

# The Glu69Asp Polymorphism of *EME1* Gene is Associated with an Increased Risk of Hepatocellular Carcinoma in Guangxi Population, China

Youxin Wang<sup>1,\*</sup>, Xinglei Huang<sup>1,\*</sup>, Zhaohui Su<sup>1,\*</sup>, Junquan He<sup>1</sup>, Na Zhao<sup>1</sup>, Liyun Nie<sup>1</sup>, Yanmei Tang<sup>1</sup>, Huiliu Zhao<sup>2</sup>, Qingqing Nong<sup>1,3</sup>

<sup>1</sup>Department of Environmental Health, School of Public Health, Guangxi Medical University, Nanning, 530021, People's Republic of China;

<sup>2</sup>Department of Clinical Laboratory, The Affiliated Tumor Hospital of Guangxi Medical University, Nanning, 530021, People's Republic of China;

<sup>3</sup>Guangxi Colleges and Universities Key Laboratory of Prevention and Control of Highly Prevalent Diseases, Guangxi Medical University, Nanning, 530021, People's Republic of China

\*These authors contributed equally to this work

Correspondence: Qingqing Nong, Department of Environmental Health, School of Public Health, Guangxi Medical University, Nanning, 530021, People's Republic of China, Tel +86 771-5358146, Fax +86 771-5350823, Email nnqq@gxmu.edu.cn

**Background:** The dysfunction of Essential meiotic endonuclease 1 homolog 1 (*EME1*) can lead to genomic instability and tumorigenesis. Single nucleotide polymorphisms (SNPs) in the *EME1* gene have been reported to be associated with the risk of several cancers, but its association with hepatocellular carcinoma (HCC) has not been investigated. This study aimed to determine the association between *EME1* SNPs and the risk of HCC.

**Methods:** This study included 645 HCC patients and 649 healthy controls from a Guangxi population of Southern China, and genotyped three functional SNPs (Glu69Asp: rs3760413A>C, Ile350Thr: rs12450550T>C, and rs11868055A>G) of the *EME1* gene utilizing the Agena MassARRAY platform.

**Results:** The rs3760413C variant genotypes (AC+CC: Glu/Asp+Asp/Asp) conferred a 1.419-fold risk of HCC compared to the AA (Glu/Glu) genotype (adjusted OR = 1.419, 95% CI = 1.017–1.980), and the allele C increased the risk of HCC in a dose-dependent manner ( $P_{\text{trend}} = 0.017$ ). Moreover, the effects of the rs3760413C variant genotypes were more pronounced in individuals who drank pond/ditch water (adjusted OR = 3.956, 95% CI = 1.413–11.076) than in those who never drank ( $P = 0.033$ ). We further observed that a potential carcinogen microcystin-LR induced more DNA oxidative damages in peripheral blood mononuclear cells from the carriers of rs3760413C variant genotypes than those from the subjects with AA genotype ( $P = 0.006$ ). A nomogram was also constructed combining the rs3760413A>C polymorphism and environmental risk factors for predicting HCC risk with a good discriminatory ability (concordance index = 0.892, 95% CI: 0.874–0.911) and good calibration (mean absolute error = 0.005).

**Conclusion:** Our data suggest that the Glu69Asp missense polymorphism (rs3760413) of *EME1* gene is associated with the risk of HCC, which may be a susceptible biomarker of HCC in the Guangxi population.

**Keywords:** *EME1*, hepatocellular carcinoma, missense polymorphism, pond/ditch water, nomogram

## Introduction

Liver cancer is the sixth most commonly diagnosed cancer and the third leading cause of cancer-related deaths worldwide.<sup>1</sup> China is a major contributor to the global burden of liver cancer, accounting for approximately half of cases and deaths worldwide.<sup>2</sup> The most frequent type of liver cancer is hepatocellular carcinoma (HCC), which accounts for 75–85% of all cases. The etiology of HCC is multifactorial and complex, and epidemiological studies of HCC have established many etiologic factors including alcohol, hepatitis B virus (HBV), hepatitis C virus (HCV), aflatoxin B1, microcystin-LR (MC-LR), and various carcinogens.<sup>3–5</sup> HBV infection, exposure to aflatoxin B1 or microcystins (MCs)-contaminated drinking water has been reported to contribute to the unusually high incidence of HCC in the Guangxi

Province of China.<sup>6,7</sup> These factors can potentially modify the DNA structure and cause genomic instability, thus leading to a high risk of HCC.<sup>8,9</sup> Efficient and accurate repairs of DNA damage could maintain the stability of the genome. However, deficiencies in the repair of DNA double-strand breaks (DSBs) might sensitize carcinogens, which can easily induce chromosomal aberrations, lead to genomic instability, and ultimately cause cancer development.<sup>10</sup>

In humans, there are two different pathways in the repair of DSBs, including the homologous recombination (HR) and non-homologous end-joining (NHEJ) pathways.<sup>11</sup> HR is the main pathway involved in the repair of DSBs caused by the replication machinery in eukaryotes, which synthesizes across a single-strand break or an unrepaired lesion, leading to the collapse of the replication fork through several DNA repair molecules, such as essential meiotic endonuclease 1 homolog 1 (EME1). EME1 is an important member of DNA repair genes, and its encoded protein is involved in the HR repair process of DSBs. EME1 protein binds to methyl methanesulfonate-sensitive UV-sensitive 81 (MUS81) protein to form a heterodimeric complex (MUS81-EME1), which is a structure-selective endonuclease that plays a critical role in the resolution of recombination intermediates during DNA repair after inter-strand cross-links, replication fork collapse, or DNA DSBs.<sup>12–14</sup>

The human EME1 gene is located on chromosome 17q21.33 and contains 9 exons with a total length of 20 kb. This gene encodes a protein consisting of 583 amino acids with a full length of 65 kDa.<sup>15</sup> The EME1 protein consists of a central nuclease domain, two repeated C-terminal helix–hairpin–helix (HhH) motifs, a connecting helix, and a 36-residue flexible domain linker composition, which is essential for DNA recognition.<sup>16</sup> Currently, it is believed that the differences in the DNA repair ability of different individuals are determined by genetics, probably due to the combined effects of single nuclear polymorphisms (SNPs) of some core genes in the DNA repair pathway and environmental factors. Recent studies found that the SNPs in EME1 were associated with elevated risk of glioblastoma<sup>17</sup> and pediatric brain tumors.<sup>18</sup> Most of the analyses focused on a common missense polymorphism resulting in the amino acid change (isoleucine to threonine, Ile350Thr). Zhao et al reported that the exon variant Ile350Thr of EME1 was significantly associated with an increased risk of early onset of breast cancer in southern Chinese women.<sup>19</sup> However, no study has yet tested the associations between genetic variants in EME1 gene and the risk of HCC.

