	108/314	42/93	1.57 (0.96–2.59)	0.076	20/314	10/93	1.94 (0.81–4.64)	0.138
2	67/233	28/70	1.63 (0.89–3.01)	0.115	11/233	9/70	2.87 (1.01–8.17)	**0.049**
Chemotherapy regimen 1								
Platinum-docetaxel/paclitaxel	70/180	27/57	1.05 (0.54–2.05)	0.879	4/180	3/57	2.20 (0.42–11.68)	0.354
Platinum-gemcitabine	37/103	17/29	2.80 (1.09–7.21)	**0.033**	11/103	6/29	2.54 (0.57–11.39)	0.224
Platinum-pemetrexed	56/231	25/69	1.86 (1.02–3.39)	**0.042**	14/231	10/69	2.85 (1.14–7.12)	**0.025**
Chemotherapy regimen 2								
Cisplatin-combinations	127/444	52/129	1.76 (1.16–2.67)	**0.008**	27/444	14/129	1.99 (0.97–4.08)	0.059
Carboplatin-combinations	48/103	18/34	0.90 (0.37–2.19)	0.812	4/103	5/34	7.41 (0.92–60.10)	0.061
Histologic grade								
Poorly differentiated	106/318	45/97	1.77 (1.09–2.90)	**0.022**	17/318	9/97	1.91 (0.76–4.80)	0.169
Well-moderately differentiated	69/229	25/66	1.31 (0.69–2.48)	0.414	14/229	10/66	2.80 (1.11–7.07)	**0.030**
Histological type								
Squamous cell carcinoma	44/112	12/26	1.28 (0.49–3.32)	0.613	4/112	2/26	1.64 (0.12–21.85)	0.707
Adenocarcinoma	103/365	50/113	2.05 (1.30–3.23)	**0.002**	22/365	16/113	3.00 (1.46–6.20)	**0.003**
Others[2]	21/49	6/16	0.56 (0.12–2.51)	0.447	2/49	1/16	NA	NA
Palliative radiotherapy								
No	123/378	47/113	1.43 (0.91–2.27)	0.125	23/378	15/113	2.62 (1.23–5.59)	**0.013**
Yes	52/169	23/50	1.63 (0.79–3.37)	0.189	8/169	4/50	1.69 (0.38–7.43)	0.490
TKI treatment								
No	113/356	42/97	1.61 (0.99–2.61)	0.057	16/356	10/97	2.27 (0.90–5.71)	0.081
Yes	62/191	28/66	1.74 (0.92–3.30)	0.087	15/191	9/66	2.26 (0.86–5.94)	0.100

[1]Adjusted by age at diagnose, sex, smoking status, BMI, TNM status, ECOG performance status, chemotherapy regimens, histological type, histologic grade, palliative radiotherapy (the stratified factor in each stratum excluded).
[2]Others include adenosquamous, mixed cell, neuroendocrine, and undifferentiated carcinoma. The results are in **bold** if $p < 0.05$. Abbreviations: SNP-single nucleotide polymorphism; OR-odds ratio; chemotherapy regimen 1- types of non-platinum drug; chemotherapy regimen 2- types of platinum drug.

can be speculated that the reduction in *ERCC4* expression levels could be associated with a higher toxicity and thus an unfavorable clinical outcome.

Discussion

PBC in conjunction with other cytotoxic agents has been used as the first-line therapy for advanced NSCLC,

Table 4. Association of NER rSNPs with hematologic toxicity (leukocytopenia and thrombocytopenia) in NSCLC patients

