A Distinct and Replicable Squamous Cell Carcinoma Gene *INPPA5* Variant Modifies Susceptibility of Arsenic-Associated Skin Lesions in Bangladesh

Wei Jie Seow, M.S., Sc.D.[1], **Wen-Chi Pan, M.S., Sc.D.**[1], **Molly L. Kile, Sc.D.**[2], **Tong Lin, M.S.**[3], **Andrea Baccarelli, M.D., M.P.H., Ph.D.**[1], **Quazi Quamruzzaman, M.D.**[4], **Mahmuder Rahman, M.D.**[4], **Golam Mostofa, M.Com.**[4], **Muhammad Rakibuz-Zaman, M.D.**[5], **Muhammad Kibriya, M.D., Ph.D.**[3], **Habibul Ahsan, M.D.**[3], **Xihong Lin, Ph.D.**[6], and **David C. Christiani, M.D., M.S., M.P.H.**[1,*]

[1]Department of Environmental Health, Harvard School of Public Health, Boston, Massachusetts, USA

[2]College of Public Health and Human Sciences, Oregon State University, Corvallis, Oregon, USA

[3]Department of Health Studies, University of Chicago, Chicago, Illinois, USA

[4]Dhaka Community Hospital, Dhaka, Bangladesh

[5]UChicago Research Bangladesh, Dhaka, Bangladesh

[6]Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts, USA

Abstract

Background—Single nucleotide polymorphisms (SNPs) in inflammation, one-carbon metabolism and skin cancer genes might influence susceptibility to arsenic-induced skin lesions.

Methods—A case-control study was conducted in Pabna, Bangladesh (2001-2003) and drinking water arsenic concentration was measured for each participant. A panel of twenty-five candidate SNPs was analyzed in 540 cases and 400 controls. Logistic regression was used to estimate the association between each SNP and the potential for gene-environment interactions in skin lesion risk adjusting for relevant covariates. Replication testing was conducted in an independent Bangladesh population with 488 cases and 2,794 controls.

Results—In the discovery population, genetic variants in the one-carbon metabolism genes *PEMT* (rs2278952, *P* for interaction = 0.004; rs897453, *P* for interaction = 0.05) and *DHFR* (rs1650697, *P* for interaction = 0.02), inflammation gene *IL10* (rs3024496, *P* for interaction = 0.04), and skin cancer genes *INPP5A* (rs1133400, *P* for interaction = 0.03) and *XPC* (rs2228000, *P* for interaction = 0.01) significantly modified the association between arsenic and skin lesions after adjusting for multiple comparisons. The significant gene-environment interaction between a

Corresponding author: David C. Christiani, MD, MPH, MS, Department of Environmental Health, Harvard School of Public Health, 665 Huntington Avenue, Boston, MA 02115, phone: 617-432-3323, fax number: 617-432-3441; dchris@hsph.harvard.edu.

SNP in *INPP5A* gene (rs1133400) and water arsenic on skin lesion risk was successfully replicated in an independent population (*P* for interaction = 0.03).

Conclusion—Minor allele carriers of skin cancer gene *INPP5A* modified odds of arsenic-induced skin lesions in both main and replicative populations. Genetic variation in *INPP5A* appears to have a role in susceptibility to arsenic toxicity.

Keywords

Arsenic; Environmental Health; Genetic polymorphisms; Skin cancer; Susceptibility

Introduction

Arsenic is classified as a Group 1 human carcinogen by the International Agency for Research on Cancer due to sufficient evidence of carcinogenicity in humans[1] and is associated with cancers of the urinary bladder, lung and skin. Globally, millions of individuals are exposed to arsenic from contaminated water via ingestion and dermal contact. However, Bangladesh is particularly affected by arsenic contaminated groundwater and it is estimated that approximately 73% of the shallow wells used for potable water in this country exceed the World Health Organization's provisional arsenic drinking water limit of 10 μg/L[2]. Arsenic-induced skin lesions are one of the first visible signs of arsenic toxicity [3-5]. The presence of these premalignant skin lesions is highly associated with risk of cancer later in life, especially non-melanoma skin cancers [5-9]. It was previously reported that 71.7% of skin cancer patients had keratosis and 89.7% also had hyperpigmentation [5]. The most common arsenic-induced skin cancers are squamous cell carcinoma (SCC), basal cell carcinoma (BCC) and Bowen's disease [5, 10]. Skin cancer is a complex disease that has many contributing factors. Most cancers are caused by a cascade of accumulated mutations of genes involved in many different pathways over long periods of time [11]. This is most likely also true for arsenic-induced carcinogenesis [12, 13].

