Determinants and Consequences of Arsenic Metabolism Efficiency among 4,794 Individuals: Demographics, Lifestyle, Genetics, and Toxicity

Rick J. Jansen1, Maria Argos2, Lin Tong1, Jiabei Li1, Muhammad Rakibuz-Zaman3, Md. Tariqul Islam3, Vesna Slavkovich4, Alauddin Ahmed3, Ana Navas-Acien5, Faruque Parvez4, Yu Chen6, Mary V. Gamble4, Joseph H. Graziano4, Brandon L. Pierce1,7, and Habibul Ahsan1,7,8

Abstract

Background: Exposure to inorganic arsenic (iAs), a class I carcinogen, affects several hundred million people worldwide. Once absorbed, iAs is converted to monomethylated (MMA) and then dimethylated forms (DMA), with methylation facilitating urinary excretion. The abundance of each species in urine relative to their sum (iAs%, MMA%, and DMA%) varies across individuals, reflecting differences in arsenic metabolism capacity.

Methods: The association of arsenic metabolism phenotypes with participant characteristics and arsenical skin lesions was characterized among 4,794 participants in the Health Effects of Arsenic Longitudinal Study (Araihazar, Bangladesh). Metabolism phenotypes include those obtained from principal component (PC) analysis of arsenic species.

Results: Two independent PCs were identified: PC1 appears to represent capacity to produce DMA (second methylation step), and PC2 appears to represent capacity to convert iAs to MMA (first methylation step). PC1 was positively associated (P < 0.05) with age, female sex, and BMI, while negatively associated with smoking, arsenic exposure, education, and land ownership. PC2 was positively associated with age and education but negatively associated with female sex and BMI. PC2 was positively associated with skin lesion status, while PC1 was not. 10q24.32/AS3MT region polymorphisms were strongly associated with PC1, but not PC2. Patterns of association for most variables were similar for PC1 and DMA%, and for PC2 and MMA% with the exception of arsenic exposure and SNP associations.

Conclusions: Two distinct arsenic metabolism phenotypes show unique associations with age, sex, BMI, 10q24.32 polymorphisms, and skin lesions.

Impact: This work enhances our understanding of arsenic metabolism kinetics and toxicity risk profiles.

Introduction

Inorganic arsenic (iAs) exposure is considered toxic and carcinogenic (1). Exposure is estimated to affect several hundred million people worldwide with highest levels occurring in areas of South America and Asia (2). Arsenic exists in the air, soil, and water, and routes of exposure include breathing, contact with skin, and diet, with the most common source being drinking water.

Chronic exposure to iAs concentrations in drinking water that exceed 50 to 100 μg/L poses risk of adverse health effects, including cancer, developmental effects, neurotoxicity, cardiovascular disease, and diabetes mellitus (2, 3). Skin lesions are often one of the first signs of high, chronic iAs exposure, and risk for lesions increases with increasing arsenic exposure (4, 5). Additional risk factors such as age, sex, smoking, and low-protein intake, in conjunction with iAs exposures, have shown associations with cancer risk (6). Multiple mechanisms have been suggested for how iAs leads to the development of disease including oxidative stress, genetic aberrations, and epigenetic alterations (7–11).

Under the Challenger model of arsenic metabolism, inorganic arsenic enters the body as arsenite (iAsIII) or arsenate (iAsV) which can be reduced to iAsIII. Methylation and additional reduction reactions then produce monomethylarsonic acid (MMA) and dimethylarsinic acid (DMA). Methylation of arsenic is catalyzed by arsenic (+3 oxidation state) methyltransferase (AS3MT) with S-adenosylmethionine (SAM) as the methyl group donor (12, 13). Although there is some uncertainty regarding the arsenic metabolism pathway (14), it is generally agreed that there are two methylation steps: (i) iAs is methylated to produce MMA and (ii) MMA is methylated to DMA.

Relative concentrations of arsenic metabolites measured in urine represent an individual’s capacity to metabolize arsenic.
Typical urine metabolite percentages in humans are 10% to 30% for iAs, 10% to 20% for MMA, and 60% to 80% for DMA (15–21). Both pentavalent and trivalent forms of these arsenic species can be present in urine (20. 22–25). The trivalent forms of iAs, MMA, and DMA are likely to be the more toxic forms (20, 26–29) with MMA\textsuperscript{III} having the highest toxicity followed by iAs\textsuperscript{III} (30, 31). Previous studies have shown that iAs%, MMA%, and MMA/iAs are positively associated and DMA%, and DMA/MMAs are negatively associated with skin lesions with higher levels of MMA% associated with highest risk of skin lesions (32, 33).

