Biological Trace Element Research
https://doi.org/10.1007/s12011-025-04515-2

RESEARCH

AS3MT Gene Variant Shows Association with Skin Lesions in an Arsenic Exposed Population of India

Soma Ghosh[1,5] · Arijit Chakraborty[1] · Neelotpal Das[2] · Subhamoy Bhowmick[1,6] · Kunal Kanti Majumdar[3] · Samsiddhi Bhattacharjee[4] · Mouli Mukherjee[1] · Nilabja Sikdar[2,7] · Sreemanta Pramanik[1]

Received: 7 November 2024 / Accepted: 3 January 2025

Abstract

AS3MT, *GSTO2*, and *GSTP1* genes play important roles in the arsenic biotransformation pathway, while *CYP2E1* gene has a prominent role in the metabolic activation of xenobiotics. Hence, polymorphisms of these genes might have an effect on arsenic biotransformation and could impact susceptibility to arsenical skin lesions in individuals of chronic arsenic toxicity. The present case–control study, comprising 148 subjects, attempted to evaluate genetic association between nine polymorphisms of *AS3MT*, *GSTO2*, *GSTP1* and *CYP2E1* genes and arsenical skin lesions in a West Bengal (WB) population. A statistically significant association was found between rs11191439 (*AS3MT*) and arsenical skin lesions (OR = 5.50, *P*-value = 0.01) using logistic regression with age and gender as covariates. Among non-genetic risk factors, age and groundwater arsenic were found to be significantly associated with skin lesions (*P*-value < 0.05). When haplotypes among the intragenic polymorphisms of *AS3MT*, *CYP2E1* and *GSTO2* genes were analyzed, 'ATA' and 'ACG' haplotypes of the *AS3MT* gene showed significant difference between the case and control. Multifactor dimensionality reduction (MDR) analysis was performed on the nine polymorphisms and groundwater and urinary arsenic for studying gene-environment interactions. Strong association was observed between groundwater arsenic and skin lesions relative to the SNPs (*P*-value < 10^{-5}). The best model with maximum testing accuracy included one SNP from the *AS3MT* (rs11191439) and groundwater arsenic (*P*-value < 0.0001). The present study documents the first report about the association of *AS3MT* gene variant with skin lesions in an arsenic exposed population of WB. Presumably, this is also the first study that has used MDR to investigate gene-environment interactions in arsenic-induced toxicity.

Keywords Arsenic · Skin lesion · Genetic association · Polymorphism · Haplotype · Multifactor dimensionality reduction

✉ Soma Ghosh
ghoshsoma1987@gmail.com

✉ Sreemanta Pramanik
sreemanta@gmail.com; sr_pramanik@neeri.res.in

[1] Kolkata Zonal Centre, CSIR-National Environmental Engineering Research Institute, I-8 Sector-C, East Kolkata Township, Kolkata 700107, India

[2] Human Genetics Unit, Indian Statistical Institute, 203 B. T. Road, Baranagar, Kolkata 700108, India

[3] Dept. of Community Medicine, KPC Medical College and Hospital, 1F Raja S. C. Mullick Road, Jadavpur, Kolkata 700032, India

[4] Biotechnology Research and Innovation Council-National Institute of Biomedical Genomics, P.O.: N.S.S, Kalyani 741251, West Bengal, India

[5] Present Address: Dept. of Biotechnology, School of Life Science and Biotechnology, Adamas University, Kolkata 700126, West Bengal, India

[6] Academy of Scientific and Innovative Research (AcSIR), Ghaziabad 201002, India

[7] Estuarine and Coastal Studies Foundation, Howrah 711101, West Bengal, India

Published online: 20 January 2025

Introduction

Arsenic (As) has long been recognized as a potent environmental toxicant and type-I carcinogen [1]. An estimated 230 million people in 108 countries worldwide are at risk of As contamination from groundwater, out of which 180 million are from Asian countries [2]. Groundwater As contamination in West Bengal (WB), India, is very well documented [3, 4], where around 26 million people are at risk of As toxicity through drinking water [5]. Chronic exposure to As, more than the permissible limit of 10 µg/L (WHO, 2018) through drinking water, could lead to adverse health effects in the form of cancerous (e.g. skin, bladder, lung, liver) and non-cancerous outcomes (e.g. skin lesions, vascular diseases, hepatotoxicity, peripheral neuropathy, reproductive abnormalities, diabetes) [6]. Several studies suggest that there is inter-individual variation in susceptibility to the health effects of As and this variation could be influenced by age, gender, nutritional status, arsenic metabolism and genetics [7–13].

Arsenical skin lesions such as melanosis, keratosis, leucomelanosis and pigmentation are considered the earliest manifestations of classical As toxicity that appear with shorter periods of As exposure and could be a precursor for arsenic-induced skin cancers [14]. Several reports from Asian countries have observed a dose-dependent relationship between arsenic exposure and risk of skin lesions [15–17]. Many studies pointed out that only a fraction of people who were exposed to As exhibited arsenical skin lesions [18, 19] and a very minor fraction of individuals with lesions might eventually develop malignancies. Although the exact pathophysiological mechanisms through which arsenic induces cancer remains elusive, enhanced oxidative stress, immune dysregulation and chromosomal abnormalities are thought to be the possible mechanisms of arsenic-induced carcinogenesis [20]. Genetic variations in the key enzymes of the As biotransformation pathway, such as arsenite methyltransferase (*AS3MT*), glutathione S-transferases (*GST*), purine nucleoside phosphorylase (*PNP*) and methylenetetrahydrofolate reductase (*MTHFR*), could contribute in influencing susceptibility to skin lesions induced by As toxicity [10, 12, 16, 21].

Arsenite methyltransferase (*AS3MT*) is an S-adenosyl methionine (SAM)–dependent key enzyme that plays fundamental roles in As metabolism and catalyzes the transfer of a methyl group to As^{III} species during the As biotransformation process [22]. The Met287Thr polymorphism in the *AS3MT* gene has been implicated in arsenical skin lesions [23, 24], cancer [25, 26] and diabetes [27].

Glutathione S-transferases (*GST*s) are a group of enzymes in the phase 2 detoxification pathway that detoxify xenobiotics by conjugation with reduced glutathione. *GSTO1* and *GSTO2* catalyze reduction of methyl arsenic in As metabolism pathway as the rate-limiting enzymes [28]. *GSTP1* plays a role in detoxifying arsenic-based drugs by reacting with cysteine in the presence of glutathione [29]. Chanda et al. reported that *GSTM1* and *GSTT1* gene variants were significantly associated with arsenical skin manifestations in a WB population [30].

Cytochrome P-450 2E1 (*CYP2E1*) is an important enzyme in the metabolic activation of a number of xenobiotics such as tobacco-derived nitrosamines, ethanol, benzene, tolune and some drugs like acetaminophen [31]. RsaI (− 1053 C > T) and PstI (− 1293 G > C) are two of the most studied polymorphisms of the *CYP2E1* gene that can alter transcription rates and enhance enzymatic activity and have been studied in relation to many types of carcinomas like head and neck, cervical, oesophageal, colorectal, gastric, lung and urothelial [32]. A 96-bp insertion/deletion (indel) in the *CYP2E1* regulatory region was correlated with higher transcriptional activity and elevated levels of enzyme activity in alcoholic and obese subjects [33]. So far, no studies have attempted to evaluate the relationship between *CYP2E1* polymorphisms and skin lesions induced by arsenic toxicity.

Haplotype-based genetic association studies could offer vital clues in identifying common haplotypes that can predict disease susceptibility [34]. Many studies on As toxicity around the world have performed haplotype analysis to identify specific haplotypes or haplotype blocks in the populations that can confer or protect against disease risk [12, 35–37]. A protective *AS3MT* haplotype linked with efficient methylation was found to be very common (70%) among the Andes population; however, it was found to be less frequent (17%) among a Bangladeshi population [35].

Multifactor dimensionality reduction (MDR) is a powerful statistical method for detecting gene–gene and gene-environment interactions in a non-parametric manner [38]. Presumably, no studies have so far carried out MDR analysis to detect gene-environment interactions in arsenic-induced toxicity. A few studies have tried to detect gene-environment interactions in As toxicity through logistic regression [39] and linear regression modeling [40].

