Environmental arsenic exposure and DNA methylation of the tumor suppressor gene p16 and the DNA repair gene MLH1: effect of arsenic metabolism and genotype.

Hossain, Bakhtiar; Vahter, Marie; Concha, Gabriela; Broberg Palmgren, Karin

Published in:
Metallomics

DOI:
10.1039/c2mt20120h

2012

Environmental arsenic exposure and DNA methylation of the tumor suppressor gene *p16* and the DNA repair gene *MLH1*; effect of arsenic metabolism and genotype

Mohammad Bakhtiar Hossain[1,2,3], Marie Vahter[3], Gabriela Concha[4], Karin Broberg[1, 3*]

[1] Department of Laboratory Medicine, Division of Occupational and Environmental Medicine, Lund University, Lund, SE-22185 Sweden

[2] International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR,B), Dhaka, Dhaka-1212, Bangladesh

[3] Section for Metals and Health, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, SE - 171 77, Sweden

[4] Department of Risk Benefit Assessment, National Food Agency, Uppsala, SE - 751 26, Sweden

[*] To whom correspondence should be addressed. Tel: +46 (0) 46 173819. Fax: +46 (0) 46 143201. E-mail: karin.broberg_palmgren@med.lu.se.

Short Running Head: Arsenic, its metabolism, and DNA methylation

Abstract

Arsenic is carcinogenic, possibly partly through epigenetic mechanisms. We evaluated the effects of arsenic exposure and metabolism on DNA methylation. Arsenic exposure and methylation efficiency in 202 women in the Argentinean Andes were assessed from concentrations of arsenic metabolites in urine (inorganic arsenic, methylarsonic acid [MMA], and dimethylarsinic acid [DMA]), measured by HPLC-ICPMS. Methylation of CpGs of the tumor suppressor gene *p16*, the DNA repair gene *MLH1*, and the repetitive elements *LINE1* were measured by PCR pyrosequencing of blood DNA. Genotyping (N=172) for *AS3MT* was performed with Sequenom^{TM}, and gene expression (N=90) with Illumina DirectHyb HumanHT-12 v3.0. Median arsenic concentration in urine was 230 µg/L (range 10.1–1,251). In linear regression analysis, $\log_2$-transformed urinary arsenic concentrations were positively associated with methylation of *p16* (β=0.14 *P*=0.0028) and *MLH1* (β=0.28, *P*=0.0011), but not with *LINE1*. Arsenic concentrations were of borderline significance negatively correlated with expression of *p16* (r_s=-0.20; *P*=0.066)), but not with *MLH1*. The fraction of inorganic arsenic was positively (β=0.026; *P*=0.010) and DMA was negatively (β=-0.017, *P*=0.043) associated with *p16* methylation with no effect of MMA. Carriers of the slow-metabolizing *AS3MT* haplotype were associated with more *p16* methylation (*P*=0.022). Arsenic exposure was correlated with increased methylation in blood of genes encoding enzymes that suppress carcinogenesis, and the arsenic metabolism efficiency modified the degree of epigenetic alterations.

Introduction

Inorganic arsenic, a commonly occurring pollutant of drinking water, is a potent toxicant and carcinogen, causing cancer in the lung, skin, and bladder, and possibly also in the liver, kidney, and prostate.[1-3] Arsenic is clastogenic,[4] but only a weak inducer of point mutations.[5-7] Thus, other mechanisms are probably involved in the carcinogenic effects of arsenic.

Recently, it has been suggested that arsenic acts as a carcinogen by modifying epigenetic mechanisms.[8-9] Indeed, hypermethylation of the tumor suppressor genes *p53* and *p16* was observed in arsenic-induced skin tumors.[10-11] In urothelial tumors from patients living in an endemic arsenic-exposed area in Taiwan, low expression of the p16 protein was shown to be much more common (81%) than in tumors from patients in a non-exposed area (25%), indicating that p16 is important in arsenic-related carcinogenesis.[12]

Expression of *hMSH2*, a key gene involved in the DNA mismatch repair system, has been found to be down-regulated with increasing severity of arsenic-related skin symptoms,[13] but whether or not this is mediated through modification of DNA methylation of the mismatch repair genes is not known. MLH1 is an important component in the mismatch repair system and promoter methylation of the *MLH1* gene has been implicated in the development of gastric and colorectal carcinoma.[14-15]

The Long Interspersed Nuclear Element 1 (*LINE1*) retrotransposable elements make up about 17% of human DNA[16] and methylation status of *LINE1* is often used as a proxy to represent global DNA methylation.[17] Treatment with arsenic results in global DNA hypomethylation in cell cultures[18] and in mice.[19] The CpG sites in *LINE1* are usually heavily methylated, and genome-wide loss of methylation from these sites has been regarded as a common epigenetic event in malignancy.[20]

Susceptibility to arsenic toxicity varies widely between individuals and populations. The efficiency of arsenic metabolism, measured as the formation and urinary excretion of

arsenic metabolites, may be one important susceptibility factor. Inorganic arsenic is metabolized in the body by a series of reduction and methylation reactions that produce methylarsonic acid (MMA) and dimethylarsinic acid (DMA), both of which are excreted in the urine.[21] Trivalent MMA is considered to be the most toxic metabolite[22] and a higher fraction of MMA in urine is associated with more toxic effects.[23-24] In line with that, a higher fraction of MMA or a lower fraction of DMA in urine is associated with a lower rate of arsenic excretion in urine.[21, 25] The main methyltransferase in arsenic metabolism is arsenic (+3 oxidation state) methyltransferase (AS3MT)[26], which can methylate both inorganic arsenic and MMA. We have previously shown that *AS3MT* genotype determines the fractions of the different arsenic metabolites in urine.[27]

The aim of this study was to evaluate epigenetic effects of arsenic exposure in cancer-related genes and repetitive DNA sequences in individuals who have been exposed through drinking water, and to determine the possible influence of arsenic metabolism. We chose to measure epigenetic changes in the *p16* and *MLH1* genes, as *p16* and mismatch repair genes have previously been associated with arsenic-related malignancy, and both are important for tumor suppression by regulating cell proliferation and maintain genome stability. We hypothesized that if there is an epigenetic effect of arsenic on these genes, they may be altered in blood cells as well. *LINE1* was measured as a proxy for global DNA methylation.

Methods and materials

Subjects

We studied 202 women living on the Andean plateau (3,800 m above sea level) in Northern Argentina, an area that has minimal industrial and vehicle-derived pollution. However, in some villages there are elevated concentrations of arsenic in drinking water.[28] Most of the study participants (N = 161) were from the village San Antonio de los Cobres, which has

about 5,000 inhabitants, and 200 µg/L arsenic in the drinking water. The remaining participants (N = 41) were from small surrounding villages with lower concentrations of arsenic in the water.

The study individuals were recruited with the assistance of the medical personnel, except in a small mining village, Tolar Grande, where we went from house to house to explain the project and invite the adults to participate. We included only women, as men were often away from home for longer periods for work, and therefore had another pattern of exposure to toxic elements in food and drinking water. We invited all women who were interested to participate in our study. Only four of the study participants smoked cigarettes and one reported drinking alcohol, but almost half of the women (46.5%) often chewed coca leaves. As reported previously, interviews with the study participants revealed that almost all women drank public drinking water exclusively, and that their diets consisted mainly of corn, beans, chicken, and pork.[29] Only 3 women reported taking any medication at the time of the study; one was being treated for gastritis and two for hypertension. The questionnaire data showed that the women were homogenous with regard to factors that may influence the peripheral blood methylation, i.e., most of the women did not use any medication, smoke tobacco or drink alcohol. Further, their diet was very similar. The women were asked if they had had any history of or current diseases and their hands were inspected for arsenic-related skin lesions.

Informed consent, both oral and written, was provided by all the study participants. The study was approved by the Ministry of Health in Salta, Argentina, and the Regional Ethical Committee of Karolinska Institutet, Sweden.

Blood and urine collection

All the biological samples were collected during daytime as non-fasting spot samples at the local health clinics and at the hospital in San Antonio de los Cobres. Peripheral blood for DNA extraction (N = 202) was collected in K_2EDTA tubes (Vacuette©; Greiner, Germany), blood for element analysis (N = 202) was collected in lithium heparin tubes (Vacuette©), and blood for RNA extraction (N = 122; the first women studied) was collected in PAX tubes (Beckton Dickinson, Franklin Lakes, NJ). Spot urine samples were collected and processed as described previously.[30] Blood and urine samples were kept at -20°C before and after transport (with cooling blocks) to Sweden for analysis.