Not all individuals with the same level of iAs exposure develop skin lesions, so genetic variation has been hypothesized to play a role. The most extensively studied gene in relation to arsenic metabolism has been AS3MT, and variation in the region of this gene has shown consistent association with arsenic metabolite percentages across multiple studies (34). A recent GWAS study observed two independent association signals for DMA\% in the AS3MT region, as represented by rs9527 and rs1191527. For both these SNPs, the allele associated with higher DMA\% is associated with lower MMA% and iAs% (35).

In this study, we characterize associations between participant characteristics (demographics, lifestyle, exposures, and genetics) and arsenic metabolism phenotypes (measured in urine) among 4,794 Bangladeshi individuals. We also examine the association between these metabolism phenotypes and arsenical skin lesion risk. Although most prior studies assess arsenic metabolism efficiency as relative concentrations of three correlated arsenic species in urine (iAs, MMA, and DMA), we attempted to create independent variables that represent arsenic metabolism efficiency using principal component analysis (PCA) of arsenic species.

**Materials and Methods**

**Study participants**

The current study included 4,794 men and women who participated in the Health Effects of Arsenic Longitudinal Study (HEALS; ref. 36) and whose urine samples were analyzed and had detectable levels for all three arsenic species. This subgroup of HEALS has been oversampled for skin lesions and high arsenic exposure. Details of the HEALS study methods have been described previously (36, 37). Briefly, the HEALS population-based prospective cohort study was designed to assess the association of arsenic exposure with health outcomes in the rural region of Araihazar, Bangladesh. Between October 2000 and May 2002, 11,746 participants between the ages of 18 to 75 years old were recruited, and 11,022 participants provided urine samples. A total of 10,970 wells in the area were tested for arsenic. Using the same set of study procedures, a second group of 8,287 new participants were added between July 2006 and August 2008.

**Metabolite measurement**

Urine samples collected at the time of baseline HEALS recruitment were assayed for arsenic concentrations using a graphite furnace atomic absorption method (Perkin-Elmer Analyst 600 graphite furnace system; limit of detection = 5 µg/L) in the Columbia University Trace Metal Core Laboratory (38). Then urinary arsenic metabolites [arsenobetaine, arsenocholine, arsenite (As\textsuperscript{III}), arsenate (As\textsuperscript{V}), monomethylarsonic acid (MMA), and dimethylarsinic acid (DMA)] were separated and detected as described by Heitkemper and colleagues (39) using high-performance liquid chromatography and inductively coupled plasma-mass spectrometry, respectively (40). During sample transport, As\textsuperscript{III} can oxidize to As\textsuperscript{V} and therefore, the sum (iAs = As\textsuperscript{III} + As\textsuperscript{V}) is presented here. Using a colorimetric diagnostics kit (Sigma), a measure of urinary creatinine was obtained and used to create a creatinine-adjusted total arsenic concentration (μg/g creatinine; ref. 41). Total urinary arsenic was created by summing the three individual metabolites and was used as the denominator when calculating the percentage of iAs, MMA, and DMA.

**Skin lesion assessment**

At baseline, HEALS study physicians, trained in the detection and diagnosis of skin lesions, recorded if the following conditions were present and their location: melanosis (hyperpigmentation), leucolmelanosis (hypopigmentation), or keratosis (hyperkeratotic thickening of skin). A participant was identified as having baseline skin lesions if any of these three conditions was present and was identified as having incident skin lesions if any of the three conditions was present at any of the follow-up visits. Of the 4,814 participants with metabolite data, only 3,555 had data on skin lesions status at both baseline and follow-up that allowed us to confidently assign skin lesions status (present/absent at any time during baseline and follow-up).

**Genotyping and quality control**

Using the Flexigene DNA Kit (Qiagen, Cat#51204), DNA was extracted from clotted blood with the concentration and quality checked using Nanodrop 1000. Using an Illumina Infinium HD SNP array with 250 ng of starting DNA, samples were processed on HumanCytoSNP-12 v2.1 chips (299,140 markers) and read on the BeadArray Reader and imaged in BeadStudio to create genotype calls.

There were 3,454 participants for whom genotype data are available. Quality control and genotype procedures have been described previously (35, 42, 43). We then removed those with low call rates (<97%) and poorly called or monomorphic SNPs (<95%). Additional participants were removed because of: gender mismatches, a lack of technical replicate DNA sample. There were no participants excluded because of outlying autosomal heterozygosity or inbreeding. Additional SNPs were excluded because of HWE P values < 10\textsuperscript{-2}. PLINK was used to perform all quality control (44). The final sample included 2,060 genotyped participants who had available data on urinary arsenic metabolites.

