Original Article

Association of positively selected eIF3a polymorphisms with toxicity of platinum-based chemotherapy in NSCLC patients

Ji-ye YIN^{1, 2, 3, *}, Xiang-guang MENG⁴, Chen-yue QIAN^{1, 2, 3}, Xiang-ping LI⁵, Juan CHEN^{1, 2, 3}, Yi ZHENG^{1, 2, 3}, Rong LIU^{1, 2, 3}, Hong-hao ZHOU^{1, 2, 3}, Zhao-qian LIU^{1, 2, 3, *}

¹Department of Clinical Pharmacology, Xiangya Hospital, Central South University, Changsha 410008, China; Institute of Clinical Pharmacology, Central South University; ²Hunan Key Laboratory of Pharmacogenetics, Changsha 410078, China; ³Hunan Province Cooperation Innovation Center for Molecular Target New Drug Study, Hengyang 421001, China; ⁴Laboratory of Cardiovascular Disease and Drug Research, Zhengzhou No. 7 People's Hospital, Zhengzhou 450016, China; ⁵Department of Pharmacy, Xiangya Hospital, Central South University, Changsha 410008, China

Aim: Eukaryotic translation initiation factor 3 subunit A (eIF3a) plays critical roles in regulating the initiation of protein translation, and eIF3a is highly expressed in lung cancer. In this study, we investigated the association of the positively selected SNPs of eIF3a with the response to and toxicity of platinum-based chemotherapy in Chinese patients with non-small cell lung cancer (NSCLC).

Methods: SNP data for eIF3a locus were downloaded from HapMap database. For each SNP, haplotype, LD profile and population differentiation were analyzed. The long-range haplotype (LRH) test was employed to identify positively selected SNPs of eIF3a. A total of 325 NSCLC patients were enrolled and genotyped for these SNPs.

Results: Five positively selected (rs1409314, rs4752219, rs4752220, rs7091672 and rs10510050) and 5 non-positively selected SNPs (rs10886342, rs11198804, rs2275112, rs10787899 and rs4752269) were identified in the LRH test. However, none of them was correlated with the platinum-based chemotherapy response. In contrast, 4 of the positively selected SNPs (rs1409314, rs4752219, rs4752220 and rs7091672) were significantly correlated with the toxicities tested (neutropenia, anemia, thrombocytopenia, emesis and hepatotoxicity). In addition, rs10510050 was significantly correlated with thrombocytopenia, emesis and hepatotoxicity. None of the 5 non-positively selected SNPs was correlated with the 5 toxicities.

Conclusion: The positively selected SNPs of eIF3a are significantly correlated with platinum-based chemotherapy toxicities in Chinese NSCLC patients.

Keywords: eIF3a; positive selection; polymorphism; lung cancer; platinum-based chemotherapy; chemotherapy toxicity; Chinese NSCLC patients

Acta Pharmacologica Sinica (2015) 36: 375–384; doi: 10.1038/aps.2014.160

Introduction

Translation is one of the major mechanisms of gene expression regulation. It consists of three major steps: initiation, elongation and termination. Among these, initiation is the rate-limiting step and involves at least 11 translation initiation factors (eIFs)^[1]. EIF3 is the largest and most complex, consisting of at least 13 subunits named eIF3a to eIF3m^[2, 3]. EIF3a is the largest subunit with a molecular weight of 170 kDa, and the gene

encoding it has 24 exons spanning approximately 45.8 kb of chromosome 10q26^[4].

EIF3a plays critical roles in all steps of cap-dependent translation initiation, including binding with the 40S ribosomal subunit, scanning and recognizing the start codon^[2]. In addition, it has been reported to be involved in internal ribosomal entry site (IRES)-mediated cap-independent translation initiation^[5, 6]. However, recent studies have shown that eIF3a is overexpressed in many cancers, including breast, lung, cervix, stomach and esophagus^[2]. Our previous study also showed that eIF3a regulates drug response by altering the expression of some DNA repair proteins^[7]. These studies thus implicate eIF3a as a potential oncogene.

* To whom correspondence should be addressed.

E-mail liuzhaoqian63@126.com (Zhao-qian LIU);

yinjiye2009@hotmail.com (Ji-ye YIN)

Received 2014-08-30 Accepted 2014-11-13

Single nucleotide polymorphisms (SNPs) are the most abundant human genome variations. Our previous study demonstrated that two *eIF3a* SNPs were correlated with drug response and toxicity^[8, 9]. We thus hypothesized that additional *eIF3a* SNPs may correlate with response and toxicity of platinum-based chemotherapy in non-small cell lung cancer (NSCLC) patients. It has been reported that screening SNPs with evidence of positive selection is a valid strategy for identifying functional SNPs^[10–12]. This method is based on the theory that phenotypic differences between individuals are the functional consequences of genetic variants that facilitated the survival of ancestral human populations under different environmental stressors^[13]. Therefore, SNPs under positive selection are likely to be functionally important for their linked genes. This strategy has been used successfully to identify functional SNPs associated with many genes, including glucose-6-phosphate dehydrogenase (*G6PD*)^[13], multidrug resistance 1 (*MDR1*)^[10], cytochrome P450, family 3, subfamily A (*CYP3A*)^[11] and proprotein convertase subtilisin-like kexin type 9 (*PCSK9*)^[12].