Analysis of arsenic and other elements in urine

The concentration of arsenic metabolites, i.e. MMA, DMA, and remaining unmethylated inorganic arsenic in urine were measured using high performance liquid chromatography (HPLC) coupled with hydride generation and inductively coupled plasma mass spectrometry (ICPMS) (Agilent 7500ce; Agilent Technologies, Japan) employing appropriate quality control.[31] The sum of metabolite concentrations was used to assess arsenic exposure. The fractions of the different metabolites in urine were used to assess the efficiency of arsenic metabolism. We also measured cadmium in urine and selenium and zinc in blood. Cadmium has been reported to influence DNA methylation.[32-34] Selenium and zinc represent nutritional status. Cadmium in urine and selenium and zinc in blood were analyzed using ICPMS (Agilent 7500ce) with a collision/reaction cell system (Agilent 7500ce) with analytical accuracy ascertained from commercially available reference materials as described previously.[35-36] Concentrations of elements in urine were adjusted to the mean specific gravity of urine (1.020 g/mL) measured with a digital refractometer (EUROMEX RD 712 clinical refractometer; EUROMEX, Arnhem, the Netherlands).[37]

DNA isolation and pyrosequencing

DNA was isolated using the QIAGEN DNA Blood Midi kit (Qiagen, Hilden, Germany). One µg of DNA (50 ng/µL) was bisulfite-treated using an EpiTect® kit (Qiagen). Bisulfite-treated DNA (0.6–1 µl) was used in a 15- to 25-µL PCR reaction using the Pyromark PCR kit (Qiagen). One of the PCR primers was biotinylated. The PCR product was purified using Streptavidin Sepharose High Performance beads (Amersham Biosciences, Uppsala, Sweden). The Sepharose beads containing immobilized PCR products were purified, washed, and denatured using 0.2 N NaOH, and washed again using a vacuum prep tool (Pyrosequencing Inc., Westborough, MA). Twelve µL of pyrosequencing primer (0.3 µmol/L) was annealed to the purified single-stranded PCR product and pyrosequencing was done using the PSQ HS96 Pyrosequencing System (Qiagen). The degree of methylation was expressed as percentage of methylated cytosines divided by the sum of methylated and unmethylated cytosines. Bisulfite conversion was verified using non-CpG cytosine residues as built-in controls, and at least >95% conversion of cytosine at a non-CpG site in the pyrosequencing amplicon ensured successful bisulfite conversion.

Commercially available kits (Qiagen) were used to measure the methylation of *p16* at 7 CpG sites (between +148 and +174 in exon 1; GenBank accession no. L27211, http://www.ncbi.nlm.nih.gov/genbank/GenbankSearch.html) specific for the *p16* transcript, *MLH1* at 4 CpG sites (between -209 and -188; GenBank accession no U07418) (Figure 1), and *LINE1* at 4 CpG sites (between 305 and 331 of GenBank accession no. X58075), following the manufacturer's instructions. A single PCR fragment spanning part of each genetic element was amplified, and the degree of methylation was analyzed in a single pyrosequencing reaction using 3 µL PCR products for *p16*, 4 µL for *MLH1*, and 20 µL for *LINE1*. The samples were analyzed in triplicate for *p16* and *MLH1* and as singlets for *LINE1*. The repeatability of the method, expressed as the variation in coefficients, was 36.7, 26.4, and

2.0%, for *p16*, *MLH1*, and *LINE1* methylation, respectively. We performed amplification bias tests[38] for *p16* and *MLH1* using different proportions (0:100, 25:75, 50:50, 75:25, and 100:0) of methylated and unmethylated human DNA (Qiagen), and found both the assays as acceptable (amplification bias, b = 0.48 and 0.46 for *p16* and *MLH1*, respectively). We analyzed positive controls for *p16* and *MLH1* methylation obtained from Qiagen. The average degree of methylation was 88% (range 68-100%) for *p16* and 87% (range 84-90%) for *MLH1*. We also analyzed negative controls (whole genome amplified DNA and unmethylated DNA from Qiagen). The average methylation for these samples ranged between 1-2%.

Genotyping and gene expression analysis

We have previously shown that *AS3MT* haplotype 1, constituted by considering the polymorphisms of rs7085104, rs3740400, rs3740393, rs3740390, rs11191439, rs11191453, rs10748835, and rs1046778 (haplotype frequency 16.5% in this study population), is associated with more inorganic arsenic, more MMA and less DMA in urine (relative amounts), i.e. a slower metabolism of arsenic.[27] The most common haplotype in this population (haplotype 2) is associated with a more efficient metabolism. We therefore focused on hapotype 1 in this study. All polymorphisms, but one (rs11191439), are intronic. Some of the intronic polymorphisms are functional, as they are associated with altered expression of the *AS3MT* gene.[39] Genotyping for polymorphisms in the *AS3MT* was performed in peripheral blood DNA using Sequenom™ technology (Sequenom Inc.,San Diego, CA) at the Swegene DNA facility at Malmö University Hospital, Malmö, Sweden, as described previously.[27]

Blood was collected in PAX tubes and stored at -80°C until the RNA was extracted using the PAXgene Blood RNA kit (PreAnalytiX, Hombrechtikon, Switzerland). RNA

concentration and purity were evaluated on a NanoDrop spectrophotometer (NanoDrop Products,Wilmington, DE) and integrity (RNA integrity number, RIN) was evaluated on a Bioanalyzer 2100 (Agilent, Santa Clara, CA), and the results showed good RNA quality for the large majority of the samples (RIN>7.5). Of 122 available PAX tubes for RNA extraction, 90 samples with a sufficient amount of good quality RNA were selected for gene expression analysis (characteristics of the individuals in Table 1). DirectHyb HumanHT-12 v3.0 (Illumina, San Diego, CA) was used for gene expression analysis according to the manufacturer's instructions, at the SCIBLU facility of Lund University. Gene expression data were filtered with BioArray Software Environment (BASE) software. On the HT-12 arrays, one valid transcript for *p16* (exon 3) was present according to a BLAST search (blast.ncbi.nlm.nih.gov/Blast) with NM_000077.4 (www.ncbi.nlm.nih.gov/gene) as a reference RNA; one transcript was present for *MLH1* with NM_000249.3 as a reference RNA, but there was no transcript for *LINE1*.

Statistical analysis

We used Spearman's rank (r_s) correlation analysis to assess bivariate correlations between epigenetic markers (individual C-positions in *p16* and *MLH1* and also averages of all CpG sites measured for *p16, MLH1,* and *LINE1*), arsenic concentration and the fractions of arsenic metabolites in urine, *p16* and *MLH1* gene expression, and the characteristics of the women. Checking of the linearity of associations with outcome variables by P-P plots and visual inspection of scatter plots indicated that there were linear relationships between arsenic exposure on one hand and DNA methylation and gene expression on the other. In the further analyses we selected the CpG showing the strongest correlation and the best fit with linearity to arsenic exposure, in order to avoid too many comparisons. The linear fit improved with $\log_2$-transformed urinary arsenic concentration (Supplementary Figures 1and 2) and,

therefore, $\log_2$-transformed values were used in the linear regression analysis to estimate the associations (β-coefficients) of urinary arsenic concentrations with epigenetic markers. There was a strong correlation between individual concentrations of arsenic metabolites and total urinary arsenic (r_s = 0.81–0.99) and therefore only total urinary arsenic was used in relation to DNA methylation. We performed adjusted regression analysis, controlling for age and coca use, since these factors were weakly associated with the outcomes (univariate linear regression analysis, $p<0.2$). In addition, we controlled for concentrations of cadmium in urine, which we has previously found to be associated with epigenetic modifications in vitro[33] and selenium and zinc in blood, which are important for anti-oxidative defense. Cadmium concentration in urine was $\log_2$-transformed to obtain normal distributions of the residuals. We did not adjust for BMI, as it was highly correlated with age (r_s = 0.46). The same analysis was performed for *p16* and *MLH1* gene expression as was performed for *p16* and *MLH1* methylation.