In this study, we performed a case-control study that included a Chinese Han or Zhuang population of 645 HCC cases and 649 cancer-free controls to investigate the association between three SNPs (Glu69Asp: rs3760413A>C, Ile350Thr: rs12450550T>C, and rs11868055A>G) of the EME1 gene and the risk of HCC in the Guangxi Province of China. Furthermore, we also developed a risk prediction nomogram of HCC based on EME1 SNPs and environmental risk factors.

## Materials and Methods

### Study Subjects

A total of 645 patients with histopathologically proven HCC confirmed by at least two pathologists (Mean age: 50.1 ± 10.5 years old; 547 males and 98 females) were recruited between 2018 and 2021 in the Affiliated Tumor Hospital of Guangxi Medical University, and 649 sex and age frequency-matched cancer-free individuals (mean age: 49.2 ± 10.6 years old; 541 males and 108 females) were enrolled as a control group in the same period. The patients were all newly diagnosed with HCC, but had not been treated with radiation, chemotherapy, or surgical therapy before enrollment in the study. All the participants were ethnically Han or Zhuang Chinese who have lived in Guangxi for more than 10 years and were excluded from this study if they were affected by other types of cancer. All participants signed a written informed consent, then donated 5 mL of blood, and were interviewed using a structured questionnaire to collect their personal information, such as alcohol drinking, tobacco smoking, family history of cancer, drinking pond/ditch water, HBV and HCV infection, amongst others. After the survey, to ensure the authenticity and accuracy of the information collected by the questionnaires, this study arranged for reviewers to randomly select 5% of the participants from the survey questionnaires for a questionnaire return visit. The results of the return visit were basically the same as those of the previous questionnaire. Participants who had cumulatively smoked at least 100 cigarettes in their lifetime were considered as smokers, otherwise as non-smokers.<sup>21</sup> Similarly, participants who had consumed alcohol at least once a week for more than one year were defined as drinkers of alcohol, otherwise as non-drinkers of alcohol.<sup>20,21</sup> A pond/ditch was defined as a water storage place, approximately 5–6 m in width, 100 m in length, and 1–2 m in depth, usually located close to residential areas.<sup>6</sup> Participants who had used water from ponds/ditches as the source of drinking water for

20 years or more were defined as drinkers of pond/ditch water.<sup>6</sup> This study was approved by the Medical Ethics Committee of Guangxi Medical University (GXMU-20160303-9), and the study was conducted in accordance with the Declaration of Helsinki.

## SNP Selection

By comprehensively considering the functional regions where the SNPs are located, potential biological functions, the Minor Allele Frequency (MAF), and reviewing existing studies in the literature, three putative functional SNPs related to the DNA repair gene EME1 were finally selected. After searching the NCBI dbSNP database (<http://www.ncbi.nlm.nih.gov/>), they were found to all have a MAF > 5% in the Chinese Han population. These three SNPs are rs11868055A>G, located in the 2000 bp upstream promoter of the EME1 gene, and Glu69Asp and Ile350Thr, located in the coding region of the EME1 gene.

## Genotype Analysis

The EME1 polymorphisms were genotyped by using a matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometer, and the design of primers for PCR and extension was performed using the software MassARRAY Assay Design 3.1 (Agena Bioscience, San Diego, CA). For rs3760413A>C, genomic DNA was amplified using the primers 5'-ACGTA CGAAC ACACG CACAC TTACA CACAC-3' (forward) and 5'-ACGTA CGAAC AGATA CACAA GCCTC CACTC-3'(reverse), and the single-base extension reaction was promoted using the UEP primer 5'-GGGTT TACAC ACACA GTTTC AGCTA T-3'. For rs12450550T>C, the primers were 5'-ACGTA CGAAC CACAA TCACC AGACA CAGAG-3' (forward) and 5'-ACGTA CGAAC AAGGG AAGGA AACGC TTCAG-3'(reverse), and the UEP primer in the single-base extension reaction was 5'- CCCTC ACCAC TCTTA CCACA C-3'. For rs11868055A>G, the primers were 5'-ACGTA CGAAC CTTAC CTCTT CACCC TACTC-3' (forward) and 5'-ACGTA CGAAC GGTTT CACCC TAAGC AACAC-3'(reverse), and the UEP primer in the single-base extension reaction was 5'-TCTCC TCCTC AAAGT AAAAC GT-3'. The specific method has been described previously.<sup>6</sup>

## Bioinformatic Analysis

To support the biological plausibility of the variant, we performed bioinformatics analysis with the UniProt database (<https://www.uniprot.org/uniprot/Q96AY2>) to explore the potential function of the SNP in the EME1 gene. Also, expression quantitative trait loci (eQTL) were checked with Genotype-Tissue Expression (GTEx) project (<https://gtexportal.org/home/>), and certain tissues were selected to present normalized effect size and *P* value.

## Detection of Oxidative DNA Damage

To further explore the contents of DNA damage caused by MC-LR in individuals with different EME1 rs3760413 genotypes, we determined the basal and MC-LR-induced DNA oxidative damage by 8-hydroxy-2'-deoxyguanosine (8-OHdG) analysis. An additional 150 control participants were recruited to donate 6 mL of peripheral blood for isolating peripheral blood mononuclear cells (PBMCs), which were subsequently used to evaluate MC-LR exposure sensitivities. The experimental process of PBMCs isolation has been described in a previous study.<sup>6</sup> The PBMCs were treated with 3 µg/mL MC-LR (treated group) or medium alone (untreated group) for 24 h, and then the content of 8-OHdG was assayed by the kits obtained from Nanjing Jiancheng Biology Engineering Institute (Jiangsu, China). The exact 8-OHdG value of each participant was expressed as the 8-OHdG value of MC-LR-treated group minus the 8-OHdG value of the untreated group.