In the present study, we employed the long-range haplotype (LRH) test to identify positively selected SNPs of *eIF3a* and explored whether they were correlated with response and toxicity of platinum-based chemotherapy in Chinese NSCLC patients.

Materials and methods

Study population, acquisition of clinical information and genotyping

The study protocol was approved by the Ethics Committee of Xiangya School of Medicine, Central South University. A total of 325 lung cancer patients were enrolled in this study. All individuals provided written informed consent in compliance with the code of ethics of the World Medical Association (Declaration of Helsinki) before this study was initiated. Eligible subjects were treated at the Affiliated Cancer Hospital or Xiangya Hospital of Central South University (Changsha, Hunan, China) between November 2011 and May 2013. For this study, patients had to meet the following criteria: (1) histologic or cytologic confirmation of lung cancer and (2) presence of a primary tumor. Exclusion criteria included (1) pregnancy or lactation, (2) active infection, (3) symptomatic brain or leptomeningeal metastases and (4) previous or concomitant other malignancies. All demographic and clinical information was obtained from the two hospitals. The clinical characteristics of all subjects are summarized in Table 1. We applied this study for clinical admission in the Chinese Clinical Trial Register (registration number: ChiCTR-RC-12002830).

The chemotherapy response was evaluated by the response evaluation criteria in solid tumor (RECIST) guidelines. Complete response (CR) and partial response (PR) were defined as drug response, whereas stable disease (SD) and progressive disease (PD) were defined as drug resistance. In the current study, hematologic toxicity (anemia, neutropenia and thrombocytopenia) and gastrointestinal toxicity (emesis and hepatotoxicity) were evaluated. Toxicity was classified according

to the National Cancer Institute Common Toxicity Criteria 3.0 (NCI-CTC 3.0) at each cycle of chemotherapy. For each patient, the most severe grade during the entire study interval was evaluated. In brief, for all side effects, grade 0 and grade ≥ 1 were considered non-toxicity and toxicity, respectively. Genomic DNA of all subjects was isolated from the peripheral blood sample using a FlexiGene DNA Kit (Qiagen, Hilden, Germany) and stored at 4 °C until use. Genotyping was conducted by Sequenom's MassARRAY system (Sequenom, San Diego, California, USA).

Polymorphism data and SNP marker selection

Genotype data of SNPs were retrieved from the International HapMap Project database (Phase III Release #28, www.hapmap.org). Six populations were included in this study, including CEU, CHB, JPT, YRI, ASW, and CHD. The ancestral allele of each SNP was defined by aligning the sequence of each SNP to a chimpanzee (*Pan troglodyte*) sequence from the UCSC database (genome.ucsc.edu).

SNPs spanning approximately 45.8 kb of the *eIF3a* locus were selected. To generate a more evenly distributed and denser SNP map, we extended the selection region to 200 kb upstream and downstream of the *eIF3a* locus. SNPs in the following categories were excluded from the study: (1) SNPs with a minor allele frequency < 0.05 , (2) SNPs that were not genotyped in at least one population and (3) SNPs departing from Hardy-Weinberg equilibrium. In total, 127 SNPs were included, and a flow chart describing SNP selection and information regarding included SNPs are provided in Figure S1 and Table S1, respectively.

Data analysis

Data analysis was performed using R and Bioconductor. In addition, the Haploview program was utilized to perform genetic analysis, including allele frequency description and selection, LD and haplotype block analysis^[14]. Lewontin's coefficient $|D'|$ and Pearson's correlation r^2 ^[15] were employed to estimate LD between SNP pairs. Statistical significance of LD between SNP pairs was analyzed by Arlequin^[16]. The classical LD decay equation $D = D_0(1 - \theta)^t$ was utilized to calculate LD decay trend line, which evaluated the overall LD in the populations^[17]. Half-LD ($LD_{0.5}$) was defined as the distance at which $|D'| = 0.5$. Haplotype reconstruction was conducted in the PHASE program^[18]. An unbiased estimator of F_{ST} (F-statistic/fixation index) from pairwise population comparisons was calculated. Randomization method (1000 permutations) was used to test the statistical significance of F_{ST} in the comparison using the FSTAT program^[19]. A modified LRH test was used to assess for positive selection, and single SNP rather than haplotype was referred as "core"^[13]. Haplotype homozygosity (HH) was calculated in a stepwise manner as extended haplotype homozygosity (EHH)^[13]. Relative EHH (REHH) was the ratio of the EHH on the test core allele compared with the EHH of the grouped set of core alleles at the region not including the core allele tested^[13]. Haplotype bifurcation diagram (HBD) was used to graphically evaluate the LD decay of both

ancestral and derived alleles for a given core SNP. To test the statistical significance of the LRH test, an empirical distribution of core haplotype frequencies versus REHH was obtained by screening the entire chromosome 10 HapMap data in ASN, CEU and YRI populations. The HBD and LRH statistical tests were performed using Sweep software.