The mechanism by which arsenic exposure induces skin lesions remains unclear. Moreover, epidemiological studies have shown that only a fraction of the exposed population develops skin lesions or other arsenic-induced disease [14]. It has been posited that this heterogeneity could be due to individual genetic differences [15]. Previous studies examining the role of genetic susceptibility in arsenic-related skin lesions found that SNPs in glutathione S-transferase (*GST*) genes, methyltransferase (*AS3MT*), X-ray repair cross-complementing group 1 (*XRCC1*) and methylenetetrahydrofolate reductase (*MTHFR*) influence individual susceptibility for skin lesions [16-19].

Identifying genetic loci that increase susceptibility to skin lesions should improve our understanding of the underlying biological mechanisms of arsenic toxicity. In this study, we utilized a hypothesis-driven candidate gene approach that focused on SNPs in inflammation, one-carbon metabolism and skin cancer genes that have previously been shown to be associated with arsenic metabolism and/or skin carcinogenesis. This panel of candidate genes is listed in the supplemental material (Table S1). Our rationale for focusing on these pathways is the *a priori* understanding that inflammation and one-carbon metabolism are involved in carcinogenesis [20, 21]. For example, the inflammatory gene IL10 has been

reported to help human papillomavirus (HPV) evade immune surveillance[22], which together with arsenic-induced immune suppression may increase the risk of arsenic-induced carcinogenesis[23], whereas the one-carbon metabolism gene *DHFR*, or dihydrofolate reductase, was shown to be associated with acute lymphoblastic leukemia (ALL)[24]. Genes that have been previously reported to be associated with skin cancers such as *INPP5A*, or inositol polyphosphate-5-phosphatase, inhibit the proliferation of cells and may function as a tumor suppressor for SCC[25].

In our case-control study, we evaluated the main effects of SNPs and arsenic on skin lesion risk, as well as the interaction between SNPs and arsenic on skin lesions to evaluate differential susceptibility to arsenic exposures by genotype. The main findings were replicated in an independent Bangladesh population.

Material and Methods

Study populations

Main population—In 2001-2003, 900 case-control pairs matched on age and sex were recruited from 23 villages within the Pabna district of Bangladesh by Dhaka Community Hospital (DCH) to examine genetic factors of susceptibility to arsenic-related skin lesions. This study has been previously described in detail [17]. Among these individuals, 940 participants (540 cases and 400 controls) had sufficient quantities of DNA in biostorage for additional genotyping. These participants became the basis for this analysis. Arsenic-induced skin lesion status was assessed by an experienced physician. The physician was blinded to the arsenic concentration in drinking water. Cases were defined as participants who were at least 16 years of age (in 2001) and were diagnosed with one or more types of arsenic-induced skin lesions (melanosis, keratosis or hyperkeratosis). Controls were healthy individuals without visible skin lesions living in the same general community as the cases. The protocol for this study was approved by the Institutional Review Boards at the Harvard School of Public Health and Dhaka Community Hospital. Informed consent was obtained from every participant prior to participation.

Replication population—The replication study population consisted of 488 skin lesion cases and 2,794 healthy controls from Araihazar, Bangladesh, who participated in The Health Effects of Arsenic Longitudinal Study (HEALS) and provided DNA for a genome-wide association study (GWAS) of arsenic-related metabolism and toxicity [26]. In 2000-2002, healthy individuals who were at least 18 years of age, were residents of the study area for at least five years and were drinking water from a local well were recruited. Trained study physicians, who were blinded to the arsenic measurements, recorded the presence/absence of skin lesions (melanosis, keratosis or leukomelanosis) for each individual at each follow-up. The study protocol was approved by the Institutional Review Boards of The University of Chicago, Columbia University, and the Bangladesh Medical Research Council. Informed consent was obtained from all participants.

Questionnaires and interviews

Trained interviewers administered questionnaires to collect socio-demographic information, drinking water history, medical history, lifestyle factors and residential history including identification of the primary water source (tube well). Interviewers in the study were blinded to the participants' arsenic exposure to minimize the potential for bias.

Exposure assessment

Personal drinking water samples were analyzed for arsenic. Water sample from the participant's primary drinking source were collected in a 50ml falcon tube and 0.1ml of pure metal grade nitric acid was added to preserve the samples following the United States Environmental Protection Agency's guidelines for collecting and analyzing natural waters for metals. The samples were stored at room temperature out of direct contact with ultraviolet light prior to analysis. Total arsenic concentration was measured using Environmental Protection Agency (EPA) method 200.8 with Inductively Coupled Plasma Mass Spectroscopy (ICP-MS) (Environmental Laboratory Services, North Syracuse, New York). The limit of detection (LOD) was 1μg As/L. Samples below the LOD were assigned a value of 0.5μg As/L (n = 125). To validate the water arsenic measurements, we also collected and measured toenail arsenic concentrations for each participant using methods as previously described[27].