**Statistical analysis**

To describe our sample, we calculated percentages of selected baseline characteristic categories and calculated mean values for each of the three metabolite groups (iAs%, MMA%, and DMA%). Arsenic metabolism phenotypes were created using PCA where each component represents a different linear combination of the three metabolite percentages (iAs%, MMA%, and DMA%). PCA was performed using both transformed (using sqrt, and probit function) and untransformed metabolite percentage variables (DMA%, MMA%, and iAs%) and using various rotation methods, including no rotation. Linear regression models among those with no baseline skin lesions were used to estimate associations between arsenic metabolism phenotypes and participant characteristics [age, sex, PC, BMI (17.5–18, 18.5–23, >23), cigarette or bidi smoking (current, former, never), years of formal education, land ownership (yes, no), TV ownership (yes, no), level of arsenic in primary well water (<10, 10–50, 50–150, >150), and HEALS enrollment period (2000–2002 and 2006–2008)]. Logistic
regression models where used to estimate associations between arsenic metabolism phenotypes and skin lesion status (prevalent + incident vs. none), adjusting for age and sex (reduced model) as well as all participant characteristics used in the linear regression model (full model). All PCA and regression analyses were performed using SAS (version 9.3) and figures were created using R (version 3.0.2). The genetic association analyses were conducted by using GEMMA (45) adjusting for age, sex, and log-transformed total water arsenic exposure, accounting for relatedness.

**Results**

The means (SD) in the analyzed sample for iAs%, MMA%, and DMA% were 15.2 (6.5), 13.5 (5.0), and 71.3 (8.5), respectively. DMA% was higher among females, individuals with higher BMI, never smokers, individuals with lower arsenic exposure, and those without skin lesions (Table 1). MMA% was higher among males, older individuals, individuals with lower BMI, current and former smokers, individuals with higher arsenic exposure, and those with skin lesions. iAs% was higher among females, younger individuals, individuals with lower BMI, and higher arsenic exposure. For most of these characteristics, MMA% and iAs% show similar associations that are opposite those of DMA%. The joint distributions of iAs%, MMA%, and DMA% for nonzero values are shown in Fig. 1, with most individuals having a higher DMA% measured in their urine as compared with MMA% and iAs%. A few individuals have much higher, outlying values for iAs%, thereby reducing the percentages for other two metabolites. In addition, there were 20 individuals who have values of iAs% and/or MMA% that were below the detection limits, and these individuals were dropped from the analysis.

**PCA of the metabolite percentages**

We observed very little variation in the results based on different transformations and rotations; thus, we present the untransformed, unrotated results. Two principal components (PC) explained all the variation in the data, and this is expected as the three metabolite percentages sum to 100%. DMA loaded negatively (−100) on PC1 [eigenvalue = 2.08], and MMA (67) and iAs (80) both loaded positively. For PC2 [eigenvalue = 0.93], MMA loaded positively (75), iAs loaded negatively (−60), and DMA had a loading near zero (2). We multiplied PC1 by −1 so that higher PCs scores represent more methylation (i.e., more DMA% and MMA%, respectively). The variance explained by PC1 was 68.9% and by PC2 was 31.1%. There was no correlation between PC1 and PC2 (Fig. 2). Our interpretation is that PCA was able to identify two independent arsenic metabolism phenotypes in our study population: principal component 1 (PC1), which could represent one’s capacity to produce DMA (second methylation step), and PC2, which could represent one’s capacity to convert iAs to MMA (first methylation step).

**Correlates of arsenic metabolism efficiency**

To compare our results for the PCs with other commonly used measures of arsenic metabolism, additional models among those with no baseline skin lesions (n = 4,073) were run with the outcome as metabolite percentages (iAs%, MMA%, and DMA%) and metabolite ratios (PMI = MMA/iAs and SMI = DMA/MMA; Table 2). Metabolite ratios were also analyzed as log-transformed values, but as P values and interpretation remained constant, we chose to present results based on untransformed ratios. With PC1 as the outcome, we observed significant positive associations (+) with age, female sex, BMI, and TV ownership; we observed significant negative associations (−) with betel nut use, education categories, land ownership, and well water arsenic exposure. When PC2 as the outcome (+): age, current betel nut use, current cigarette smoking, and education categories; (−): females, BMI, and TV ownership. The general patterns of association are similar for PC1, DMA%, and SMI. The general patterns of association are similar for PC2, MMA%, and PMI with the exception of water arsenic exposure which is not significant when PC2 is the outcome, has a positive dose-response relationship when MMA% is the outcome, and has a negative dose-response