Statistical analysis of association study

All continuous variables were presented as means $\pm$ SD and confidence intervals (95% CI), and normally distributed data were analyzed by the two-sample *t* test. Noncontiguous variables, genotype distributions, and allelic frequencies in different groups were compared using the χ^2 test. Unconditional logistic regression was performed to estimate the association of drug response and toxicity with *eIF3a* polymorphisms by calculating odds ratios (OR) and their 95% confidence intervals (CI). The *P* value was two-sided and *P*<0.05 was considered statistically significant. The aforementioned statistical analyses were performed using PLINK^[20] and SPSS 18.0 software (SPSS Inc, Chicago, Illinois, USA). Bonferroni correction was used to correct multiple testing.

Results

Haplotype and LD profile

We first detected the haplotype and LD profile of the involved SNPs, and the pair-wise $|D'|$ and r^2 were calculated to determine the *eIF3a* LD pattern. As shown in Figure 1, the patterns of LD structure were highly different across the entire locus between Africans and non-Africans. There are 18, 16, 17, 14, 29, and 29 blocks observed in CHD, CHB, JPT, CEU, ASW, and

YRI populations, respectively. The longest blocks spanned 91 kb in CHD, 91 kb in CHB, 91 kb in JPT, 94 kb in CEU, 48 kb in ASW, and 26 kb in YRI.

We then plotted the $|D'|$ and r^2 between SNP pairs against their physical distance, as shown in Figure 2. The LD generally decreased with physical distance. A curve was then fitted to determine the half LD (LD_{50}) in all populations. The genetically older populations were expected to have lower LDs, and this was the case in our study. The LD_{50} s of African populations YRI (44.5 kb) and ASW (61.5 kb) were shorter than those of non-African populations, while the longest LD_{50} s presented in the Chinese population were CHD (78.3 kb) and CHB (75.4 kb).

Furthermore, we calculated the proportion of SNP pairs with significant LD in these populations and observed clear differences. As shown in Figure S2A, African populations had fewer SNP pairs with significant LDs: the proportions were 23.66% and 19.03% in YRI and ASW populations. For non-African populations, however, 39.07%, 38.72%, 36.71%, and 35.88% SNP pairs showed significant LD in the CHD, CEU, CHB, and JPT populations, respectively. We also plotted the proportion of SNP pairs with insignificant LDs against physical distance for ASN (East Asian population; combined CHB and JPT populations), CEU and YRI populations (Figure S2B). In accordance with previous observation, the African YRI population showed lower LD levels, while the ASN population showed higher allelic association.

Together, the haplotype and LD pattern displayed high diversity among the 6 populations analyzed, especially between African (ASW and YRI) and non-African (CHD, CHB,

Figure 1. LD pattern of *eIF3a* polymorphisms in the CHD, CHB, JPT, CEU, ASW, and YRI populations. Pairwise LD between polymorphisms was measured by $|D'|$. Haplotype blocks partitioned in each population were also numbered above them. The analysis was performed in the Haploview program (version 4.2).

Figure 2A. LD decay profile of *eIF3a* polymorphisms in the CHD, CHB, JPT, CEU, ASW, and YRI populations. (A) $|D'|$ between *eIF3a* SNP pairs were calculated and plotted against their physical distance in the CHD, CHB, JPT, CEU, ASW, and YRI populations. A classical LD decay equation was utilized to calculate LD decay trend line, which is indicated by a red line in the figure. The LD_{50} was also shown. Details of the calculation were described in the Materials and methods section.

CEU, and JPT) populations.

Population differentiation

To further test for differences in *eIF3a* allelic frequency among populations, we calculated F_{ST} values. Considering that the CHB and JPT populations exist in the same region and have substantial similarity in their genetic makeup, they were combined to make one ASN population in further analyses. Additionally, the following analyses were conducted on only the ASN, CEU, and YRI populations, which reflected a substantial amount of the genetic variation found throughout the world^[21].

Based on the pairwise F_{ST} values, *eIF3a* showed high popula-

tion differentiation. SNPs with a proportion of 21.71% in CEU versus ASN, 43.41% in CEU versus YRI and 33.33% in ASN versus YRI showed significantly high F_{ST} , respectively. We next plotted the F_{ST} values against physical distance in Figure S3. As shown, 7 SNPs in CEU versus YRI and 5 SNPs in ASN versus YRI had F_{ST} values above the threshold of 0.45, which was considered the threshold value of selection pressure^[22, 23]. However, none of the SNPs in CEU versus ASN had F_{ST} values above 0.45, indicating that *eIF3a* had low levels of differentiation between these two populations.