To evaluate the effect of efficiency of arsenic metabolism on epigenetic markers, we analyzed the associations between the fractions (percentages; as continuous variables) of the different arsenic metabolites and epigenetic markers, taking total arsenic concentration in urine into account, since the efficiency of metabolism is affected by the level of exposure.[21] We also evaluated the effect of the major arsenic metabolizing gene *AS3MT* on the association between arsenic and *p16* and *MLH1* methylation (in individuals who were not first-degree relatives, N = 172). First, we analyzed the *p16* and *MLH1* methylation by genotype (the degree of methylation for groups with 0, 1, or 2 copies). Then, we evaluated the effect of arsenic on gene methylation, stratifying for haplotype copy number. Third, *AS3MT* haplotype was entered as a categorical variable in the analysis. In order to obtain an effect estimate in carriers with different numbers of haplotype copies, we performed

regression analysis of urinary arsenic and *p16* and *MLH1* methylation, stratified by *AS3MT* haplotype. All these analyses were adjusted for age and coca chewing.

Any *P* value < 0.05 was considered statistically significant. All statistical analyses were performed using SPSS (version 18; SPSS, Chicago, IL).

Results

Descriptive data for the study participants, the concentrations of elements in blood and urine, and the fractions of urinary metabolites are shown in Table 1. The arsenic concentration ranged from 10.1 to 1,251 μg/L (median was 230 μg/L) in urine. None of the women reported history of malignancy and none of them had arsenic-associated skin lesions such as hyperkeratosis on their hands. Comparisons of the all pairwise combinations showed that the *p16* methylations at the 7 different CpG sites were highly correlated (with r_s-values of 0.76–0.90, Supplementary Table 1), and also showed correlations with the average degree of *p16* methylation (r_s = 0.88–0.97). Similarly, methylations of different CpG sites in *MLH1* were highly correlated (r_s = 0.32–0.87, Supplementary Table 1). Although the *p16* and *MLH1* genes are located in two different chromosomes, their methylation levels were highly correlated (Supplementary Table 1).

The arsenic concentrations in urine were positively correlated with the degree of average *p16* and *MLH1* methylation (r_s-values of 0.21 and 0.20, respectively; Table 2), but not with *LINE1* methylation. Position 1 in both *p16* and *MLH1* (the first CpG site in each amplicon) was used in the subsequent analyses as this showed the strongest linearity, although rather weak, with arsenic concentrations in urine (Supplemental Figures 1 and 2), compared with the average degree of methylation (Table 2) and the other individual positions (data not shown). The fraction of inorganic arsenic in urine showed a positive correlation and

the fraction of DMA showed a negative correlation with *p16* methylation (r_s-values of 0.24 and -0.20, respectively), but not with *MLH1* methylation.

Arsenic concentrations showed a negative correlation (of borderline significance) with the expression of *p16* transcript in exon 3 (ILMN_1717714) in peripheral blood (r_s = -0.20; *P* = 0.066), but not with the expression of *MLH1* transcript from exons 19 and 20 (r_s = 0.0016; *P* = 0.99). None of the nutritional markers (BMI, selenium and zinc) showed any significant correlations with the epigenetic markers (r_s < 0.11).

In the linear regression analysis, urinary arsenic concentrations were found to be positively associated with *p16* methylation (position 1: β = 0.14; *P* = 0.0028) (Table 3). Adjustments for age and coca use, and also urinary cadmium, selenium and zinc concentrations in blood, did not alter the association (β = 0.14; *P* = 0.0039). The fractions of inorganic arsenic and DMA in urine, but not MMA, were associated with *p16* methylation (position 1) (Table 3). The percentage of inorganic arsenic showed the strongest association (β_1 = 0.026; *P* = 0.010), and the association of the total concentrations of arsenic in urine remained essentially the same (β_2 = 0.12; *P* = 0.013). Also in the model with percentage of DMA as main predictor (β_1 = -0.017; *P* = 0.043), the association of the total concentrations of arsenic remained the same (β_2 = 0.12; *P* = 0.013) (Table 3). Further adjustments for age and coca use decreased the association of inorganic arsenic with *p16* methylation by about 10%; the association with DMA became statistically non-significant (*P* = 0.082).

Arsenic concentrations were also found to be positively associated with *MLH1* methylation (position 1: β = 0.28, *P* = 0.0011), and the effect estimate became even stronger after adjustments (Table 3). However, the metabolite fractions were not associated with *MLH1* methylation (Table 3). The average degree of both *p16* and *MLH1* methylation showed similar, but somewhat weaker, associations with urinary arsenic and arsenic fractions (data not shown). The linear regression analysis showed no association between arsenic

concentrations in urine, or arsenic metabolite fractions, and *LINE1* methylation (Supplemental Table 2).

Evaluation of the influence of the major arsenic metabolizing gene *AS3MT* showed that the slow arsenic-metabolizing *AS3MT* haplotype 1 had an effect on *p16* methylation (Table 4). Individuals carrying two or one copies of the *AS3MT* haplotype 1 showed more methylation of *p16* than null carriers of the haplotype ($P = 0.022$ for the main effect of the haplotype; adjusted $P = 0.043$) (Table 4). A haplotype dosage effect was observed with on average, 13% more *p16* methylation per haplotype copy. No effect of the *AS3MT* haplotype 1 was observed for *MLH1* methylation ($P = 0.86$ for the main effect of the haplotype; adjusted $P = 0.82$).

As shown in Table 1, the descriptive data for the 90 women who had their blood samples used for gene expression were similar to those for the whole cohort of 202 women. There was a negative association between expression of *p16* (N = 90) and $\log_2$-transformed arsenic in urine ($\beta = -1.45$; $P = 0.0094$ and $\beta = -1.48$; $P = 0.0091$ when adjusted for age and coca use) (Figure 2). There was one influential outlier with very high *p16* expression. After removal of this individual, the association became much weaker ($\beta = -0.81$; $P = 0.081$). The outlier did not show any characteristics that deviated from the rest of the population. Urinary arsenic concentrations were not associated with the degree of *MLH1* expression ($\beta = -0.0082$; $P = 1.00$). There was no correlation between *p16* or *MLH1* methylation and *p16* or *MLH1* expression

Discussion

In this study we found that women who had long-term exposure to arsenic through drinking water demonstrated a correlation with increased DNA methylation of both the tumor suppressor gene *p16* and the DNA mismatch repair gene *MLH1*. We did not find any

association between methylation and gene expression in a subset of samples. However, arsenic exposure was found to be associated with reduced expression of *p16* mRNA in peripheral blood, but this result needs to be interpreted cautiously, as it was partly influenced by an outlier. The efficiency of arsenic metabolism showed relationship with the degree of *p16* gene methylation. Women with a higher fraction of the dimethylated arsenic metabolite (DMA) in urine, which is associated with faster arsenic excretion and less toxicity,[21] had a lower degree of *p16* methylation. Surprisingly, only the fraction of inorganic arsenic in urine, and not that of MMA, showed a positive association with *p16* methylation. This indicates that a high percentage of MMA in urine, which is generally associated with more toxic effects, including genotoxicity,[23, 40] may be less relevant to epigenetic changes. Further support for the role of arsenic metabolism in susceptibility to arsenic-related epigenetic effects came from the observed association of genotype of the major arsenic-metabolizing gene *AS3MT*: i.e., the effect of arsenic on *p16* methylation increased with increasing number of copies of the slow metabolism haplotype.

We conducted this study on a homogeneous population in which subjects differed very little in terms of diet, as revealed from questionnaire response, and exposure to industrial pollution. Also, there were very few smokers and only one subject reported alcohol consumption. A weakness of the study was the analysis of DNA methylation in the DNA isolated from whole blood only, and not from different populations of blood cells. Despite the fact that methylation in whole blood cells has been used successfully for the study of epigenetic changes in relation to disease,[41] the effects of exposure to toxicants on DNA methylation may be different in different types of blood cells.[42] The women did not report any malignancy and it was thus not possible to analyze epigenetic changes in arsenic-related tumors.

