Identification of Novel Gene Targets and Putative Regulators of Arsenic-Associated DNA Methylation in Human Urothelial Cells and Bladder Cancer

ARTICLE *in* CHEMICAL RESEARCH IN TOXICOLOGY · JUNE 2015

Impact Factor: 3.53 · DOI: 10.1021/tx500393y · Source: PubMed

Chemical Research in Toxicology

Article

pubs.acs.org/crt

Identification of Novel Gene Targets and Putative Regulators of Arsenic-Associated DNA Methylation in Human Urothelial Cells and Bladder Cancer

Julia E. Rager,[†] Sloane K. Tilley,[†] Samantha E. Tulenko,[†] Lisa Smeester,[†] Paul D. Ray,[†,‡] Andrew Yosim,[†] Jenna M. Currier,[‡] María C. Ishida,[§] Maria del Carmen González-Horta,[§] Blanca Sánchez-Ramírez,[§] Lourdes Ballinas-Casarrubias,[§] Daniela S. Gutiérrez-Torres,[§] Zuzana Drobná,[∥] Luz M. Del Razo,[⊥] Gonzalo G. García-Vargas,[#] William Y. Kim,[○] Yi-Hui Zhou,[▽,◆] Fred A. Wright,[▽,◆] Miroslav Stýblo,[‡,∥] and Rebecca C. Fry*[,†,‡]

[†]Department of Environmental Sciences and Engineering, UNC Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27516, United States

[‡]Curriculum in Toxicology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, United States

[§]Facultad de Ciencias Químicas, Universidad Autónoma de Chihuahua, Chihuahua 31125, México

[∥]Department of Nutrition, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, United States

[⊥]Departamento de Toxicología, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, México, DF 07360, México

[#]Facultad de Medicina, Universidad Juárez del Estado de Durango, Gómez Palacio, Durango 34000, México

[○]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27514, United States

[▽]Bioinformatics Research Center, [◆]Department of Statistics, North Carolina State University, Raleigh, North Carolina 27695, United States

Supporting Information

ABSTRACT: There is strong epidemiologic evidence linking chronic exposure to inorganic arsenic (iAs) to myriad adverse health effects, including cancer of the bladder. We set out to identify DNA methylation patterns associated with arsenic and its metabolites in exfoliated urothelial cells (EUCs) that originate primarily from the urinary bladder, one of the targets of arsenic-induced carcinogenesis. Genome-wide, gene-specific promoter DNA methylation levels were assessed in EUCs from 46 residents of Chihuahua, Mexico, and the relationship was examined between promoter methylation profiles and the intracellular concentrations of total arsenic and arsenic species. A set of 49 differentially methylated genes was identified with increased promoter methylation associated with EUC tAs, iAs, and/or monomethylated As (MMAs) enriched for their roles in metabolic disease and cancer. Notably, no genes had differential methylation associated with EUC dimethylated As (DMAs), suggesting that DMAs may influence DNA methylation-mediated urothelial cell responses to a lesser extent than iAs or MMAs. Further analysis showed that 22 of the 49 arsenic-associated genes (45%) are also differentially methylated in bladder cancer tissue identified using The Cancer Genome Atlas repository. Both the arsenic- and cancer-associated genes are enriched for the binding sites of common transcription factors known to play roles in carcinogenesis, demonstrating a novel potential mechanistic link between iAs exposure and bladder cancer.

Received: September 26, 2014

DOI: 10.1021/tx500393y
Chem. Res. Toxicol. XXXX, XXX, XXX−XXX

INTRODUCTION

Exposure to arsenic in drinking water is of great concern globally, with more than 100 million people exposed to levels exceeding the World Health Organization's guideline of 10 μg As/L.[1,2] Chronic exposure to arsenic has been linked to a host of detrimental health effects in humans, including impaired memory and intellectual function, heart and respiratory system disease, liver hypertrophy, and diabetes.[3] In addition to these noncancer end points, arsenic is a known carcinogen, associated with liver, lung, prostate, skin, and urinary bladder cancers.[4] Bladder cancer is of particular concern, as an estimated 571 500 individuals are currently impacted by this disease in the U.S. alone.[5]

Decades of research have uncovered several mechanisms that likely underlie arsenic-induced carcinogenesis. Specifically, oxidative stress,[6] altered DNA repair capacity,[7] and alterations in epigenetic-mediated gene regulation[8] have all been suggested as potential causal events in arsenic-induced carcinogenesis. In relationship to epigenetic modifications, the following changes have been observed in response to arsenic exposure: differential microRNA expression,[9] post-translational histone modifications,[10] and changes in both global and gene-specific DNA methylation patterns.[8,11] These arsenic-associated epigenetic shifts may alter the expression of target genes/pathways that influence disease state such as cancer.[8] Of relevance to the current study, *in vitro* evidence suggests that arsenic exposure may cause alterations in the expression levels of critical genes, such as regulators of cell proliferation and the carcinogenic process, through changes in the DNA methylation maintenance machinery in human uroepithelial cells.[12,13] Similarly, genes related to cell death and proliferation have also been identified as mutated in human urothelial bladder cancer.[14]

Responses to arsenic exposure, including interindividual differences in disease susceptibility, are known to be tightly associated with metabolism.[15] *In vitro* evidence has also clearly demonstrated that the toxicity associated with arsenic differs from its metabolites.[16] The capacity to metabolize arsenic into trivalent and pentavalent monomethylated and dimethylated arsenicals (MMAs and DMAs, respectively) differs among individuals even with the same exposure levels, influencing the relative amounts of iAs, MMAs, and DMAs excreted in urine.[15] Low ratios of urinary DMAs/MMAs, which is thought to be an indicator of low arsenic methylation capacity, have been correlated with an increased risk of some of the adverse health effects.[17] For example, higher percentages of urinary MMAs and thus lower percentages of DMAs have been associated with increased risk of lung cancer[18] and urothelial bladder cancer,[19] among others diseases.[17] A recent study in Chihuahua, Mexico, showed that arsenic species retained in exfoliated urothelial cells (EUCs) can also serve as indicators of adverse health effects associated with exposure.[20]

In the present study, we set out to investigate cytosine methylation associated with arsenic exposure in an exposed human population and to contrast these findings with biomarkers of an arsenic-associated disease (i.e., bladder cancer). The ultimate goal was the identification of key genes and biological pathways potentially involved in arsenic-induced disease. Specifically, inorganic arsenic and its metabolites were analyzed in relationship to promoter DNA methylation profiles in EUCs of subjects from a recently established cohort based in the Chihuahua, Mexico.[20] This region is an area of concern, as it is estimated that more than 450 000 people are exposed to arsenic levels exceeding 50 μg/L in Mexico.[21] The current research aimed to examine associations between DNA methylation profiles in EUCs with trivalent and pentavalent forms of total arsenic (tAs), iAs, MMAs, and DMAs in the same EUCs. Upon relating these arsenic-associated epigenetic changes to changes observed in bladder cancer tissue, a common set of differentially methylated genes was identified. Mechanistic similarities were identified between the arsenic-associated and bladder cancer-associated genes, where common transcription factors with known involvement in carcinogenesis are known to bind to the promoter regions of the genes. This finding may have functional implications for arsenic-induced human carcinogenesis, as these data represent a direct linkage between DNA methylation patterns and levels of arsenic/arsenic metabolites within the same target cells.

MATERIALS AND METHODS

Study Population and Sample Collection. All procedures involving human subjects were approved by IRBs at UNC−Chapel Hill, U.S., and Cinvestav−IPN, Mexico. Individuals participating in the present study ($n = 46$) represent a subcohort of a larger study ($n = 374$) in Chihuahua, Mexico.[20] Study participants were required to be at least 18 years of age and have at least 5 years of uninterrupted residency in the study area. All participants selected for the current investigation were Hispanic females, as the EUC counts in urine from females are significantly higher than those in males,[20] thus producing adequate DNA sample amounts for DNA methylation analysis. Pregnant women and subjects with kidney or urinary tract infections were excluded because these medical conditions could affect arsenic metabolism or purity of EUCs isolated from urine. Individuals at risk for occupational exposure to arsenic were also excluded. Household tap water samples were collected for arsenic analysis performed at Cinvestav−IPN (Mexico City). Study subjects were evaluated through medical examination at the Universidad Autónoma de Chihuahua. Body weight and height were recorded and used to calculate the body mass index (BMI). Spot urine samples were collected for arsenic analysis and isolation of EUCs. Urine and EUC samples were frozen and shipped at −70 °C for speciation analysis at UNC−Chapel Hill.