Gene/SNPs	Genotype	Events/No.	Unadjusted OR (95%CI)	*p*	Adjusted OR[1] (95% CI)	*p*[1]
Leukocytopenia (Events = 104)						
***ERCC1*/rs3212924**	AA	40/342	1.00 (ref.)		1.00 (ref.)	
	AG	56/305	**1.71 (1.10–2.65)**	**0.017**	**1.91 (1.23–2.98)**	**0.004**
	GG	8/61	1.15 (0.510 2.59)	0.741	1.09 (0.47–2.56)	0.839
	Dominant	64/366	**1.60 (1.05– 2.45)**	**0.031**	**1.77 (1.13–2.76)**	**0.013**
	Recessive	96/647	0.87 (0.40–1.88)	0.717	0.78 (0.35–1.75)	0.549
***ERCC1*/rs3212986**	CC	36/306	1.00 (ref.)		1.00 (ref.)	
	CA	61/328	**1.72 (1.10–2.68)**	**0.017**	**1.87 (1.17–2.99)**	**0.009**
	AA	7/75	0.78 (0.33–1.82)	0.558	0.78 (0.32–1.89)	0.583
	Dominant	68/403	1.52 (0.99–2.35)	0.058	**1.63 (1.03–2.57)**	**0.036**
	Recessive	97/634	0.57 (0.25–1.28)	0.172	0.55 (0.24–1.28)	0.165
***ERCC3*/rs13385611**	TT	83/553	1.00 (ref.)		1.00 (ref.)	
	CT	15/143	0.67 (0.37–1.19)	0.173	0.69 (0.37–1.27)	0.229
	CC	6/12	**5.69 (1.79–18.06)**	**0.003**	**5.32 (1.58–17.95)**	**0.007**
	Dominant	21/155	0.89 (0.53–1.49)	0.65	0.92 (0.54–1.58)	0.763
	Recessive	98/696	**6.10 (1.93–19.30)**	**0.002**	**5.64 (1.68–18.96)**	**0.005**
***RAD23*/rs7041137**	CC	50/434	1.00 (ref.)		1.00 (ref.)	
	CT	46/243	**1.80 (1.16–2.78)**	**0.008**	**2.07 (1.31–3.27)**	**0.002**
	TT	8/32	**2.57 (1.09–6.02)**	**0.030**	2.23 (0.91–5.47)	0.079
	Dominant	54/275	**1.88 (1.24–2.85)**	**0.003**	**2.08 (1.34–3.24)**	**0.001**
	Recessive	96/677	2.02 (0.88–4.62)	0.097	1.69 (0.71–4.02)	0.239
Thrombocytopenia (Events=26)						
***ERCC5*/rs873601**	AA	11/177	1.00 (ref.)		1.00 (ref.)	
	GA	9/352	0.40 (0.16–0.97)	**0.044**	**0.35 (0.13–0.91)**	**0.031**
	GG	6/181	0.52 (0.19–1.43)	0.204	0.46 (0.16–1.35)	0.159
	Dominant	15/533	0.44 (0.20–0.97)	**0.042**	**0.39 (0.17–0.91)**	**0.029**
	Recessive	20/529	0.87 (0.35–2.21)	0.774	0.84 (0.32–2.22)	0.731

[1]Adjusted by age at diagnose, sex, smoking status, BMI, TNM status, ECOG performance status, chemotherapy regimens, histological type, histologic grade and palliative radiotherapy. The results are in **bold** if $p < 0.05$.
Abbreviations: SNP-single nucleotide polymorphism; OR-odds ratio.

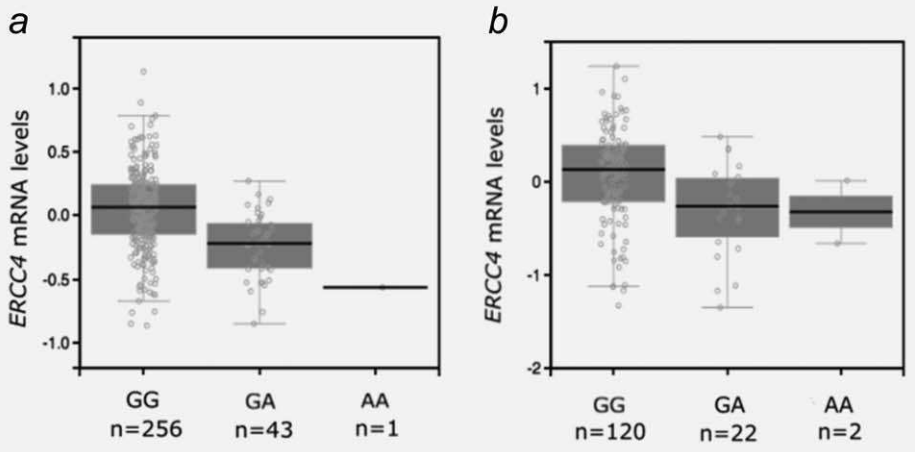

Figure 3. eQTL analysis of *ERCC4* rs1799798 from GTEx database. (a) mRNA expression levels of ERCC4 by rs1799798 geneotypes in transformed fibroblast cells from 300 individuals ($P = 0.81 \times 10^{-8}$, effect size = −0.35); (b) mRNA expression levels of ERCC4 by rs1799798 genotypes in brain caudate basal ganglia tissues from 146 individuals ($P = 0.0013$, effect size = −0.40).

accompanied by variation in severe side effects among the patients, which hamper treatment outcome.[23–25] The variation in these toxicities is likely due to variation in the NER pathway that is responsible for the removal of PBC-induced DNA damage, leading to variation in response to toxic platinum agents. DNA repair capacity, particularly of the NER pathway, has been associated with the PBC efficacy and toxicity, and numerous studies have indicated that genetic variations, such as SNPs in the NER pathway genes, may modulate repair functions and cellular apoptosis, prompting individual variation in toxicity outcomes of PBC.[26,27]