### Table 1. Mean urinary arsenic species percentages (reflecting arsenic methylation capacity) for 4,794 HEALS participants stratified by baseline characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>%</th>
<th>iAs%</th>
<th>MMA%</th>
<th>DMA%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>54.7</td>
<td>14.9</td>
<td>15.1</td>
<td>70.0</td>
</tr>
<tr>
<td>Female</td>
<td>45.3</td>
<td>15.4</td>
<td>11.6</td>
<td>73.0</td>
</tr>
<tr>
<td>Age, y²</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;30</td>
<td>17.4</td>
<td>16.4</td>
<td>11.8</td>
<td>71.8</td>
</tr>
<tr>
<td>30–39</td>
<td>28.0</td>
<td>16.0</td>
<td>13.1</td>
<td>70.9</td>
</tr>
<tr>
<td>40–49</td>
<td>29.4</td>
<td>15.1</td>
<td>13.8</td>
<td>71.2</td>
</tr>
<tr>
<td>≥50</td>
<td>25.2</td>
<td>13.4</td>
<td>14.7</td>
<td>71.9</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;17</td>
<td>21.0</td>
<td>15.5</td>
<td>14.2</td>
<td>70.3</td>
</tr>
<tr>
<td>17–18.5</td>
<td>23.3</td>
<td>15.3</td>
<td>14.0</td>
<td>70.7</td>
</tr>
<tr>
<td>18.6–23</td>
<td>42.1</td>
<td>15.2</td>
<td>13.2</td>
<td>71.5</td>
</tr>
<tr>
<td>&gt;24</td>
<td>13.7</td>
<td>14.0</td>
<td>12.2</td>
<td>73.8</td>
</tr>
<tr>
<td>Betel nut use</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>54.2</td>
<td>15.4</td>
<td>13.0</td>
<td>71.6</td>
</tr>
<tr>
<td>Current</td>
<td>42.3</td>
<td>14.9</td>
<td>14.0</td>
<td>71.1</td>
</tr>
<tr>
<td>Former</td>
<td>3.5</td>
<td>15.0</td>
<td>14.6</td>
<td>70.4</td>
</tr>
</tbody>
</table>

**NOTE:** Bold values represent significant (P < 0.05) difference between category and the lowest category.

- Lower limits are inclusive and upper limits are exclusive.
- Creatinine adjusted values.
- a Lower limits are missing, categories based on WHO safety standards.
- b Two missing at baseline and 843 missing values for all follow-up periods; comparisons for skin lesion status were: (i) baseline vs. no baseline and (ii) incident vs. no incident.
relationship when PMI is the outcome. When iAs% is the outcome (+): land ownership, and well water arsenic exposure; (−): age, and BMI.

Genetic polymorphisms associated with urinary arsenic metabolism phenotypes

Associations between the arsenic metabolism measures and AS3MT SNPs rs9527 and rs11191527 were assessed using a regression model adjusting for age, sex, enrollment period of HEALS, water arsenic, and relatedness (Table 3). DMA% and PC1 have the most significant relationship with each SNP (P values <1E−10), whereas PC2 and PMI have the least significant associations with each SNP (P > 1E−03). Previously, a significant interaction between total drinking water exposure (tertiles) and DMA% was observed (35). So we additionally ran these models with an interaction term for quartiles of total drinking water exposure (<10, 10–50, 50–150, and ≥150) and each SNP, but there were no significant results (not shown).

Association between arsenic metabolism phenotypes and skin lesion status

Multivariate logistic regression models were run to analyze the association between skin lesions status (prevalent + incident vs. none) and potential risk factors including PC1 and PC2. Significant (P < 0.05) associations with increased risk were observed for increasing age, PC2, TV ownership, and arsenic exposure from primary well water. Associations with decreased risk were observed for female sex and increasing education (Supplementary Table S2). We repeated this analysis using quartiles of our PC variables (Supplementary Table S3), and PC2’s third and fourth quartiles (compared with the first) showed suggestive evidence (P < 0.1) of positive association with skin lesions status.

A logistic regression model was used to assess the association between skin lesion status and the various arsenic metabolism phenotypes one at a time (Table 4). For the age and sex-adjusted model, PC1 and DMA% show significant (P < 0.05) associations with decreased risk, whereas PC2 and MMA% show an association with increased risk. In a fully adjusted model, the associations remain significant for PC2 and MMA%. We additionally ran these models with an interaction term for quartiles of total drinking water exposure (<10, 10–50, 50–150, and ≥150) and each arsenic metabolism phenotype, but no interaction terms were significant (not shown).

These logistic regression analyses were repeated stratifying by arsenic exposure from primary well water (<50 and ≥50 μg/L). In the age- and sex-adjusted model among individuals with low exposure, PC2 and MMA% were associated with a significant (P < 0.05) increase in skin lesion risk, however, none of the metabolism phenotypes were associated with risk among individuals with high exposure (≥50 μg/L). In the fully adjusted model, the associations for PC2 and MMA% remained significant in the low exposure group (Table 4), whereas none of the metabolite

Figure 1.
A triangle scatter plot of iAs%, MMA%, and DMA% for 4,794 HEALS participants. The horizontal lines indicate a constant level of iAs%, the lines that slope downward to the right indicate a constant level of MMA%, and the lines that slope downward to the left indicate a constant level of DMA%. Those with values of zero for iAs% and MMA% have been removed (n = 20). Right triangle represents a zoomed in view of the highlighted portion from triangle on the left side of the figure.
Arsenic Metabolism Phenotypes

measures were significant among individuals with higher exposure (≥50 μg/L; Table 4).