Determination of Arsenic Species in Household Water and Urine. Concentrations of arsenic in acid-digested water samples were determined in Cinvestav−IPN (Mexico City) using hydride generation atomic absorption spectrometry (HG-AAS) with cryotrapping (CT), as previously described.[22,23] Two water samples were not available from the 46 individuals in the study; thus, a total of 44 water samples was available for household water analysis. Arsenic species in urine, including inorganic arsenic (iAs), MMAs, and DMAs, were analyzed by HG-CT-AAS.[23] Limits of detections (LODs) were determined based upon urine volumes of 200 μL per sample, resulting in 0.05 ng As/mL for MMAs, 0.05 ng As/mL for DMAs, and 0.1 ng As/mL for iAs. Total arsenic (tAs) in urine was calculated as the sum of iAs, MMAs, and DMAs. Urinary creatinine concentration was determined by colorimetric assay (Cayman Chemical Company, Ann Arbor, MI), and specific gravity was measured using a digital Atago PAL refractometer (Atago USA, Bellevue, WA).

EUC Isolation. The isolation and purity of EUCs was described in detail in our previous report from the Chihuahua cohort.[20] Briefly, EUCs were isolated from urine (100 mL per subject) by centrifugation at 4 °C, washed with phosphate buffered saline (PBS), and again centrifuged. EUCs were resuspended in PBS and counted. Resuspended EUCs were checked for the presence of bacteria, yeast, and red/white blood cells using a light microscope. All EUCs used in the present study were free of microbial contamination and contained <5% of red and white blood cells. The cell counts of the collected EUC pellets ranged from 269 500 to 6 049 600 (mean = 944 991, median = 704 300).

For the purpose of this investigation, all cells isolated from bacteria- and yeast-free urine containing <5% of red and white blood cells were defined as EUCs. As described in detail in our previous report,[20] these cells originate from the epithelial lining of the urinary tract, the urothelium. Because the routine microscopy used in this study to assess the collected EUCs cannot reliably distinguish between cells of various

DOI: 10.1021/tx500393y
Chem. Res. Toxicol. XXXX, XXX, XXX−XXX

origins, it is not possible to further characterize the types and origins of EUCs. It is important to note that the majority of the EUCs originate in the bladder epithelium, as EUCs from female urine have been shown to originate mainly from the vesical trigone area of the bladder.[24,25] For this reason, EUCs are also used in clinical practice to diagnose bladder cancer.[26] Further substantiating the use of EUCs, a study specifically compared gene-specific promoter methylation patterns in urine sediment DNA versus human bladder cancer tissue and found that, for all patients evaluated, the DNA methylation patterns in urine matched those in bladder cancer tissue.[27] Isolated EUCs were stored at −80 °C and used for analyses 4 weeks after collection.

Determination of Arsenic Species in EUCs. Trivalent and pentavalent arsenic species were measured in EUC lysates using HG-CT-inductively coupled plasma-mass spectrometry (ICP-MS), as previously detailed.[20,28] Concentrations of arsenite (iAs^{III}), arsenate (iAs^{V}), methylarsonite ($MMAs^{III}$), methylarsonate ($MMAs^{V}$), dimethylarsinite ($DMAs^{III}$), and dimethylarsinate ($DMAs^{V}$) were determined and expressed in ng As/10 000 cells. EUC pellets were lysed in deionized water. The trivalent arsenic species, iAs^{III}, $MMAs^{III}$, and $DMAs^{III}$, were directly measured using an aliquot of cell lysate. A separate cell lysate aliquot was treated with 2% cysteine and analyzed for total inorganic arsenic (iAs^{III} + iAs^{V}), MMAs ($MMAs^{III}$ + $MMAs^{V}$), and DMAs ($DMAs^{III}$ + $DMAs^{V}$). Subtracting the values of cysteine-treated samples from those of the untreated samples presented the concentrations of iAs^{V}, $MMAs^{V}$, and $DMAs^{V}$. Total trivalent arsenic (tAs^{III}) was defined as the sum of iAs^{III}, $MMAs^{III}$, and $DMAs^{III}$. Total pentavalent arsenic (tAs^{V}) was defined as the sum of iAs^{V}, $MMAs^{V}$, and $DMAs^{V}$. The trivalent and pentavalent total arsenic species were summed to generate $tAs^{(III+V)}$. The instrumental LODs of the HG-CT analysis ranged from 0.04 to 2.0 pg of As. Concentrations of EUC tAs, iAs, MMAs, and DMAs were correlated with each EUC arsenic species as well as urinary tAs, iAs, MMAs, and DMAs using the Spearman rank correlation test (TIBCO Spotfire, v5.0.0).

DNA Extraction and Methylation Analysis. DNA was extracted from the exfoliated EUCs of 46 subjects using the QIAamp DNA blood mini kit (Qiagen, Valenica, CA) according to manufacturer's instructions. Enrichment of CpG-methylated DNA was performed using the MethylMiner methylated DNA enrichment kit (Invitrogen/Life Technologies, Grand Island, NY) and 250 μg of DNA. Enriched DNA was amplified and hybridized to Affymetrix Human Promoter 1.0R arrays (Affymetrix, Santa Clara, CA) as previously described.[11] The Affymetrix Human Promoter 1.0R arrays represent >25 500 human promoter regions, ∼13 000 of which contain CpG islands, known targets of DNA methylation. Data were normalized using robust multichip average[29] and bioinformatically summarized at the CpG island level based on the Human Genome 18 (HG18) assembly. Average methylation abundance levels mapped to gene promoter regions were used in this analysis, as previously defined.[11] Specifically, in order to calculate the average methylation abundance for each gene, the methylation abundance levels for all CpG sites mapped to the same promoter region were summed and then divided by the number of CpG sites for that gene. Microarray data have been submitted to National Center for Biotechnology Information (NCBI) Gene Expression Omnibus repository[30] and are available under accession number GSE58499 (www.ncbi.nlm.nih.gov/geo).

Analysis of the Association between DNA Methylation Levels and Arsenic Species Concentrations in EUCs. The associations between DNA methylation levels and arsenic metabolite concentrations in EUCs were evaluated using separate multivariable models for each arsenic species. In these statistical models, DNA methylation levels were the dependent variables and EUC arsenical levels (i.e., iAs^{III}, iAs^{V}, $iAs^{(III+V)}$, $MMAs^{III}$, $MMAs^{V}$, $MMAs^{(III+V)}$, $DMAs^{III}$, $DMAs^{V}$, $DMAs^{(III+V)}$, tAs^{III}, tAs^{V}, and $tAs^{(III+V)}$) were the independent variables, as previously described.[9] The models included covariates that are plausibly related to EUC DNA methylation profiles, specifically, age (continuous variable) and BMI (continuous variable). Differential DNA methylation was defined as a significant association between DNA methylation levels and EUC metabolite levels, where the following requirements were set: (i) p-value < 0.05, and (ii) q-value < 0.10. In addition, as a result of publications that suggest that likelihood of functional change at the mRNA level is linked to methylation abundance,[8] a 2-fold change in methylation abundance was required. Fold changes were calculated using the following metric: (average DNA methylation levels of the highest exposed quartile (n = 11))/(average DNA methylation levels of the lowest exposed quartile (n = 11)). All genes that passed the statistical filters above were identified as differentially methylated genes associated with EUC arsenic levels. These statistical calculations were carried out using Partek Genomics Suite software (St. Louis, MO).

An additional robust regression analysis of tAs vs methylation levels was performed using Winsorized regression, as implemented using *lmWinsor* in the R *fda* package (version 3.1.2), with the trim = 0.1 option. Two analyses were carried out: one using age and BMI as covariates and the other as a direct regression of residuals for tAs vs residuals for methylation, after correction for the covariates. p-values for robust regression can be more significant than for standard linear regression if there are skewed observations, which can inflate the standard errors for regression coefficients.