## Statistical Analysis

The Hardy-Weinberg equilibrium (HWE) was tested by a goodness-of-fit chi-square test to compare the expected genotype frequencies with observed genotype frequencies in controls. Student's *t*-test and chi-square ( $\chi^2$ ) tests were performed on selected variables in the case and control groups to determine their significance. Potential associations between genotype and HCC risk were examined by unconditional logistic regression under respective codominant, dominant, and recessive genetic models. Crude and adjusted odds ratios (ORs) and their 95% confidence intervals (CI) were analyzed by unconditional logistic regression unadjusted or adjusted for age, sex, ethnicity, family history of cancer, tobacco smoking, alcohol drinking, HBV

infection, and drinking pond/ditch water. The logistic model was also used for the trend test of variant genotypes of the EME1. Student's *t*-test and one-way analysis of variance (ANOVA) test were used to examine the difference in the levels of 8-OHdG between the different subgroups. Data analysis was performed using SPSS version 23.0 (IBM, Chicago, USA) software and GraphPad Prism 7 (GraphPad Software Inc., California, USA). Furthermore, the sample size and statistical power were calculated using the PASS software (<https://www.ncss.com>). The heterogeneity of different subgroups was calculated by Stata 14.0 software (StataCorp, Texas, USA). A two-sided *P* value < 0.05 was considered statistically significant.

Selected variables were incorporated in the nomogram to predict the probability of HCC risk using statistical software (rms in R, version 4.1.3; <http://www.r-project.org>). For allocating points in the nomogram, regression coefficients were applied to each individual observation to define the linear predictor. The model performance was evaluated based on the predictive accuracy for individual outcomes (discriminating ability) and the accuracy of point estimates (calibration). The model was validated using bootstrapped resampling.

## Results

### Basic Characteristics of Study Subjects

The basic characteristics of the 645 HCC cases and 649 healthy controls are shown in Table 1. There were no significant differences in the distribution of age, sex, ethnicity, Body Mass Index (BMI), tobacco smoking, alcohol drinking, HCV infection, and history of diabetes between cases and controls (*P* > 0.05 for all). However, the proportions for family history of cancer, HBsAg, HBeAg, HBV-DNA and drinking pond/ditch water were significantly different between the cases and controls (*P* < 0.001 for all).

### Distributions of the Genotypes of EME1 SNPs and Their Risks with HCC

To determine whether the rs3760413, rs12450550, and rs11868055 polymorphisms of the EME1 gene are associated with HCC, we examined the genotypic distribution of these three polymorphisms of EME1 in both cases and controls. As shown in Table 2, the genotype frequencies of the three SNPs among controls are all in agreement with the Hardy-Weinberg equilibrium (*P* > 0.05 for all). The chi-square test showed that the genotype distribution of rs3760413 was significantly different between the cases and controls (*P* = 0.019). However, for the other two SNPs, no significant differences in genotype distribution were found (*P* > 0.05 for all). We determined the association of different genotype models of the three EME1 SNPs with HCC risk. Under the codominant model, the logistic regression analysis showed that the rs3760413CC (Asp/Asp) genotype conferred a 2.206-fold risk of HCC compared to the common AA (Glu/Glu) genotype (adjusted OR = 2.206, 95% CI = 1.046–4.654, *P* = 0.038), but not the AC (Glu/Asp) genotype (adjusted OR = 1.321, 95% CI = 0.930–1.876, *P* = 0.120). Meanwhile, the trend analysis indicated that the detrimental effect of EME1 rs3760413 to increase HCC risk was exhibited in a C allele dose-dependent manner ( $P_{\text{trend}} = 0.017$ ). Moreover, under the dominant model, the rs3760413C variant genotypes (AC+CC) increased the HCC risk by 41.9% when compared to the AA genotype (adjusted OR = 1.419, 95% CI = 1.017–1.980, *P* = 0.039). However, under the recessive model, no significant association with HCC risk was found (CC vs AA/AC: adjusted OR = 2.028, 95% CI = 0.969–4.242, *P* = 0.061). In addition, no statistical correlations were found in the other two SNPs under any genetic models.

### Stratification Analysis of the Association Between EME1 rs3760413 Variant Genotypes and HCC Risk

We performed stratified analysis to evaluate the effects of other factors on associations between EME1 genotypes and HCC risk. According to the smallest Akaike's information criterion (AIC), the effect of rs3760413A>C best fitted the dominant model, and only the result of rs3760413A>C was presented because the other SNPs had no significant findings. As shown in Table 3, the increased HCC risk caused by the rs3760413C variant genotypes (AC+CC) was more pronounced among individuals who were drinking pond/ditch water (adjusted OR = 3.956, 95% CI = 1.413–11.076) than those who were not ( $P_{\text{heterogeneity}} = 0.033$ ). In contrast, there were no significant differences in the stratified ORs according to sex, age, ethnicity, tobacco smoking, alcohol drinking, HBsAg, and family history of cancer (*P* > 0.05 for all).

**Table 1** Demographic and Selected Variables in HCC Cases and Controls

Variables	Cases (n = 645) N (%)	Controls (n = 649) N (%)	$\chi^2_{/t}$	P <sup>a</sup>
Age (years)				
≤50	339(52.56)	349(53.78)	0.192	0.661
>50	306(47.44)	300(46.22)		
Sex				
Male	547(84.81)	541(83.36)	0.506	0.477
Female	98(15.19)	108(16.64)		
BMI (kg/m <sup>2</sup> ), Mean±SD	23.47±2.94	23.70±3.10	-1.419	0.156
Ethnicity				
Han	381(59.07)	356(54.85)	2.346	0.126
Zhuang	264(40.93)	293(45.15)		
Family history of cancer				
Yes	152(23.57)	46(7.09)	67.776	<0.001
No	493(76.43)	603(92.91)		
Tobacco smoking				
Ever	254(39.38)	239(36.83)	0.895	0.344
Never	391(60.62)	410(63.17)		
Alcohol drinking				
Ever	237(36.74)	214(32.97)	2.025	0.155
Never	408(63.26)	435(67.03)		
HBsAg				
Positive	528(81.86)	81(12.48)	625.023	<0.001
Negative	117(18.14)	568(87.52)		
HBeAg				
Positive	200(31.01)	21(3.24)	176.182	<0.001
Negative	445(68.99)	628(96.76)		
HBV-DNA (copies/mL)				
≤500	278(43.10)	617(95.07)	409.66	<0.001
>500	367(56.90)	32(4.93)		
HCV infection				
Yes	31(4.81)	19(2.93)	3.073	0.080
No	614(95.19)	630(97.07)		
Drinking pond/ditch water				
Ever	145(22.48)	83(12.79)	20.934	<0.001
Never	500(77.52)	566(87.21)		
History of diabetes				
Yes	55(8.53)	46(7.09)	0.931	0.335
No	590(91.47)	603(92.91)		