In the present investigation, we assessed genetic association between nine polymorphisms of *AS3MT*, *GSTO2*, *GSTP1* and *CYP2E1* genes and arsenical skin lesions in a WB population through odd's ratio, logistic regression and haplotype analysis. We also evaluated gene-environment interactions between nine polymorphisms and groundwater and urinary arsenic through MDR analysis.

Materials and Methods

Study Areas

Chakraborti et al. reported in their study that 95% of the blocks in the districts of Nadia, Murshidabad and North 24-Parganas in West Bengal had As concentration of more than 50 µg/L, whereas Purulia, Bankura, Birbhum, East Midnapur and West Midnapur were unaffected with As concentration of less than 3 µg/L [5]. The subjects of the present study were selected from Nadia, an As-affected district, and Paschim Medinipur, an As-unaffected district of West Bengal (WB). Treated piped water has been provided to the local populations of Nadia district by the Government. However, our field survey shows that local people are still using groundwater for drinking and also for their daily activities such as cooking, washing and bathing. The local populations of Paschim Medinipur district are mostly dependent on groundwater as piped water supply are almost non-existent. Prior permission was taken from appropriate administrative authorities for carrying out sampling from these two districts. For the As-affected Nadia district, subjects were recruited from villages under the Chakdaha and Haringhata blocks. For the As-unaffected Paschim Medinipur district, subjects were recruited from villages under the Garhbeta block 1. The subjects chosen from these two districts were of similar socio-economic status.

Ethical Approval, Enrolment and Clinical Examination of Subjects

The Institutional Ethics Committee of CSIR-National Environmental Engineering Research Institute (CSIR-NEERI), Nagpur, India, had accorded the approval to carry out the study (Eth.Com./002/IEC/EISD/05/2019 dated 07/05/2019). An age limit between 18 and 60 years, having characteristic skin lesions or no lesions typically associated with arsenic-induced toxicity and a history of residing in the affected or unaffected areas for at least 5 years were set up as the eligibility criteria. A prior sensitization survey was done in the study areas to ascertain exposure and socio-economic status of the villagers and also to disseminate knowledge among the local populations about the perils of As toxicity through groundwater and foodstuff. A total of 148 eligible individuals were then enrolled as subjects through local health camps. Informed consents were obtained from all the participants who voluntarily agreed to register themselves as subjects for the present investigation after the objectives and benefits of the study were explained to them. Each subject was asked to complete a questionnaire that contained general information, demographic, lifestyle and socio-economic variables, medical history, exposure etc.

The clinical examination of subjects was carried out by a medical doctor who specializes in treating As-affected individuals. Body parameters such as height, weight and age were recorded for all the participants. The subjects were divided into two separate groups, case and control. Subjects manifesting characteristic skin lesions typically associated with As toxicity (e.g. melanosis, keratosis, pigmentation) were selected as case ($n = 50$) from the As-affected Nadia district and were designated as GM-L. On the other hand, subjects manifesting no characteristic skin lesions in spite of living in the As-affected Nadia district were selected as control ($n = 50$) and were designated as GM-WL. As the primary goal of the study was genetic association and gene-environment interaction analyses, we included additional control subjects ($n = 48$) from an As-unaffected district (Paschim Medinipur), designated as GM-C, to increase the power to detect genetic variants and their interactions. The cases and controls were unrelated.

Collection of Groundwater, Urine and Blood Samples

Groundwater samples used for drinking or household purposes were collected from the study areas to assess level of As exposure. A total of 100 ml of groundwater samples was collected from tubewells in polypropylene bottles, acidified with nitric acid and transferred to laboratory to be stored at 4°C [41]. Six milliliters of blood samples was drawn from the enrolled subjects by experienced professionals under the supervision of a physician at the health camps. Blood samples were collected in sterile BD vacutainer EDTA tubes and were immediately stored in ice boxes until transferred to laboratory. Analysis of blood samples has been described later in the sub-section, 'SNPs used in the study and genotyping'. Sterile containers were used to collect spot urine samples from subjects. Around 100 ml of urine was collected and acidified with 0.2% concentrated HNO_3 and stored in an ice box [41]. All the biological samples were preserved in − 80°C freezer until they were used for experimental analyses.

Estimation of Total Arsenic Concentration in Water and Urine Samples

Acidified groundwater and urine samples were thawed at room temperature after taking out from the freezer and filter sterilized through 0.2-µm PTFE membrane filters (Whatmann). The sample preparation and analysis were done according to the protocol mentioned in Biswas et al. [41].

Total arsenic in groundwater and urine was measured using inductively coupled plasma mass spectrometry (ICP-MS, Perkin Elmer, NexION 300X) at the central research facility of NEERI, Nagpur.

SNPs Used in the Study and Genotyping

Eight SNPs from four genes (*AS3MT*, *CYP2E1*, *GSTO2* and *GSTP1*) and one indel (insertion/deletion polymorphism) from the *CYP2E1* gene were selected and used for the study based on literature information and variant types. Supplementary Table 1 lists eight SNPs and one indel (insertion/deletion) and their details such as SNP ID, chromosomal position, risk allele, minor allele frequency (MAF) and types of variants. QIAamp DNA Blood Midi Kit (Qiagen) was used to extract genomic DNA from whole blood. These DNA samples were used then for genotyping of the desired locus either through PCR (for the indel) or PCR–RFLP technique (for the SNPs). Supplementary Table 2 shows the primers that were used to amplify the respective polymorphic loci, annealing temperatures of PCR reaction, restriction enzymes used for genotyping through PCR–RFLP, incubation temperatures for restriction digestion and sizes of allelic fragments [42–45]. The PCR and RFLP products were visualized in agarose gels.

Statistical Analysis of Data

Mean concentrations of groundwater and urinary arsenic were compared between the case and control of As-affected Nadia district using unpaired *t*-test (https://www.graphpad.com). Polymorphisms were evaluated for deviation from the Hardy–Weinberg equilibrium through chi-square test by using an online software tool, Hardy–Weinberg calculator (www.dr-petrek.eu/documents/HWE.xls). The marginal distribution for both genetic and non-genetic variables were obtained separately for cases and controls and then they were tested for equality using either unpaired *t*-test or binomial test depending on the nature of variables. Logistic regression was used to examine the relationship between arsenical skin lesions and non-genetic factors like age, gender, urinary arsenic and groundwater arsenic. Odd's ratio (OR) and 95% confidence interval (CI) for disease risk estimation were calculated by PLINK software v1.07 (whole-genome association analysis toolset (www.harvard.edu)) using allele frequencies of the polymorphisms. To further address confounding due to age and gender, logistic regression was used on each SNP, keeping age and gender as covariates. PLINK software v1.07 was used for analysis of haplotypes for the intragenic SNPs of *AS3MT*, *GSTO2* and *CYP2E1* genes. Multifactorial dimensionality reduction (MDR) analysis was carried out among 11 attributes (nine polymorphisms and groundwater arsenic as environmental factor and urinary arsenic as biomarker) by using software package MDR (http://sourceforge.net/projects/mdr). There were 21 and 14 instances of missing values of groundwater arsenic in cases and controls respectively. Also, there were four instances of missing values of urinary arsenic (two each in cases and controls). These instances were removed from the data to perform MDR. MDR operates by finding the best combination of factors for each *n*th order model. To help us choose the best model among different order models, it outputs three metrics for each *n*th order model—Balanced Accuracy Cross Validation Training, Balanced Accuracy Cross Validation Testing and Cross Validation Consistency. Their formula are described in the following equations. True positives are the total number of cases correctly predicted by the model as diseased individuals, while false positives are the number of cases incorrectly predicted as controls or non-diseased individuals. The opposite of true positives are called true negatives and the opposite of false positives are called false negatives. The number 'k' denotes the number of cross-validation sets, which in our case is 10. The number of cross-validation training sets in which a specific model has the highest training accuracy determines the cross-validation consistency (CVC). Further description of the method can be found in Ritchie et al. (2001) [38] and Chakraborty et al. (2022) [46].