Published studies on NER pathway variants and PBC toxicity in NSCLC patients have mainly focused on missense variants, with very few investigating a group of core genes in the pathway or considering LD between the variants or non-coding variants.[24,28,29] In the present study, we employed a hypothesis-based approach with a main focus on independent regulatory variants with predicted biological functions. Based on a two-phase screening design, we observed that under a dominant genetic model the *ERCC4* rs1799798 SNP had a significant effect on both overall and GI toxicity outcomes in advanced Chinese NSCLC patients treated with PBC, it is possible that the rs1799798 A allele associated with a higher risk of overall toxicity was driven by its association with GI toxicity (Table 2). The association with this risk SNP seems to be robust in discovery, replication and the combined datasets, but the prediction model for overall toxicity incorporating this risk SNP exhibited a moderate increase in the AUC value in all patients dataset, compared to the clinical variable-based model (Supporting Information, Table S6). This SNP has previously been found to be associated with the overall survival of these patients (Supporting Information, Fig. S2). It is possible that the increased risk of death observed in those patients carrying the risk allele was exacerbated by the consequence of the higher toxicity. *ERCC4* participates in the removal of DNA inter-strand cross-links and double-stranded breaks, and therefore SNPs in *ERCC4* may impact the repair capacity and thus affect chemotherapy outcome. The rSNP rs1799798 resides in an intron of the *ERCC4* gene without a linkage to any other SNPs within a 1 Mb window (Supporting Information, Fig. S1). Bioinformatics analysis predicted a functional role of this SNP to be in the ELF-1 binding site.[30]

The present study also suggested that *ERCC4* variant might affect overall, GI and hematologic toxicities based on different demographic or clinical characteristics. For instance, overall and GI toxicities of patients carrying *ERCC4* rs1799798 GA/AA genotypes were more evident in subgroups of older, non-smokers, adenocarcinoma and recipients of gemcitabine, pemetrexed chemotherapies. Such effect of *ERCC4* rs1799798 was more apparent in subgroups of gemcitabine, cisplatin-based chemotherapy and low histologic grade for overall toxicity but subgroups of high histologic grade and no radiotherapy for GI toxicities. The association of *ERCC4* rs1799798 in the non-radiotherapy subgroup indicates that this SNP might contribute to the observed GI toxicity independent of radiotherapy (Table 3). We also identified additional NER variants to be associated with various types of hematologic toxicity (Table 4). Due to the relatively small sample size in some subgroups or toxicity incidence, these results need to be further validated in larger study populations.

Finally, there are some inherent limitations in the present study. First, the recruitment of patients treated in the same hospital may lead to selection bias and may not be generalizable to other patient populations; therefore interpretation of our results needs to be dealt with caution, and additional replication studies in different patient populations are desired. Second, with the aim of studying rSNPs, we did not incorporate the coding region SNPs with known effects. This might weaken the power of our ROC prediction models. Third, multiple testing correction was not conducted in the present study, because this was an exploratory study with limited study power. Last, because it is a retrospective study with a relatively small sample size, further prospective studies with large populations are warranted to validate our findings.

In conclusion, the present study provided some evidence that rSNP in the NER core genes may modulate PBC-induced toxicity. This explorative study may serve as the basis for further replicative and functional studies to unravel the molecular mechanisms underlying PBC-induced toxicity, which may lead to development of useful biomarkers for predicting platinum-based chemotherapy toxicity in patients with advanced NSCLC.

Acknowledgements

The authors acknowledge Dr. Haiquan Chen at the Department of Thoracic Surgery for contributing lung cancer tissue samples, the staff members at the Tissue Bank of the Fudan University Shanghai Cancer Centre for their tremendous support in providing stored blood samples.

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:E359–86.
2. Herbst RS, Heymach JV, Lippman SM. Lung cancer. *N Engl J Med* . 2008;359:1367–80.
3. Reck M, Heigener DF, Mok T, et al. Management of non-small-cell lung cancer: recent developments. *Lancet*. 2013;382:709–19.
4. Azzoli CG, Baker S, Jr., Temin S, et al. American Society of Clinical Oncology Clinical Practice Guideline update on chemotherapy for stage IV non-small-cell lung cancer. *J Clin Oncol*. 2009;27:6251–66.
5. Schiller JH, Harrington D, Belani CP, et al. Comparison of four chemotherapy regimens for advanced non-small-cell lung cancer. *N Engl J Med*. 2002;346:92–8.
6. Scagliotti GV, Parikh P, von Pawel J, et al. Phase III study comparing cisplatin plus gemcitabine with cisplatin plus pemetrexed in chemotherapy-naive patients with advanced-stage non-small-cell lung cancer. *J Clin Oncol*. 2008;26:3543–51.
7. Di Francesco AM, Ruggiero A, Riccardi R. Cellular and molecular aspects of drugs of the future: oxaliplatin. *Cell Mol Life Sci*. 2002;59:1914–27.