Discussion

In this large study of arsenic metabolism phenotypes (n > 4,500), we used PCA to obtain two independent arsenic metabolism phenotypes and evaluated their associations with genetic and nongenetic participant characteristics, as well as arsenic toxicity outcomes. Our interpretation is that PC1 observed in our study may represent one’s capacity to produce DMA (second methylation step), and PC2 may represent one’s capacity to produce MMA (first methylation step). Our results suggest that, women, nonsmokers, individuals with higher BMI, and individuals with lower arsenic exposure have the reduced skin lesion risk profile (i.e., higher PC1 and reduced PC2), under the assumption that MMA% (particularly MMAIII) is the most toxic arsenic species relative to iAs and DMA.

PCA seems to have some advantages compared with the various highly correlated measures of arsenic metabolism (correlation table in Fig. 2) used in prior studies (i.e., iAs%, MMA%, DMA%, PMI, and SMI). By creating PCs, we are able to isolate two independent underlying measures of arsenic metabolism; thus, we can avoid the use of correlated outcomes when assessing associations between participant characteristics and metabolism phenotypes. This allows us to demonstrate that our two metabolism phenotypes are quite distinct with respect to their associations with participant characteristics. For example, we show that 10q24.32 SNPs known to influence arsenic metabolism are strongly related to PC1, but essentially unrelated to PC2. In addition, PC2 is the only metabolite measure that shows no association with arsenic exposure. One interpretation of this finding could be that PC2 represents metabolic reactions that do not become saturated when exposure is high, and are thus not a rate-limiting step in the conversion of iAs to DMA. An additional interpretation of what PC2 might represent is differences in individuals’ ability to transport arsenic across cell membranes, as variants in the SLCO1B1 anion transporter gene have been reported to be associated with arsenic metabolite percentages in urine (46).

Generally, our results are consistent with previous epidemiologic studies of arsenic metabolism across various populations and exposure levels. Decreased levels of iAs% are seen among women compared with men and with increasing age (6, 47–51), but associations with decreasing age (52) and increasing BMI (48, 51, 53) have also been reported. Regarding MMA%, lower levels have been observed among females compared with males (47, 49–51, 54), among never smokers (50), and with increasing age (47) and increasing BMI (49, 51, 53, 55), but lower levels have also been observed with decreasing age (6, 52). Higher levels of DMA% have been observed among females compared with males (6, 49, 50), with increasing age (6, 47, 51, 52), with increasing BMI (49, 51, 53, 56), and among never smokers (51).

In analyses stratified by exposure level, PC2 and MMA% are only associated with skin lesion risk at the lower exposure level (<50 μg/L). This suggests that variation in iAs metabolism, specifically MMA%, may be more important at lower exposure levels. In other words, at higher levels of exposure, the risk of developing skin lesions does not depend heavily on the efficiency of the first methylation step. This finding has potential implications for future prevention efforts among populations with low to moderate exposure, in which interventions could potentially be targeted toward individuals with ‘high-risk’ metabolism profiles.

We observe strong differences in metabolism of arsenic by sex. Such differences have been hypothesized to be due to hormonal differences, and several prior studies have examined pregnancy and menopause in relationship to arsenic metabolism and skin lesions. Pregnant women have increased methylation which increases with weeks of gestation (57, 58) and women who are highly exposed to arsenic have an earlier onset of menopause (59). The positive dose–response relationship between BMI and PC1 (i.e., DMA) could be explained by more intake of nutrients that help in arsenic metabolism (60–62). The inconsistency in associations between age and metabolites across studies could be a result of sample selection techniques resulting in different age ranges and frequencies and/or healthy survivor effect.

Our results indicated that the second methylation step is a rate-limiting step as exposure to arsenic in drinking water is not associated with PC2, but exposure is associated with PC1 in a dose–response relationship. This lack of association with PC2 was not evident when examining the association of water arsenic with PMI, iAs%, or MMA% (presumably because these phenotypes are correlated with DMA% and PC1). Previous studies suggest that the second methylation step is inhibited at increased exposure levels (particularly elevated levels of iAsIII and MMAIII) both in the experimental setting (63, 64) and in human observational studies (65), consistent with the associations observed in this study.

Across both the full and reduced logistic regression models with skin lesion status as the outcome, only PC2 and MMA% are significantly associated with increased risk of skin lesions. Consistent with our result, in prior studies of metabolite percentage measures, only MMA% is consistently associated with having skin lesions regardless of model adjustments (17, 32, 49, 54, 66, 67).

Figure 2.
Scatter plot of the PC scores and correlations between each PC and metabolite percentage (n=4,794). The variance explained by PC1 is 68.95%, and it captures the inverse correlation between DMA% and both iAs% and MMA% (based on correlation structure). PC2 explains 30.77% of the variance in the data, and it captures the inverse relationship between MMA% and iAs% (based on the residual correlation after adjusting for DMA%). PC1 has been multiplied by −1.
Jansen et al.