Analysis of the Association between DNA Methylation Levels and Arsenic Species Concentrations in Urine. To identify genes with differential methylation associated with arsenic in urine, similar statistical analyses were carried out as with arsenic in EUCs. Similar multivariable models were used as previously described, in which the associations between DNA methylation levels and arsenic species in urine were evaluated using age and BMI as covariates. Separate models were used to assess each of the associations between DNA methylation levels and the summed trivalent and pentavalent arsenical levels, specifically, urinary $iAs^{(III+V)}$, $MMAs^{(III+V)}$, $DMAs^{(III+V)}$, and $tAs^{(III+V)}$. The same statistical filters were used in these analyses as the analyses with EUC arsenicals, using a multiple test correction q-value filter requirement of 0.002.

Network and Disease/Functional Enrichment Analysis. Network analysis was performed to identify enriched biological pathways among the genes with differential methylation associated with one or more EUC arsenicals.[9] Networks were algorithmically constructed based on connectivity, enabled through Ingenuity Pathway Analysis (Ingenuity Systems, Redwood City, CA). Disease signatures and biological functions within the constructed networks were identified using the right-tailed Fisher's Exact test, as detailed previously.[9,31] Over-represented diseases/functions were defined as those that contain more targets than expected by chance using a p-value cutoff of 0.01.

Comparison of EUC Arsenic-Associated Genes to the Comparative Toxicogenomics Database. In order to evaluate whether there is evidence that the genes with differential methylation associated with EUC arsenic have altered gene and/or protein expression associated with arsenic, the CTD was queried. This is a manually curated database specific for environmental contaminants and their relationships to genes, including alterations at the mRNA and protein level, collected from published toxicological and epidemiological studies. At the time of the current analysis, the database included over 95 000 studies to derive over 15 million toxicogenomic relationships between approximately 11 000 chemicals and 27 000 genes.[32]

Identifying Bladder Cancer-Associated Genes. In order to identify genes with differential methylation in human bladder cancer, a separate data set was analyzed from The Cancer Genome Atlas (http://cancergenome.nih.gov/), the product of a collaborative effort between the National Cancer Institute and the National Human Genome Research Institute to provide comprehensive data sets for use in cancer research. Publically available data for bladder urothelial carcinoma were analyzed.

DNA methylation data were obtained for 18 individuals (the maximum number available) with both tumor and matched nontumor tissue from the bladder epithelium. Data were normalized on probes using quantile normalization and filtered for probes with missing data or probes that represent known single nucleotide polymorphisms, as detailed previously,[33] after which 335 570 probes remained for analysis. The final DNA methylation files included β-values for 20 256 genes, generated using the Infinium HumanMethylation450 BeadChip (Illumina, San Diego, CA). The β-value is defined as the ratio of methylated probe intensity to the intensity from both methylated and

DOI: 10.1021/tx500393y
Chem. Res. Toxicol. XXXX, XXX, XXX−XXX

unmethylated probes, where a β-value of 1 indicates that every copy of that CpG site was methylated.[34] Similar to the genome-wide analysis previously detailed, the association between DNA methylation patterns and bladder cancer disease status was evaluated using a multivariable model. DNA methylation levels were the dependent variables and were contrasted against tumor or nontumor status as the dichotomous independent variable. Covariates included in the model were age at initial diagnosis, diagnosis subtype (papillary or nonpapillary), gender, and smoking status (binary variable). A significant relationship between tumor status and DNA methylation levels was defined for each gene by the following four requirements, which varied slightly from the arsenic-associated gene requirements due to differences in the DNA methylation array platforms: (i) p-value < 0.05, (ii) q-value < 0.05, (iii) average β difference (average β value in tumors−average β value in nontumors) ≥ 0.1 or ≤ -0.1, and (iv) a concordant β difference ≥ 0.05 or ≤ -0.05 for ≥50% of the patients. β difference was calculated as (β value of the tumor specimen − β value of the nontumor specimen) for each of the patients.

For statistical testing, a permutation test was performed comparing the n = 11 837 genes evaluated for differential methylation using the Affymetrix Human Promoter assay and the Illumina Methylation BeadChip platform. A comparison between the number of total genes tested on each platform versus those that were identified as associated with EUC arsenic and bladder cancer were compared using randomly generated lists of the same size using a Chi-squared test.

Transcription Factor Binding Enrichment Analysis. Analysis for enrichment of transcription factor binding sites within the EUC arsenic-associated genes and the bladder cancer-associated genes was carried out using Genomatix (Genomatix Software Inc., Ann Arbor, MI) database version ElDorado 12-2013. Promoter regions were designated as 500 bp upstream and 100 bp downstream of each gene's transcriptional start site. All differentially methylated genes and their corresponding promoter sequences were analyzed with both a minimum core and a matrix similarity of 0.95, the second highest level of sensitivity possible. Analysis of transcription factor binding enrichment was restricted to transcription factor families associated with vertebrates or general core promoter elements.

Gene-Specific Promoter Methylation Validation. In order to evaluate CpG methylation using an alternate approach, EUC DNA samples were analyzed using a CpG-methylation enrichment followed by gene-specific quantitative real-time PCR (qPCR). DNA samples from five subjects within the study cohort were selected with varying levels of EUC tAs. The CpG methylation levels for two genes, specifically, proteasome (prosome, macropain) subunit, beta type, 2 (*PSMB2*) and solute carrier family 12 (potassium/chloride transporter), member 7 (*SLC12A7*), were evaluated. Enrichment of CpG-methylated DNA was analyzed using the MethylMiner methylated DNA enrichment kit (Invitrogen/Life Technologies, Grand Island, NY). For each sample, 200 ng of extracted DNA was subjected to methylated DNA immunoprecipitation. Input and immunoprecipitated methylated DNA were assessed with qPCR with iQ SYBR Green supermix (Bio-Rad) using primer pairs designed within the promoter regions for *PSMB2* (F: CCCCGACAAACTCCTTCTG; R: ACAGAGAAACTGGGCGTCAC) and *SLC12A7* (F: TCCTGGGACTTGTTGAGGAC; R: GCGGTTTGTGTTGTGTGACT) plated in technical triplicate. qPCR cycle threshold (C_T) results were generated for the following fractions: (1) bound DNA, or gene regions that were bound to the beads with methyl-binding protein, representing methylated DNA, and (2) supernatant and wash DNA, or the fraction of DNA that was unbound, representing unmethylated DNA. In order to calculate a relative measure of methylation levels, the following equation was used: methylation level = 100 − (bound DNA C_T − (supernatant DNA C_T + wash DNA C_T)). These methylation levels were compared to EUC $tAs^{(III+V)}$ levels using Spearman rank correlation tests.

RESULTS

Characteristics of Study Population and Arsenic Concentrations in EUCs and Urine. Drinking water samples, spot urine, and EUC samples from 46 Hispanic women of the larger Chihuahua cohort[20] were analyzed within this subcohort. The average age of these women was 45, and the average BMI was 30.2. Arsenic in drinking water samples ranged from <LOD to 275.4 ng As/mL, with an average of 64.1 ng As/mL and a median of 50.7 ng As/mL. Of particular note, 70% (31/44) of the drinking water samples collected were above the WHO's recommendation of 10 μg As/L (Table 1).