These results indicated that there was significantly different population differentiation among the ASN, CEU, and YRI populations, and that some SNPs may be under

Figure 2B. LD decay profile of *eIF3a* polymorphisms in the CHD, CHB, JPT, CEU, ASW, and YRI populations. (B) r^2 between *eIF3a* SNP pairs were calculated and plotted against their physical distance in the CHD, CHB, JPT, CEU, ASW, and YRI populations. Details of the calculation were described in the Materials and methods section.

selective pressure.

Signature of positive selection of *eIF3a*

To test for evidence of positive selection occurring at these SNPs with *eIF3a*, we performed an LRH test. We first plotted the relative EHH (REHH) of core SNPs against their allele frequencies. Based on the empirical distribution, 41, 33, and 13 SNPs showed significant departure from evolutionary neutrality in the YRI, ASN, and CEU populations, respectively. The top 5 SNPs with lowest P values were then involved in further study. As shown in Figure 3, A allele of SNP rs1409314 ($P=0.0012$), C allele of SNP rs4752219 ($P=0.0025$), A allele of SNP rs4752220 ($P=0.0025$), T allele of SNP rs7091672 ($P=0.0043$) and A allele of SNP rs10510050 ($P=0.0046$) deviated significantly from evolutionary neutrality. To assess this result, the EHH decay of these 5 SNPs was plotted against the physical distance. Generally, the overall EHH of the rs1409314A, rs4752219C, rs4752220A, rs7091672T, and rs10510050A alleles decayed more slowly than their corresponding ancestral/

derived alleles in all three populations (Figure S4). To confirm this result, we plotted haplotype branching diagrams (HBD) for these 5 SNPs, as indicated in Figure S5. Consistent with previous results, the rs1409314A, rs4752219C, rs4752220A, rs7091672T, and rs10510050A alleles had obvious long-range LD across all populations, especially in YRI, as indicated by a single distinctly predominant thick branch. Taken these results together, we concluded that rs1409314, rs4752219, rs4752220, rs7091672, and rs10510050 showed a signature of recent positive selection.

As a control, we further analyzed 5 randomly selected SNPs (rs10886342, rs11198804, rs2275112, rs10787899, and rs4752269), which didn't show evidence of positive selection. These 5 SNPs were selected as representative non-positively selected SNPs of *eIF3a* for use in the following study.

Association of selected SNPs and platinum-based chemotherapy response

Our previous study demonstrated that *eIF3a* may act as an

Figure 3. LRH test for 5 positively selected SNPs. The REHH value of 5 core alleles was plotted against its frequency in the ASN, CEU and YRI populations. The points of REHH against allele frequency were based on the calculation of chromosome 10 phased data downloaded from the HapMap database (data point not shown in the figure). The 50th, 75th, 90th, 95th, 99th, and 99.9th percentile lines were obtained by binning the empirical data by allele frequencies into 20 bins with equal intervals of 0.05. All analysis was performed in the Sweep programme.

oncogene and may be correlated with lung cancer chemotherapy response and toxicity^[4, 6, 7]. Thus, we tested whether these positively selected SNPs were associated with platinum-based chemotherapy response in Chinese NSCLC patients. A total of 325 lung cancer patients were recruited for genotyping of the aforementioned 5 positively selected and 5 non-positively selected SNPs. There were significant differences between responders and non-responders in terms of age and chemotherapy regimen, suggesting that these factors need to be adjusted in the statistical analysis (Table 1). All tested SNPs were consistent with Hardy-Weinberg equilibrium in both responders and non-responders.

Table 2 summarizes the results of association study of 10 SNPs in additive, dominant and recessive models. As indicated, neither the 5 positively selected SNPs (rs1409314, rs4752219, rs4752220, rs7091672 and rs10510050) nor the 5 non-positively selected SNPs (rs10886342, rs11198804, rs2275112, rs10787899, and rs4752269) were significantly associated with NSCLC patients' platinum-based chemotherapy response.

Association of selected SNPs and platinum-based chemotherapy toxicity

We next tested whether positively selected SNPs were correlated with chemotherapy toxicity in these patients. In the present study, we investigated 5 toxicities: neutropenia, anemia, thrombocytopenia, emesis and hepatotoxicity. As indicated in Table 3, each was divided into two groups: toxicity (0 grade) and non-toxicity (≥ 1 grade). In general, 64%–68% patients showed at least one toxicity and only 32%–36% patients did not suffer from any toxicity.

It is interesting to note that 4 (rs1409314, rs4752219, rs4752220, and rs7091672) of the 5 positively selected SNPs were significantly correlated with all tested toxicities (Table 4). In addition, a significant association was also detected between rs10510050 and thrombocytopenia (OR=4.34, 95% CI=1.47–12.81, $P=0.00$), emesis (OR=0.39, 95% CI=0.18–0.83, $P=0.01$) and hepatotoxicity (OR=0.58, 95% CI=0.34–0.99, $P=0.04$). In contrast, of the 5 non-positively selected SNPs (rs10886342, rs11198804, rs2275112, rs10787899, and rs4752269), none

Table 1. Baseline characteristic of NSCLC patients treated with platinum-based chemotherapy.