In the replication population, personal drinking water samples were collected in 50ml acid-washed tubes. Briefly, water arsenic concentrations were analyzed in the Geochemistry Research Laboratory of Columbia University LDEO using graphite furnace atomic absorption (GFAA). Water samples with arsenic concentration below the LOD (5μg As/L) (n = 682) were reanalyzed using ICP-MS [28], which has a detection limit of 0.1μg/L (26). Samples below the LOD were assigned a value of 0.1μg/L (n = 33).

SNPs selection

SNPs were selected based on the following criteria: 1) a minor allele frequency (MAF) of at least 5% based on the publicly available Hapmap population data from Gujarati Indians in Houston, Texas (GIH) who share similar genetics with the Bengali population in Bangladesh (Table S2); and either 2) potentially functional SNPs that are either non-synonymous SNPs, or lie in the 3′-Untranslated Region (UTR) or 5′-UTR, or were as previously reported to be functional; or 3) in previous studies have been found to be associated with either inflammation, one-carbon metabolism, and skin cancers. A total of 25 SNPs from 23 genes were selected for genotyping, of which 3 SNPs were from 3 inflammatory genes, 11 SNPs were from 9 one-carbon metabolism genes, and 11 SNPs were from 11 skin cancer genes (Table S1).

Genotyping

Genomic DNA was extracted from 6mL human whole blood using the Puregene® DNA purification reagents according to the manufacturer's instructions (Genera Systems, Minneapolis). The concentrations of genomic DNA were determined by absorptions at 260nm, and quality checks were done by measuring the ratio of absorbance at 260 and 280. Extracted DNA was genotyped using Sequenom MassARRAY iPLEX genotyping system

(San Diego, CA) by Partners Center for Personalized Genetic Medicine (Boston, MA). Genotyping procedures were validated by repeating 10% of the samples and all SNPs achieved a concordance rate of 100%. For each SNP, the average genotyping call rates were at least 95% (Table S1). SNPs that had genotyping efficiencies of < 90% and failed the Hardy-Weinberg Equilibrium (HWE) test (P < 0.05) were also excluded from the analysis (n = 3).

In the replication population, genomic DNA was extracted from clot blood using Flexigene DNA kit (Qiagen, Venlo, Netherlands), as described previously [26]. Nanodrop 1000 was used to check the concentration and quality of all extracted DNA. Extracted DNA was genotyped using the Illumina Infinium HD SNP array according to Illumina's protocol and read on the BeadArray Reader. SNPs that had call rates of < 95% (n = 103) and HWE p-values of < 10^{-7} (n = 164) were excluded from the analysis, resulting in 259,597 SNPs for 2,879 individuals.

Statistical analysis

Individual characteristics and known risk factors for skin lesions were compared between cases and controls using Fisher's exact test for categorical variables and Wilcoxon rank sum test for continuous variables. Dominant genetic effect models were used in all analyses with dichotomous outcomes for presence or absence of the minor allele. Drinking water arsenic variable was transformed using the natural logarithm. After modeling the concentration-response relationship between arsenic and skin lesions, we noticed a non-linear pattern. We, therefore, opted to use piecewise linear regression with 3 degrees of freedom for water arsenic (knots at 2 and 5) to account for this curvature. Multiple logistic regression models were fitted for each SNP to estimate the main association of each individual SNP within the gene on skin lesions. All models were adjusted for age (continuous), sex (male or female), education (none to primary, secondary to college, or graduate or above), body mass index (continuous), smoking status (never versus ever or current smoker), chewing betel nut (never versus ever or current user) and drinking water arsenic (3 degrees of freedom in piecewise linear fashion), because these variables are risk factors of skin lesions and/or change the coefficients of SNPs by at least 10%. Odds ratio (OR) and 95% confidence interval (CI) were calculated to estimate the magnitude of the association between SNPs and skin lesions.

We added interaction terms as a product between each SNP and arsenic exposure to estimate gene-environment interaction. Likelihood ratio tests with 3 degrees of freedom were used to test the significance of the interaction terms. The number of degrees of freedom was selected based on the curvature of the dose-response relationship. Similar results were seen with additive genetic models. A dose-response figure was plotted using odds ratio estimates from cubic regression splines.

All statistical analyses were carried out using R version 3.1.0 (www.r-project.org/), and SAS version 9.2 (SAS Institute Inc., Cary, NC, USA). A two-sided *P*-value of less than 0.05 was considered statistically significant. The 'qvalue' package in R was used to adjust for multiple comparisons, and we set false discovery rate (FDR) *q*-values of less than 0.10 to reach statistical significance for all analyses [29]. The same analyses were performed on

significant SNPs in an independent Bangladesh population (HEALS) for replication to minimize possible false positives findings.