Because arsenic metabolism plays a large role in arsenic-associated disease etiology,[17] the concentrations of various arsenic species in EUCs were determined, namely, arsenite (iAs^{III}), arsenate (iAs^{V}), methylarsonite ($MMAs^{III}$), methylarsonate ($MMAs^{V}$), dimethylarsinite ($DMAs^{III}$), and dimethylarsinate ($DMAs^{V}$) were determined and expressed in ng As/10 000 cells. Total trivalent arsenic (tAs^{III}) was defined as the sum of iAs^{III}, $MMAs^{III}$, and $DMAs^{III}$. Total pentavalent arsenic (tAs^{V}) was defined as the sum of iAs^{V}, $MMAs^{V}$, and $DMAs^{V}$. The trivalent and pentavalent total arsenic species were summed to generate $tAs^{(III+V)}$. The arsenic species were measured in EUCs for each of the 46 study subjects (Table 2). Notably, the EUC concentration of $tAs^{(III+V)}$, an indicator of arsenic exposure, ranged from 0.9 to 46.1 pg As/10 000 cells (mean = 14.3 pg As/10 000 cells). iAs^{III}, iAs^{V}, and $MMAs^{III}$ were the major As species in EUCs; $MMAs^{V}$, $DMAs^{V}$, and $DMAs^{III}$ were only minor metabolites. The average ratios of EUC arsenicals were as follows: MMAs/iAs = 0.21, DMAs/MMAs = 1.64, and DMAs/iAs = 0.37. In contrast, DMAs was the major metabolite found in urine, followed by MMAs and iAs. The urine metabolite ratios were as follows: MMAs/iAs = 6.2, DMAs/MMAs = 6.3, and DMAs/iAs = 34.1. Despite the large differences in arsenical ratios, EUC arsenical concentrations were significantly ($p < 0.05$) correlated with urine arsenical concentrations for each respective species when using either unadjusted, creatinine-adjusted, or specific gravity-adjusted urine arsenical measures to allow for cross-study comparisons (Table S1, Supporting Information). Interestingly, the most significant correlations were found between EUC arsenic concentrations and the unadjusted urine arsenic concentrations for all arsenic species evaluated: tAs, iAs, MMAs, and DMAs. Also of note, the creatinine-adjusted urinary arsenical concentrations showed the least significant correlation to arsenical concentrations in EUCs. Among the EUC arsenic species, there were significant ($p < 0.05$) correlations with only one exception, namely, the comparison between $MMAs^{III}$ and $DMAs^{V}$ ($p = 0.05$).

Gene-Specific DNA Methylation Levels Are Associated with Concentrations of Arsenic Species in EUCs. The relationships between EUC arsenic species and promoter DNA methylation levels were assessed across >14 000 genes. A total of 49 differentially methylated genes were identified, all of which were hypermethylated in relationship to one or more arsenic species. A total of 47 genes had promoter methylation levels associated with arsenic exposure, as characterized by EUC tAs levels, in which 14 genes had promoter methylation levels associated with EUC tAs^{V} and 43 with $tAs^{(III+V)}$. In addition, 3 genes had promoter methylation levels associated with EUC iAs^{V}, 8 with $iAs^{(III+V)}$, 4 with $MMAs^{III}$, 0 with $MMAs^{V}$, and 11 with $MMAs^{(III+V)}$ (Tables 3 and 4 and Table S2 and S3, Supporting Information). There were no genes with methylation levels associated with EUC tAs^{III}, iAs^{III}, $DMAs^{III}$, $DMAs^{V}$, or $DMAs^{(III+V)}$.

There was considerable overlap among the genes with differential methylation associated with EUC arsenic in relationship to the arsenic species (Figure 1 and Table S2, Supporting Information). Specifically, 16% (7 out of 43) of the genes

DOI: 10.1021/tx500393y
Chem. Res. Toxicol. XXXX, XXX, XXX−XXX

Table 1. Demographic Information and Arsenic Levels in Drinking Water, Urine, and EUCs for the Study Cohort[a]

	mean, median (range)
age (years)	45, 45 (20−71)
BMI (kg/m^2)	30.2, 29.1 (19.4−49.1)
arsenic in drinking water (ng As/mL)[b]	64.1, 50.7 ((<LOD[c])−275.4)
urine creatinine (mg/dL)	114.3, 110.5 (8.2−397.4)
urine specific gravity	1.01, 1.01 (1.00−1.04)
Urine Arsenicals (ng/mL)	
urine iAs	6.4, 3.8 ((<LOD)−55.3)
urine MMAs	9.4, 5.2 (0.3−58.2)
urine DMAs	49.0, 41.6 (1.7, 260.3)
urine tAs	64.7, 54.2 (2.2−373.9)
Creatinine-Adjusted Urine Arsenicals (ng As/mg Creatinine)	
iAs	5.6, 3.8 ((<LOD)−45.6)
MMAs	8.4, 6.1 (0.4−48.0)
DMAs	44.6, 36.2 (2.7−214.6)
tAs	58.6, 48.6 (3.4−308.2)
Specific Gravity-Adjusted Urinary Arsenicals (ng As/SG)	
iAs	8.0, 6.7 ((<LOD)−39.1)
MMAs	12.2, 8.9 (0.8−50.0)
DMAs	64.7, 70.1 (3.8−223.1)
tAs	84.9, 93.1 (7.3−278.9)
Urine Percent Arsenicals (%)	
iAs	8.7, 7.6 (0.1−22.5)
MMAs	14.5, 14.0 (5.0−29.3)
DMAs	76.9, 79.6 (52.5, 90.8)
Urine Arsenical Ratios	
MMAs/iAs	6.2, 1.6 (0.5, 111.4)
DMAs/MMAs	6.3, 5.9 (1.8−17.0)
DMAs/iAs	34.1, 10.6 (2.6−690.0)
EUC Percent Arsenicals (%)	
iAs	69.8, 74.6 (21.0−89.5)
MMAs	13.6, 12.6 (6.8−28.7)
DMAs	16.6, 9.8 (2.6−71.5)
EUC Arsenical Ratios	
MMAs/iAs	0.21, 0.18 (0.08−0.44)
DMAs/MMAs	1.64, 0.66 (0.17−9.57)
DMAs/iAs	0.37, 0.13 (0.03−3.41)
Lowest Exposed Quartile	
EUC tAs$^{(III+V)}$ (pg As/10 000 cells)	2.14, 2.21 (0.86−3.20)
age (years)	50, 50 (27−71)
BMI (kg/m^2)	29.3, 29.0 (19.4−48.6)
Highest Exposed Quartile	
EUC tAs$^{(III+V)}$ (pg As/10 000 cells)	33.84, 34.44 (22.38−46.13)
age (years)	42, 46 (20−71)
BMI (kg/m^2)	30.8, 26.0 (20.9−42.3)

[a]Population characteristics and EUC arsenic levels are also detailed for the 11 subjects with the lowest EUC tAs$^{(III+V)}$ levels and the 11 subjects with the highest EUC tAs$^{(III+V)}$ levels, representing the lowest and highest exposed quartiles when considering EUC tAs levels, respectively. Adjustments by creatinine and specific gravity were performed to allow for cross-study comparisons. [b]Arsenic levels in drinking water were determined using n = 44 samples. [c]LOD refers to limit of detection.

Table 2. Arsenic Species Concentrations in Human EUCs[a]

	III	V	III + V
EUC arsenic species	mean, median (range in pg As/10 000 cells)		
tAs	10.9, 8.6 (0.4−58.6)	3.7, 2.4 (0.4−15.4)	14.3, 11.2 (0.9−46.1)
iAs	8.9, 6.7 (0.2−53.6)	2.3, 1.1 ((<LOD)−9.6)	10.9, 8.0 (0.5−38.9)
MMAs	1.7, 1.3 (0.1−8.3)	0.4, 0.2 ((<LOD)−2.1)	2.1, 1.4 (0.1−10.3)
DMAs	0.3, 0.2 ((<LOD)−2.5)	1.0, 0.7 (0.1−4.6)	1.4, 1.0 (0.2−5.3)

[a]Means and medians were calculated across all subjects (n = 46).

Table 3. Number of Genes with Differential Methylation Associated with Arsenic Species in EUCs

arsenic species	III	V	III + V
tAs	0	14	43
iAs	0	3	8
MMAs	4	0	11
DMAs	0	0	0

associated with tAs$^{(III+V)}$ were also associated with iAs$^{(III+V)}$, and 21% (9 out of 43) that were associated with tAs$^{(III+V)}$ were also associated with MMAs$^{(III+V)}$. Eleven genes were associated with two or more arsenical groups with summed trivalent and pentavalent species: tAs$^{(III+V)}$, iAs$^{(III+V)}$, and MMAs$^{(III+V)}$. These 11 genes included alanyl (membrane) aminopeptidase (*ANPEP*), laminin, alpha 2 (*LAMA2*), methyltransferase like 13 (*METTL13*), RGPD4 antisense RNA 1 (*RGPD4-AS1*, also known as *AK097754*), *SLC12A7*, and ubiquitin specific peptidase 7 (*USP7*), among others (Table 4). These 11 genes had fold changes in methylation greater than 2 or, when calculated as percent change in methylation, represent estimated changes ranging from 5 to 35% (average = 16%) associated with tAs$^{(III+V)}$ (Table S3, Supporting Information). To note, although all genes associated with EUC arsenic were hypermethylated, a less stringent statistical filter (p < 0.10) would identify both hypo- and hypermethylated genes as associated with EUC arsenic (data not shown). Because the largest number of genes had methylation levels associated with EUC tAs$^{(III+V)}$, descriptive statistics were reported for this arsenic exposure as well as the covariates used in the regression model for each quartile of arsenic exposure (Table 1).