Characteristics	All patients (%)	Responders (%)	Non-responders (%)	P value
Total	325 (100)	132 (40.62)	193 (59.38)	
Age (mean±SD)	53.82±9.24	55.35±9.27	52.78±9.09	0.010
Gender				
Male	209 (64.31)	82 (62.12)	127 (65.80)	
Female	116 (35.69)	50 (37.88)	66 (34.20)	0.556
Histology				
NSCLC	246 (75.69)	97 (73.48)	149 (77.20)	
SCLC	79 (24.31)	35 (26.52)	44 (22.80)	0.511
Smoking status				
Smokers	169 (52.00)	63 (47.73)	106 (54.92)	
Non-smokers	156 (48.00)	69 (52.27)	87 (45.08)	0.215
ECOG PS				
0–1	311 (95.69)	123 (93.18)	188 (97.41)	
≥2	14 (4.31)	9 (6.82)	5 (2.59)	0.093
Chemotherapy				
Platinum+gemcitabine (GP)	94 (28.92)	32 (24.24)	62 (32.12)	
Platinum+etoposide (EP)	78 (24.00)	46 (34.85)	32 (16.58)	
Platinum+paclitaxel (TP)	16 (4.92)	3 (2.27)	13 (6.74)	
Platinum+pemetrexed (PP)	79 (24.31)	27 (20.45)	52 (26.94)	
Other	58 (17.85)	24 (18.19)	34 (17.62)	0.002

showed significant association with any of the 5 tested toxicities. These results together suggest that the aforementioned 5 positively selected *eIF3a* SNPs are significantly correlated with platinum-based chemotherapy toxicities in Chinese NSCLC patients.

Discussion

In the present study, we first characterized the haplotype and LD profile of *eIF3a* in data obtained from the HapMap database, then analyzed the population differentiation in three populations (ASN, CEU, and YRI). Moreover, we identified 5 SNPs (rs1409314, rs4752219, rs4752220, rs7091672, and rs10510050) of *eIF3a* that had undergone recent positive selection using the LRH test. Finally, we demonstrated that these SNPs were correlated with platinum-based chemotherapy toxicities in Chinese NSCLC patients.

With the explosion of newly identified SNPs, it is becoming impractical to investigate every SNP individually for correlation with a disease or drug response. Although new methods of DNA sequencing and genotyping have been developed, these are time consuming and costly. Thus, bioinformatics methods have been developed to identify potential functional SNPs. Based on the theories of natural selection and the aforementioned “Out-of-Africa” hypothesis, searching for signatures of recent positive selection can be used to identify potential functional SNPs^[24]. With the dispersal of humans from Africa to the other regions of the world, different popu-

lations faced various selective pressures presented by climate, pathogens and sources of food. Thus, functional SNPs that facilitated survival in ancestral human populations may result in phenotypic differences among modern individuals or groups^[25]. These SNPs may be retained under the pressure of positive selection; for example, a classic G6PD-202A variant that confers malaria resistance showed a strong evidence of recent positive selection^[13]. Other groups have successfully utilized this method to detect functional SNPs in *PCSK9*^[12] and ABC drug transporter genes^[26]. In the current study, we employed this method to identify 5 positively selected *eIF3a* SNPs and confirmed that they were significantly correlated with platinum-based chemotherapy toxicities in Chinese NSCLC patients. Importantly, another 5 non-positively selected control SNPs did not show evidence of association. Although our previous studies identified a number of platinum-based chemotherapy responses and toxicities associated SNPs^[27–29], this time- and cost-saving method is a potentially powerful tool for specifically identifying SNPs that correlate with drug response in the future.

Although *eIF3a* is the largest subunit of *eIF3*, which is itself the largest and most complex initiation factor, it has been shown to be dispensable for the function of *eIF3*-mediated translation initiation^[2, 30]. By contrast, previous studies by our group and others have suggested that *eIF3a* is a potential oncogene. Thus, it is important to understand the contributions of *eIF3a* polymorphisms to its correlation with cancers.

Table 2. Association of *eIF3a* SNPs with platinum-based chemotherapy response in the Chinese patients.