Results

Participants' demographics and SNPs genotyped

Genotyped skin lesion cases and controls in the main population had similar characteristics (Table 1) except water arsenic levels was significantly higher in cases compared to controls (mean of 208 μg/L versus 61.9 μg/L for water arsenic, *P*-value < 0.001). To check for potential selection bias, we compared the individual characteristics of participants and non-participants, but did not find any significant differences between the two groups [30]. Individual characteristics of the replication population can be found in Supplementary Material, Table S3. Similarly, individuals with skin lesion were more highly exposed to arsenic than controls in the replication population (mean of 122.7 μg/L versus 76.4 μg/L for water arsenic, *P*-value < 0.001).

Main effect association of SNPs with skin lesions

In the main population, four SNPs were found to have nominally significant associations with arsenic-induced skin lesions, after adjusting for water arsenic, age, sex, smoking status, betel nut chewing, education and BMI. Two SNPs passed the 10% FDR cutoff, and they were SNPs located in the inflammation gene *IL10* (rs3024496, adjusted OR = 0.68, *P*-value = 0.007) and one-carbon metabolism gene *PEMT* (rs7946, adjusted OR = 1.46, *P*-value = 0.01) (Table 2). Main effect association results for all SNPs in the main population can be found in Supplementary Material, Table S4. No significant main effects (FDR < 10%) were observed between these two SNPs and odds of skin lesions in the replication population (Table 2).

Gene-environment interactions with arsenic exposure

In the main population, six SNPs were observed to significantly interact with water arsenic levels on the odds of skin lesions after adjusting for age, sex, smoking status, betel nut consumption, education and BMI (Table 3). They were rs2278952 (*P* for interaction = 0.004) and rs897453 (*P* for interaction = 0.05) in the *PEMT* gene, rs2228000 (*P* for interaction = 0.01) in the *XPC* gene, rs1650697 (*P* for interaction = 0.02) in the *DHFR* gene, rs3024496 (*P* for interaction = 0.04) in the *IL10* gene and rs1133400 (*P* for interaction = 0.03) in the *INPP5A* gene. All six SNPs passed multiple comparisons at 10% FDR level. In the replication population, we observed consistent and significant interaction effects with water arsenic on odds of skin lesion for SNP rs1133400 (*P* for interaction in main population = 0.03; *P* for interaction in replication population = 0.03) in the *INPP5A* gene (Table 3). The dose-response relationship of arsenic and skin lesions in minor allele (G) carriers versus non-carriers of SNP rs1133400 is shown in Figure 1. Minor allele carriers (participants with at least one copy of G allele) had lower odds of skin lesion (Figure 1A), compared to non-minor allele carriers (Figure 1B). Interaction results for all SNPs in the main population can be found in Supplementary Material, Table S5.

Discussion

We analyzed a panel of 25 candidate SNPs based on literature data in 540 arsenic-induced skin lesions cases and 400 controls in Bangladesh and found six genetic variants in the skin cancer genes *INPP5A* and *XPC*, one-carbon metabolism genes *PEMT* and *DHFR,* and inflammation gene *IL10* to significantly interact with arsenic exposure on skin lesion risk after adjusting for multiple comparisons. Similar findings were obtained using toenail arsenic concentrations in the main population (data not shown). The interaction between rs1133400 on skin cancer gene *INPP5A* and arsenic was successfully replicated in an independent population.

Genetic variant rs1133400 is a non-synonymous SNP located on the *INPP5A* gene (inositol polyphosphate-5-phosphatase) which functions as a negative regulator of inositol signaling [31, 32]. A study has reported the loss of *INPP5A* in primary SCC tumors (24% *INPP5A* gene deletion in SCC tumors versus 0% in normal skin) and progressive reduction of *INPP5A* levels is observed as the tumor progresses to the metastatic stage [25]. Both *in vitro* and *in vivo* studies have also reported that reduced *INPP5A* expression leads to transformation and tumor growth, suggesting that *INPP5A* may play a crucial role as a tumor suppressor by inhibiting the proliferation of tumor cells [33-35]. SCC is one of the most common types of skin cancer that is associated with chronic arsenic exposures and is highly correlated with arsenic-induced skin lesions [5]. The loss of tumor suppressor *INPP5A* gene in an arsenic-exposed population might lead to increased vulnerability to skin lesions and potentially skin cancers later on. This significant interaction between *INPP5A* SNP rs1133400 and arsenic is consistent in both our main population and the replication population and suggest increased susceptibility to arsenic exposures by *INPP5A* genotype.