$$\text{Balanced Accuracy} = \left(\frac{\text{True Positives}}{\text{Positives}} + \frac{\text{True Negatives}}{\text{Negatives}}\right)/2$$

$$\text{Bal.Acc.CV Training} = \left(\sum_{i=1}^{k} \text{Bal.Acc}_i\right)/k$$

$$\text{Bal.Acc.CV Testing} = \left(\sum_{j=1}^{k} \text{Bal.Acc}_j\right)/k$$

Results

Demographic Details and Arsenic Concentrations in Groundwater and Urine

A total of 148 individuals between 18 and 60 years were enrolled as subjects in the present study. Table 1 presents the demographic details about the enrolled subjects and concentrations of arsenic in groundwater and urine among the cases ($n=50$) and controls ($n=50$) from Nadia and controls ($n=48$) from Paschim Medinipur districts. Mean age of subjects were (52.32 ± 6.17), (45.68 ± 8.33) and (43.13 ± 9.41) years among the case (Nadia), control (Nadia) and control (Paschim Medinipur) groups respectively. Mean height,

Table 1 Demographic details of subjects and arsenic concentrations in groundwater and urine

Parameter	Case (Nadia Dist) ($n=50$)	Control (Nadia Dist) ($n=50$)	Control (Paschim Midnapore Dist) ($n=48$)
Gender			
Male	33	23	14
Female	17	27	34
Age (year)			
Mean ± SD	52.32 ± 6.17	45.68 ± 8.33	43.13 ± 9.41
Range	40–60	28–60	28–59
Height (meter)			
Mean ± SD	1.59 ± 0.12	1.59 ± 0.08	1.57 ± 0.13
Range	1.14–1.74	1.37–1.80	1.24–1.83
Weight (kg)			
Mean ± SD	53.52 ± 8.57	54.78 ± 9.56	54.69 ± 13.17
Range	35–74	30–72	34–89
BMI (kg/m^2)			
Mean ± SD	21.0 ± 4.26	22.0 ± 3.46	22.17 ± 4.94
Range	14–37	12–31	14–39
GW As (μg/L)			
Mean ± SD	191.92 ± 176.84	155.24 ± 145.90	1.51 ± 1.21
Range	BDL–743.08	8.0–557.76	0.268–5.52
Urinary As (μg/L)			
Mean ± SD	121.17 ± 105.08	116.08 ± 91.14	0.59 ± 0.42
Range	13.7–535.0	14.55–392.0	0.00048–1.73

weight and BMI were similar among the three groups. Mean groundwater As concentrations did not differ much between the case and control of As-affected Nadia district (191.92 ± 176.84 μg/L in the case vs 155.24 ± 145.90 μg/L in the control) and the difference was not statistically significant (two-tailed *P*-value = 0.261). Mean urinary arsenic concentrations also did not differ much among the case and control of As-affected Nadia district (121.17 ± 105.08 μg/L in the case vs 116.08 ± 91.14 μg/L in the control) and the difference was not statistically significant either (two-tailed *P*-value = 0.796). As expected, mean groundwater and urinary arsenic concentrations were drastically lower among the controls of As-unaffected Paschim Medinipur district, i.e. (1.51 ± 1.21) μg/L and (0.59 ± 0.42) μg/L respectively. Majority of the participants in both districts were from the low-income group.

Association of Genetic and Non-genetic Risk Factors with Arsenical Skin Lesions

The allele and genotype frequencies of the nine polymorphisms (eight SNPs and one indel) in the case and control are represented in Supplementary Table 3. The observed genotype frequencies of the nine polymorphisms did not deviate from the Hardy–Weinberg equilibrium except for rs7085104 (*AS3MT*) in the case. HWE P-value for the *CYP2E1* SNP (rs3813867) could not be calculated as the minor allele ('C' allele) was absent in the control.

The marginal distribution of minor allele frequencies of the nine polymorphisms in the case and control is represented in Supplementary Table 4. They were tested for equality using binomial test for proportions. The mean of minor allele frequencies of the two SNPs of the *AS3MT* gene (rs11191439 and rs11191459) was found to be different among the case and control (*P*-value = 0.029 and *P*-value = 0.033 respectively). The marginal distribution of non-genetic variables like age, gender, urinary arsenic and groundwater arsenic was also obtained for case and control as shown in Supplementary Table 5. They were tested for equality using binomial test for proportions or unpaired *t*-test, depending on the nature of the variable. All the four non-genetic risk factors were found to be significantly different among the case and control. As age and gender influenced susceptibility in development of skin lesions after exposure to arsenic according to Lindberg et al. [47], they were treated as confounders and were adjusted in the subsequent statistical analysis.

Table 2 shows the OR with 95% CI among the minor vs major alleles in the case and control for estimation of disease risk. Statistically significant associations were found between rs11191439 (*AS3MT*), rs11191459 (*AS3MT*) and

Table 2 Odd's ratio and confidence interval in minor versus major allele frequencies of nine polymorphisms among the case and control

Gene (SNP ID)	Case ($n=50$)		Control ($n=98$)		Odd's ratio (95% CI)	P-value (χ^2)
	Minor allele (n)	Major allele (n)	Minor allele (n)	Major allele (n)		
AS3MT (rs7085104)	G (31)	A (69)	G (54)	A (142)	1.18 (0.70–2.00)	0.53 (0.38)
AS3MT (rs11191439)	C (09)	T (91)	C (05)	T (191)	3.77 (1.23–12.94)	**0.01** (11.59)
AS3MT (rs11191459)	A (27)	G (73)	A (79)	G (117)	0.55 (0.32–0.92)	**0.02** (5.10)
CYP2E1 [96-bp ins (I)/del (D)]*	I (22)	D (78)	I (62)	D (134)	0.61 (0.34–1.23)	0.08 (3.02)
CYP2E1 (rs3813867)	C (03)	G (97)	C (0)	G (196)	- (-)	**0.03**[#] (-)
CYP2E1 (rs2031920)	T (04)	C (96)	T (02)	C (194)	4.04 (0.72–22.46)	0.08 (2.96)
GSTO2 (rs2297235)	G (21)	A (79)	G (36)	A (160)	1.18 (0.64–2.157)	0.58 (0.29)
GSTO2 (rs156697)	G (37)	A (63)	G (76)	A (120)	0.93 (0.563–1.52)	0.76 (0.08)
GSTP1 (rs1695)	G (29)	A (71)	G (46)	A (150)	1.33 (0.77–2.29)	0.30 (1.07)

*I/I-Insertion homozygous, I/D-Heterozygous, D/D-Deletion homozygous for the *CYP2E1* 96-bp Indel polymorphism

#Fisher's exact test was used to calculate the P-value for the *CYP2E1* (rs3813867) polymorphism as minor allele frequency(MAF) in Control was zero/low

P-value of ≤ 0.05 considered statistically significant and indicated in bold for significant association

Table 3 Association of nine polymorphisms with skin lesions using logistic regression with age and gender as covariates

Gene (SNP)	OR (95% CI)	P-value
AS3MT (rs7085104)	1.19 (0.68–2.07)	0.54
AS3MT (rs11191439)	5.50 (1.60–21.04)	**0.01**
AS3MT (rs11191459)	0.65 (0.36–1.11)	0.12
CYP2E1 [96-bp ins (I)/del (D)]	0.59 (0.30–1.12)	0.12
CYP2E1 (rs2031920)	4.02 (0.53–38.92)	0.19
GSTO2 (rs2297235)	1.15 (0.57–2.30)	0.69
GSTO2 (rs156697)	0.78 (0.44–1.38)	0.4
GSTP1 (rs1695)	1.43 (0.79–2.64)	0.24

CYP2E1 rs3813867 was not included further in analyses as it was close to monomorphic (MAF < 0.03) in our samples

P-value of ≤ 0.05 considered statistically significant and indicated in bold for significant association

Table 4 Association between non-genetic risk factors and skin lesions using logistic regression

Non-genetic factor	OR (95% CI)	P-value
Age	1.10 (1.03–1.19)	**0.01**
Gender	2.70 (0.98–7.81)	0.06
Urinary arsenic	1.0 (0.99–1.00)	0.78
Groundwater arsenic	1.00 (1.00–1.01)	**0.03**

P-value of ≤ 0.05 considered statistically significant and indicated in bold for significant association

rs3813867 (*CYP2E1*) SNPs and risk of arsenical skin lesions. Further, we considered the probability of spurious associations due to sampling of controls from two different geographical locations (Nadia and Paschim Medinipur districts). So, all the polymorphisms were tested for association with the population of origin among the controls from two geographical locations. rs11191459 (*AS3MT*) and rs2297235 (*GSTO2*) were found to be associated with population labels (Supplementary Table 6) and, hence, the association of rs11191459 of *AS3MT* was not pursued further. For the *AS3MT* rs11191439 SNP, we could predict an increased risk of arsenical skin lesions for the 'C' allele as OR was greater than 1 (OR = 3.77) and P-value was less than 0.05 (P-value = 0.01) (Table 2). The 96-bp ins/del and rs2031920 SNP of the *CYP2E1* gene also showed small P-values suggesting potential association (P-value < 0.1).