Polymorphisms	Genotype	Responders <i>n</i> (%)	Non-responders <i>n</i> (%)	Call rate (%)	MAF	Additive OR (95% CI)	<i>P</i>	Dominant OR (95% CI)	<i>P</i>	Recessive OR (95% CI)	<i>P</i>
rs1409314	AA	27 (20.93)	38 (20.32)	97.74	0.43	0.83 (0.61–1.12)	0.23	0.66 (0.41–1.05)	0.08	0.96 (0.55–1.68)	0.90
	AG	62 (48.06)	73 (39.04)								
	GG	40 (31.01)	76 (40.64)								
rs4752219	CC	45 (36.29)	64 (35.75)	96.20	0.42	1.16 (0.85–1.58)	0.35	1.02 (0.64–1.65)	0.92	1.60 (0.90–2.85)	0.11
	CT	58 (46.77)	71 (39.66)								
	TT	21 (16.94)	44 (24.59)								
rs4752220	AA	48 (36.92)	67 (35.83)	98.62	0.42	1.16 (0.86–1.57)	0.33	1.05 (0.66–1.67)	0.84	1.56 (0.88–2.75)	0.13
	AG	60 (46.15)	75 (40.11)								
	GG	22 (16.93)	45 (24.06)								
rs7091672	TT	19 (15.45)	29 (16.20)	96.03	0.39	0.89 (0.64–1.23)	0.47	0.76 (0.47–1.21)	0.24	1.06 (0.56–2.00)	0.86
	CT	58 (47.15)	71 (39.66)								
	CC	46 (37.40)	79 (44.14)								
rs10510050	AA	25 (19.23)	34 (18.28)	98.45	0.40	0.83 (0.61–1.12)	0.23	0.68 (0.43–1.08)	0.10	0.94 (0.53–1.67)	0.83
	AG	59 (45.38)	69 (37.10)								
	GG	46 (35.39)	83 (44.62)								
rs10886342	AA	27 (20.93)	45 (23.94)	98.27	0.44	1.06 (0.78–1.42)	0.73	1.00 (0.63–1.60)	1.00	1.19 (0.69–2.04)	0.53
	AG	56 (43.41)	76 (40.43)								
	GG	46 (35.66)	67 (35.63)								
rs11198804	TT	29 (22.31)	53 (28.19)	98.79	0.49	1.12 (0.83–1.52)	0.46	1.01 (0.62–1.66)	0.96	1.37 (0.81–2.30)	0.24
	TC	64 (49.23)	82 (43.62)								
	CC	37 (28.46)	53 (28.19)								
rs2275112	GG	82 (65.60)	128 (71.11)	96.55	0.16	0.83 (0.52–1.31)	0.42	0.77 (0.47–1.27)	0.31	2.10 (0.22–20.44)	0.52
	GA	42 (33.60)	49 (27.22)								
	AA	1 (0.80)	3 (1.67)								
rs10787899	AA	19 (15.32)	32 (17.88)	96.03	0.43	1.09 (0.78–1.53)	0.60	1.08 (0.67–1.75)	0.76	1.20 (0.65–2.24)	0.56
	AG	62 (50.00)	88 (49.16)								
	GG	43 (34.68)	59 (32.96)								
rs4752269	TT	4 (3.23)	7 (3.89)	96.37	0.19	1.44 (0.94–2.20)	0.09	1.58 (0.97–2.56)	0.06	1.21 (0.35–4.24)	0.76
	TC	34 (27.42)	67 (37.22)								
	CC	86 (69.35)	106 (58.89)								

Table 3. Groups of graded platinum-based chemotherapy toxicities.

Side effect	0 grade <i>n</i> (%)	≥1 grade <i>n</i> (%)
Neutropenia	112 (34.46)	213 (65.54)
Anemia	114 (35.08)	211 (64.92)
Thrombocytopenia	114 (35.08)	211 (64.92)
Emesis	116 (35.69)	209 (64.31)
Hepatotoxicity	104 (32.00)	221 (68.00)

Our previous study investigated the correlation of *eIF3a* polymorphisms and platinum-based chemotherapy resistance and severe toxicity in the Chinese population^[8, 9]. Olson *et al* found that *eIF3a* rs10787899 was significantly correlated with breast cancer risk in Caucasian female populations^[31]. In the current study, we further identified 5 polymorphisms which correlated with lung cancer platinum-based chemotherapy toxicities in the Chinese population. While the mechanisms of the correlation remain unknown, we speculate that these

Table 4. Association of eIF3a SNPs with platinum-based chemotherapy toxicities in the Chinese patients. ^bP<0.05.