The p16 is a key regulator of the cell cycle and hypermethylation of *p16* is found in tumors of the colon, liver, and skin,[43] as well as in premalignant lesions.[44-46] Previous studies on *p16* alterations in relation to arsenic exposure were smaller and had no information about individual arsenic exposure and/or influence of arsenic metabolism. Fifty-one subjects exposed to arsenic through drinking water in India showed hypermethylation of *p16* at high levels of arsenic in water (>250 ug/L).[11] Studies of individuals exposed to arsenic-rich coal in unventilated homes in China indicated trends of increasing *p16* methylation in peripheral blood lymphocytes (N = 103)[13], and reduced amounts of p16 protein in skin samples (N = 51) [47], with increasing severity of arsenic-related skin changes (hyperpigmentation, hypopigmentation, or squamous adenocarcinoma). However, the comparison between our study and the two reports by Zhang et al.[13, 47] is not straightforward, as the arsenic-rich coal probably contained several other carcinogenic compounds that could induce alterations in *p16.*

MLH1 is an important component in the maintenance of genomic stability, both for repair of small DNA mismatches and for proper recombination during meiosis.[48] Promoter methylation of the *MLH1* gene has been implicated in the development of gastric and colorectal carcinoma.[14-15] Increasing DNA damage with increasing arsenic exposure and impaired expression of genes involved in different DNA repair systems, including the mismatch repair gene *hMSH2,* have been reported in humans. Again, the arsenic exposure came from arsenic-rich coal, and compounds other than arsenic may have had a negative effect on the repair systems.[13] We could not show reduced *MLH1* expression in relation to arsenic exposure, and arsenic metabolism did not significantly influence *MLH1* methylation. The association of arsenic exposure and *MLH1* methylation we observed should be treated with caution as there was a high degree of correlation between *p16* and *MLH1* methylation.

Arsenic is metabolized mainly by *AS3MT,* and different polymorphisms in this gene influence the efficiency of the arsenic metabolism.[27] The population under study has a very efficient arsenic metabolism compared to other populations.[28] This is probably largely determined by the low frequency (16.5%) of the *AS3MT* slowly metabolizing haplotype 1 and a very high frequency (70%) of the quickly metabolizing haplotype 2 in this population compared to most other populations.[27, 49] The efficient arsenic metabolism may explain that the average level of *p16* methylation in peripheral blood was low (on average 3%) in spite of the fairly high arsenic exposure. The low level of *p16* methylation detected here are consistent with unmethylated CpG islands, since a low level of DNA methylation (~3%) is typically found in the promoter-region CpG islands[50] and similar to *p16* methylation in blood samples from non-exposed Chinese individuals (2%). In contrast, Chinese patients with arsenic-related diseases demonstrated much higher *p16* methylation levels in blood (42%).[47] It should be mentioned that one reason behind the discrepancies in the previous studies may be due to differences in *AS3MT* genotype distribution which was not reported before. However, despite the low *p16* methylation levels we could still show a significant effect of arsenic on methylation of *p16* and *MLH1*. Still, since the levels of methylation were low in this population, this warrants some caution and future studies are needed to follow up our findings.

Despite the fact that there was a wide range of arsenic exposure, we found no indications that arsenic affected global DNA methylation. This contrasts with the results of a study in Bangladesh where arsenic exposure through drinking water was found to be positively associated with increased global DNA methylation.[51] This difference in findings may be attributed to the differences in *AS3MT* genotypes and in the methods used to measure global DNA methylation. The investigators used [^{3}H]-methyl incorporation assay assessing

methylation at all genomic CpG sites and we used bisulfite-PCR pyrosequencing to quantitate *LINE1* methylation, which served as a surrogate for global DNA methylation.

In conclusion, we have shown that arsenic levels in individuals with prolonged arsenic exposure were correlated with increased DNA methylation in tumor suppressor genes in peripheral blood. Importantly, changes seen with inorganic arsenic suggest it may be more toxic for the epigenome than the monomethylated MMA metabolite.

Supplementary material

Supplementary Tables 1 and 2, and Figures 1 and 2 can be found online.

Funding

This research was supported by the Swedish Council for Working Life and Social Research, Karolinska Institutet, Kungliga fysiografiska sällskapet, and by the EU (FP6; "PHIME" FOOD-CT-2006-016253).

Acknowledgements

The authors thank the Department of Oncology, Lund University for technical assistance.

Conflict of Interest Statement: None declared.

References

1. IARC *Some drinking-water disinfectants and contaminants, including arsenic.*; IARC Monographs on the Evaluation of Carcinogenic Risks to Humans International Agency for Research on Cancer: Lyon, 2004.

2. NRC *Arsenic in Drinking Water: 2001 update*; Washington DC, 2001.

3. K. Straif, L. Benbrahim-Tallaa, R. Baan, Y. Grosse, B. Secretan, F. El Ghissassi, V. Bouvard, N. Guha, C. Freeman, L. Galichet and V. Cogliano, A review of human carcinogens--part C: metals, arsenic, dusts, and fibres. *Lancet Oncol* 2009, **10**, 453-4.

4. A. Basu, J. Mahata, S. Gupta and A. K. Giri, Genetic toxicology of a paradoxical human carcinogen, arsenic: a review. *Mutat Res* 2001, **488**, 171-94.

5. T. G. Rossman, D. Stone, M. Molina and W. Troll, Absence of arsenite mutagenicity in E coli and Chinese hamster cells. *Environ Mutagen* 1980, **2**, 371-9.

6. Y. Noda, T. Suzuki, A. Kohara, A. Hasegawa, T. Yotsuyanagi, M. Hayashi, T. Sofuni, K. Yamanaka and S. Okada, In vivo genotoxicity evaluation of dimethylarsinic acid in MutaMouse. *Mutat Res* 2002, **513**, 205-12.

7. K. Harrington-Brock, M. Cabrera, D. D. Collard, C. L. Doerr, R. McConnell, M. M. Moore, H. Sandoval and J. C. Fuscoe, Effects of arsenic exposure on the frequency of HPRT-mutant lymphocytes in a population of copper roasters in Antofagasta, Chile: a pilot study. *Mutat Res* 1999, **431**, 247-57.

8. X. Ren, C. M. McHale, C. F. Skibola, A. H. Smith, M. T. Smith and L. Zhang, An emerging role for epigenetic dysregulation in arsenic toxicity and carcinogenesis. *Environ Health Perspect* 2011, **119**, 11-9.

9. T. G. Rossman and C. B. Klein, Genetic and epigenetic effects of environmental arsenicals. *Metallomics* 2011, **3**, 1135-41.

10. M. J. Mass and L. Wang, Arsenic alters cytosine methylation patterns of the promoter of the tumor suppressor gene p53 in human lung cells: a model for a mechanism of carcinogenesis. *Mutat Res* 1997, **386**, 263-77.

11. S. Chanda, U. B. Dasgupta, D. Guhamazumder, M. Gupta, U. Chaudhuri, S. Lahiri, S. Das, N. Ghosh and D. Chatterjee, DNA hypermethylation of promoter of gene p53 and p16 in arsenic-exposed people with and without malignancy. *Toxicol Sci* 2006, **89**, 431-7.

12. Y. C. Huang, W. C. Hung, W. T. Chen, W. H. Jiang, H. S. Yu and C. Y. Chai, Effects of MEK and DNMT inhibitors on arsenic-treated human uroepithelial cells in relation to Cyclin-D1 and p16. *Toxicol Lett* 2011, **200**, 59-66.

13. A. Zhang, H. Feng, G. Yang, X. Pan, X. Jiang, X. Huang, X. Dong, D. Yang, Y. Xie, L. Peng, L. Jun, C. Hu, L. Jian and X. Wang, Unventilated indoor coal-fired stoves in Guizhou province, China: cellular and genetic damage in villagers exposed to arsenic in food and air. *Environ Health Perspect* 2007, **115**, 653-8.

14. J. M. Wheeler, A. Loukola, L. A. Aaltonen, N. J. Mortensen and W. F. Bodmer, The role of hypermethylation of the hMLH1 promoter region in HNPCC versus MSI+ sporadic colorectal cancers. *J Med Genet* 2000, **37**, 588-92.

15. K. Imai and H. Yamamoto, Carcinogenesis and microsatellite instability: the interrelationship between genetics and epigenetics. *Carcinogenesis* 2008, **29**, 673-80.

16. E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. D. McPherson, M. A. Marra, E. R. Mardis, L. A. Fulton, A. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chissoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, A. Delehaunty, J. B. Kramer, L. L. Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J. F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R. A. Gibbs, D. M. Muzny, S. E. Scherer, J. B. Bouck, E. J. Sodergren, K. C. Worley, C. M. Rives, J. H. Gorrell, M. L. Metzker, S. L. Naylor, R. S. Kucherlapati, D. L.