To investigate robustness in our findings of association of tAs with methylation levels, we also performed Winsorized regression using age and BMI as covariates, with 10% trimming

DOI: 10.1021/tx500393y
Chem. Res. Toxicol. XXXX, XXX, XXX−XXX

Table 4. Genes (n = 49) with Differential Methylation Associated with at Least One EUC Arsenic Species and Their Relationship to Metabolic Disease and Cancer[a]

gene symbol(s)	gene name	associated EUC arsenical(s)	metabolic disease signaling (p = 0.001)	cancer signaling (p = 0.002)	gene promoter hyper-/hypomethylation in bladder cancer
AK095619/DLX6-AS1	DLX6 antisense RNA 1	tAsV, tAsIII+V			
AK097754/RGPD4-AS1	RGPD4 antisense RNA 1 (head to head)	tAsIII+V, iAsIII+V, MMAsIII, MMAsIII+V			
ANKFY1	Ankyrin repeat and FYVE domain containing 1	tAsIII+V		×	hypermethylation
ANPEP	Alanyl (membrane) aminopeptidase	tAsIII+V, iAsIII+V, MMAsIII, MMAsIII+V		×	hypomethylation
ASAP1	ArfGAP with SH3 domain, ankyrin repeat and PH domain 1	tAsV, tAsIII+V		×	hypermethylation
ASTN2	Astrotactin 2	tAsIII+V		×	
ATP5D	ATP synthase, H+ transporting, mitochondrial F1 complex, delta subunit	tAsIII+V			hypomethylation
AX748071/RIIAD1	Regulatory subunit of type II PKA R-subunit (RIIa) domain containing 1	tAsIII+V			
BC041459/TARID	TCF21 antisense RNA inducing promoter demethylation	tAsIII+V			
C10orf139/LINC00200	Long intergenic nonprotein coding RNA 200	tAsIII+V			
C5orf39/ANXA2R	Annexin A2 receptor	tAsV, tAsIII+V			hypermethylation
CALHM1	Calcium homeostasis modulator 1	tAsIII+V		×	hypomethylation
COL12A1	Collagen, type XII, alpha 1	tAsIII+V		×	
CR605298/FENDRR	FOXF1 adjacent noncoding developmental regulatory RNA	tAsIII+V			
EOMES	Eomesodermin	tAsIII+V, MMAsIII+V		×	hypermethylation
FBXW8	F-box and WD repeat domain containing 8	tAsIII+V			hypermethylation
GGA2	Golgi-associated, gamma adaptin ear containing, ARF binding protein 2	tAsV		×	
GPNMB	Glycoprotein (transmembrane) nmb	tAsV, tAsIII+V	×	×	
HIST1H3G	Histone cluster 1, H3g	tAsIII+V, MMAsIII+V			
IGFBP5	Insulin-like growth factor binding protein 5	tAsV, tAsIII+V	×	×	hypomethylation
IGSF21	Immunoglobin superfamily, member 21	tAsIII+V			hypomethylation
IQWD1/DCAF6	DDB1 and CUL4 associated factor 6	tAsV			
KCNT2	Potassium channel, subfamily T, member 2	tAsIII+V		×	
LAMA2	Laminin, alpha 2	tAsIII+V, iAsIII+V, MMAsIII+V		×	hypomethylation
LRP5L	Low density lipoprotein receptor-related protein 5-like	tAsIII+V			
LRRC40	Leucine rich repeat containing 40	MMAsIII+V		×	
MAFF	V-maf avian musculoaponeurotic fibrosarcoma oncogene homologue F	tAsIII+V			
MAPK15	Mitogen-activated protein kinase 15	tAsV, tAsIII+V, iAsV			hypomethylation
MEIS2	Meis homeobox 2	tAsV, tAsIII+V, iAsIII+V		×	
METTL13	Methyltransferase like 13	tAsV, tAsIII+V, iAsV, iAsIII+V, MMAsIII, MMAsIII+V			
NCAN	Neurocan	tAsIII+V		×	hypomethylation
NRSN1	Neurensin 1	tAsIII+V	×		hypomethylation
OR4Q3	Olfactory receptor, family 4, subfamily Q, member 3	tAsIII+V		×	
OR8S1	Olfactory receptor, family 8, subfamily S, member 1	tAsV			hypomethylation
PCDHGC5	Protocadherin gamma subfamily C, 5	tAsV, tAsIII+V, iAsV, iAsIII+V, MMAsIII, MMAsIII+V		×	
PRDM2	PR domain containing 2, with ZNF domain	tAsIII+V		×	hypomethylation
PRIC285/HELZ2	Helicase with zinc finger 2, transcriptional coactivator	tAsIII+V			
PSMB2	Proteasome (prosome, macropain) subunit, beta type, 2	tAsIII+V	×	×	hypomethylation
RALGPS1	Ral GEF with PH domain and SH3 binding motif 1	tAsIII+V		×	hypomethylation

DOI: 10.1021/tx500393y
Chem. Res. Toxicol. XXXX, XXX, XXX−XXX

Table 4. continued

gene symbol(s)	gene name	associated EUC arsenical(s)	metabolic disease signaling (p = 0.001)	cancer signaling (p = 0.002)	gene promoter hyper-/hypomethylation in bladder cancer
SLC12A7	Solute carrier family 12 (potassium/chloride transporter), member 7	tAsV, tAsIII+V, MMAsIII+V	×	×	hypermethylation
TAOK2	TAO kinase 2	tAsIII+V			
TMSB15A	Thymosin beta 15a	tAsIII+V, iAsIII+V			
TRAF3IP1	TNF receptor-associated factor 3 interacting protein 1	tAsIII+V		×	
UBE3C	Ubiquitin protein ligase E3C	tAsIII+V		×	
USP7	Ubiquitin specific peptidase 7	tAsIII+V, MMAsIII+V	×	×	hypomethylation
VWA3A	Von Willebrand factor A domain containing 3A	MMAsIII+V			hypomethylation
YJEFN3	YjeF N-terminal domain containing 3	tAsIII+V			hypermethylation
ZBTB39	Zinc finger and BTB domain containing 39	tAsIII+V		×	
ZNF14	Zinc finger protein 14	tAsV, iAsIII+V		×	

[a]× indicates a relationship of a gene to metabolic disease and/or cancer signaling. Genes are also specified as being hyper- or hypomethylated in bladder cancer tissue.

of the observations. The results (Figure S1, Supporting Information) showed even stronger association than the original regression-based findings, indicating that the findings were not driven by a few influential observations.

Gene-Specific DNA Methylation Levels Are Associated with Concentrations of Arsenic Species in Urine. In order to compare differentially methylated genes identified using EUC arsenic versus urinary arsenic, the relationships between urinary arsenic species and promoter DNA methylation levels were assessed across the 49 genes with differential methylation associated with EUC arsenic. Of these 49 genes, 13 were associated with one or more urinary arsenic species. Specifically, three genes showed promoter methylation levels associated with urinary $tAs^{(III+V)}$, and 13 genes showed methylation levels associated with urinary $DMAs^{(III+V)}$ (Table S2, Supporting Information). The three genes associated with urinary $tAs^{(III+V)}$, namely, *METTL13*, *SLC12A7*, and *RGPD4-AS1* (also known as *AK095619*), were also associated with urinary $DMAs^{(III+V)}$. Similar to the analysis using the EUC arsenic concentrations, all of the differentially methylated genes were hypermethylated in association with urinary arsenic.