SNPs	Leukopenia		Anemia		Thrombocytopenia		Emesis		Hepatotoxicity	
	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P
rs1409314	2.36 (1.23–4.53)	0.01 ^b	2.44 (1.27–4.67)	0.01 ^b	2.44 (1.27–4.67)	0.01 ^b	2.47 (1.29–4.74)	0.01 ^b	1.58 (1.14–2.61)	0.01 ^b
rs4752219	1.64 (1.12–2.39)	0.01 ^b	1.58 (1.09–2.30)	0.02 ^b	1.64 (1.07–2.75)	0.02 ^b	1.55 (1.07–2.25)	0.02 ^b	1.52 (1.02–2.53)	0.02 ^b
rs4752220	1.64 (1.12–2.39)	0.01 ^b	1.58 (1.09–2.30)	0.02 ^b	1.64 (1.07–2.75)	0.02 ^b	1.55 (1.07–2.25)	0.02 ^b	1.52 (1.02–2.53)	0.02 ^b
rs7091672	3.44 (1.32–8.98)	0.01 ^b	2.84 (1.16–6.92)	0.02 ^b	1.60 (1.05–2.68)	0.02 ^b	2.88 (1.18–7.02)	0.02 ^b	1.64 (1.07–2.75)	0.02 ^b
rs10510050	1.01 (0.47–2.21)	0.97	0.78 (0.36–1.69)	0.52	4.34 (1.47–12.81)	0.00 ^b	0.39 (0.18–0.83)	0.01 ^b	0.58 (0.34–0.99)	0.04 ^b
rs10886342	1.13 (0.53–2.42)	0.75	1.31 (0.63–2.74)	0.46	0.51 (0.16–1.65)	0.25	0.89 (0.48–1.66)	0.72	0.95 (0.59–1.53)	0.84
rs11198804	1.42 (0.66–3.06)	0.36	1.21 (0.58–2.53)	0.61	1.05 (0.37–3.04)	0.92	1.40 (0.77–2.57)	0.27	1.06 (0.66–1.69)	0.81
rs2275112	1.24 (0.46–3.37)	0.67	0.85 (0.29–2.52)	0.76	2.11 (0.65–6.87)	0.21	0.91 (0.48–1.82)	0.85	1.27 (0.69–2.34)	0.44
rs10787899	0.50 (0.21–1.21)	0.12	1.06 (0.49–2.27)	0.89	0.78 (0.26–2.34)	0.65	1.26 (0.90–1.77)	0.18	1.39 (0.87–2.25)	0.17
rs4752269	1.48 (0.61–3.61)	0.39	0.85 (0.32–2.29)	0.75	2.24 (0.74–6.81)	0.14	0.68 (0.28–1.66)	0.39	0.84 (0.46–1.56)	0.59

SNPs may be potential genetic markers for drug response and toxicity.

In the present investigation, we extended the SNP selection region to 200 kb upstream and downstream of the *eIF3a* locus to generate a more even distribution. However, some regions selected overlapped with other genes, including family with sequence similarity 45, member A (*FAM45A*) in the *eIF3a* 5' untranslated region (5' UTR) and nanos homolog 1 (*Drosophila*) (*NANOS1*) in the *eIF3a* 3' UTR regions. Indeed, the 5 positively selected SNPs (rs1409314, rs4752219, rs4752220, rs7091672, and rs10510050) were located in the overlap region of *NANOS1* and *eIF3a*. Thus, they may be considered genetic markers for both genes. This point is especially important for *in vitro* functional study, where both genes should be considered in the investigation. It should be noted that some limitations need to be considered when interpreting results of this study. First, although we confirmed the function of these positively selected SNPs in clinical patients, we did not conduct a study *in vitro*. Therefore, how these SNPs affected *eIF3a* function still remains unknown. Another limitation came from our sample size. The current study had a relatively small sample size and lacked confirmation from other validation stage populations. Additionally, we only detected the correlation in the Chinese population. Thus, this result needs to be further confirmed in a larger sample size and in different ethnic populations.

In summary, we identified 5 positively selected *eIF3a* SNPs using the LRH test and confirmed that they were significantly correlated with platinum-based chemotherapy toxicities in Chinese NSCLC patients. Our results imply that this strategy of SNP selection may be an inexpensive, effective way to discovery of SNPs correlated to drug response and toxicities.

Acknowledgements

This work was supported by National High-tech R&D Program of China (863 Program Grant, No 2012AA02A517 and 2012AA02A518), National Natural Science Foundation of China (No 81173129, 81202595, 81373490, and 81273595), and Hunan Provincial Natural Science Foundation of China (No 2015JJ1024).

Author contribution

Ji-ye YIN, Hong-hao ZHOU, and Zhao-qian LIU conceived and designed the experiments. Ji-ye YIN, Xiang-guang MENG, Juan CHEN, and Yi ZHENG performed the experiments. Ji-ye YIN, Xiang-guang MENG, and Rong LIU analyzed the data. Chen-yue QIAN and Xiang-ping LI contributed reagents, materials and analysis tools. Ji-ye YIN and Zhao-qian LIU wrote the paper.

Supplementary Information

Supplementary Tables and Figures are available on the Acta Pharmacologica Sinica's website.

References

- 1 Sonenberg N, Hershey JWB, Mathews MB. Translational control of gene expression. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2000.
- 2 Yin JY, Dong Z, Liu ZQ, Zhang JT. Translational control gone awry: a new mechanism of tumorigenesis and novel targets of cancer treatments. Biosci Rep 2011; 31: 1–15.
- 3 Dong Z, Zhang JT. Initiation factor eIF3 and regulation of mRNA translation, cell growth, and cancer. Crit Rev Oncol Hematol 2006; 59: 169–80.
- 4 Pincheira R, Chen Q, Zhang JT. Identification of a 170-kDa protein over-expressed in lung cancers. Br J Cancer 2001; 84: 1520–7.
- 5 Buratti E, Tisminetzky S, Zotti M, Baralle FE. Functional analysis of the interaction between HCV 5'UTR and putative subunits of eukaryotic translation initiation factor eIF3. Nucleic Acids Res 1998; 26: 3179–87.
- 6 Yin JY, Dong ZZ, Liu RY, Chen J, Liu ZQ, Zhang JT. Translational regulation of RPA2 via internal ribosomal entry site and by eIF3a. Carcinogenesis 2013; 34: 1224–31.
- 7 Yin JY, Shen J, Dong ZZ, Huang Q, Zhong MZ, Feng DY, et al. Effect of eIF3a on response of lung cancer patients to platinum-based chemotherapy by regulating DNA repair. Clin Cancer Res 2011; 17: 4600–9.
- 8 Xu X, Han L, Duan L, Zhao Y, Yang H, Zhou B, et al. Association between eIF3alpha polymorphism and severe toxicity caused by platinum-based chemotherapy in non-small cell lung cancer patients. Br J Clin Pharmacol 2013; 75: 516–23.
- 9 Xu X, Han L, Yang H, Duan L, Zhou B, Zhao Y, et al. The A/G allele of eIF3a rs3740556 predicts platinum-based chemotherapy resistance