Nelson, G. M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D. R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H. M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R. W. Davis, N. A. Federspiel, A. P. Abola, M. J. Proctor, R. M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D. R. Cox, M. V. Olson, R. Kaul, N. Shimizu, K. Kawasaki, S. Minoshima, G. A. Evans, M. Athanasiou, R. Schultz, B. A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W. R. McCombie, M. de la Bastide, N. Dedhia, H. Blocker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J. A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D. G. Brown, C. B. Burge, L. Cerutti, H. C. Chen, D. Church, M. Clamp, R. R. Copley, T. Doerks, S. R. Eddy, E. E. Eichler, T. S. Furey, J. Galagan, J. G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L. S. Johnson, T. A. Jones, S. Kasif, A. Kaspryzk, S. Kennedy, W. J. Kent, P. Kitts, E. V. Koonin, I. Korf, D. Kulp, D. Lancet, T. M. Lowe, A. McLysaght, T. Mikkelsen, J. V. Moran, N. Mulder, V. J. Pollara, C. P. Ponting, G. Schuler, J. Schultz, G. Slater, A. F. Smit, E. Stupka, J. Szustakowski, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y. I. Wolf, K. H. Wolfe, S. P. Yang, R. F. Yeh, F. Collins, M. S. Guyer, J. Peterson, A. Felsenfeld, K. A. Wetterstrand, A. Patrinos, M. J. Morgan, P. de Jong, J. J. Catanese, K. Osoegawa, H. Shizuya, S. Choi and Y. J. Chen, Initial sequencing and analysis of the human genome. *Nature* 2001, **409**, 860-921.

17. A. Baccarelli, R. O. Wright, V. Bollati, L. Tarantini, A. A. Litonjua, H. H. Suh, A. Zanobetti, D. Sparrow, P. S. Vokonas and J. Schwartz, Rapid DNA methylation changes after exposure to traffic particles. *Am J Respir Crit Care Med* 2009, **179**, 572-8.

18. C. Q. Zhao, M. R. Young, B. A. Diwan, T. P. Coogan and M. P. Waalkes, Association of arsenic-induced malignant transformation with DNA hypomethylation and aberrant gene expression. *Proc Natl Acad Sci U S A* 1997, **94**, 10907-12.

19. H. Chen, S. Li, J. Liu, B. A. Diwan, J. C. Barrett and M. P. Waalkes, Chronic inorganic arsenic exposure induces hepatic global and individual gene hypomethylation: implications for arsenic hepatocarcinogenesis. *Carcinogenesis* 2004, **25**, 1779-86.

20. S. H. Choi, S. Worswick, H. M. Byun, T. Shear, J. C. Soussa, E. M. Wolff, D. Douer, G. Garcia-Manero, G. Liang and A. S. Yang, Changes in DNA methylation of tandem DNA repeats are different from interspersed repeats in cancer. *Int J Cancer* 2009, **125**, 723-9.

21. M. Vahter, Mechanisms of arsenic biotransformation. *Toxicology* 2002, **181-182**, 211-7.

22. K. Piatek, T. Schwerdtle, A. Hartwig and W. Bal, Monomethylarsonous acid destroys a tetrathiolate zinc finger much more efficiently than inorganic arsenite: mechanistic considerations and consequences for DNA repair inhibition. *Chem Res Toxicol* 2008, **21**, 600-6.

23. A. L. Lindberg, E. C. Ekstrom, B. Nermell, M. Rahman, B. Lonnerdal, L. A. Persson and M. Vahter, Gender and age differences in the metabolism of inorganic arsenic in a highly exposed population in Bangladesh. *Environ Res* 2008, **106**, 110-20.

24. C. J. Chung, Y. M. Hsueh, C. H. Bai, Y. K. Huang, Y. L. Huang, M. H. Yang and C. J. Chen, Polymorphisms in arsenic metabolism genes, urinary arsenic methylation profile and cancer. *Cancer Causes Control* 2009, **20**, 1653-61.

25. R. M. Gardner, B. Nermell, M. Kippler, M. Grander, L. Li, E. C. Ekstrom, A. Rahman, B. Lonnerdal, A. M. Hoque and M. Vahter, Arsenic methylation efficiency increases during the first trimester of pregnancy independent of folate status. *Reprod Toxicol* 2011, **31**, 210-8.

26. S. Lin, Q. Shi, F. B. Nix, M. Styblo, M. A. Beck, K. M. Herbin-Davis, L. L. Hall, J. B. Simeonsson and D. J. Thomas, A novel S-adenosyl-L-methionine:arsenic(III) methyltransferase from rat liver cytosol. *J Biol Chem* 2002, **277**, 10795-803.

27. K. Engstrom, M. Vahter, S. J. Mlakar, G. Concha, B. Nermell, R. Raqib, A. Cardozo and K. Broberg, Polymorphisms in arsenic(+III oxidation state) methyltransferase (AS3MT) predict gene expression of AS3MT as well as arsenic metabolism. *Environ Health Perspect* 2011, **119**, 182-8.

28. M. Vahter, G. Concha, B. Nermell, R. Nilsson, F. Dulout and A. T. Natarajan, A unique metabolism of inorganic arsenic in native Andean women. *Eur J Pharmacol* 1995, **293**, 455-62.

29. K. Broberg, G. Concha, K. Engstrom, M. Lindvall, M. Grander and M. Vahter, Lithium in drinking water and thyroid function. *Environ Health Perspect* 2011, **119**, 827-30.

30. G. Concha, B. Nermell and M. V. Vahter, Metabolism of inorganic arsenic in children with chronic high arsenic exposure in northern Argentina. *Environ Health Perspect* 1998, **106**, 355-9.

31. A. L. Lindberg, W. Goessler, M. Grander, B. Nermell and M. Vahter, Evaluation of the three most commonly used analytical methods for determination of inorganic arsenic and its metabolites in urine. *Toxicol Lett* 2007, **168**, 310-8.

32. M. B. Hossain, M. Vahter, G. Concha and K. Broberg, Low-level environmental cadmium exposure is associated with DNA hypomethylation in Argentinean women. *Environ Health Perspect* 2012, **120**, 879-84.

33. D. Huang, Y. Zhang, Y. Qi, C. Chen and W. Ji, Global DNA hypomethylation, rather than reactive oxygen species (ROS), a potential facilitator of cadmium-stimulated K562 cell proliferation. *Toxicol Lett* 2008, **179**, 43-7.

34. K. H. Hong, C. L. Keen, Y. Mizuno, K. E. Johnston and T. Tamura, Effects of dietary zinc deficiency on homocysteine and folate metabolism in rats. *J Nutr Biochem* 2000, **11**, 165-9.

35. G. Concha, K. Broberg, M. Grander, A. Cardozo, B. Palm and M. Vahter, High-level exposure to lithium, boron, cesium, and arsenic via drinking water in the Andes of northern Argentina. *Environ Sci Technol* 2010, **44**, 6875-80.

36. M. Kippler, W. Goessler, B. Nermell, E. C. Ekstrom, B. Lonnerdal, S. El Arifeen and M. Vahter, Factors influencing intestinal cadmium uptake in pregnant Bangladeshi women--a prospective cohort study. *Environ Res* 2009, **109**, 914-21.

37. Y. Suwazono, A. Akesson, T. Alfven, L. Jarup and M. Vahter, Creatinine versus specific gravity-adjusted urinary cadmium concentrations. *Biomarkers* 2005, **10**, 117-26.

38. P. M. Warnecke, C. Stirzaker, J. R. Melki, D. S. Millar, C. L. Paul and S. J. Clark, Detection and measurement of PCR bias in quantitative methylation analysis of bisulphite-treated DNA. *Nucleic Acids Res* 1997, **25**, 4422-6.

39. K. S. Engstrom, M. Vahter, C. Lindh, F. Teichert, R. Singh, G. Concha, B. Nermell, P. B. Farmer, U. Stromberg and K. Broberg, Low 8-oxo-7,8-dihydro-2'-deoxyguanosine levels and influence of genetic background in an Andean population exposed to high levels of arsenic. *Mutat Res* 2010, **683**, 98-105.

40. F. Ebert, A. Weiss, M. Bultemeyer, I. Hamann, A. Hartwig and T. Schwerdtle, Arsenicals affect base excision repair by several mechanisms. *Mutat Res-Fund Mol M* 2011, **715**, 32-41.