In the main population, genetic variants of *PEMT* rs7946 was shown to increase the odds of skin lesions. A significant gene-environment interaction was observed between *PEMT* SNPs rs2278952, a 5′-UTR SNP and rs897453, a potentially functional non-synonymous SNP that might result in loss of function, with arsenic exposure on skin lesions. These associations, however, were not significant in the replication study. Phosphatidylethanolamine N-methyltransferase (*PEMT*) is responsible for de novo choline synthesis and hence regulates the amount of choline and its metabolites available for S-adenosyl methionine (SAM) production, which in turn determines the availability of methyl groups for arsenic metabolism [36, 37]. *PEMT* rs7946 is a non-synonymous SNP that leads to substitution of the amino acid valine to methionine at position 175 in the *PEMT* protein (V175M), and is a loss of function variant [38]. Expression of one-carbon metabolism genes such as *DNMT1* has been shown to be significantly associated with urinary arsenic levels [39]. Together, this evidence suggests that genetic variation in the one-carbon metabolism pathway genes may determine susceptibility to develop skin lesions and subsequently squamous cell cancers in arsenic-exposed individuals. Interestingly, we observed significant main and interaction effects of inflammation gene *IL10* SNP rs3024496 and arsenic-induced skin lesions, however, this was not replicated in the replication population.

Our study had several strengths including accurate measurement of individual environmental exposure levels from drinking water and biological relevance of the SNPs

genotyped. It is noteworthy that we replicated some of our significant findings in an independent population with similar ethnicity and arsenic exposure levels. Replication of molecular epidemiological data is essential for a reliable study and reducing false positives [40]. Having reproducible results in an independent population confirmed that the significant associations we observed in our population can be transferred across different geographical regions of Bangladesh and West Bengal, among people who are drinking high levels of arsenic contaminated water. Some potential limitations of this study include possible arsenic exposure misclassification that might result from subjects drinking from different water sources which were unaccounted for. However, this is unlikely since we collected water samples from their main drinking well. Toenail arsenic was also used as an internal biomarker of arsenic exposure to correct this possible bias and we observed consistency with the findings from drinking water (data not shown). The LODs used for the measurement of arsenic exposures in the two populations were different (0.5μg As/L in main population versus 0.1μg As/L in replication population), however, we did a sensitivity analysis by using the same minimum water arsenic cutoff of 0.5μg As/L in the replication population (n = 88 samples removed) and observed the same results (data not shown). A relatively small sample size and multiple testing among 25 SNPs might result in possible false positives. However, we adjusted all our results using a stringent FDR cutoff of 10% to reduce the probability of potential false positives due to chance.

Conclusions

Our findings demonstrated differential susceptibility to arsenic exposures on skin lesion risk by genetic variation in *INPP5A* in our study population. This finding was replicated in a second, Bangladesh population, both of which were exposed to wide range of arsenic from drinking contaminated water. This suggests possible biological mechanisms that might link arsenic exposures to skin lesions and identifies genetically susceptible sub-populations which may be more vulnerable to arsenic exposure thus requiring more stringent environmental protection or screening. Future studies are warranted to confirm the observed associations in functional studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. International Agency for Research on Cancer. Some drinking-water disinfectants and contaminants, including arsenic. IARC monographs on the evaluation of carcinogenic risks to humans. 2004; 84:1–477. [PubMed: 15645577]

2. Survey BG. Final Report: London, UK. Mott MacDonald Ltd; 1999. Groundwater studies for arsenic contamination in Bangladesh.

3. Rahman MM, Chowdhury UK, Mukherjee SC, et al. Chronic arsenic toxicity in Bangladesh and West Bengal, India--a review and commentary. J Toxicol Clin Toxicol. 2001; 39:683–700. [PubMed: 11778666]

4. Chakraborty AK, Saha KC. Arsenical dermatosis from tubewell water in West Bengal. Indian J Med Res. 1987; 85:326–334. [PubMed: 2956191]

5. Tseng WP, Chu HM, How SW, Fong JM, Lin CS, Yeh S. Prevalence of skin cancer in an endemic area of chronic arsenicism in Taiwan. J Natl Cancer Inst. 1968; 40:453–463. [PubMed: 5644201]

6. Hsu LI, Chen GS, Lee CH, et al. Use of arsenic-induced palmoplantar hyperkeratosis and skin cancers to predict risk of subsequent internal malignancy. Am Journal Epidemiol. 2013; 177:202–212. [PubMed: 23299695]

7. Yu HS, Liao WT, Chai CY. Arsenic carcinogenesis in the skin. J Biomed Sci. 2006; 13:657–666. [PubMed: 16807664]

8. Sun G, Xu Y, Li X, Jin Y, Li B, Sun X. Urinary arsenic metabolites in children and adults exposed to arsenic in drinking water in Inner Mongolia, China. Environ Health Perspect. 2007; 115:648–652. [PubMed: 17450238]

9. Cuzick J, Sasieni P, Evans S. Ingested arsenic, keratoses, and bladder cancer. Am Journal Epidemiol. 1992; 136:417–421. [PubMed: 1415161]

10. Leonardi G, Vahter M, Clemens F, et al. Inorganic arsenic and basal cell carcinoma in areas of Hungary, Romania, and Slovakia: a case-control study. Environ Health Perspect. 2012; 120:721–726. [PubMed: 22436128]