- in lung cancer patients. *Lung Cancer* 2013; 79: 65–72.
- 10 Tang K, Wong LP, Lee EJ, Chong SS, Lee CG. Genomic evidence for recent positive selection at the human MDR1 gene locus. *Hum Mol Genet* 2004; 13: 783–97.
 - 11 Chen X, Wang H, Zhou G, Zhang X, Dong X, Zhi L, *et al*. Molecular population genetics of human CYP3A locus: signatures of positive selection and implications for evolutionary environmental medicine. *Environ Health Perspect* 2009; 117: 1541–8.
 - 12 Ding K, Kullo IJ. Molecular population genetics of PCSK9: a signature of recent positive selection. *Pharmacogenet Genomics* 2008; 18: 169–79.
 - 13 Sabeti PC, Reich DE, Higgins JM, Levine HZ, Richter DJ, Schaffner SF, *et al*. Detecting recent positive selection in the human genome from haplotype structure. *Nature* 2002; 419: 832–7.
 - 14 Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; 21: 263–5.
 - 15 Lewontin RC. Annotation: the analysis of variance and the analysis of causes. *Am J Hum Genet* 1974; 26: 400–11.
 - 16 Excoffier L, Lischer HE. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 2010; 10: 564–7.
 - 17 Schulze TG, Chen YS, Akula N, Hennessy K, Badner JA, McInnis MG, *et al*. Can long-range microsatellite data be used to predict short-range linkage disequilibrium. *Hum Mol Genet* 2002; 11: 1363–72.
 - 18 Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 2001; 68: 978–89.
 - 19 JG FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from <http://www.unil.ch/izea/software/fstat.html>. Updated from Goudet (1995). 2001.
 - 20 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; 81: 559–75.
 - 21 International HapMap Consortium. The International HapMap Project. *Nature* 2003; 426: 789–96.
 - 22 Akey JM, Zhang G, Zhang K, Jin L, Shriver MD. Interrogating a high-density SNP map for signatures of natural selection. *Genome Res* 2002; 12: 1805–14.
 - 23 Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, Zhivotovsky LA, *et al*. Genetic structure of human populations. *Science* 2002; 298: 2381–5.
 - 24 Biswas S, Akey JM. Genomic insights into positive selection. *Trends Genet* 2006; 22: 437–46.
 - 25 Bamshad M, Wooding SP. Signatures of natural selection in the human genome. *Nat Rev Genet* 2003; 4: 99–111.
 - 26 Wang Z, Wang J, Tantoso E, Wang B, Tai AY, Ooi LL, *et al*. Signatures of recent positive selection at the ATP-binding cassette drug transporter superfamily gene loci. *Hum Mol Genet* 2007; 16: 1367–80.
 - 27 Li XP, Yin JY, Wang Y, He H, Li X, Gong WJ, *et al*. The ATP7B genetic polymorphisms predict clinical outcome to platinum-based chemotherapy in lung cancer patients. *Tumour Biol* 2014; 35: 8259–65.
 - 28 Wang Y, Yin JY, Li XP, Chen J, Qian CY, Zheng Y, *et al*. The association of transporter genes polymorphisms and lung cancer chemotherapy response. *PLoS One* 2014; 9: e91967.
 - 29 Wang Y, Li XP, Yin JY, Zhang Y, He H, Qian CY, *et al*. Association of HMGB1 and HMGB2 genetic polymorphisms with lung cancer chemotherapy response. *Clin Exp Pharmacol Physiol* 2014; 41: 408–15.
 - 30 Chaudhuri J, Chakrabarti A, Maitra U. Biochemical characterization of mammalian translation initiation factor 3 (eIF3). Molecular cloning reveals that p110 subunit is the mammalian homologue of *Saccharomyces cerevisiae* protein Prt1. *J Biol Chem* 1997; 272: 30975–83.
 - 31 Olson JE, Wang X, Goode EL, Pankratz VS, Fredericksen ZS, Vierkant RA, *et al*. Variation in genes required for normal mitosis and risk of breast cancer. *Breast Cancer Res Treat* 2010; 119: 423–30.