41. M. B. Terry, L. Delgado-Cruzata, N. Vin-Raviv, H. C. Wu and R. M. Santella, DNA methylation in white blood cells: association with risk factors in epidemiologic studies. *Epigenetics* 2011, **6**, 828-37.

42. H. C. Wu, L. Delgado-Cruzata, J. D. Flom, M. Kappil, J. S. Ferris, Y. Liao, R. M. Santella and M. B. Terry, Global methylation profiles in DNA from different blood cell types. *Epigenetics* 2011, **6**, 76-85.

43. D. Deng, Z. Liu and Y. Du, Epigenetic alterations as cancer diagnostic, prognostic, and predictive biomarkers. *Adv Genet* 2010, **71**, 125-76.

44. Z. Jicai, Y. Zongtao, L. Jun, L. Haiping, W. Jianmin and H. Lihua, Persistent infection of hepatitis B virus is involved in high rate of p16 methylation in hepatocellular carcinoma. *Mol Carcinog* 2006, **45**, 530-6.

45. T. Kukitsu, T. Takayama, K. Miyanishi, A. Nobuoka, S. Katsuki, Y. Sato, R. Takimoto, T. Matsunaga, J. Kato, T. Sonoda, S. Sakamaki and Y. Niitsu, Aberrant crypt foci as precursors of the dysplasia-carcinoma sequence in patients with ulcerative colitis. *Clin Cancer Res* 2008, **14**, 48-54.

46. J. D. Licchesi, W. H. Westra, C. M. Hooker and J. G. Herman, Promoter hypermethylation of hallmark cancer genes in atypical adenomatous hyperplasia of the lung. *Clin Cancer Res* 2008, **14**, 2570-8.

47. A. H. Zhang, H. H. Bin, X. L. Pan and X. G. Xi, Analysis of p16 gene mutation, deletion and methylation in patients with arseniasis produced by indoor unventilated-stove coal usage in Guizhou, China. *J Toxicol Environ Health A* 2007, **70**, 970-5.

48. J. E. Stone and T. D. Petes, Analysis of the proteins involved in the in vivo repair of base-base mismatches and four-base loops formed during meiotic recombination in the yeast Saccharomyces cerevisiae. *Genetics* 2006, **173**, 1223-39.

49. J. Fujihara, T. Yasuda, H. Kato, I. Yuasa, A. Panduro, T. Kunito and H. Takeshita, Genetic variants associated with arsenic metabolism within human arsenic (+3 oxidation state) methyltransferase show wide variation across multiple populations. *Arch Toxicol* 2011, **85**, 119-25.

50. A. K. Maunakea, R. P. Nagarajan, M. Bilenky, T. J. Ballinger, C. D'Souza, S. D. Fouse, B. E. Johnson, C. Hong, C. Nielsen, Y. Zhao, G. Turecki, A. Delaney, R. Varhol, N. Thiessen, K. Shchors, V. M. Heine, D. H. Rowitch, X. Xing, C. Fiore, M. Schillebeeckx, S. J. Jones, D. Haussler, M. A. Marra, M. Hirst, T. Wang and J. F. Costello, Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* 2010, **466**, 253-7.

51. J. R. Pilsner, X. Liu, H. Ahsan, V. Ilievski, V. Slavkovich, D. Levy, P. Factor-Litvak, J. H. Graziano and M. V. Gamble, Genomic methylation of peripheral blood leukocyte DNA: influences of arsenic and folate in Bangladeshi adults. *Am J Clin Nutr* 2007, **86**, 1179-86.

52. M. Fatemi, M. M. Pao, S. Jeong, E. N. Gal-Yam, G. Egger, D. J. Weisenberger and P. A. Jones, Footprinting of mammalian promoters: use of a CpG DNA methyltransferase revealing nucleosome positions at a single molecule level. *Nucleic Acids Res* 2005, **33**, e176.

53. C. E. Pardo, I. M. Carr, C. J. Hoffman, R. P. Darst, A. F. Markham, D. T. Bonthron and M. P. Kladde, MethylViewer: computational analysis and editing for bisulfite sequencing and methyltransferase accessibility protocol for individual templates (MAPit) projects. *Nucleic Acids Res* 2011, **39**, e5.

Table 1. Descriptive data on study participants and variables measured in blood and urine

	N = 202[a]		N = 90[a]	
Variable	Median	5–95 %	Median	5–95 %
Age (years)	34.0	18.0–64.0	32.0	15.0–60.5
Body mass index (BMI) (kg/m^2)	24.7	18.8–35.0	24.0	18.0–33.1
p16 (average) methylation (%)[b]	3.4	1.7–5.8	3.2	1.7–5.8
p16 (position 1) methylation (%)	2.8	1.2–4.2	2.8	1.2–4.2
MLH1 (average) methylation (%)[c]	4.2	1.2–7.1	3.6	1.1–7.1
MLH1 (position 1) methylation (%)	1.9	0.0–4.7	1.2	0.0–5.1
LINE1 (average) methylation (%)[d]	86.2	83.5–89.3	85.7	83.6–89.1
Urinary arsenic (µg/L)[e]	230	20.8–545	203	15.6–622
Fraction of urinary inorganic arsenic (%)	11.6	4.3–24.3	13.0	3.6–23.2
Fraction of urinary MMA (%)	7.7	3.9–14.1	7.7	3.5–16.4
Fraction of urinary DMA (%)	80.2	65.7–91.2	78.8	67.0–92.3
Blood zinc (mg/L)	6.8	5.5–8.2	6.8	5.3–8.2
Blood selenium (µg/L)	176	150–221	178	154–220
Urinary cadmium (µg/L)[e]	0.23	0.091–0.70	0.20	0.092–0.61

[a]N = 202 individuals for all variables apart from blood selenium (N = 201). Gene expression was analyzed in 90 individuals with sufficient good quality RNA.

[b]Average of relative amounts of C in 7 CpG sites (between +148 and +174) in exon 1, GenBank no. L27211.

[c]Average of relative amounts of C in 4 CpG sites (between -209 and -188), no. U07418.

[d]Average of relative amounts of C in 4 CpG sites (between +305 and +331) no. X58075.

[e]Adjusted for specific gravity of urine (1.020 g/ml).

Table 2. Spearman's rank correlations for age, epigenetic markers, arsenic concentrations and fractions (%) of metabolites in urine, and gene expression in blood[a]

	p16	*p16* (P1)	*MLH1*	*MLH1* (P1)	*LINE1*	Arsenic	iAs (%)	MMA (%)	DMA (%)	*p16* exprn[b]	*MLH1* exprn[c]
Age	-0.12	-0.13	-0.049	-0.035	-0.095	0.076	-0.15*	-0.12	0.21**	0.0037	0.12
p16 (average)		0.88**	0.78**	0.58**	0.42**	0.21**	0.17*	0.0044	-0.13	0.027	-0.080
p16 (position 1)			0.64**	0.44**	0.33**	0.21**	0.24**	0.012	-0.20**	0.020	-0.077
MLH1 (average)				0.71**	0.51**	0.20**	0.090	-0.0028	-0.060	0.032	-0.10
MLH1 (position 1)					0.44**	0.22**	0.087	-0.0076	-0.048	-0.016	-0.093
LINE1 (average)						-0.011	0.0071	0.068	-0.018	-0.048	-0.017
Arsenic (concentration)							0.17*	0.10	-0.17*	-0.20	0.0016
Inorganic arsenic, iAs (%)								0.22**	-0.93**	-0.085	-0.046
MMA (%)									-0.53**	0.0018	-0.046
DMA (%)										0.069	0.056
p16 expression[b]											-0.061

[a]**P* <0.05; ***P* <0.01; exprn, expression; iAs, inorganic arsenic; MMA, monomethylated arsenic; DMA, dimethylarsinic acid.

[b]*p16* expression was represented by the transcript ILMN_1717714 (Illumina HT12 array).

[c]*MLH1* expression was represented by the transcript ILMN_1788363.