Evaluating Gene and Protein Expression Changes Using the Comparative Toxicogenomics Database. All cells from the 100 mL urine aliquots were used for DNA isolation. Because of relatively low EUC cell counts, no cells were left for RNA or protein isolation. As a result, it was not possible to directly test whether changes in gene promoter methylation were associated with functional changes in transcription or protein expression. As an alternate approach, the Comparative Toxicogenomics Database, a rich toxicological database containing thousands of published arsenic-associated studies, was queried for known relationships between the 49 genes with differential methylation associated with EUC arsenic and perturbations at the mRNA and protein expression levels. The database contained information that 12 of the 49 (25%) genes have been shown to be modulated at the mRNA/protein level by arsenic or arsenic metabolites through *in vitro* and/or *in vivo* studies (Table S4, Supporting Information).

Metabolic Disease and Cancer Signaling Pathways are Enriched Among the Arsenic-Associated Genes. An enrichment analysis identified disease signatures/biological functions to be associated ($p < 0.01$) with the 49 genes with differential methylation related to one or more EUC arsenic species. The two most significantly enriched disease signatures were metabolic disease ($p = 0.001$) and cancer ($p = 0.002$) (Table 4). A highly significant ($p < 10^{-9}$) network was constructed using the proteins encoded by the 49 genes associated with EUC arsenic. This network contains 25 proteins encoded by hypermethylated genes related to cancer and eight proteins encoded by hypermethylated genes associated with metabolic disease (Figure S2, Supporting Information).

Overlap between EUC Genes Differentially Methylated in Response to Arsenic Exposure and Bladder Cancer Genes. In order to generate a DNA methylation signature related to a disease outcome pertinent to arsenic exposure and to EUCs, a separate database was used to identify genes differentially methylated in bladder cancer. Specifically, the Cancer Genome Atlas repository was used to identify 7042 genes with differential methylation in human bladder cancer (Table S5, Supporting Information).

Of the 49 genes with differential methylation associated with EUC arsenic, 22 (45%) were also differentially methylated in bladder cancer (Table 4 and Table S5, Supporting Information). Permutation-based analysis demonstrates that this proportion of overlap is higher than would be expected by chance ($p < 0.05$). Of the 22 common EUC arsenic- and bladder cancer-associated genes, 7 were hypermethylated and 15 were hypomethylated in bladder cancer tissue versus noncancerous tissue.

Common Transcription Factor Binding Sites Identified in the EUC Arsenic- and Bladder Cancer-Associated Genes. Transcription factor binding site enrichment analysis of the 49 genes with differential methylation associated with EUC arsenic revealed 27 transcription factor families with an enrichment ($p < 0.05$). A matched gene set was queried from the genes with differential methylation associated with bladder cancer where the binding site enrichment analysis identified 23 transcription factor families ($p < 0.05$).

Interestingly, both the EUC arsenic- and bladder cancer-associated genes shared 23 transcription factor families in common, suggesting that common transcription factors may play a role in response to arsenic exposure as well as in bladder tumors. These common transcription factor families include Sine oculis homeobox homologue 3 (SIX3), RNA polymerase II transcription factor II B (TF2B), TCF11 transcription factor

DOI: 10.1021/tx500393y
Chem. Res. Toxicol. XXXX, XXX, XXX−XXX

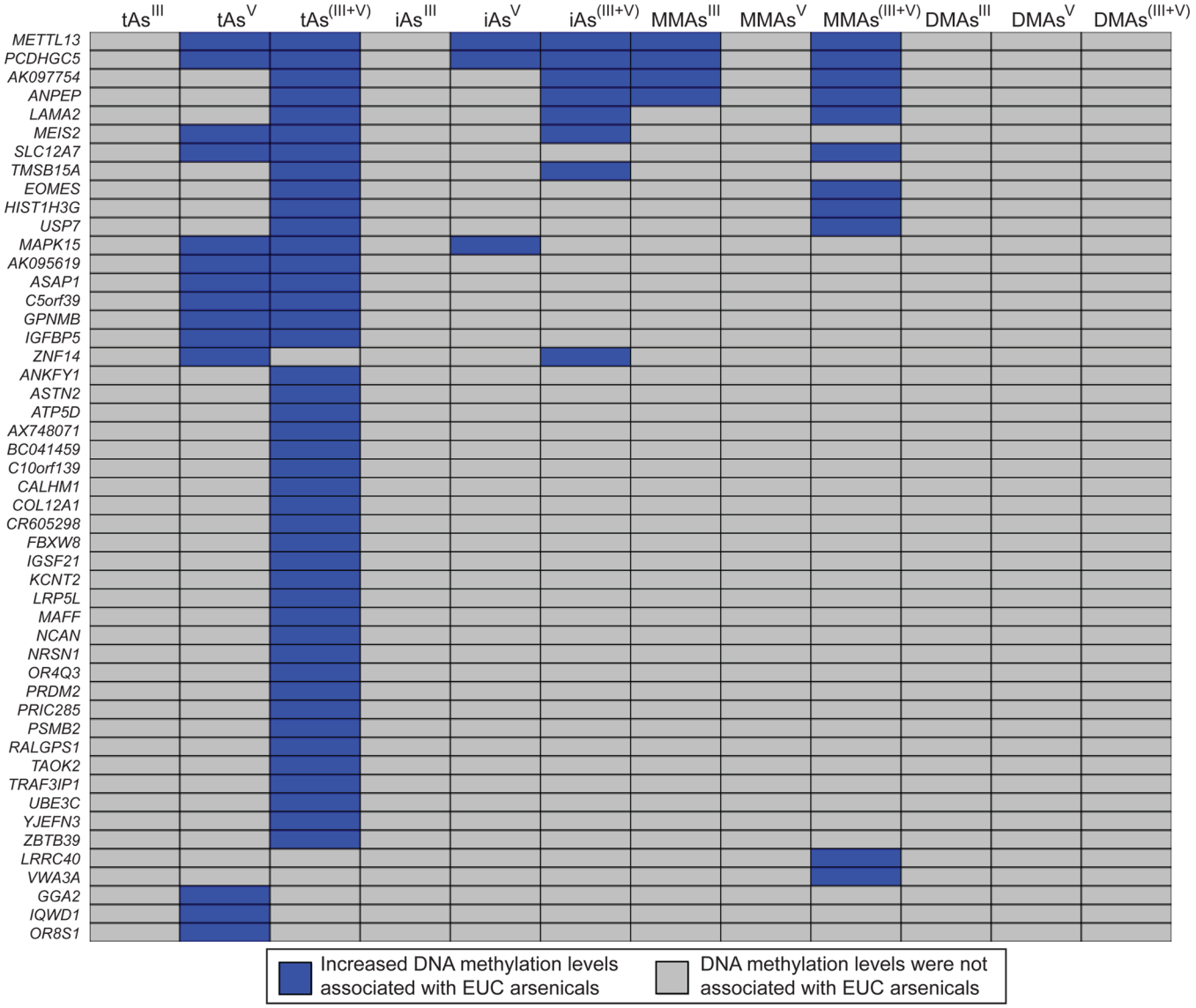

Figure 1. Heat map illustrating the differentially methylated genes associated with tAs, iAs, MMAs, and/or DMAs concentrations in EUCs. Genes are sorted according to those associated with the highest number of EUC arsenical groups (top) to the lowest number of arsenical groups (bottom).

(TCFF), and Two-handed zinc finger homeodomin transcription factors (ZFHX) (Table S6, Supporting Information).

PCR Promoter Methylation Results. The CpG methylation levels for two genes of interest, *PSMB2* and *SLC12A7*, were tested with qPCR using DNA from five subjects within the study cohort. The subjects spanned the range of exposure to arsenic. These two genes were selected for further evaluation because they were both identified as genes with differential methylation associated with EUC arsenic and bladder cancer. Similar to the genome-wide findings, the promoter methylation levels of both genes increased with increasing levels of EUC $tAs^{(III+V)}$ ($R = 0.90$ for *PSMB2* and $R = 0.60$ for *SLC12A7*), where *PSMB2* was statistically significant ($p = 0.04$) (Figure S3).