**Note:** <sup>a</sup>P values for a two-sided  $\chi^2$  test or Student's t-test.

**Abbreviations:** HCC, hepatocellular carcinoma; BMI, Body Mass Index.

**Table 2** Frequency Distribution of Genotypes in EME1 SNPs and Results of Logistic Regression Analysis for Their Associations with HCC Risk

SNPs	Genetic Model	Genotype	Cases (n=645) N (%)	Controls <sup>a</sup> (n=649) N (%)	P <sup>b</sup>	Crude OR (95% CI)	Adjusted OR (95% CI) <sup>c</sup>
rs3760413	Codominant model	AA	401(62.17)	445(68.57)	0.019	1.000(ref.)	1.000(ref.)
		AC	205(31.78)	181(27.89)		1.257(0.988–1.600)	1.321(0.930–1.876)
		CC	39(6.05)	23(3.54)		1.882(1.105–3.205)	2.206(1.046–4.654)
	*Trend test P value					0.006	0.017
	Dominant model	AA	401(62.17)	445(68.57)		1.000(ref.)	1.000(ref.)
		AC+CC	244(37.83)	204(31.43)		1.327(1.055–1.670)	1.419(1.017–1.980)
	Recessive model	AA+AC	606(93.95)	626(96.46)		1.000(ref.)	1.000(ref.)
		CC	39(6.05)	23(3.54)		1.752(1.034–2.968)	2.028(0.969–4.242)
	rs12450550	Codominant model	TT	567(87.91)		581(89.52)	0.477
TC			75(11.63)	67(10.32)	1.147(0.809–1.626)	1.602(0.975–2.630)	
CC			3(0.46)	1(0.16)	3.074(0.319–29.640)	1.731(0.054–55.712)	
*Trend test P value					0.294	0.062	
Dominant model		TT	567(87.91)	581(89.52)	1.000(ref.)	1.000(ref.)	
		TC+CC	78(12.09)	68(10.48)	1.175(0.832–1.660)	1.604(0.981–2.623)	
Recessive model		TT+TC	642(99.54)	648(99.84)	1.000(ref.)	1.000(ref.)	
		CC	3(0.46)	1(0.16)	3.028(0.314–29.187)	1.637(0.051–52.289)	
rs11868055		Codominant model	AA	526(81.55)	546(84.13)	0.201	
	AG		111(17.21)	100(15.41)	1.152(0.857–1.548)		1.224(0.798–1.878)
	GG		8(1.24)	3(0.46)	2.768(0.730–10.490)		3.807(0.669–21.649)
	*Trend test P value				0.137		0.147
	Dominant model	AA	526(81.55)	546(84.13)	1.000(ref.)		1.000(ref.)
		AG+GG	119(18.45)	103(15.87)	1.199(0.898–1.602)		1.293(0.851–1.965)
	Recessive model	AA+AG	637(98.76)	646(99.54)	1.000(ref.)		1.000(ref.)
		GG	8(1.24)	3(0.46)	2.704(0.714–10.240)		3.683(0.650–20.885)

**Notes:** \*Trend test P value for variant genotypes of EME1 SNPs; <sup>a</sup>The observed genotype frequencies of the three SNPs among the control subjects were all in agreement with the Hardy-Weinberg equilibrium ( $P > 0.05$  for all); <sup>b</sup> $\chi^2$  test for differences in the distribution of genotype frequencies between cases and controls; <sup>c</sup>Adjusted in an unconditional logistic regression model that included age, sex, ethnicity, family history of cancer, tobacco smoking, alcohol drinking, HBsAg, and drinking pond/ditch water. **Abbreviations:** SNPs, single nucleotide polymorphisms; HCC, hepatocellular carcinoma.

## Bioinformatic Analysis of the Potential Function of EME1 rs3760413A>C Polymorphism

By the use of UniProt database (<https://www.uniprot.org/uniprot/Q96AY2>), we found that the rs3760413A>C is a functional missense polymorphism, which causes an amino acid change from glutamic acid (Glu) to aspartic acid (Asp) at codon 69 (Glu69Asp). The SNP of rs3760413 was further selected to explore the expression quantitative trait loci (eQTL) in GETx (<https://gtexportal.org/home/>). As shown in Table 4, the SNP of rs3760413 mainly *cis*-regulated the expression of EME1 and C allele reduced the expression of EME1 gene.

## Effects of the EME1 rs3760413A>C Polymorphism on the MC-LR-Induced Oxidative DNA Damage in PBMCs

We recruited 150 additional controls and then investigated whether this rs3760413 polymorphism could affect the responses of PBMCs against MC-LR-induced DNA oxidative damage. As shown in Figure 1, the average level of 8-OHdG in the 6 cases of rs3760413CC carriers was significantly higher than that of 97 cases with AA genotype or 47