All nine polymorphisms were tested further for association with skin lesions using logistic regression with potential confounders, age and gender as covariates. The results are shown in Table 3. Only rs11191439 (*AS3MT*) was found to be statistically significant (OR = 5.50, P-value = 0.01). In addition, this particular SNP was tested for interaction with groundwater arsenic using logistic regression and the main effect was found to be statistically significant (P-value = 0.004), while the interaction effect showed suggestive significance (P-value = 0.08). We further performed age-stratified association analysis with rs11191439 (*AS3MT*) polymorphism using Cochran-Mantel–Haenszel (CMH) test on three age categories, i.e. (27–40), (40–53) and (53–60), and found a statistically significant association (chi-square = 3.8112, P-value = 0.050). Thus, association of rs11191439 (*AS3MT*) with skin lesions was still significant after adjusting for age stratification as like in logistic regression (Table 3).

We also studied the associations between non-genetic risk factors, age, gender, urinary arsenic and groundwater arsenic and skin lesions using logistic regression (Table 4). All the factors except urinary arsenic and gender were found to be statistically significant. Age had the highest risk factor associated with arsenic toxicity (OR = 1.10, P-value = 0.01), followed by groundwater

arsenic (OR = 1.00, *P*-value = 0.03). Gender showed suggestive significance with high OR and *P*-value close to 0.05 (OR = 2.70, *P*-value = 0.06). All the association metrics are provided in the Table 4.

Haplotypes Among the Intragenic Polymorphisms of AS3MT, CYP2E1 and GSTO2 Genes

In the marginal genetic association tests (Table 2), polymorphisms of the *AS3MT*, *CYP2E1* and *GSTO2* showed some evidence of association with risk for skin lesions. So, haplotypes were analyzed for these three genes only. The distribution and frequencies of common haplotypes of the cases and controls are presented in Table 5. Thirteen common haplotypes were identified among the case and control (six, three and four haplotypes with three, three and two intragenic SNPs of *AS3MT*, *CYP2E1*, *GSTO2* genes respectively). 'ATA' and 'ACG' haplotypes of the *AS3MT* gene, showing *P*-values of 0.021 and 0.033 respectively, showed significant difference between the case and control (*P*-value ≤ 0.05). The haplotypes for the *GSTO2* gene did not show significant difference between the case and control. For the *CYP2E1* gene, the IGC haplotype showed suggestive significance (*P*-value < 0.1) between the case and control. AA, DGC and ATG were the most commonly observed haplotypes for the *GSTO2*, *CYP2E1* and *AS3MT* genes, respectively, among both the case and control.

Table 5 Distribution and frequencies of common haplotypes in the intragenic polymorphisms of *AS3MT*, *CYP2E1* and *GSTO2* among the case and control groups

Gene (SNP)	Haplotype	Frequency (Case)	Frequency (Control)	*P*-value
AS3MT (A1-A2-A3)	GTA	0.146	0.166	0.654
	ATA	0.124	0.237	**0.021**
	GCG	0.035	0.013	0.2
	ACG	0.055	0.012	**0.033**
	GTG	0.129	0.096	0.395
	ATG	0.511	0.475	0.555
CYP2E1 (C1-C2-C3)	DGT	0.031	0.01	0.19
	IGC	0.224	0.316	0.099
	DGC	0.745	0.673	0.211
GSTO2 (G1-G2)	GG	0.187	0.176	0.817
	AG	0.183	0.212	0.561
	GA	0.023	0.007	0.264
	AA	0.607	0.605	0.969

Abbreviations used: A1-rs7085104, A2-rs11191439, A3-rs11191459; C1-96-bp Indel (I/D), C2-rs3813867, C3- rs2031920; G1- rs2297235, G2-rs156697

P-value of ≤ 0.05 considered statistically significant and indicated in bold

MDR Analysis to Check Interaction Between the Genes and Groundwater and Urinary Arsenic

All nine polymorphisms along with groundwater and urinary arsenic were used to study gene-environment interactions in the MDR model. Each SNP for an individual was coded as 0, 1 and 2, based on the number of minor alleles one possessed. Groundwater arsenic was coded as 1 if the arsenic level exceeded the acceptable safe drinking water level in India (10 µg/L); otherwise, it was labelled as 0. In the case of the biomarker, the presence of urinary arsenic above (safe level 50 µg/L) is denoted as 1, while urinary arsenic content less than 50 µg/L is denoted as 0.

The main effects model of the MDR analysis is represented in Table 6 involving all the polymorphisms and groundwater and urinary arsenic. Strong association between groundwater arsenic and skin lesions (*P*-value < 10^{-5}) and significant association with urinary arsenic and skin lesions (*P*-value < 0.05) were observed. Suggestive association was found for the SNPs, rs11191439 (*AS3MT*), rs2031920 (*CYP2E1*) and rs3813867 (*CYP2E1*) (*P*-value < 0.1). Pairwise and higher order interactions up to *n*th order were considered in our model. The summarized results in Table 7 displayed five best models for each interaction order considered, along with their corresponding training and testing accuracy. A higher CVC value suggests greater consistency in the results. In this study, the best model was chosen based on maximizing testing accuracy, which yielded a value of 0.71. The best model included one SNP from the *AS3MT* gene (rs11191439) and the environmental factor, groundwater arsenic.

Figure 1 displays the distribution of subjects in both the case and control for the optimal model. The shaded cells in the figure indicate a higher risk of disease, while the bars on

Table 6 Main effects model of MDR analysis with nine polymorphisms and groundwater and urinary arsenic

Model	Chi-square	Chi-square *P*-value
Groundwater_arsenic	19.2	**1.17E-05**
Urinary_arsenic	4.7	**0.03**
AS3MT.rs11191459.G.A	2.81	0.25
GSTO2.rs156697.A.G	1.24	0.54
GSTP1.rs1695.A.G	1.25	0.53
AS3MT.rs11191439.T.C	2.95	0.09
CYP2E1.96 bp.ins.del	2.02	0.36
AS3MT.rs7085104.A.G	1.74	0.42
CYP2E1.rs2031920.C.T	2.91	0.09
CYP2E1.rs3813867G.C	3.07	0.08
GSTO2.rs2297235.A.G	0.13	0.94

P-value of ≤ 0.05 considered statistically significant and indicated in bold

Table 7 Best model of each order from MDR analysis

Model	Bal. Acc. CV training	Bal. Acc. CV testing	CV consistency
gw_arsenic	0.74	0.66	10/10
AS3MT.rs11191439.C, gw_arsenic	0.76	0.71	5/10
AS3MT.rs7085104.A.G, GSTO2.rs156697.A.G, gw_arsenic	0.79	0.57	3/10
AS3MT.rs7085104.A.G, AS3MT.rs11191459.G.A, CYP2E1.96 bp.ins.del, gw_arsenic	0.85	0.69	8/10
AS3MT.rs7085104.A.G, AS3MT.rs11191459.G.A, CYP2E1.96 bp.ins.del, GSTO2.rs156697.A.G, GSTP1.rs1695.A.G	0.90	0.48	6/10