Thus, our work provides additional observational evidence that MMA is the most toxic methylation state of the arsenic species present in the body.

The association between SNPs in the AS3MT region and arsenic metabolites and/or skin lesions is well-established humans (34, 35, 42, 47, 52, 68, 69), and silencing of AS3MT has been shown to dramatically reduce iAs methylation in cultured cells (70) and in knockout mice (71, 72). In this work, SNPs in the AS3MT region (rs9527 and rs11191527) showed a strong association with PC1/DMA%, but not with PC2/PMI. One potential explanation is that AS3MT SNPs play an important role in the second step of metabolism, but less so for the first step. However, the mechanism(s) by which these SNPs influence metabolism remain unclear. Kinetic studies have shown that the AS3MT binding affinity differs for the first methylation step as compared with the second step (73, 74) and there are differences in the number of binding sites required for each of the methylation steps (75). These differences in AS3MT kinetics for the first versus the second methylation step suggest that it is possible that the SNPs in this region could have different effects on these two methylation reactions this population.

Several considerations need to be taken into account when evaluating and interpreting our results. The creation of PCs is entirely dependent on the study sample and specific PC scores are not directly comparable across studies. Interpretations of what our PCs represent and association observed between our PCs and selected characteristics need to be evaluated in other populations. We did not measure differences between the valence states of MMA.

Thus, our work provides additional observational evidence that MMA is the most toxic methylation state of the arsenic species present in the body.

The association between SNPs in the AS3MT region and arsenic metabolites and/or skin lesions is well-established humans (34, 35, 42, 47, 52, 68, 69), and silencing of AS3MT has been shown to dramatically reduce iAs methylation in cultured cells (70) and in knockout mice (71, 72). In this work, SNPs in the AS3MT region (rs9527 and rs11191527) showed a strong association with PC1/DMA%, but not with PC2/PMI. One potential explanation is that AS3MT SNPs play an important role in the second step of metabolism, but less so for the first step. However, the mechanism(s) by which these SNPs influence metabolism remain unclear. Kinetic studies have shown that the AS3MT binding affinity differs for the first methylation step as compared with the second step (73, 74) and there are differences in the number of binding sites required for each of the methylation steps (75). These differences in AS3MT kinetics for the first versus the second methylation step suggest that it is possible that the SNPs in this region could have different effects on these two methylation reactions this population.

Several considerations need to be taken into account when evaluating and interpreting our results. The creation of PCs is entirely dependent on the study sample and specific PC scores are not directly comparable across studies. Interpretations of what our PCs represent and association observed between our PCs and selected characteristics need to be evaluated in other populations. We did not measure differences between the valence states of MMA.

**Table 2.** Linear regression estimates and 95% CIs for the association between baseline characteristics and each metabolite measure among those with no baseline skin lesions (n = 4,073)*

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>PC1</th>
<th>PC2</th>
<th>iAs%</th>
<th>MMA%</th>
<th>DMA%</th>
<th>MMA/iAs</th>
<th>DMA/MMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y&lt;sub&gt;60&lt;/sub&gt;</td>
<td>ref</td>
<td>ref</td>
<td>0.03</td>
<td>0.05791</td>
<td>0.13</td>
<td>0.00041</td>
<td>-0.64</td>
</tr>
<tr>
<td>BMI&lt;18.5</td>
<td>ref</td>
<td>ref</td>
<td>0.14</td>
<td>0.03035</td>
<td>0.30</td>
<td>&lt;0.0001</td>
<td>-1.92</td>
</tr>
<tr>
<td>Sex - female</td>
<td>ref</td>
<td>ref</td>
<td>0.31</td>
<td>&lt;0.0001</td>
<td>0.47</td>
<td>&lt;0.0001</td>
<td>-3.43</td>
</tr>
<tr>
<td>Cigarette smoking</td>
<td>ref</td>
<td>ref</td>
<td>0.32</td>
<td>&lt;0.0001</td>
<td>-0.50</td>
<td>&lt;0.0001</td>
<td>0.30</td>
</tr>
<tr>
<td>Education (y)</td>
<td>ref</td>
<td>ref</td>
<td>0.39</td>
<td>&lt;0.0001</td>
<td>-0.14</td>
<td>0.00085</td>
<td>-1.44</td>
</tr>
<tr>
<td>Never</td>
<td>ref</td>
<td>ref</td>
<td>0.09</td>
<td>0.0447</td>
<td>0.01</td>
<td>0.7470</td>
<td>-0.54</td>
</tr>
<tr>
<td>Current</td>
<td>ref</td>
<td>ref</td>
<td>0.15</td>
<td>0.0003</td>
<td>-0.06</td>
<td>0.0866</td>
<td>-0.03</td>
</tr>
<tr>
<td>Land ownership</td>
<td>ref</td>
<td>ref</td>
<td>0.39</td>
<td>&lt;0.0001</td>
<td>-0.14</td>
<td>0.00085</td>
<td>-1.44</td>
</tr>
<tr>
<td>TV ownership</td>
<td>ref</td>
<td>ref</td>
<td>0.09</td>
<td>0.0000</td>
<td>0.31</td>
<td>0.7184</td>
<td>0.17</td>
</tr>
<tr>
<td>Water arsenic (µg/L)&lt;sub&gt;0&lt;/sub&gt;</td>
<td>ref</td>
<td>ref</td>
<td>0.09</td>
<td>0.0059</td>
<td>0.55</td>
<td>0.0759</td>
<td>0.17</td>
</tr>
</tbody>
</table>