**Table 3** Stratified Analysis of the EME1 rs3760413 A>C Genotypes by Selected Variables in HCC Cases and Controls

Variables	Cases (n = 645)		Controls (n = 649)		Adjusted OR (95% CI) <sup>a</sup>	P <sub>Heterogeneity</sub> <sup>b</sup>
	AA N (%)	AC+CC N (%)	AA N (%)	AC+CC N (%)	AC+CC vs AA	
Age (years)						
≤50	210(32.56)	129(20.00)	237(36.52)	112(17.26)	1.487(0.915–2.418)	0.733
>50	191(29.61)	115(17.83)	208(32.05)	92(14.17)	1.321(0.820–2.128)	
Sex						
Male	335(51.94)	212(32.87)	365(56.24)	176(27.12)	1.276(0.880–1.849)	0.216
Female	66(10.23)	32(4.96)	80(12.33)	28(4.31)	2.331(0.967–5.621)	
Ethnicity						
Han	237(36.74)	144(22.33)	240(36.98)	116(17.87)	1.570(0.983–2.509)	0.627
Zhuang	164(25.43)	100(15.50)	205(31.59)	88(13.56)	1.328(0.816–2.163)	
Family history of cancer						
Yes	87(13.49)	65(10.08)	36(5.55)	10(1.54)	2.758(0.940–8.097)	0.190
No	314(48.68)	179(27.75)	409(63.02)	194(29.89)	1.293(0.907–1.843)	
Tobacco smoking						
Ever	150(23.26)	104(16.12)	164(25.27)	75(11.55)	1.415(0.841–2.383)	0.845
Never	251(38.91)	140(21.71)	281(43.30)	129(19.88)	1.515(0.972–2.361)	
Alcohol drinking						
Ever	123(19.07)	114(17.67)	140(21.57)	74(11.40)	1.464(0.837–2.560)	0.937
Never	278(43.10)	130(20.16)	305(47.00)	130(20.03)	1.423(0.930–2.179)	
HBsAg						
Positive	329(51.01)	199(30.85)	57(8.78)	24(3.70)	1.224(0.717–2.089)	0.524
Negative	72(11.16)	45(6.98)	388(59.78)	180(27.74)	1.530(0.996–2.351)	
Drinking pond/ditch water						
Ever	81(12.56)	64(9.92)	68(10.48)	15(2.31)	3.956(1.413–11.076)	0.033
Never	320(49.61)	180(27.91)	377(58.09)	189(29.12)	1.207(0.843–1.728)	

**Notes:** <sup>a</sup>ORs were adjusted for age, sex, ethnicity, family history of cancer, tobacco smoking, alcohol drinking, HBsAg, and drinking pond/ditch water; <sup>b</sup>P<sub>heterogeneity</sub> values were calculated to compare the difference of ORs within each stratum.

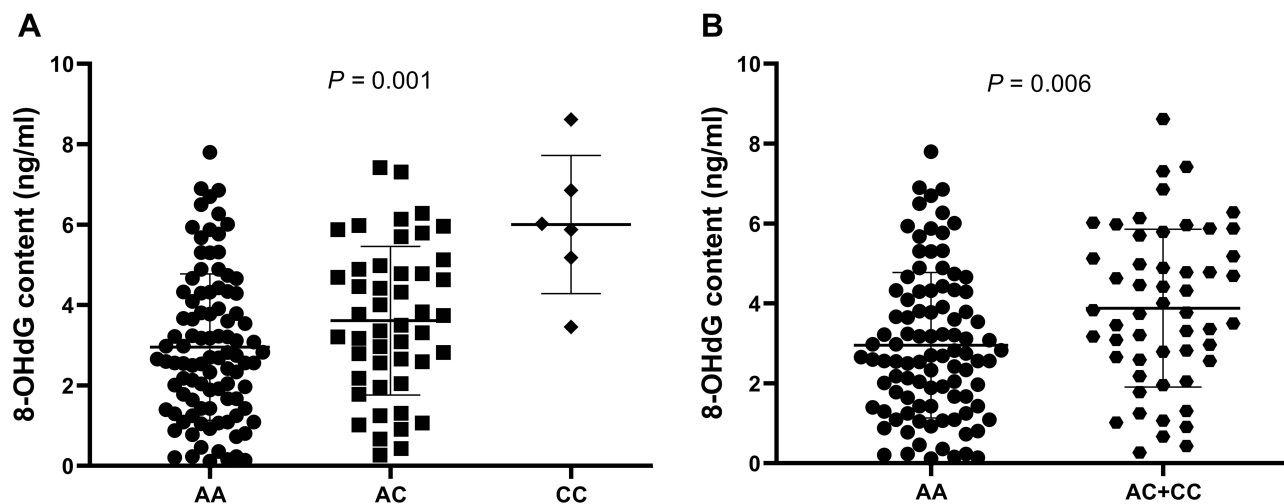
**Abbreviation:** HCC, hepatocellular carcinoma.

**Table 4** Cis-Acting Genes with Transcript Level of SNPs of EME1 Gene in Selected Human Tissues

Genes of Interest	Significant SNPs	Major/Minor Alleles	Cis-Acting Genes Symbol	Tissues	Normalized Effect Size	P value
EME1	rs3760413	A/C	EME1	Brain-Cortex	-1.2	4.3 × 10 <sup>-16</sup>
			EME1	Brain-Anterior cingulate cortex (BA24)	-1.2	1.2 × 10 <sup>-11</sup>
			EME1	Brain-Frontal Cortex (BA9)	-1.0	6.2 × 10 <sup>-11</sup>
			MRPL27	Thyroid	-0.15	3.1 × 10 <sup>-6</sup>
			MRPL27	Lung	-0.12	5.3 × 10 <sup>-6</sup>
			MRPL27	Colon-Sigmoid	-0.23	2.5 × 10 <sup>-5</sup>

**Note:** All the information derived from GTEx (<https://gtexportal.org/home/index.html>).

**Abbreviations:** SNPs, single nucleotide polymorphisms; HCC, hepatocellular carcinoma.



**Figure 1** Effect of the EME1 rs3760413A>C polymorphism on MC-LR-induced DNA oxidative damage in PBMCs from 150 healthy participants. **(A)** Comparison of 8-OHdG contents by the three rs3760413A>C genotypes carriers. **(B)** Comparison of 8-OHdG contents between the AA genotype and the other genotypes (AC+CC) carriers. The differences of the 8-OHdG contents in participants with different genotypes of rs3760413A>C were analyzed by analyzed by Student's *t*-test and one-way ANOVA test.

cases with AC genotype (ANOVA test:  $P = 0.001$ ). Furthermore, the average level of 8-OHdG of the rs3760413C variant genotypes (AC+CC) carriers was much higher than that of AA genotype (Student's *t*-test:  $P = 0.006$ ).