11. Loeb KR, Loeb LA. Significance of multiple mutations in cancer. Carcinogenesis. 2000; 21:379–385. [PubMed: 10688858]

12. Liu J, Xie Y, Ducharme DM, et al. Global gene expression associated with hepatocarcinogenesis in adult male mice induced by in utero arsenic exposure. Environ Health Perspect. 2006; 114:404–411. [PubMed: 16507464]

13. Vujcic M, Shroff M, Singh KK. Genetic determinants of mitochondrial response to arsenic in yeast Saccharomyces cerevisiae. Cancer Res. 2007; 67:9740–9749. [PubMed: 17942904]

14. Argos M, Kalra T, Pierce BL, et al. A prospective study of arsenic exposure from drinking water and incidence of skin lesions in Bangladesh. Am Journal Epidemiol. 2011; 174:185–194. [PubMed: 21576319]

15. Hernandez A, Marcos R. Genetic variations associated with interindividual sensitivity in the response to arsenic exposure. Pharmacogenomics. 2008; 9:1113–1132. [PubMed: 18681785]

16. Ahsan H, Chen Y, Kibriya MG, et al. Arsenic metabolism, genetic susceptibility, and risk of premalignant skin lesions in Bangladesh. Cancer Epidemiol Biomarkers Prev. 2007; 16:1270–1278. [PubMed: 17548696]

17. Breton CV, Zhou W, Kile ML, et al. Susceptibility to arsenic-induced skin lesions from polymorphisms in base excision repair genes. Carcinogenesis. 2007; 28:1520–1525. [PubMed: 17374727]

18. McCarty KM, Chen YC, Quamruzzaman Q, et al. Arsenic methylation, GSTT1, GSTM1, GSTP1 polymorphisms, and skin lesions. Environ Health Perspect. 2007; 115:341–345. [PubMed: 17431481]

19. Valenzuela OL, Drobna Z, Hernandez-Castellanos E, et al. Association of AS3MT polymorphisms and the risk of premalignant arsenic skin lesions. Toxicol Appl Pharmacol. 2009; 239:200–207. [PubMed: 19538983]

20. Maru GB, Gandhi K, Ramchandani A, Kumar G. The role of inflammation in skin cancer. Adv Exp Med Biol. 2014; 816:437–469. [PubMed: 24818733]

21. Locasale JW. Serine, glycine and one-carbon units: cancer metabolism in full circle. Nat Rev Cancer. 2013; 13:572–583. [PubMed: 23822983]

22. Azar KK, Tani M, Yasuda H, Sakai A, Inoue M, Sasagawa T. Increased secretion patterns of interleukin-10 and tumor necrosis factor-alpha in cervical squamous intraepithelial lesions. Hum Pathol. 2004; 35:1376–1384. [PubMed: 15668895]

23. States JC, Barchowsky A, Cartwright IL, Reichard JF, Futscher BW, Lantz RC. Arsenic toxicology: translating between experimental models and human pathology. Environ Health Perspect. 2011; 119:1356–1363. [PubMed: 21684831]

24. Gemmati D, De Mattei M, Catozzi L, et al. DHFR 19-bp insertion/deletion polymorphism and MTHFR C677T in adult acute lymphoblastic leukaemia: Is the risk reduction due to intracellular folate unbalancing? Am J Hematol. 2009; 84:526–529. [PubMed: 19536847]

25. Sekulic A, Kim SY, Hostetter G, et al. Loss of inositol polyphosphate 5-phosphatase is an early event in development of cutaneous squamous cell carcinoma. Cancer Prev Res (Phila). 2010; 3:1277–1283. [PubMed: 20876729]

26. Pierce BL, Kibriya MG, Tong L, et al. Genome-wide association study identifies chromosome 10q24.32 variants associated with arsenic metabolism and toxicity phenotypes in Bangladesh. PLoS Genet. 2012; 8:e1002522. [PubMed: 22383894]

27. Seow WJ, Pan WC, Kile ML, et al. Arsenic reduction in drinking water and improvement in skin lesions: a follow-up study in Bangladesh. Environ Health Perspect. 2012; 120:1733–1738. [PubMed: 23060367]

28. Ahsan H, Chen Y, Parvez F, et al. Health Effects of Arsenic Longitudinal Study (HEALS): description of a multidisciplinary epidemiologic investigation. J Expo Sci Environ Epidemiol. 2006; 16:191–205. [PubMed: 16160703]

29. Dabney A, S J, Warnes GR. qvalue: Q-value estimation for false discovery rate control. R package version 1.26.0 ed. 2011

30. Pan WC, Seow WJ, Kile ML, et al. Association of Low to Moderate Levels of Arsenic Exposure With Risk of Type 2 Diabetes in Bangladesh. Am J Epidemiol. 2013; 178:1563–1570. [PubMed: 24049161]