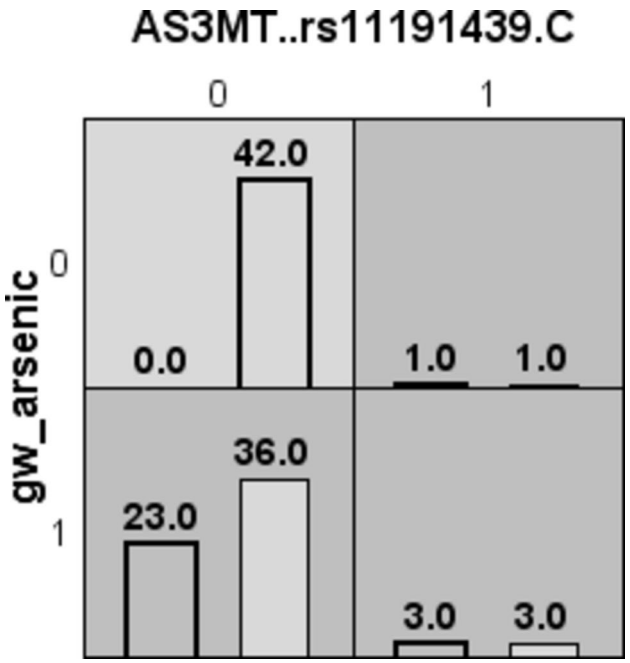

Fig. 1 Distribution of case and control in the best model involving one SNP from the *AS3MT* gene (rs11191439) and groundwater arsenic. The dark cells represented high risk for disease. The left and right bars represented the number of cases and controls respectively

the right signify the number of controls. The model achieved a *P*-value of less than 0.0001, indicating that it is highly unlikely to achieve a testing accuracy of 0.71 or higher under the assumption that there is no association. Figure 2 depicts all the interactions between variables along with their percentages, with different colors indicating different relationships. Green and blue colors denoted redundancy or correlation, yellow represented independence, while red and orange colors indicated a synergistic relationship. The redundant relationships (in blue and green) could be found between CYP2E1.96BP.ins.del and AS3MT.rs11191439, and also, between urinary arsenic and groundwater arsenic. These relationships do not help us predict the skin lesion in an individual. On the other hand, the best synergistic relationship (in red and orange) was found between AS3MT. rs11191439 and groundwater arsenic, followed by AS3MT. rs11191459 and AS3MT.rs7085104.

Discussion

As only a small fraction (15–20%) of people when exposed to arsenic toxicity manifests specific skin lesions and even a very minor fraction of the people with skin lesions eventually develops malignancies, it is strongly believed that genetic variability plays a prominent role in arsenic toxicity as well as carcinogenicity when environmental and other parameters remain the same [48]. Several studies from around the world including India and neighboring Bangladesh have established the link between genetic variations in arsenic metabolism, detoxification, DNA repair, tumor suppressor, and cytokine genes and arsenic-induced toxicity including skin lesions [12, 16, 21, 24, 48–51]. We investigated genetic association between nine polymorphisms of *AS3MT*, *GSTO2*, *GSTP1* and *CYP2E1* genes and arsenical skin lesions in a WB population through a case–control study.

In our study, odd's ratio with 95% CI indicated a statistically significant association between *AS3MT* rs11191439 polymorphism and arsenical skin lesions in the exposed population of WB. An increased risk of arsenical skin lesions could be predicted for the 'C' allele as OR was greater than 1 (OR = 5.50) and *P*-value was less than 0.05 (*P*-value = 0.01). The study by Valenzueala et al. [23] in a Mexican population also observed that the frequency of the 'C' allele of the Met287Thr (rs11191439) polymorphism was higher among the individuals with skin lesions in comparison to the individuals without skin lesions. Furthermore, the individuals harboring the 'C' allele had greater fraction of monomethylated arsenicals in their urine implying increased susceptibility to premalignant arsenical skin lesions. On the contrary, the study by De Chaudhuri et al. [21] did not find any associations between *AS3MT* polymorphisms including the Met287Thr (rs11191439) and arsenical skin lesions in their study population in WB. They found that only the three exonic polymorphisms of the *PNP* among *AS3MT*, *GSTO1*,

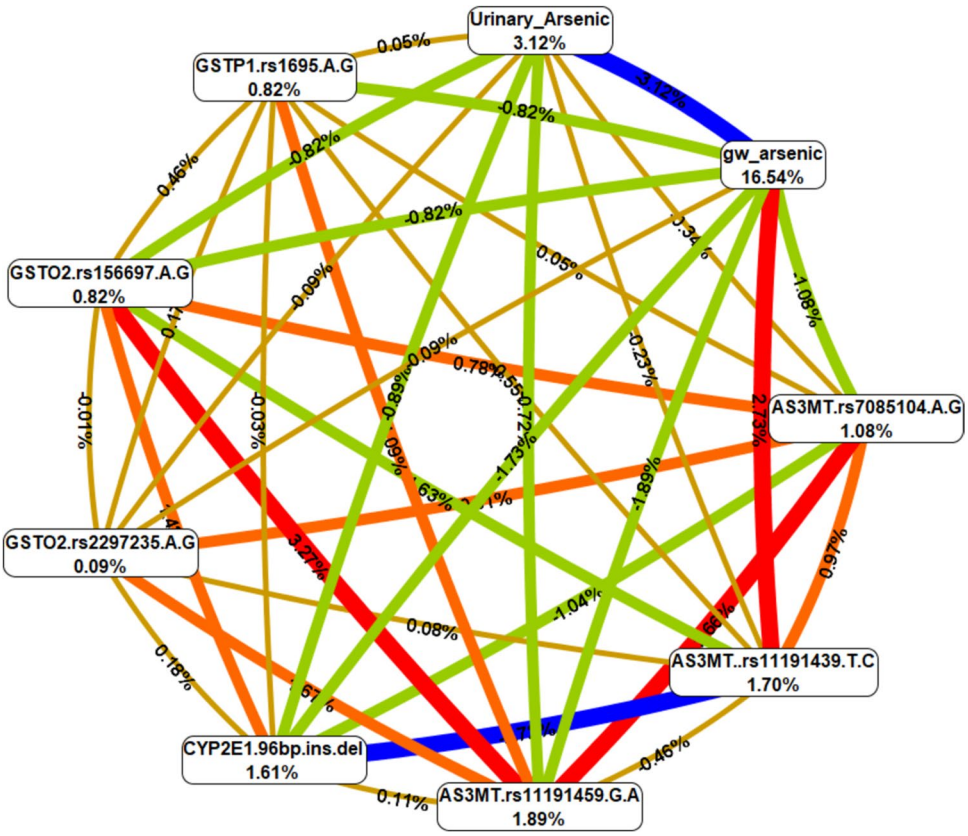

Fig. 2 Interaction graph summarizing the measures of information gain. Circle graph shows interactions along with percentages (green and blue suggest redundancy or correlation, yellow suggests independence, and red and orange suggest there is a synergistic relationship)

GSTO2 and *PNP* genes were significantly associated with arsenical skin lesions. Their study was carried out from the same geographical region as like our study.

Lindberg et al. [47] observed in their study among Bangladeshi men and women that age and gender could influence susceptibility in development of skin lesions after exposure to arsenic. They found that men had a higher risk of developing arsenical skin lesions compared to women because of less efficient methylation of arsenic. They further observed that age at first exposure could be a modifying factor for developing skin lesions later in life. So, we further tested all the nine polymorphisms for association with skin lesions using logistic regression with age and gender as covariates. Only rs11191439 (*AS3MT*) polymorphism was found to be statistically significant in our study (OR = 5.50, *P*-value = 0.01). In addition, when the same SNP was tested for interaction with groundwater arsenic using logistic regression, the main effect was found to be statistically significant. Further, age-stratified association analysis with rs11191439 (*AS3MT*) polymorphism using Cochran-Mantel–Haenszel (CMH) test on three age categories: (27–40), (40–53) and (53–60) produced a statistically significant association (chi-square = 3.8112, *P*-value = 0.050) indicating that the association of rs11191439 (*AS3MT*) with skin lesions was still significant after adjusting for age stratification. We also studied associations between non-genetic risk factors like age, gender, groundwater arsenic and urinary arsenic and skin lesions using logistic regression. We found that age had highest risk factor associated with arsenic toxicity (OR = 1.10, *P*-value = 0.01) followed by groundwater arsenic (OR = 1.00, *P*-value = 0.03). Gender also showed suggestive significance with *P*-value nearly equal to 0.05 (OR = 2.70, *P*-value = 0.057). Our results were similar to Lindberg et al. [47] who used a bigger sample size than our study and found a higher risk of arsenical skin lesions in men compared to women. This is the first instance that an *AS3MT* gene variant (rs11191439) has been found to be associated with arsenical skin lesions in an exposed population of WB. Variants in the *AS3MT* gene could alter capacity to methylate arsenic [52] and significantly increase the percentage fraction of MMAs in urine [23], which in turn could lead to adverse health effects, e.g. skin lesions induced by arsenic toxicity.