*All models included adjustment for age, sex enrollment period of HEALS, and water arsenic. Relatedness was accounted for using GEMMA. For both SNPs, the betas correspond to the per-allele association between the T allele and the metabolite (the C allele is the reference allele).
Table 4. ORs and 95% confidence intervals for the association between specific metabolite variables (entering only one in each model) and skin lesion status [prevalent and incident (n = 1,575) versus none (n = 1,780)].

<table>
<thead>
<tr>
<th>Metabolite variable used in model</th>
<th>Minimally adjusted modela</th>
<th>Fully adjusted modelb</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>Lower CI</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-----</td>
<td>----------</td>
</tr>
<tr>
<td>PC1</td>
<td>0.91</td>
<td>0.85</td>
</tr>
<tr>
<td>PC2</td>
<td>1.08</td>
<td>1.01</td>
</tr>
<tr>
<td>iAs%</td>
<td>1.01</td>
<td>0.99</td>
</tr>
<tr>
<td>MMA%</td>
<td>1.03</td>
<td>1.01</td>
</tr>
<tr>
<td>DMA%</td>
<td>0.99</td>
<td>0.98</td>
</tr>
<tr>
<td>PMI</td>
<td>1.05</td>
<td>0.93</td>
</tr>
<tr>
<td>SMI</td>
<td>0.99</td>
<td>0.96</td>
</tr>
</tbody>
</table>

Water arsenic exposure < 50 μg/L:

| PC1                              | 0.91| 0.80     | 1.05     | 0.1946  | 0.92| 0.80     | 1.06     | 0.2271  |
| PC2                              | 1.15| 1.01     | 1.33     | 0.0516  | 1.23| 1.05     | 1.43     | 0.0090  |
| iAs%                             | 1.00| 0.98     | 1.02     | 0.9346  | 0.99| 0.97     | 1.02     | 0.5772  |
| MMA%                             | 1.03| 1.01     | 1.06     | 0.0214  | 1.04| 1.01     | 1.07     | 0.0268  |
| DMA%                             | 0.99| 0.97     | 1.01     | 0.2044  | 0.99| 0.97     | 1.01     | 0.2171  |
| PMI                              | 1.1 | 0.95     | 1.29     | 0.1934  | 1.15| 0.98     | 1.35     | 0.0794  |
| SMI                              | 0.97| 0.93     | 1.01     | 0.1132  | 0.96| 0.92     | 1.00     | 0.0553  |

Water arsenic exposure ≥ 50 μg/L:

| PC1                              | 0.97| 0.88     | 1.07     | 0.5317  | 0.97| 0.88     | 1.07     | 0.5402  |
| PC2                              | 1.05| 0.95     | 1.16     | 0.3617  | 1.06| 0.96     | 1.18     | 0.2488  |
| iAs%                             | 1.00| 0.99     | 1.01     | 0.9780  | 1.00| 0.98     | 1.01     | 0.8633  |
| MMA%                             | 1.01| 0.99     | 1.03     | 0.2619  | 1.01| 0.99     | 1.03     | 0.1961  |
| DMA%                             | 1.00| 0.99     | 1.01     | 0.5407  | 1.00| 0.99     | 1.01     | 0.5516  |
| PMI                              | 1.03| 0.84     | 1.27     | 0.7866  | 1.07| 0.87     | 1.33     | 0.5222  |
| SMI                              | 1.01| 0.98     | 1.04     | 0.3914  | 1.01| 0.98     | 1.04     | 0.407   |

*Model includes adjustment for age categories and sex.
*Model includes adjustment for age categories, sex, BMI categories, betel nut use categories, smoking categories, education categories, own land, own TV, exposure to arsenic in well water categories, and enrollment period of HEALS.