DISCUSSION

This study is the first to characterize gene-specific DNA methylation levels associated with arsenic exposure in urothelial cells primarily originating in the human urinary bladder epithelium, one of the target tissues of arsenic-induced carcinogenesis. We examined the relationship between arsenic exposure, arsenic metabolism and DNA methylation profiles in EUCs in individuals selected from a recently established cohort in Chihuahua, Mexico. A total of 49 differentially methylated genes were identified, all of which were hypermethylated in relationship to arsenic or with arsenic metabolites, iAs and MMAs, retained in EUCs. Interestingly, differences in DNA methylation were most apparent when analyzed in the context of the sum of the trivalent and pentavalent arsenic species (in contrast to the individual metabolites), with the greatest number of genes associated with $tAs^{(III+V)}$. These findings provide novel information regarding arsenic exposure and metabolism and its relationship to genome-wide epigenetic modifications in urothelial cells.

Although arsenic exposure has previously been associated with both hypo- and hypermethylation in leukocyte-derived DNA;[33] here, all 49 EUC arsenic-associated genes showed CpG promoter region hypermethylation. Interestingly, if a less stringent statistical criteria is used, then genes with both hyper- and hypomethylation are identified. Nevertheless, the observation that the most significant genes show increased methylation associated

DOI: 10.1021/tx500393y
Chem. Res. Toxicol. XXXX, XXX, XXX−XXX

with arsenic exposure is consistent with prior results in which skin lesion status in an arsenic-exposed human population was associated with a general trend of gene-specific CpG hypermethylation in peripheral blood leukocytes.[11] Gene-specific CpG hypermethylation is also prevalent in DNA from tumors of subjects with diseases relevant to arsenic exposure, including cancer.[35] Further supporting this relationship to disease, a set of the bladder cancer-associated genes was also hypermethylated. This trend for gene-specific differential methylation associated with both arsenic exposure and bladder cancer may represent a mechanistic link between exposure and disease.

In the present study, there were no differentially methylated genes associated with EUC DMAs concentrations. This finding is consistent with a previous study where urinary DMA levels influenced DNA methylation patterns in peripheral leukocytes to a lesser extent than it did the urinary levels of iAs and MMAs.[36] Furthermore, a previous epidemiological study of an arsenic-exposed population in Taiwan found significantly higher percentages of MMAs and lower percentages of DMAs in the urine of patients with urothelial carcinoma in comparison to healthy residents.[37] Together, results of this study and the previously published data suggest that accumulation of iAs and/or MMAs in urine and target tissues, which may indicate an inefficient arsenic methylation, may play a significant role in both the differential DNA methylation and health risks associated with arsenic exposure.

It is important to note that the proportions and ratios of arsenic species in EUCs differed from those present in urine, a trend that was also observed in the larger cohort.[20] Specifically, $iAs^{(III+V)}$ were the most abundant species in EUCs, whereas $DMAs^{(III+V)}$ were the most abundant urinary metabolite in both the larger cohort[20] and the subcohort analyzed in the present study. Another difference between the EUC and urinary arsenic results were the number of differentially methylated genes identified. Specifically, more genes were associated with urinary DMAs than urinary tAs, whereas the opposite trend was observed in the EUC arsenic findings. Thus, the DNA methylation associations are influenced by the selection of intracellular or urinary measures of arsenic as biomarkers of exposure. These findings further suggest that the identified genes with differential methylation likely depend upon the presence and proportion of arsenic metabolites in EUCs. Regardless of these differences, it is important to note that the EUC arsenical concentrations were significantly correlated with the urinary arsenical concentrations. These findings provide important information for future investigations into DNA methylation patterns in relationship to biomarkers of arsenic exposure in humans.

The 49 genes with differential methylation associated with EUC arsenic are known to play a role in metabolic disease and cancer. This finding is intriguing, since we recently found that the accumulation of arsenic and MMAs in EUCs from subjects in the Chihuahua cohort is linked to an increased prevalence of diabetes and cardiometabolic risk factors (data not shown). Among these 49 genes was *USP7*, a regulator of diabetes-related signaling.[38] Of interest given known links between iAs exposure and diabetes,[3,39] knockdown of *USP7* in primary murine hepatocytes was shown to increase the expression of FoxO1-target gluconeogenic genes and increase glucose production.[40] Furthermore, in relationship to cancer, USP7 activity plays a role in oncogenesis where elevated *USP7* expression has been linked to cancers of the bladder, colon, prostate, liver, and lung.[41] Notably, cancers of the bladder, prostate, liver, and lung are among the known adverse effects of arsenic exposure.[4] Previous studies from our lab have linked arsenic exposure with alterations in gene methylation/signaling related to metabolic disease and cancer.[9,11,36,42,43] The current study further supports the hypothesis that arsenic and/or its metabolites may alter these key cellular signaling pathways through epigenetic modifications, specifically through the alteration of DNA methylation profiles in target tissues.

In order to determine whether the DNA methylation signature in EUCs from arsenic-exposed individuals may show commonalities with the genes altered in bladder cancer, the EUC arsenic-associated genes were compared against bladder cancer-associated genes analyzed from the Cancer Genome Atlas repository. A total of 22 of the 49 arsenic-associated genes (45%) were also differentially methylated in bladder cancer. While many studies have linked arsenic exposure with increased prevalence of bladder cancer in exposed populations,[4,37,44,45] these data provide novel evidence that the epigenetic dysregulation of cancer-related genes may play an important role in the carcinogenic effects of arsenic exposure.

Changes in DNA methylation levels do not always contribute to functional changes in gene expression. Our recent report suggests a small fraction of arsenic-associated DNA methylation changes may relate to functional changes at the transcript level.[33] However, DNA methylation within CpG islands located within gene promoter regions have been shown to be the most predictive of transcriptional changes,[33] and as such, the present analysis focused on CpG sites within gene promoter regions. As RNA and protein were not available from the limited cell number of EUCs, it was not possible to test whether the observed changes in DNA methylation were directly related to functional change in mRNA levels. Nevertheless, a total of 12 (25%) of the 49 differentially methylated genes have been previously associated with arsenic-induced changes in mRNA/protein levels. These findings provide support for potential links between arsenic-associated DNA methylation and functions changes in expression.

We have hypothesized that transcription factors, through their binding to specific regions of DNA and subsequent occupancy, may play a role in determining DNA methylation patterning that occurs in response to environment exposures, resulting in environmental footprints.[46] The current study's *in silico* evidence further supports a plausible role of transcription factor occupancy in the regulation of arsenic exposure-associated DNA methylation patterns. A total of 23 transcription factor families were identified with enriched binding sites in the promoter regions of EUC arsenic and bladder cancer-associated genes. This is of great interest because transcription factors are not only involved in the regulation of gene expression influenced by DNA methylation but also potential regulators of the DNA methylation profiles themselves.[47] Among the most significantly enriched transcription factor families were SIX3, TF2B, TCFF, and ZFHX, all of which have known associations with human cancers.[48−51] Of particular interest, dysregulation of two members of the ZFHX family, zinc finger E-box binding homeobox 1 (Zeb1) and zinc finger E-box binding homeobox 2 (Zeb2), has been associated with bladder cancer.[51,52] Overall, the extensive overlap between transcription factor binding sites within genes with differential methylation associated with EUC arsenic and bladder cancer indicates that similar transcription factors may mediate both exposure and disease-related epigenetic patterning. This finding may have clinical implication as the identified transcription factors could be targeted for disease prevention associated with environmental exposure to arsenic.

While this study provides an increased understanding of arsenic exposure and its link to epigenetic effects within cells

DOI: 10.1021/tx500393y
Chem. Res. Toxicol. XXXX, XXX, XXX−XXX

derived primarily from the bladder, it is not without limitations. The Cancer Genome Atlas data used to identify the bladder cancer-associated genes contains data from humans with differing demographics than those used in the arsenic analysis and was analyzed using a different methylation platform. The database also includes data from individuals with invasive bladder cancer, whereas the EUCs were collected from apparently healthy individuals. Despite these differences, 22 genes were identified as common among the genes with differential methylation associated with EUC arsenic and bladder cancer, representing important gene targets potentially linking arsenic exposure with disease. It is likely that even more overlaps would be identified if exposure and disease status (e.g., bladder cancer) were established in the same individual and study cohort. Nevertheless, such a comparison is currently limited by the available epigenetic repositories and the scope of medical examination that is feasible in a field study, similar to the present study involving the cohort in Chihuahua.