Polymorphisms in different members of the GST family have been reported to be associated with urinary arsenic metabolites and arsenic-induced cancers [28, 44]. *GSTO2* rs156697 SNP along with some other polymorphisms of *GSTO1*, *GSTO2* and *PNP* contributed to increased risk of arsenical skin lesions in a Chinese population exposed to high concentrations of As [12]. Our study failed to find any statistically significant association of *GSTO2* rs156697 SNP with risk of arsenical skin lesions in the WB population. The study by De Chaudhuri et al. [21] also did not find any

association between *GSTO2* Asn142Asp SNP (rs156697) and risk of arsenical skin lesions in a WB population. The Ile105Val (rs1695) has been one of the most extensively studied polymorphisms in the *GSTP1* gene. We did not observe any association between rs1695 SNP of the *GSTP1* gene and arsenic-induced skin lesions in the study population of WB. The results were consistent with McCarty et al. [53] who did not observe association between Ile105Val polymorphism and arsenic-induced skin lesions in a Bangladeshi population. Ghosh et al. [48] also did not observe any association between the Ile105Val polymorphism and arsenic-induced skin lesions in a WB population.

There have been only a few studies on associations of *CYP2E1* polymorphisms in relation to As toxicity [44]. Wang et al. [44] investigated the combined effects of environmental exposures including arsenic from drinking water and others like alcohol and smoking and genotypes of *GSTO1*, *GSTO2* and *CYP2E1* on the risk of urothelial carcinoma (UC). They found an elevated risk of UC among the individuals of environmental exposures and risk genotypes. As *CYP2E1* is a key metabolic enzyme in the bioactivation of several procarcinogens, variants in the gene have been also investigated for association with risk of cancer development [54]. Odd's ratio analysis found statistically significant association between rs3813867 SNP of the *CYP2E1* gene and risk of arsenical skin lesions in the WB population in our study. However, no statistically significant association with *CYP2E1* SNPs were found when all nine polymorphisms were tested further for association with skin lesions using logistic regression with age and gender as covariates. So far, no studies have found any potential association between *CYP2E1* gene variants and arsenical skin lesions.

Haplotype-based association studies offer a powerful approach to identify unique chromosomal segments in complex diseases that could potentially harbor disease predisposing genes [34]. As haplotypes account for a number of closely linked marker alleles, they can provide more information than those based on individual markers. 'ATA' and 'ACG' haplotypes of the *AS3MT* gene showed statistically significant difference between the case and control in our study. Although there were no significant differences in the haplotypes for the *GSTO2* gene, the IGC haplotype of the *CYP2E1* gene showed suggestive significance (P-value < 0.1) between the case and control in the present study. While carrying out a study to find association between *AS3MT*, *GSTO1*, *GSTO2* and *PNP* SNPs and arsenic-induced skin lesions among a population exposed to high concentration of As in China, Luo et al. [12] found that 'CT' and 'GCG' haplotypes in the *GSTO1* and *GSTO2* genes, respectively, conferred a significantly high risk of arsenic-induced skin lesions. Haplotype-based association studies on As toxicity have been very few among Indian populations [21]. The study by De Chaudhuri et al. [21] found that three exonic polymorphisms of the *PNP* gene contributed significantly to increased risk of arsenical skin lesions in a WB population. However, no significant difference was found between the distribution of haplotypes for the *PNP* gene among the case and control. Our study results suggest that individuals carrying the 'ATA' and 'ACG' haplotypes of the *AS3MT* gene could pose an increased risk for arsenic-induced skin lesions in the WB population.

MDR is able to detect gene–gene and gene-environment interactions in datasets with categorical independent variables. It is a big challenge to identify gene-environment interactions for understanding susceptibility to complex, multifactorial disease such as hypertension and diabetes. So, identifying these interactions could be the key in understanding disease etiology and develop intervention strategies for targeting the susceptible sub-populations in As toxicity. Groundwater arsenic showed a strong association with skin lesions (P-value $< 10^{-5}$) in the main effects model of MDR. The model with groundwater arsenic and *AS3MT* rs11191439 SNP was found to be the best model with maximum testing accuracy (P-value < 0.0001). So far, to the best of our knowledge, no studies have carried out MDR analysis for detecting gene-environment interactions in relation to arsenic-induced toxicity. A few earlier studies have used logistic regression [39] and linear regression modeling [40] to detect gene–gene and gene-environment interactions in arsenic-induced toxicity. Argos et al. [55] evaluated gene-environment interactions by combining the data of As exposure (through urine analysis), genome-wide SNP genotyping data and clinical phenotypes in relation to arsenic-induced skin lesions among a Bangladeshi population in a two-step approach using regression and linear mixed model for considering binary outcomes. They suggested that their approach could add more power to genome-wide interaction research for enhancing the understanding of the disease etiology. Our study findings obtained through MDR analysis suggest that gene-environment interactions between groundwater arsenic and rs11191439 SNP in the *AS3MT* gene could play a vital role in the development of arsenical skin lesions in the WB population. Also, when genetic variables were tested further for association using logistic regression with potential confounders like age and gender as covariates, only rs11191439 (*AS3MT*) SNP was found to be statistically significant (OR = 5.50, P-value = 0.01) among all the nine polymorphisms. Progress has been limited in identifying gene-environment interactions in epidemiological studies using existing statistical approaches for genome-wide searches [55]. Also, statistical approaches like logistic regression modelling are not flexible enough for dealing with many genetic and environmental factors as number of possible interactions could increase exponentially and this could lead to biased estimates and large standard errors [56].

A major limitation of the present study was its small sample size. Future studies could be planned with increased sample size and inclusion of other variants of the *AS3MT* gene to corroborate the findings of the present investigation in arsenic exposed populations of WB. We would also like to gain more insight on differences of As metabolism between the metabolites of urinary As and its possible relationship with genetic variants of *AS3MT* in our future studies. Genetic association studies of these kinds could help in identifying the high-risk sub-groups within a population in affected areas so that appropriate interventions could be targeted to these sub-groups for prevention and treatment of arsenic-induced toxicities.

Conclusion

In summary, we conducted a case–control study with 148 subjects and evaluated genetic association between nine polymorphisms of *AS3MT*, *GSTO2*, *GSTP1* and *CYP2E1* genes and arsenic-induced skin lesions among an arsenic exposed population of WB, India. A statistically significant association was found between rs11191439 (*AS3MT*) SNP and skin lesions (OR = 5.50, *P*-value = 0.01) using logistic regression with age and gender as covariates. When non-genetic risk factors of age, gender, groundwater arsenic and urinary arsenic were studied for association with skin lesions using logistic regression, age and groundwater arsenic were found to be significantly associated (*P*-value < 0.05). Two haplotypes (namely, 'ATA' and 'ACG') of the *AS3MT* gene showed statistically significant differences between the case and control. Groundwater arsenic showed a strong association with skin lesions (*P*-value < 10^{-5}) in the main effects model of MDR. The model with groundwater arsenic and *AS3MT* rs11191439 SNP was found to be the best model with maximum testing accuracy (*P*-value < 0.0001). The study for the first time reports about the association of rs11191439 (*AS3MT*) SNP with skin lesions in an arsenic exposed population of West Bengal in India. Further, presumably, this is the first study that used MDR to detect gene-environment interactions in arsenic-induced toxicity.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12011-025-04515-2.