Conclusion

We identify two independent arsenic metabolism phenotypes that show distinct patterns of associations with age, sex, BMI, SES, and 10q24.32 polymorphisms. Arsenic exposure shows a negative association with the first (PC1), but no association with the second (PC2), suggesting that capacity to methylate MMA to DMA is reduced with increasing lesions beyond drinking water exposure (62). Adding total rice intake to our models as a covariate (not shown) does not change our results and total rice intake was weak and inconsistently (across quartiles of rice intake) associated with PC2, MMA%, and PMI. Other dietary patterns or nutrients have shown an association with arsenic excretion and/or skin lesion risk, but not consistently (82–85). Bioaccessibility and bioavailability play a role in potential health effects from arsenic exposure. Most of the metabolism of arsenic is thought to take place in the liver; however, there is evidence that gut microbiome metabolism may also take place and could have a significant impact on absorption/excretion and associated health effects form arsenic exposure (60, 86). Unmeasured bioaccessibility/bioavailability will increase the variation in our arsenic exposure estimates and when using the metabolite variables as outcomes we would have reduced power to detect associations with these exposure factors.

Our study has several strengths including the largest sample size to date for a study of arsenic metabolism, ethnically homogeneous population to avoid population stratification bias, and a large subgroup of the population with measurement of arsenic exposure, genetic variation, and demographic characteristics.

[III vs. V] of iAs, DMA, or MMA since these states can be affected by method of sample collection and transport and trivalent methylated species are unstable and difficult to detect. The trivalent state has been indicated as the most toxic form of iAs and its metabolites. Therefore, future study designs which are able to assess the individual states of metabolites and disease outcomes would be of interest. Prior studies have looked at both blood and urine metabolites in the same participants (76) as well as arsenic exposure and measurement in hair and nail samples (77, 78) suggesting strong correlation between metabolites excreted and retained. In humans, the average half-life of inorganic metabolites in the body is about 10 days (79, 80); however, there may be variation among individuals which may be worth explicitly investigating in future studies.

Here, we interpret our results under the Challenger model of arsenic metabolism. Working under a competing model of arsenic metabolism (ref. 81; e.g., using As–GSH complexes) may require different forms of exposure measurements at the study design phase and maybe an important future research direction. Those with skin lesions at baseline were excluded from the metabolism efficiency linear regression analysis, but it is possible that underlying disease processes (e.g., altered methylation of key metabolism genes) may have influenced metabolism in those who develop skin lesions during follow-up. However, when performing a sensitivity analysis (not shown) removing those with incident skin lesions from the linear analysis, our β estimates remain relatively consistent and interpretations remain the same as presented.

Rice is likely a major dietary contributor (excluding drinking water) of inorganic arsenic exposure in the population and has been observed to be associated with an increased risk of skin lesions beyond drinking water exposure (62). Adding total rice intake to our models as a covariate (not shown) does not change our results and total rice intake was weak and inconsistently (across quartiles of rice intake) associated with PC2, MMA%, and PMI. Other dietary patterns or nutrients have shown an association with arsenic excretion and/or skin lesion risk, but not consistently (82–85). Bioaccessibility and bioavailability play a role in potential health effects from arsenic exposure. Most of the metabolism of arsenic is thought to take place in the liver; however, there is evidence that gut microbiome metabolism may also take place and could have a significant impact on absorption/excretion and associated health effects form arsenic exposure (60, 86). Unmeasured bioaccessibility/bioavailability will increase the variation in our arsenic exposure estimates and when using the metabolite variables as outcomes we would have reduced power to detect associations with these exposure factors.

Our study has several strengths including the largest sample size to date for a study of arsenic metabolism, ethnically homogeneous population to avoid population stratification bias, and a large subgroup of the population with measurement of arsenic exposure, genetic variation, and demographic characteristics.
exposures. The second metabolism phenotype showed a significant association with skin lesion status, suggesting that those with higher MMA% in their urine are at increased risk for skin lesions. This work characterizes risk profile for those exposed to arsenic and can be used in future prevention and intervention studies.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors’ Contributions
Conception and design: F. Parvez, B.L. Pierce, H. Ahsan
Development of methodology: F. Parvez, M.V. Gamble, B.L. Pierce, J.H. Graziano
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): M. Rakibuz-Zaman, F. Parvez, M.V. Gamble, H. Ahsan, J.H. Graziano
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): R.J. Jansen, M. Argos, L. Tong, J. Li, V. Slavkovich, A. Navas-Acien, M.V. Gamble, B.L. Pierce, H. Ahsan, J.H. Graziano
Writing, review, and/or revision of the manuscript: R.J. Jansen, M. Argos, J. Li, A. Navas-Acien, Y. Chen, M.V. Gamble, M.V. Gamble, B.L. Pierce, H. Ahsan, J.H. Graziano
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): M.T. Islam, F. Parvez
Study supervision: A. Ahmed, F. Parvez, H. Ahsan

Grant Support
This work was funded in part by R01ES020506 (to B.L. Pierce), R01ES023834 (to B.L. Pierce), P42ES010349 (to J.H. Graziano), R01CA107431 (to H. Ahsan), and R24TW009555 (to H. Ahsan).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received July 6, 2015; revised November 5, 2015; accepted November 18, 2015; published OnlineFirst December 16, 2015.

References