Table 3: Associations between arsenic concentration (log_2-transformed) and fraction (%) of arsenic metabolites in urine and degree of *p16* and *MLH1*methylation.

		p16 (position 1)		*MLH1* (position 1)	
Predictor		β_1 (95% CI)	*P* value	β_1 (95% CI)	*P* value
Urinary arsenic	Unadjusted [a]	0.14 (0.050 to 0.24)	0.0028	0.28 (0.11 to 0.45)	0.0011
	Model 1	0.15 (0.057 to 0.25)	0.0017	0.30 (0.13 to 0.47)	0.00047
	Model 2	0.14 (0.047 to 0.24)	0.0039	0.32 (0.15 to 0.50)	0.00030
Inorganic arsenic (%)	Analysis 1[b]	0.026 (0.0064 to 0.046)	0.010	0.00048 (-0.036 to 0.037)	0.98
	Analysis 2	0.024 (0.0040 to 0.045)	0.019	-0.0066 (-0.043 to 0.030)	0.73
MMA (%)	Analysis 1	-0.0069 (-0.047 to 0.033)	0.74	-0.037 (-0.11 to 0.035)	0.31
	Analysis 2	-0.012 (-0.053 to 0.029)	0.56	-0.051 (-0.12 to 0.021)	0.16
DMA (%)	Analysis 1	-0.017 (-0.034 to -0.00059)	0.043	0.0063 (-0.024 to 0.037)	0.69
	Analysis 2	-0.015 (-0.033 to -0.0020)	0.082	0.014 (-0.017 to 0.045)	0.37

[a]Unadjusted model (N = 202): Methylation (%) = $\alpha_1 + \beta_1$ arsenic concentrations (log_2-transformed); Model 1: model adjusted for age and coca chewing; Model 2: model with additional adjustments for cadmium in urine (log_2-transformed) and selenium and zinc in blood.

[b]Analysis 1 (N = 202): Methylation (%) = $\alpha_1 + \beta_1$ fraction of arsenic metabolite + β_2 arsenic concentrations (log_2-transformed); Analysis 2: model adjusted for age and coca chewing.

Table 4. Effects of *AS3MT* haplotype 1 on average *p16* (position 1) and average *MLH1* (position 1) methylation levels (N = 172)

Genetic variation		*p16* methylation[a]			
		Mean (SD)	β_1 (95% CI)	*P* value[b]	*P* value[c]
AS3MT haplotype 1	0 copy (N = 95)	2.57 (0.92)	0.13 (-0.0038 to 0.27)	0.022	0.043
	1 copy (N = 61)	2.89 (0.93)	0.19 (0.040 to 0.34)		
	2 copies (N = 16)	3.29 (0.89)	0.32 (-0.52 to 1.16)		
		MLH1 methylation[a]			
		Mean (SD)	β_1 (95% CI)	*P* value[b]	*P* value[c]
AS3MT haplotype 1	0 copy (N = 95)	1.90 (1.68)	0.29 (0.039 to 0.54)	0.86	0.82
	1 copy (N = 61)	2.08 (1.61)	0.29 (0.028 to 0.56)		
	2 copies (N = 16)	2.31 (1.78)	-0.75 (-2.42 to 0.92)		

[a] *AS3MT* haplotype-stratified mean values (standard deviations) for *p16* methylation, and effect estimates for β_1 from the model: *p16* and *MLH1* methylation (%) = $\alpha 1 + \beta_1$ arsenic concentration ($\log_2$-tranformed).

[b] *P* for β_1 from the model: *p16* and *MLH1* methylation (%) = $\alpha_1 + \beta_1$ *AS3MT* haplotype 1 + β_2 arsenic concentration ($\log_2$-transformed).

[c] *P* for β_1 from the model above adjusted for age and coca chewing.

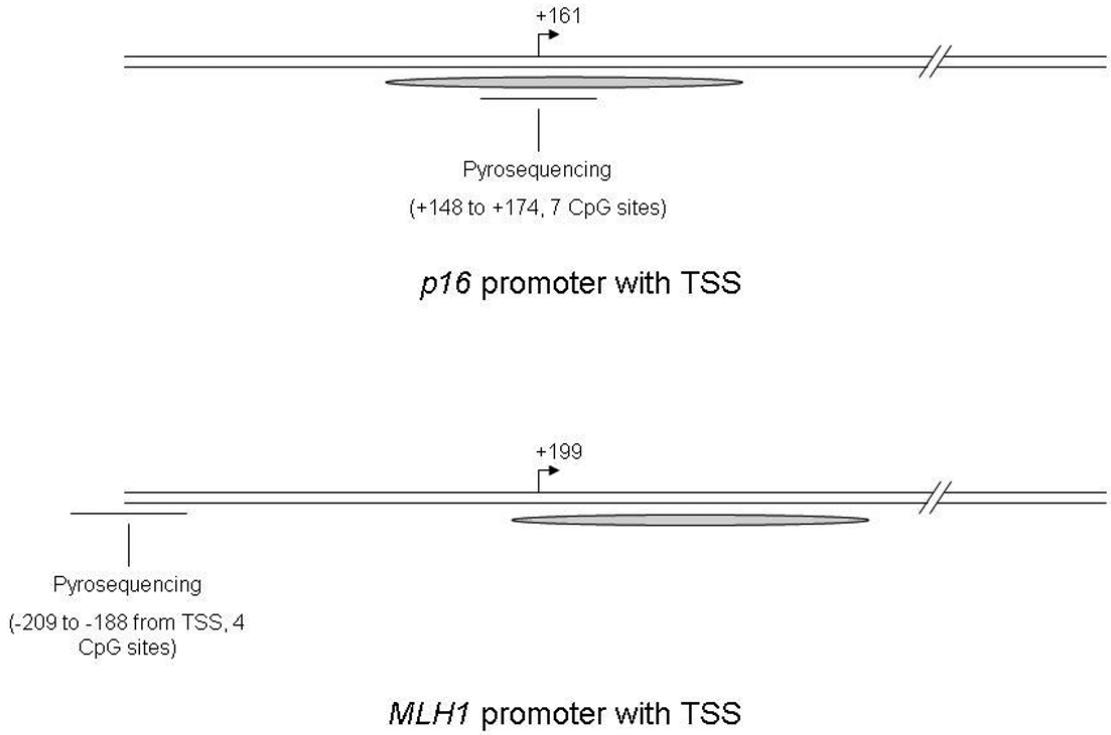


Figure 1: Pyrosequencing amplicons of *p16* and *MLH1* relative to known landmarks, including transcription start sites. Previously reported nucleosome positions (oval shape) for *p16* [52] and *MLH1* [53] are shown.

Figure 2: Scatter plot of the association between the degree of *p16* gene expression (arbitrary units) and urinary arsenic concentration (µg/L, adjusted to specific gravity 1.020 g/ml, log_2-transformed).

Supplemental material for "**Environmental arsenic exposure and DNA methylation of the tumor suppressor gene *p16* and the DNA repair gene *MLH1*; effect of arsenic metabolism and genotype**"

Mohammad Bakhtiar Hossain, Marie Vahter, Gabriela Concha, Karin Broberg

Supplemental Table 1. Spearman's rank correlations for urinary arsenic concentration and different *p16* and *MLH1* methylation positions (N = 202)[a]

	p16							*MLH1*			
	P1	P2	P3	P4	P5	P6	P7	P1	P2	P3	P4
Arsenic[b]	0.21**	0.19**	0.17*	0.16*	0.19**	0.20**	0.16*	0.22**	0.17*	0.13	0.21**
p16 P1 (position 1)		0.82**	0.81**	0.83**	0.82**	0.79**	0.76**	0.44**	0.44**	0.62**	0.50**
p16 P2			0.90**	0.90**	0.90**	0.86**	0.85**	0.55**	0.52**	0.73**	0.55**
p16 P3				0.89**	0.85**	0.80**	0.79**	0.56**	0.53**	0.72**	0.58**
p16 P4					0.84**	0.83**	0.84**	0.53**	0.55**	0.71**	0.59**
p16 P5						0.85**	0.82**	0.60**	0.42**	0.74**	0.49**
p16 P6							0.82**	0.54**	0.50**	0.73**	0.47**
p16 P7								0.57**	0.42**	0.73**	0.47**
MLH1 P1									0.32**	0.70**	0.33**
MLH1 P2										0.43**	0.74**
MLH1 P3											0.46**