Acknowledgements The authors wish to express their sincere thanks to all the subjects who participated in the study. KRC, CSIR-NEERI is duly acknowledged for checking the manuscript through the anti-plagiarism software, iThenticate (CSIR-NEERI/KRC/2023/JUNE/KZC/2 dated 23 June 2023).

Author Contribution SG: Investigation, Validation, Data curation, Supervision, Writing-Review and editing, Project administration and Funding acquisition; AC: Investigation, Data curation and Validation; ND: Software, Formal analysis, Writing-Review and editing; SB: Supervision, Investigation, Writing-Review and editing; KKM: Investigation and Supervision; SB: Software, Formal analysis, Writing-Review and editing; MM: Investigation NS: Formal analysis; SP: Conceptualization, Supervision, Investigation, Formal analysis, Resources, Writing-Original draft, Writing-Review and editing and Project administration.

Funding This work received financial support from the Department of Biotechnology, Government of India (Project No. BT/PR31826BIC/101/1211/2019) and the Department of Biotechnology, Government of India (Project No. BT/PR25850/BCE/8/1423/2017).

Data Availability Datasets generated and/or analyzed during the current study are available from the corresponding authors on reasonable request.

Declarations

Ethics Approval The questionnaire and methodology for this study was approved by The Institutional Ethics Committee of CSIR-National Environmental Engineering Research Institute (CSIR-NEERI), Nagpur, India (Eth.Com./002/IEC/EISD/05/2019 dated 07/05/2019).

Consent to Participate Informed consent was obtained from all participants enrolled in the study.

Consent to Publish All the enrolled participants gave their consent to publish the data included in the study.

Competing Interests The authors declare no competing interests.

References

1. (2012) Humans IWG on the E of CR to Arsenic and arsenic compounds. In: Arsenic, Metals, Fibres and Dusts. International Agency for Research on Cancer
2. Shaji E, Santosh M, Sarath KV et al (2021) Arsenic contamination of groundwater: a global synopsis with focus on the Indian Peninsula. Geosci Front 12(3):101079. https://doi.org/10.1016/j.gsf.2020.08.015
3. Bhowmick S, Pramanik S, Singh P et al (2018) Arsenic in groundwater of West Bengal, India: a review of human health risks and assessment of possible intervention options. Sci Total Environ 612:148–169. https://doi.org/10.1016/j.scitotenv.2017.08.216
4. Das A, Majumder S, Barman S et al (2021) Influence of basin-wide geomorphology on arsenic distribution in Nadia district. Environ Res 192:110314. https://doi.org/10.1016/j.envres.2020.110314
5. Chakraborti D, Das B, Rahman MM et al (2009) Status of groundwater arsenic contamination in the state of West Bengal, India: A 20-year study report. Mol Nutr Food Res 53:542–551. https://doi.org/10.1002/mnfr.200700517
6. Brown KG, Boyle KE, Chen CW, Gibb HJ (1989) A dose-response analysis of skin cancer from inorganic arsenic in drinking water. Risk Anal 9:519–528. https://doi.org/10.1111/j.1539-6924.1989.tb01263.x
7. Vahter M (2002) Mechanisms of arsenic biotransformation. Toxicology 181:211–217. https://doi.org/10.1016/s0300-483x(02)00285-8
8. Lindberg AL, Kumar R, Goessler W et al (2007) Metabolism of low-dose inorganic arsenic in a central European population: influence of sex and genetic polymorphisms. Environ Health Perspect 115:1081–1086. https://doi.org/10.1289/ehp.10026

9. Hernández A, Marcos R (2008) Genetic variations associated with interindividual sensitivity in the response to arsenic exposure. Pharmacogenomics 9:1113–1132. https://doi.org/10.2217/14622416.9.8.1113
10. Antonelli R, Shao K, Thomas DJ, Sams R II, Cowden J (2014) AS3MT, GSTO, and PNP polymorphisms: impact on arsenic methylation and implications for disease susceptibility. Environ Res 132:156–167. https://doi.org/10.1016/j.envres.2014.03.012
11. Minatel BC, Sage AP, Anderson C et al (2018) Environmental arsenic exposure: from genetic susceptibility to pathogenesis. Environ Int 112:183–197. https://doi.org/10.1016/j.envint.2017.12.017
12. Luo L, Li Y, Gao Y et al (2018) Association between arsenic metabolism gene polymorphisms and arsenic-induced skin lesions in individuals exposed to high-dose inorganic arsenic in northwest China. Sci Rep 8:413. https://doi.org/10.1038/s41598-017-18925-3
13. Chi L, Gao B, Tu P et al (2018) Individual susceptibility to arsenic-induced diseases: the role of host genetics, nutritional status, and the gut microbiome. Mamm Genome 29:63–79. https://doi.org/10.1007/s00335-018-9736-9
14. Yu HS, Liao WT, Chai CY (2006) Arsenic carcinogenesis in the skin. J Biomed Sci 13:657–666. https://doi.org/10.1007/s11373-006-9092-8
15. Rahman M, Vahter M, Sohel N et al (2006) Arsenic exposure and age and sex-specific risk for skin lesions: a population-based case-referent study in Bangladesh. Environ Health Perspect 114(12):1847–1852. https://doi.org/10.1289/ehp.9207
16. Ahsan H, Chen Y, Kibriya MG et al (2007) Arsenic metabolism, genetic susceptibility, and risk of premalignant skin lesions in Bangladesh. Cancer Epidemiol Biomarkers Prev 16:1270–1278. https://doi.org/10.1158/1055-9965.EPI-06-0676
17. Chen Y, Parvez F, Gamble M et al (2009) Arsenic exposure at low-to-moderate levels and skin lesions, arsenic metabolism, neurological functions, and biomarkers for respiratory and cardiovascular diseases: review of recent findings from the Health Effects of Arsenic Longitudinal Study (HEALS) in Bangladesh. Toxicol Appl Pharmacol 239(2):184–192. https://doi.org/10.1016/j.taap.2009.01.010
18. Chakraborti D, Rahman MM, Paul K et al (2002) Arsenic calamity in the Indian subcontinent: what lessons have been learned? Talanta 58:3–22. https://doi.org/10.1016/s0039-9140(02)00270-9
19. Argos M, Kalra T, Pierce BL et al (2011) A prospective study of arsenic exposure from drinking water and incidence of skin lesions in Bangladesh. Am J Epidemiol 174:185–194. https://doi.org/10.1093/aje/kwr062
20. Lee CH, Yu HS (2016) Role of mitochondria, ROS, and DNA damage in arsenic induced carcinogenesis. Front Biosci 8(2):312–320. https://doi.org/10.2741/s465
21. De Chaudhuri S, Ghosh P, Sarma N et al (2008) Genetic variants associated with arsenic susceptibility: study of purine nucleoside phosphorylase, arsenic (+ 3) methyltransferase, and glutathione S-transferase omega genes. Environ Health Perspect 116:501–505. https://doi.org/10.1289/ehp.10581
22. Lin S, Shi Q, Nix FB et al (2002) A novel S-adenosyl-L-methionine: arsenic (III) methyltransferase from rat liver cytosol. J Biol Chem 277:10795–10803. https://doi.org/10.1074/jbc.M110246200
23. Valenzuela OL, Drobná Z, Hernández-Castellanos E et al (2009) Association of AS3MT polymorphisms and the risk of premalignant arsenic skin lesions. Toxicol Appl Pharmacol 239:200–207. https://doi.org/10.1016/j.taap.2009.06.007
24 Pierce BL, Kibriya MG, Tong L et al (2012) Genome-wide association study identifies chromosome 10q24. 32 variants associated with arsenic metabolism and toxicity phenotypes in Bangladesh. PLoS Genet 8(2):e1002522. https://doi.org/10.1371/journal.pgen.1002522
25. Beebe-Dimmer JL, Iyer PT, Nriagu JO et al (2012) Genetic variation in glutathione S-transferase omega-1, arsenic methyltransferase and methylene-tetrahydrofolate reductase, arsenic exposure and bladder cancer: a case–control study. Environ Health 11:1–15. https://doi.org/10.1186/1476-069X-11-43
26. Engström KS, Vahter M, Fletcher T et al (2015) Genetic variation in arsenic (+ 3 oxidation state) methyltransferase (AS3MT), arsenic metabolism and risk of basal cell carcinoma in a European population. Environ Mol Mutagen 56:60–69. https://doi.org/10.1002/em.21896
27. Drobná Z, Del Razo LM, García-Vargas GG et al (2013) Environmental exposure to arsenic, AS3MT polymorphism and prevalence of diabetes in Mexico. J Expo Sci Environ Epidemiol 23:151–155. https://doi.org/10.1038/jes.2012.103
28. Paiva L, Hernández A, Martínez V et al (2010) Association between GSTO2 polymorphism and the urinary arsenic profile in copper industry workers. Environ Res 110:463–468. https://doi.org/10.1016/j.envres.2009.12.005
29. Parker LJ, Bocedi A, Ascher DB et al (2017) Glutathione transferase P1–1 as an arsenic drug-sequestering enzyme. Protein Sci 26(2):317–326. https://doi.org/10.1002/pro.3084
30. Chanda S, Roy J, Chakraborty T, Mazumder DG (2020) Skin lesions associated with chronic arsenic exposure through food chain is associated with GST and CYP 450 polymorphism. Int J Chem Environ Sc 1(4):18–29. https://doi.org/10.15864/ijcaes.1402
31. Lieber CS (1997) Cytochrome P-4502E1: its physiological and pathological role. Physiol Rev 77:517–544. https://doi.org/10.1152/physrev.1997.77.2.517
32. Trafalis DT, Panteli ES, Grivas A, Tsigris C, Karamanakos PN (2010) CYP2E1 and risk of chemically mediated cancers. Expert Opin Drug Metab Toxicol 6(3):307–319. https://doi.org/10.1517/17425250903540238
33. McCarver DG, Byun R, Hines RN, Hichme M, Wegenek W (1998) A genetic polymorphism in the regulatory sequences of human CYP2E1: association with increased chlorzoxazone hydroxylation in the presence of obesity and ethanol intake. Toxicol Appl Pharmacol 152:276–281. https://doi.org/10.1006/taap.1998.8532
34. Liu PY, Zhang YY, Lu Y et al (2005) A survey of haplotype variants at several disease candidate genes: the importance of rare variants for complex diseases. J Med Genet 42:221–227. https://doi.org/10.1136/jmg.2004.024752
35. Engström K, Vahter M, Mlakar SJ et al (2011) Polymorphisms in arsenic (+ III oxidation state) methyltransferase (AS3MT) predict gene expression of AS3MT as well as arsenic metabolism. Environ Health Perspect 119:182–188. https://doi.org/10.1289/ehp.1002471
36. Schlebusch CM, Lewis CM Jr, Vahter M et al (2013) Possible positive selection for an arsenic-protective haplotype in humans. Environ Health Perspect 121:53–58. https://doi.org/10.1289/ehp.1205504
37. Schlebusch CM, Gattepaille LM, Engström K et al (2015) Human adaptation to arsenic-rich environments. Mol Biol Evol 32:1544–1555. https://doi.org/10.1093/molbev/msv046
38. Motsinger AA, Ritchie MD (2006) Multifactor dimensionality reduction: an analysis strategy for modelling and detecting gene-gene interactions in human genetics and pharmacogenomics studies. Hum Genomics 2:1–11. https://doi.org/10.1186/1479-7364-2-5-318
39. Pierce BL, Tong L, Argos M et al (2013) Arsenic metabolism efficiency has a causal role in arsenic toxicity: Mendelian randomization and gene-environment interaction. Int J Epidemiol 42:1862–1872. https://doi.org/10.1093/ije/dyt182