8. Siddik ZH. Cisplatin: mode of cytotoxic action and molecular basis of resistance. *Oncogene.* 2003;22:7265–79.
9. Wu W, Zhang W, Qiao R, et al. Association of XPD polymorphisms with severe toxicity in non-small cell lung cancer patients in a Chinese population. *Clin Cancer Res.* 2009;15: 3889–95.
10. Han B, Gao G, Wu W, et al. Association of ABCC2 polymorphisms with platinum-based chemotherapy response and severe toxicity in non-small cell lung cancer patients. *Lung Cancer.* 2011;72:238–43.
11. Rabik CA, Dolan ME. Molecular mechanisms of resistance and toxicity associated with platinating agents. *Cancer Treat Rev.* 2007;33:9–23.
12. Kang S, Ju W, Kim JW, et al. Association between excision repair cross-complementation group 1 polymorphism and clinical outcome of platinum-based chemotherapy in patients with epithelial ovarian cancer. *Exp Mol Med.* 2006;38: 320–4.
13. de Boer J, Hoeijmakers JH. Nucleotide excision repair and human syndromes. *Carcinogenesis.* 2000;21:453–60.
14. Li L, Peterson CA, Lu X, et al. Mutations in XPA that prevent association with ERCC1 are defective in nucleotide excision repair. *Mol Cell Biol.* 1995; 15:1993–8.
15. Gerard M, Fischer L, Moncollin V, et al. Purification and interaction properties of the human RNA polymerase B(II) general transcription factor BTF2. *J Biol Chem.* 1991;266:20940–5.
16. Shivji MK, Podust VN, Hubscher U, et al. Nucleotide excision repair DNA synthesis by DNA polymerase epsilon in the presence of PCNA, RFC, and RPA. *Biochemistry.* 1995;34: 5011–7.
17. Jia M, Xu Y, Zhu M, et al. The P38alpha rs3804451 variant predicts chemotherapy response and survival of patients with non-small cell lung cancer treated with platinum-based chemotherapy. *Transl Oncol.* 2016;9:531–9.
18. Carbone DP, Minna JD. Chemotherapy for non-small cell lung cancer. *Bmj.* 1995;311:889–90.
19. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2010;26:589–95.
20. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25:2078–9.
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;22:568–76.
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. McTiernan A, Jinks RC, Sydes MR, et al. Presence of chemotherapy-induced toxicity predicts improved survival in patients with localised extremity osteosarcoma treated with doxorubicin and cisplatin: a report from the European Osteosarcoma Intergroup. *Eur J Cancer.* 2012;48: 703–12.
24. Suk R, Gurubhagavatula S, Park S, et al. Polymorphisms in ERCC1 and grade 3 or 4 toxicity in non-small cell lung cancer patients. *Clin Cancer Res.* 2005;11:1534–8.
25. Banerji U, Ashley S, Coward J, et al. The association of chemotherapy induced neutropenia on treatment outcomes in small cell lung cancer. *Lung Cancer.* 2006;54:371–7.
26. Powrozek T, Mlak R, Krawczyk P, et al. The relationship between polymorphisms of genes regulating DNA repair or cell division and the toxicity of platinum and vinorelbine chemotherapy in advanced NSCLC patients. *Clin Transl Oncol.* 2016;18:125–31.
27. Shiraishi K, Kohno T, Tanai C, et al. Association of DNA repair gene polymorphisms with response to platinum-based doublet chemotherapy in patients with non-small-cell lung cancer. *J Clin Oncol.* 2010;28:4945–52.
28. Chu TQ, Li R, Shao MH, et al. RAD18 polymorphisms are associated with platinum-based chemotherapy toxicity in Chinese patients with non-small cell lung cancer. *Acta Pharmacol Sin* . 2016; 37:1490–8.
29. Perez-Ramirez C, Canadas-Garre M, Alnatsha A, et al. Pharmacogenetic predictors of toxicity to platinum based chemotherapy in non-small cell lung cancer patients. *Pharmacol Res.* 2016;111:877–84.
30. Yu H, Liu Z, Huang YJ, et al. Association between single nucleotide polymorphisms in ERCC4 and risk of squamous cell carcinoma of the head and neck. *PloS One.* 2012;7:e41853.