[a]**P* <0.05; ***P* <0.01

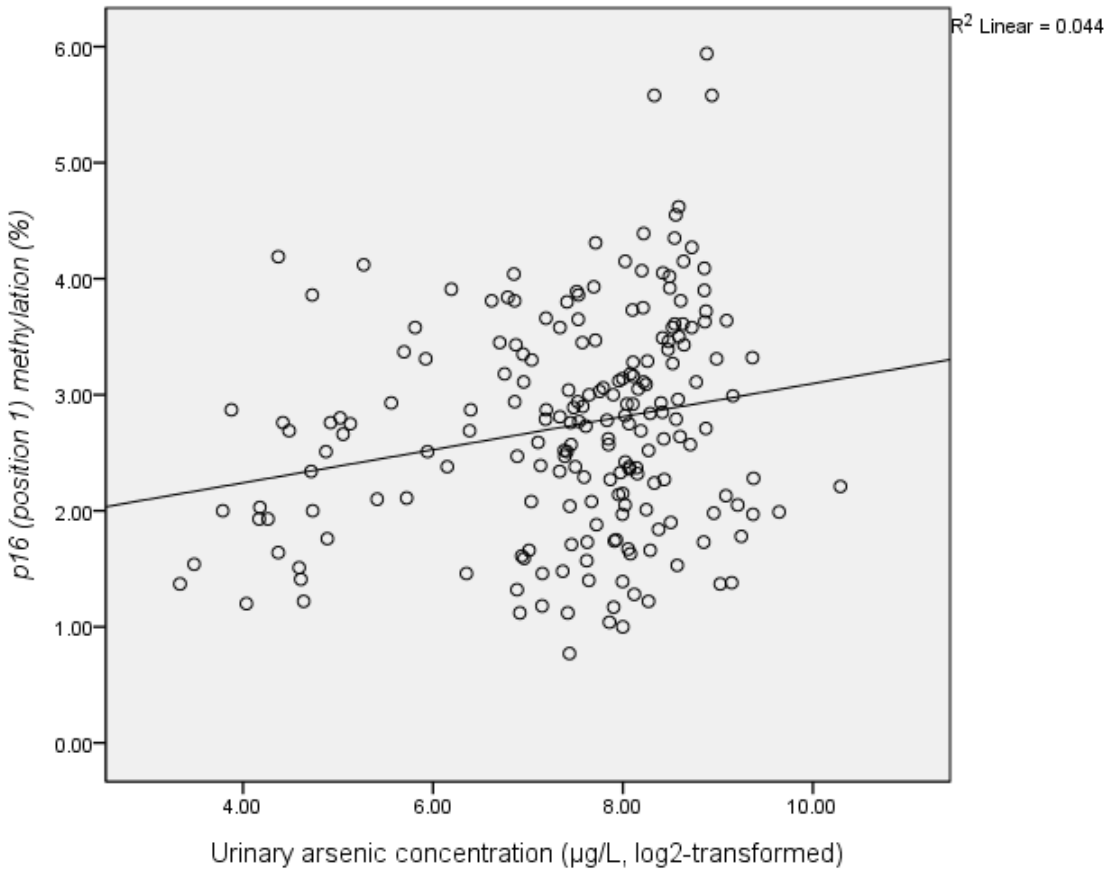
[b]Urinary arsenic concentration adjusted to specific gravity (1.015 g/ml).

Supplemental Table 2: Associations between concentrations of urinary arsenic (log_2-transformed) and fractions (%) of urinary arsenic metabolites and degree of *LINE1* methylation.

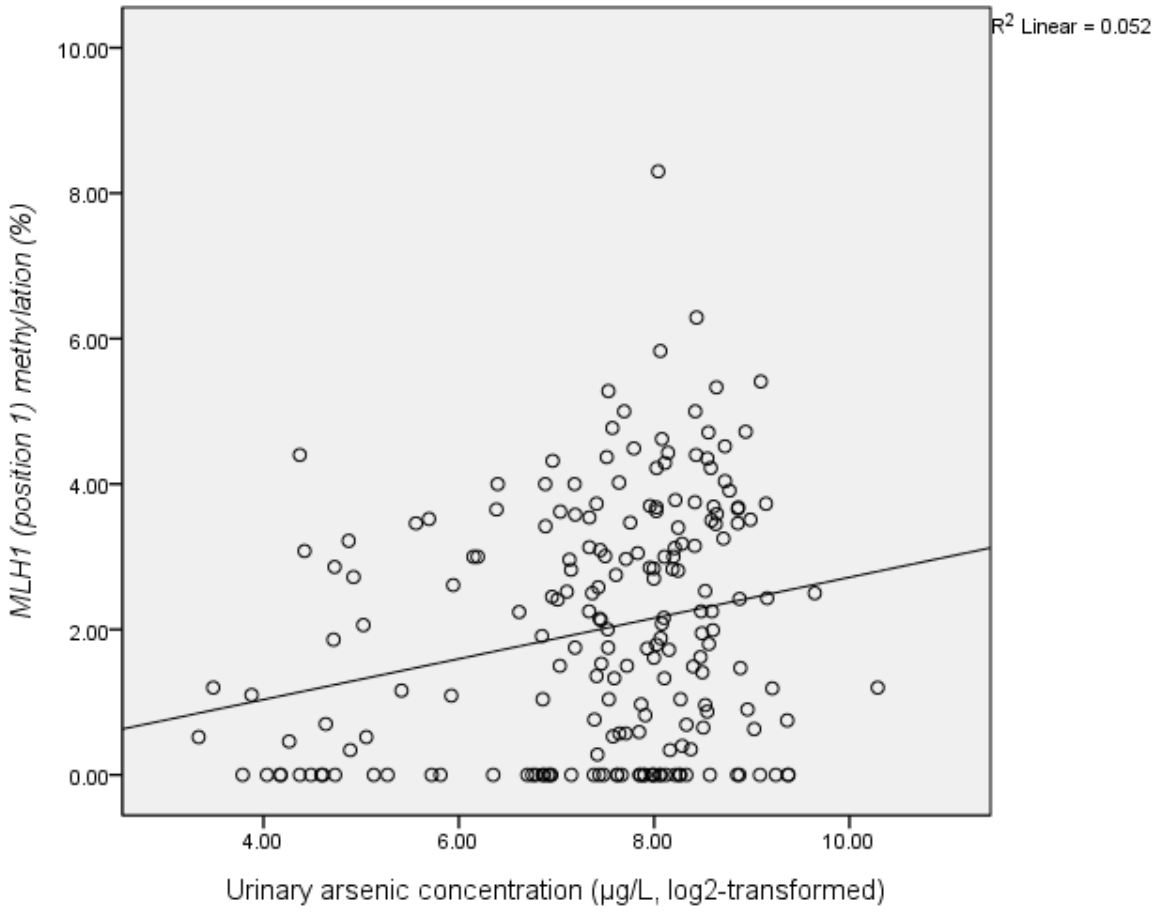
Predictor		β_1 (95% CI)	p-value
Urinary arsenic	Unadjusted[a]	0.094 (-0.082 to 0.27)	0.30
	Model 1	0.12 (-0.059 to 0.30)	0.19
	Model 2	0.14 (-0.039 to 0.33)	0.12
Inorganic arsenic (%)	Analysis 1[b]	-0.0075 (-0.046 to 0.031)	0.70
	Analysis 2	-0.016 (-0.055 to 0.022)	0.41
MMA (%)	Analysis 1	0.023 (-0.053 to 0.099)	0.56
	Analysis 2	0.0088 (-0.068 to 0.086)	0.82
DMA (%)	Analysis 1	0.0012 (-0.031 to 0.033)	0.94
	Analysis 2	0.010 (-0.023 to 0.043)	0.54

[a]Unadjusted model (N=202): Methylation (%) = $\alpha_1 + \beta_1$ arsenic concentrations (log_2-transformed); Model 1: model adjusted for age and coca chewing; Model 2: model with additional adjustments for cadmium in urine (log_2-transformed) and selenium and zinc in blood.

[b]Analysis 1 (N=202): Methylation (%) = $\alpha_1 + \beta_1$ fraction of arsenic metabolite + β_2 arsenic concentrations (log_2-transformed); Analysis 2: model adjusted for age and coca chewing.



Supplemental Figure 1: Scatter plot of the association between the degree of *p16* methylation (%) and urinary arsenic concentration (µg/L, log_2-transformed) ($\beta = 0.14$; $p = 0.0028$).



Supplemental Figure 2: Scatter plot of the association between the degree of *MLH1* methylation (%) and urinary arsenic concentration (µg/L, log_2-transformed) ($\beta = 0.28$; $p = 0.0011$).