40. Gao S, Mostofa MG, Quamruzzaman Q et al (2019) Gene-environment interaction and maternal arsenic methylation efficiency during pregnancy. Environ Int 125:43–50. https://doi.org/10.1016/j.envint.2019.01.042
41. Biswas B, Chakraborty A, Chatterjee D et al (2021) Arsenic exposure from drinking water and staple food (rice): a field scale study in rural Bengal for assessment of human health risk. Ecotoxicol Environ Saf 228:113012. https://doi.org/10.1016/j.ecoenv.2021.113012
42. Sikdar N, Paul RR, Panda CK, Banerjee SK, Roy B (2003) Loss of heterozygosity at APC and MCC genes of oral cancer and leukoplakia tissues from Indian tobacco chewers. J Oral Pathol Med 32(8):450–454. https://doi.org/10.1034/j.1600-0714.2003.00132.x
43. Agusa T, Iwata H, Fujihara J et al (2009) Genetic polymorphisms in AS3MT and arsenic metabolism in residents of the Red River Delta. Vietnam Toxicol Appl Pharmacol 236(2):131–141. https://doi.org/10.1016/j.taap.2009.01.015
44. Wang YH, Yeh SD, Shen KH et al (2009) A significantly joint effect between arsenic and occupational exposures and risk genotypes/diplotypes of CYP2E1, GSTO1 and GSTO2 on risk of urothelial carcinoma. Toxicol Appl Pharmacol 241:111–118. https://doi.org/10.1016/j.taap.2009.08.008
45. Agusa T, Iwata H, Fujihara J et al (2010) Genetic polymorphisms in glutathione S-transferase (GST) superfamily and arsenic metabolism in residents of the Red River Delta Vietnam. Toxicol Appl Pharmacol 242:352–362. https://doi.org/10.1016/j.taap.2009.11.007
46. Chakraborty A, Pramanik S, Datta K et al (2022) Possible association between polymorphisms in ESR1, COL1A2, BGLAP, SPARC, VDR, and MMP2 genes and dental fluorosis in a population from an endemic region of West Bengal. Biol Trace Elem Res 200(11):4641–4653. https://doi.org/10.1007/s12011-021-03072-8
47. Lindberg AL, Rahman M, Persson LA, Vahter M (2008) The risk of arsenic induced skin lesions in Bangladeshi men and women is affected by arsenic metabolism and the age at first exposure. Toxicol Appl Pharmacol 230:9–16. https://doi.org/10.1016/j.taap.2008.02.001
48. Ghosh P, Basu A, Mahata J et al (2006) Cytogenetic damage and genetic variants in the individuals susceptible to arsenic-induced cancer through drinking water. Int J Cancer 118:2470–2478. https://doi.org/10.1002/ijc.21640
49. Meza MM, Yu L, Rodriguez YY et al (2005) Developmentally restricted genetic determinants of human arsenic metabolism: association between urinary methylated arsenic and CYT19 polymorphisms in children. Environ Health Perspect 113:775–781. https://doi.org/10.1289/ehp.7780
50. Engström KS, Broberg K, Concha G et al (2007) Genetic polymorphisms influencing arsenic metabolism: evidence from Argentina. Environ Health Perspect 115:599–605. https://doi.org/10.1289/ehp.9734
51. Das N, Giri A, Chakraborty S, Bhattacharjee P (2016) Association of single nucleotide polymorphism with arsenic-induced skin lesions and genetic damage in exposed population of West Bengal, India. Mutat Res Genet Toxicol Environ Mutagen 809:50–56. https://doi.org/10.1016/j.mrgentox.2016.09.006
52. Drobná Z, Waters SB, Walton FS et al (2004) Interindividual variation in the metabolism of arsenic in cultured primary human hepatocytes. Toxicol Appl Pharmacol 201(2):166–177. https://doi.org/10.1016/j.taap.2004.05.004
53. McCarty KM, Chen YC, Quamruzzaman Q et al (2007) Arsenic methylation, GSTT1, GSTM1, GSTP1 polymorphisms, and skin lesions. Environ Health Perspect 115:341–345. https://doi.org/10.1289/ehp.9152
54. Danko IM, Chaschin NA (2005) Association of CYP2E1 gene polymorphism with predisposition to cancer development. Exp Oncol 27:248–256
55. Argos M, Tong L, Roy S et al (2018) Screening for gene–environment (G× E) interaction using omics data from exposed individuals: an application to gene-arsenic interaction. Mamm Genome 29:101–111. https://doi.org/10.1007/s00335-018-9737-8
56. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR (1996) A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 49:1373–1379. https://doi.org/10.1016/s0895-4356(96)00236-3