31. Kisseleva MV, Wilson MP, Majerus PW. The isolation and characterization of a cDNA encoding phospholipid-specific inositol polyphosphate 5-phosphatase. J Biol Chem. 2000; 275:20110–20116. [PubMed: 10764818]

32. Zhang Q, Claret FX. Phosphatases: the new brakes for cancer development? Enzyme Res. 2012; 2012:659649. [PubMed: 22121480]

33. Speed CJ, Little PJ, Hayman JA, Mitchell CA. Underexpression of the 43 kDa inositol polyphosphate 5-phosphatase is associated with cellular transformation. Embo Journal. 1996; 15:4852–4861. [PubMed: 8890159]

34. Ye Y, Li Q, Hu WL, et al. Loss of PI (4,5) P2 5-Phosphatase A contributes to resistance of human melanoma cells to RAF/MEK inhibitors. Transl Oncol. 2013; 6:470–481. [PubMed: 23908690]

35. Ye Y, Jin L, Wilmott JS, et al. PI (4,5) P2 5-phosphatase A regulates PI3K/Akt signalling and has a tumour suppressive role in human melanoma. Nat Commun. 2013; 4

36. Niculescu MD, Zeisel SH. Diet, methyl donors and DNA methylation: Interactions between dietary folate, methionine and choline. J Nutr. 2002; 132:2333s–2335s. [PubMed: 12163687]

37. Vance DE, Walkey CJ, Cui Z. Phosphatidylethanolamine N-methyltransferase from liver. Biochim Biophys Acta. 1997; 1348:142–150. [PubMed: 9370326]

38. Song JN, da Costa KA, Fischer LM, et al. Polymorphism of the PEMT gene and susceptibility to nonalcoholic fatty liver disease (NAFLD). FASEB J. 2005; 19:1266–1271. [PubMed: 16051693]

39. Mo JY, Xia YJ, Wade TJ, DeMarini DM, Davidson M, Mumford J. Altered gene expression by low-dose arsenic exposure in humans and cultured cardiomyocytes: Assessment by Real-Time PCR Arrays. Int J Environ Res Public Health. 2011; 8:2090–2108. [PubMed: 21776218]

40. Dahlman I, Eaves IA, Kosoy R, et al. Parameters for reliable results in genetic association studies in common disease. Nat Genet. 2002; 30:149–150. [PubMed: 11799396]
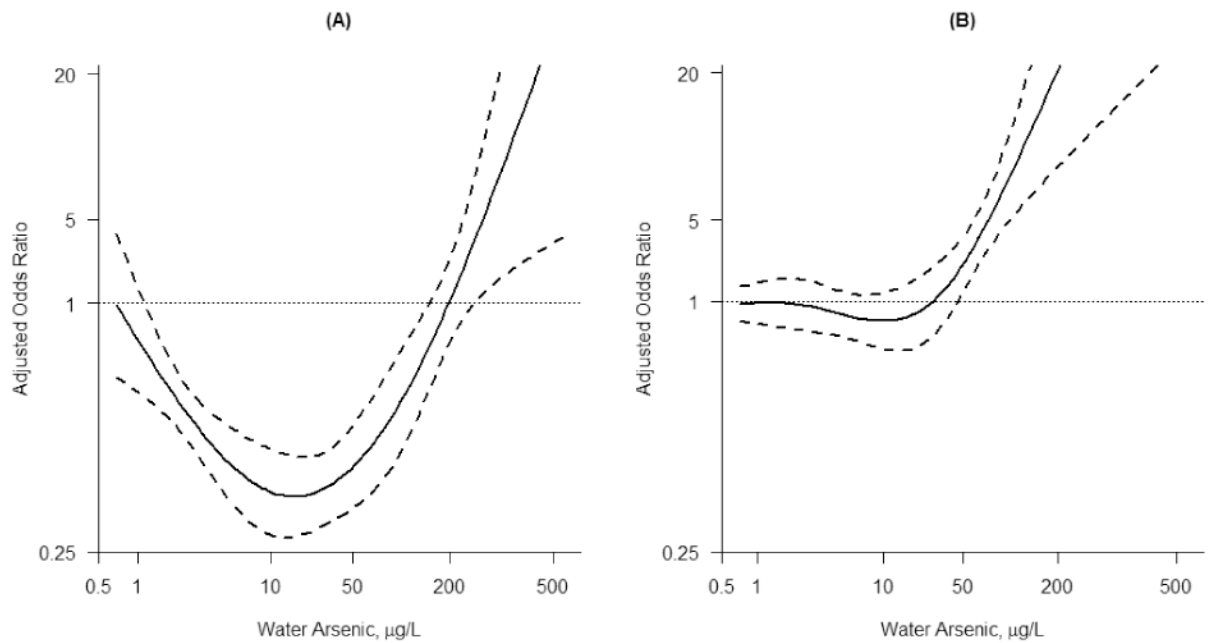
Figure 1. Odds ratios of *INPP5A* SNP rs1133400 in minor allele carriers (A) and non-minor allele carriers (B) in the main population