## Predictive Nomogram in HCC Risk Based on rs3760413A>C Polymorphism and Environmental Risk Factors

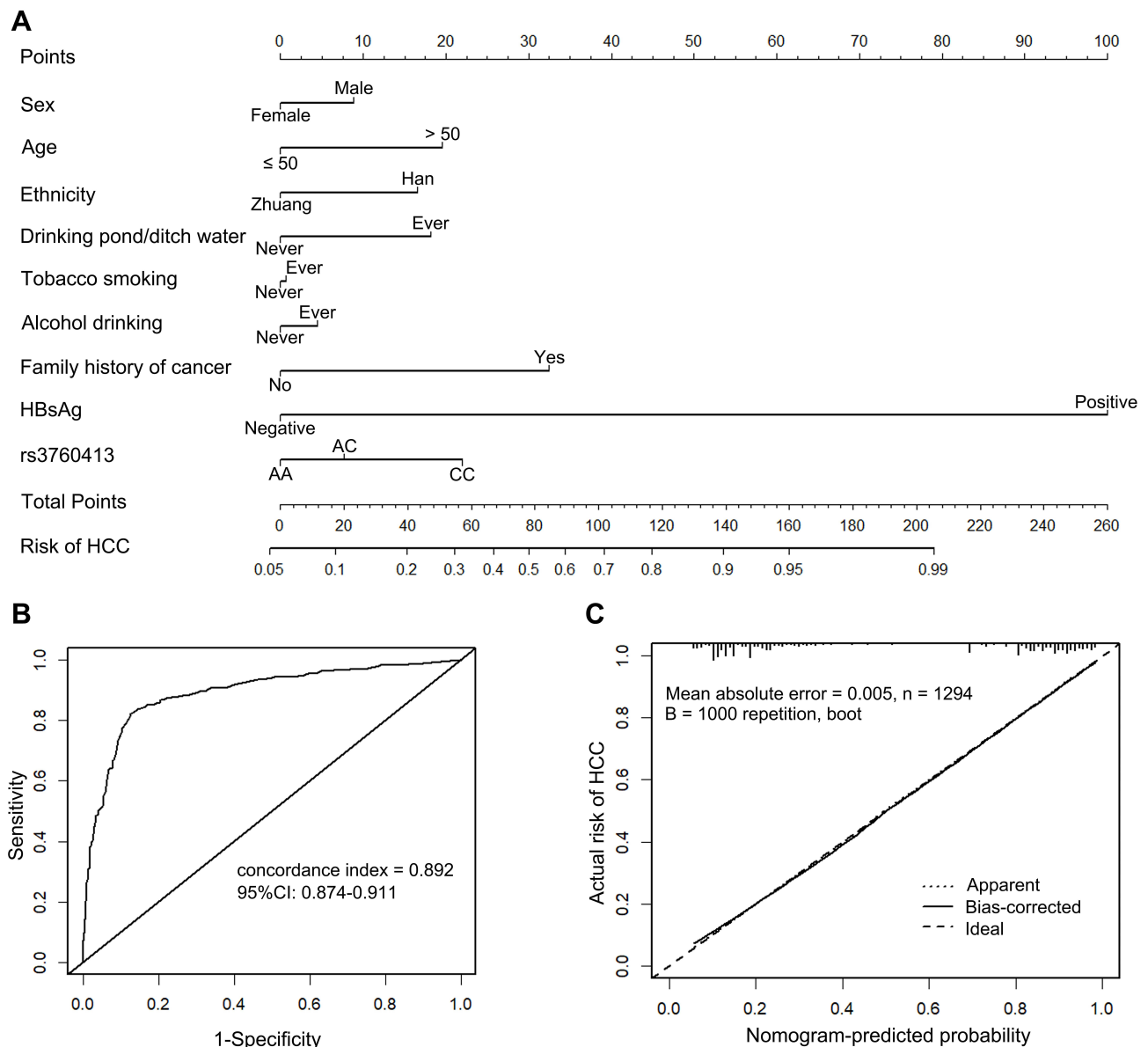
A nomogram calculator is presented in Figure 2. Through univariate analysis and comprehensive consideration of the risk factors of HCC, all the selected parameters (age, sex, ethnicity, rs3760413A>C polymorphism, tobacco smoking, alcohol drinking, HBsAg, family history of cancer, drinking pond/ditch water) were used to develop an individualized risk score to identify HCC. As shown in Figure 2A, higher total points based on the sum of the assigned number of points for each factor in the nomograms were associated with a higher HCC risk. The nomogram demonstrated good accuracy in estimating the risk of HCC, with a concordance index of 0.892 (95% CI: 0.874–0.911), as is displayed in Figure 2B. The calibration curve for the probability of HCC showed good agreement between the prediction of the nomogram and actual observation (mean absolute error = 0.005), and was confirmed through bootstrapping validation (Figure 2C).

## Discussion

In the current hospital-based case-control study of 645 patients and 649 controls, we analyzed the associations between three functional SNPs (rs3760413A>C, rs12450550T>C, and rs11868055A>G) of the EME1 gene and HCC risk. We found that the rs3760413 polymorphism conferred a significantly increased risk of HCC in a C allele dose-dependent manner. The increased risk of HCC was more evident in participants who were drinking pond/ditch water. In contrast, we did not observe significant associations between any other two SNPs in EME1 and HCC risk. To our knowledge, this is the first study to assess the association between EME1 SNPs and HCC risk, especially in individuals who used water from ponds or ditches as the source of drinking water.

EME1 is an essential participator of the HR pathway by being part of the MUS81-EME1, which is essential for the recognition of DNA and DNA DSBs. Deficiency of the EME1 can lead to genomic instability and promote tumorigenesis in the early stage of tumor initiation.<sup>22</sup> Several studies have reported the associations between the EME1 exon variant rs12450550T>C (Ile350Thr) and cancer risks. Adel Fahmideh et al found that EME1 rs12450550 polymorphism confers an increased risk of non-astrocytoma.<sup>18</sup> In addition, Chang et al revealed that EME1 rs12450550 was potentially associated with glioblastoma multiforme susceptibility in the Caucasian population.<sup>17</sup> However, EME1 rs12450550T>C is not always associated with cancer risk. For example, a study found that the rs12450550 variant





**Figure 2** Nomogram predicting HCC risk. **(A)** To calculate the probability of HCC, the points identified on the scale for the 9 variables were summed and a vertical line from the total points scale to the probability scale drawn. **(B)** Concordance index of the nomogram for predicting HCC risk. **(C)** The calibration curve of the nomogram was used to compare actual and predicted outcomes. The diagonal dotted line is a reference line, which represents the perfect prediction of the ideal model. The solid line shows the performance of the nomogram after 1000 bootstrap correction, and the closer it is to the diagonal dotted line, the better the prediction.

genotype was not associated with the risk of glioblastoma multiforme in an Arab Jordanian population.<sup>23</sup> Similarly, our present results also showed that this rs12450550 polymorphism was not associated with HCC risk in the Guangxi population of southern China. The reason for such discrepancies in results may be that our study is only related to Han and Zhuang Chinese.

Interestingly, we revealed that a novel SNP, rs3760413A>C, is associated with HCC risk in a Guangxi population. This rs3760413 locus is in the N-terminal region of EME1. We used UniProt database (<https://www.uniprot.org/uniprot/Q96AY2>) to explore the potential function of the rs3760413, and found that this functional missense polymorphism causes an amino acid change from glutamic acid to aspartic acid at codon 69. Bioinformatics analysis with the Genotype-Tissue Expression (GTEx) project (<https://gtexportal.org/home/>) showed that the SNP of rs3760413 mainly cis-regulated the expression of EME1 and C allele reduced the expression of EME1, which may result in reduced DNA repair capacity of EME1. This is consistent with our observation of rs3760413C variant genotypes with increased the

risk of HCC. After being exposed to environmental carcinogens that damage DNA, the rs3760413C variant genotypes (AC+CC) carriers may exert weaker DNA repair capacity than AA genotype carriers, making them more susceptible to HCC.

In addition, the SNP rs3760413A>C can interact with drinking pond/ditch water to increase HCC risk as that the adverse effect of the rs3760413 variant genotypes was more pronounced in individuals with had ever drunk pond-ditch water. Epidemiological studies have reported that the high incidence of HCC in Guangxi province was partially attributed to the high consumption of pond/ditch water, and much higher levels of MC-LR were observed in pond/ditch water than in tap water.<sup>6,24–26</sup> MC-LR is the most commonly encountered microcystin with strong hepatotoxicity and tumor promotion activity.<sup>27</sup> Several studies have demonstrated that MC-LR can cause DNA oxidative damage in human cells and promote cancer occurrence.<sup>28–31</sup> In the present study, we treated PBMCs with MC-LR and found that the cells from carriers of the rs3760413C variant genotypes (AC+CC) were found to induce more DNA oxidative damages than those from the AA genotype carriers, which further supports the result of this study that the EME1 rs3760413C variants may increase susceptibility to DNA oxidative damage from MC-LR. Thus, our novel finding of gene-environment interaction between the EME1 variant genotypes and drinking pond/ditch water on increasing cancer risk is biologically acceptable. We also constructed a genetic-environment interaction model to predict HCC risk. A predictive nomogram on the SNP rs3760413, drinking pond/ditch water, and other risk factors showed a better predictive capability (concordance index = 0.892,  $P < 0.001$ ).

In the current study, our selection of subjects met the requirement of random sampling in that the genotype frequencies among controls fitted the Hardy-Weinberg disequilibrium. In addition, we achieved over 95% study power (two-sided test,  $\alpha = 0.05$ ) to detect an OR of 1.419 for the rs3760413C variant genotypes compared with the rs3760413AA genotype, suggesting that this finding is noteworthy. However, some limitations still need to be considered. First, the sample size of our study is relatively small. Second, some selection biases were unavoidable because this was a hospital-based case-control study, and our study subjects came only from the Chinese Han or Zhuang population. Moreover, a considerable proportion of patients lacked of clinical stages and cancer differentiation status, thus limiting our analysis on the associations between EME1 variants and clinical features.

## Conclusions

In conclusion, our study found that the Glu69Asp missense polymorphism (rs3760413) of EME1 gene was significantly associated with an increased risk of HCC in Guangxi population, China. Particularly, the association was more pronounced in individuals who used water from ponds/ditches as the source of drinking water. We also constructed a nomogram model based on the Glu69Asp polymorphism and environmental risk factors that can better predict HCC. These findings all suggest that the Glu69Asp polymorphism of EME1 gene may be a susceptible biomarker for HCC in the Guangxi population.

## Institutional Review Board Statement

This study was approved by the Medical Ethics Committee of Guangxi Medical University (GXMU-20160303-9). Written informed consent was obtained from all included subjects.

## Informed Consent Statement

Informed consent was obtained from all subjects involved in the study.

## Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

## Funding

This study was supported by grants from the National Natural Science Foundation of China (No. 81960582 and No. 81660529).

## Disclosure

The authors declare no conflicts of interest in relation to this work.