Dose-response of water arsenic and skin lesions in *INPP5A* SNP rs1133400 A) minor allele G carriers (*P*-value < 0.001) and B) non-minor allele carriers in the main population (*P*-value < 0.001), adjusting for age, sex, smoking status, betel nut consumption, education and BMI. Minor allele carriers have lower odds of skin lesions compared to non-minor allele carriers. Dashed lines denote the 95% confidence interval.

Table 1

Characteristics of the main study population in Pabna, Bangladesh in 2001-2003

Variables[a]	Cases (n = 540)	Controls (n = 400)	P-value[b]
Sex			0.69
Male	331 (61.3%)	240 (60%)	
Female	209 (38.7%)	160 (40%)	
Age, years			0.62
Mean (SD)	34.0 (11.8)	34.4 (11.7)	
Median (IQR)	34.0 (18.0)	34.5 (18.0)	
Education level			0.23
None to Primary	479 (88.7%)	345 (86.3%)	
Secondary to College	45 (8.3%)	38 (9.5%)	
Graduate or above	16 (3.0%)	17 (4.2%)	
BMI, kg/m^2			0.11
Mean (SD)	20.2 (3.13)	20.5 (3.27)	
Median (IQR)	19.6 (3.69)	19.7 (3.56)	
Smoking status (Males only[c])			0.80
Never	177 (71.5%)	125 (71%)	
Ever or Current	154 (28.5%)	115 (29%)	
Chew betel nuts			0.61
Never	382 (70.7%)	289 (72.3%)	
Ever or Current	158 (29.3%)	111 (27.8%)	
Water arsenic, μg/L			< 0.001
Mean (SD)	208 (302)	61.9 (127)	
Median (IQR)	26.5 (388)	9.43 (45.7)	

Abbreviations: SD, standard deviation; IQR, interquartile range; BMI, body mass index

[a] Data were shown as mean (SD) and median (IQR) for continuous variables or n (%) for categorical variables. There were 3 missing values from water arsenic, and these were excluded from analysis.

[b] P-values were obtained from Wilcoxon rank sum test with continuity correction for continuous variables and Fisher's exact test for categorical variables.

[c] Only males provided smoking status in the main population.

Table 2

Significant main associations between minor allele of SNPs and skin lesion status compared with non-minor allele carriers, adjusted for log-transformed water arsenic, with *P*-value ≤ 0.05 and *q*-value < 0.10 in the main population

Gene	rs number	Genotype	Main population				Replication population			
			Ca/Co	OR[a] (95% CI)	*P*-value	*q*-value[b]	Ca/Co	OR[a] (95% CI)	*P*-value	
IL10	rs3024496[c]									
		TT	342/217	1.00	-	-	301/1787	1.00	-	
		TC/CC	180/170	0.68 (0.40-0.96)	0.007	0.06	187/1007	1.04 (0.83-1.29)	0.75	
PEMT	rs7946[d]									
		GG	136/133	1.00	-	-	167/913	1.00	-	
		GA/AA	378/247	1.46 (1.17-1.75)	0.01	0.06	321/1881	0.95 (0.76-1.19)	0.69	

Abbreviations: Ca, case; Co, control; OR, odds ratio; CI, confidence interval

[a] Adjusted for age, sex, smoking status, betel nut consumption, education, BMI and water arsenic exposure.

[b] None of the SNPs passed 10% false discovery rate (FDR).

[c] 3′ UTR SNP

[d] Non-synonymous SNP

Table 3

Significant interactions between minor allele of SNPs and arsenic exposures in determining skin lesion status, with *P* for interaction ≤ 0.05 in the main population and their replication results

Gene	rs number	SNP type	Main population		Replication population
			P for interaction[a]	*q*-value	*P* for interaction[a]
PEMT	rs2278952	5′ UTR	0.004	0.01	0.97
XPC	rs2228000	non-synonymous	0.01	0.02	0.95
DHFR	rs1650697	5′ UTR	0.02	0.02	0.48
INPP5A	rs1133400*	non-synonymous	0.03	0.02	0.03
IL10	rs3024496	3′ UTR	0.04	0.03	0.76
PEMT	rs897453	non-synonymous	0.05	0.03	0.59

Abbreviations: FDR, false discovery rate; UTR, untranslated region

[a] *P* for interaction was obtained from model adjusted for age, sex, smoking status, betel nut consumption, education, BMI, SNP and arsenic exposure.

*
P for interaction ≤ 0.05 in both main and replication populations