## References

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021;71(3):209–249. doi:10.3322/caac.21660
2. Im PK, Millwood IY, Kartsonaki C, et al. Alcohol drinking and risks of liver cancer and non-neoplastic chronic liver diseases in China: a 10-year prospective study of 0.5 million adults. *BMC Med.* 2021;19(1):216. doi:10.1186/s12916-021-02079-1
3. Akinyemiju T, Abera S, Ahmed M, et al. The burden of primary liver cancer and underlying etiologies from 1990 to 2015 at the global, regional, and national level: results from the global burden of disease study 2015. *JAMA Oncol.* 2017;3(12):1683–1691. doi:10.1001/jamaoncol.2017.3055
4. Liu Z, Jiang Y, Yuan H, et al. The trends in incidence of primary liver cancer caused by specific etiologies: results from the Global Burden of Disease Study 2016 and implications for liver cancer prevention. *J Hepatol.* 2019;70(4):674–683.
5. Chen T, Liu J, Li Y, Wei S. Burden of disease associated with dietary exposure to aflatoxins in China in 2020. *Nutrients.* 2022;14(5):1027. doi:10.3390/nu14051027
6. Wang Y, Huang Q, Huang X, et al. Genetic variant of PP2A subunit gene confers an increased risk of primary liver cancer in Chinese. *Pharmgenomics Pers Med.* 2021;14:1565–1574. doi:10.2147/PGPM.S335555
7. Xiang X, Qin HG, You XM, et al. Expression of P62 in hepatocellular carcinoma involving hepatitis B virus infection and aflatoxin B1 exposure. *Cancer Med.* 2017;6(10):2357–2369. doi:10.1002/cam4.1176
8. Long XD, Zhao D, Wang C, et al. Genetic polymorphisms in DNA repair genes XRCC4 and XRCC5 and aflatoxin B1-related hepatocellular carcinoma. *Epidemiology.* 2013;24(5):671–681. doi:10.1097/EDE.0b013e31829d2744
9. Diez-Quijada L, Hercog K, Štampar M, et al. Genotoxic effects of cylindrospermopsin, microcystin-LR and their binary mixture in human Hepatocellular Carcinoma (HepG2) cell line. *Toxins.* 2020;12:12.
10. Ui A, Chiba N, Yasui A. Relationship among DNA double-strand break (DSB), DSB repair, and transcription prevents genome instability and cancer. *Cancer Sci.* 2020;111(5):1443–1451. doi:10.1111/cas.14404
11. Ensminger M, Löbrich M. One end to rule them all: non-homologous end-joining and homologous recombination at DNA double-strand breaks. *Br J Radiol.* 2020;93(1115):20191054. doi:10.1259/bjr.20191054
12. Hanada K, Budzowska M, Modesti M, et al. The structure-specific endonuclease Mus81-Eme1 promotes conversion of interstrand DNA crosslinks into double-strands breaks. *EMBO J.* 2006;25(20):4921–4932. doi:10.1038/sj.emboj.7601344
13. Hope JC, Cruzata LD, Duvshani A, Mitsumoto J, Maffahi M, Freyer GA. Mus81-Eme1-dependent and -independent crossovers form in mitotic cells during double-strand break repair in *Schizosaccharomyces pombe*. *Mol Cell Biol.* 2007;27(10):3828–3838. doi:10.1128/MCB.01596-06
14. Gwon GH, Jo A, Baek K, et al. Crystal structures of the structure-selective nuclease Mus81-Eme1 bound to flap DNA substrates. *EMBO J.* 2014;33(9):1061–1072. doi:10.1002/emboj.201487820
15. Ciccía A, Constantinou A, West SC. Identification and characterization of the human mus81-eme1 endonuclease. *J Biol Chem.* 2003;278(27):25172–25178. doi:10.1074/jbc.M302882200
16. Chang JH, Kim JJ, Choi JM, Lee JH, Cho Y. Crystal structure of the Mus81-Eme1 complex. *Genes Dev.* 2008;22(8):1093–1106.
17. Chang JS, Yeh RF, Wiencke JK, et al. Pathway analysis of single-nucleotide polymorphisms potentially associated with glioblastoma multiforme susceptibility using random forests. *Cancer Epidemiol Biomarkers Prev.* 2008;17(6):1368–1373. doi:10.1158/1055-9965.EPI-07-2830
18. Adel Fahmideh M, Lavebratt C, Schüz J, et al. Common genetic variations in cell cycle and DNA repair pathways associated with pediatric brain tumor susceptibility. *Oncotarget.* 2016;7(39):63640–63650. doi:10.18632/oncotarget.11575
19. Zhao J, Liu L, Zhang A, et al. Effect of EME1 exon variant Ile350Thr on risk and early onset of breast cancer in southern Chinese women. *J Biomed Res.* 2013;27(3):193–201. doi:10.7555/JBR.27.20130013
20. Petrick JL, Campbell PT, Koshiol J, et al. Tobacco, alcohol use and risk of hepatocellular carcinoma and intrahepatic cholangiocarcinoma: the liver cancer pooling project. *Br J Cancer.* 2018;118(7):1005–1012. doi:10.1038/s41416-018-0007-z
21. Millwood IY, Li L, Smith M, et al. Alcohol consumption in 0.5 million people from 10 diverse regions of China: prevalence, patterns and socio-demographic and health-related correlates. *Int J Epidemiol.* 2013;42(3):816–827. doi:10.1093/ije/dyt078
22. Abraham J, Lemmers B, Hande MP, et al. Eme1 is involved in DNA damage processing and maintenance of genomic stability in mammalian cells. *EMBO J.* 2003;22(22):6137–6147. doi:10.1093/emboj/cdg580
23. Al-Khatib SM, Abdo N, Al-Eitan LN, et al. The impact of the genetic polymorphism in DNA repair pathways on increased risk of glioblastoma multiforme in the Arab Jordanian population: a case-control study. *Appl Clin Genet.* 2020;13:115–126. doi:10.2147/TACG.S248994
24. Duy TN, Lam PK, Shaw GR, Connell DW. Toxicology and risk assessment of freshwater cyanobacterial (blue-green algal) toxins in water. *Rev Environ Contam Toxicol.* 2000;163:113–185. doi:10.1007/978-1-4757-6429-1\_3
25. Ueno Y, Nagata S, Tsutsumi T, et al. Detection of microcystins, a blue-green algal hepatotoxin, in drinking water sampled in Haimen and Fusui, endemic areas of primary liver cancer in China, by highly sensitive immunoassay. *Carcinogenesis.* 1996;17(6):1317–1321. doi:10.1093/carcin/17.6.1317
26. Xiao CC, Chen MJ, Mei FB, et al. Influencing factors and health risk assessment of microcystins in the Yongjiang river (China) by Monte Carlo simulation. *Peer J.* 2018;6:e5955. doi:10.7717/peerj.5955
27. Zegura B. An overview of the mechanisms of microcystin-LR genotoxicity and potential carcinogenicity. *Mini Rev Med Chem.* 2016;16(13):1042–1062. doi:10.2174/1389557516666160308141549

28. Chen L, Yang S, Wen C, et al. Regulation of microcystin-LR-Induced DNA damage by miR-451a in HL7702 cells. *Toxins*. 2019;11(3):164.
29. Liu W, Wang L, Zheng C, et al. Microcystin-LR increases genotoxicity induced by aflatoxin B1 through oxidative stress and DNA base excision repair genes in human hepatic cell lines. *Environ Pollut*. 2018;233:455–463. doi:10.1016/j.envpol.2017.10.067
30. Jia X, Guan B, Liao J, et al. Down-regulation of GCLC is involved in microcystin-LR-induced malignant transformation of human liver cells. *Toxicology*. 2019;421:49–58. doi:10.1016/j.tox.2019.03.010
31. Wen C, Zheng S, Yang Y, et al. Effects of microcystins-LR on genotoxic responses in human intestinal epithelial cells (NCM460). *J Toxicol Environ Health*. 2019;82(21):1113–1119. doi:10.1080/15287394.2019.1698498

International Journal of General Medicine

Dovepress

## Publish your work in this journal

The International Journal of General Medicine is an international, peer-reviewed open-access journal that focuses on general and internal medicine, pathogenesis, epidemiology, diagnosis, monitoring and treatment protocols. The journal is characterized by the rapid reporting of reviews, original research and clinical studies across all disease areas. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-general-medicine-journal>