In summary, the results from the present study demonstrate that arsenic exposure and specific arsenic metabolites retained in EUCs are associated with the altered promoter methylation of genes involved in cancer and metabolic disease. Future studies will further examine biomarkers of arsenic exposure and metabolism and disease status while minimizing population variability. Taken together this research provides novel evidence of associations between arsenic and its metabolites with DNA methylation profiles within EUCs that primarily originate from a human tissue directly targeted by arsenic exposure, the bladder epithelium. Results from this study provide important knowledge of potential mechanistic links between environment exposure to arsenic and human health outcomes.

■ ASSOCIATED CONTENT

Supporting Information

Figures showing additional genome-wide statistical results, network signaling related to EUC arsenic, and PCR validation results. Tables detailing further arsenic measures, differentially methylated genes, comparative toxicogenomics results, and associated transcription factors. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/tx500393y.

■ AUTHOR INFORMATION

Corresponding Author

*Tel: 919-843-6864. Fax: 919-966-7991. E-mail: rfry@unc.edu.

Funding

This project was supported by the National Institutes of Health, specifically from the National Institute of Environmental Health Sciences (T32-ES007018, P30-ES010126, P42-ES005948, RO1-ES019315, R01-ES015326). Support was also provided by the UNC Nutrition Obesity Research Center (grant no. DK056350) and the Center for Environmental Health and Susceptibility (grant no. P30ES010126). This project was partially supported by the academic network "International Consortium of Research on Environmental Contaminants and their Health Effects" from PROMEP-SEP.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors would like to thank Nikhil Umesh and Elizabeth Martin for their assistance with the statistical analyses and Katie A. Bailey for her contributions to the sample preparation and processing.

■ ABBREVIATIONS

As, arsenic; As^{III}, arsenite; As^{V}, arsenate; BMI, body mass index; EUCs, exfoliated urothelial cells; DMAs, dimethylated arsenic; $DMAs^{III}$, dimethylarsinite; $DMAs^{V}$, dimethylarsinate; EOMES, eomesodermin; iAs, inorganic arsenic; LAMA2, laminin alpha 2; METTL13, methyltransferase like 13; MMAs, monomethylated arsenic or methylarsenic; $MMAs^{III}$, methylarsonite; $MMAs^{V}$, methylarsonate; PSMB2, proteasome (prosome, macropain) subunit, beta type, 2; RGPD4-AS1, RGPD4 antisense RNA 1; SLC12A7, solute carrier family 12 (potassium/chloride transporter), member 7; SIX3, sine oculis homeobox homologue 3; tAs, total arsenic; TCFF, TCF11 transcription factor; TF2B, RNA polymerase II transcription factor II B; USP7, ubiquitin specific peptidase 7; ZFHX, two-handed zinc finger homeodomain transcription factors

■ REFERENCES

(1) Uddin, R., and Huda, N. H. (2011) Arsenic poisoning in bangladesh. *Oman Med. J. 26*, 207.

(2) (2006) *Guidelines for Drinking Water Quality*, first addendum to 3rd addition, World Health Organization Press, Geneva, Switzerland.

(3) Kapaj, S., Peterson, H., Liber, K., and Bhattacharya, P. (2006) Human health effects from chronic arsenic poisoning—a review. *J. Environ. Sci. Health, Part A: Toxic/Hazard. Subst. Environ. Eng. 41*, 2399−2428.

(4) (2011) Arsenic and inorganic arsenic compounds, in *Report on Carcinogens*, 12th ed., U.S. Department of Health and Human Services, Public Health Service, National Toxicology Program, Research Triangle Park, NC.

(5) (2014) *SEER Stat Fact Sheets: Bladder Cancer*, National Cancer Institute, Bethesda, MD, http://seer.cancer.gov/statfacts/html/urinb.html.

(6) Kitchin, K. T., and Conolly, R. (2010) Arsenic-induced carcinogenesis—oxidative stress as a possible mode of action and future research needs for more biologically based risk assessment. *Chem. Res. Toxicol. 23*, 327−335.

(7) Shen, S., Lee, J., Weinfeld, M., and Le, X. C. (2008) Attenuation of DNA damage-induced p53 expression by arsenic: a possible mechanism for arsenic co-carcinogenesis. *Mol. Carcinog. 47*, 508−518.

(8) Reichard, J. F., and Puga, A. (2010) Effects of arsenic exposure on DNA methylation and epigenetic gene regulation. *Epigenomics 2*, 87−104.

(9) Rager, J. E., Bailey, K. A., Smeester, L., Miller, S. K., Parker, J. S., Laine, J. E., Drobná, Z., Currier, J., Douillet, C., Olshan, A. F., Rubio-Andrade, M., Stýblo, M., García-Vargas, G., and Fry, R. C. (2014) Prenatal arsenic exposure and the epigenome: altered microRNAs associated with innate and adaptive immune signaling in newborn cord blood. *Environ. Mol. Mutagen. 55*, 196−208.

(10) Chervona, Y., Hall, M. N., Arita, A., Wu, F., Sun, H., Tseng, H. C., Ali, E., Uddin, M. N., Liu, X., Zoroddu, M. A., Gamble, M. V., and Costa, M. (2012) Associations between arsenic exposure and global posttranslational histone modifications among adults in Bangladesh. *Cancer Epidemiol., Biomarkers Prev. 21*, 2252−2260.

(11) Smeester, L., Rager, J. E., Bailey, K. A., Guan, X. J., Smith, N., García-Vargas, G., Del Razo, L. M., Drobná, Z., Kelkar, H., Stýblo, M., and Fry, R. C. (2011) Epigenetic changes in individuals with arsenicosis. *Chem. Res. Toxicol. 24*, 165−167.

(12) Huang, Y. C., Hung, W. C., Chen, W. T., Yu, H. S., and Chai, C. Y. (2009) Sodium arsenite-induced DAPK promoter hypermethylation and autophagy via ERK1/2 phosphorylation in human uroepithelial cells. *Chem.−Biol. Interact. 181*, 254−262.

(13) Huang, Y. C., Hung, W. C., Chen, W. T., Jiang, W. H., Yu, H. S., and Chai, C. Y. (2011) Effects of MEK and DNMT inhibitors on

DOI: 10.1021/tx500393y
Chem. Res. Toxicol. XXXX, XXX, XXX−XXX

arsenic-treated human uroepithelial cells in relation to cyclin-D1 and p16. *Toxicol. Lett. 200*, 59−66.
(14) Cancer Genome Atlas Research Network (2014) Comprehensive molecular characterization of urothelial bladder carcinoma. *Nature 507*, 315−322.
(15) Vahter, M. (2002) Mechanisms of arsenic biotransformation. *Toxicology 181−182*, 211−217.
(16) Naranmandura, H., Carew, M. W., Xu, S., Lee, J., Leslie, E. M., Weinfeld, M., and Le, X. C. (2011) Comparative toxicity of arsenic metabolites in human bladder cancer EJ-1 cells. *Chem. Res. Toxicol. 24*, 1586−1596.
(17) Tseng, C. H. (2007) Arsenic methylation, urinary arsenic metabolites and human diseases: current perspective. *J. Environ. Sci. Health, Part C: Environ. Carcinog. Ecotoxicol. Rev. 25*, 1−22.
(18) Steinmaus, C., Yuan, Y., Kalman, D., Rey, O. A., Skibola, C. F., Dauphine, D., Basu, A., Porter, K. E., Hubbard, A., Bates, M. N., Smith, M. T., and Smith, A. H. (2010) Individual differences in arsenic metabolism and lung cancer in a case-control study in Cordoba, Argentina. *Toxicol. Appl. Pharmacol. 247*, 138−145.
(19) Chen, Y. C., Su, H. J., Guo, Y. L., Hsueh, Y. M., Smith, T. J., Ryan, L. M., Lee, M. S., and Christiani, D. C. (2003) Arsenic methylation and bladder cancer risk in Taiwan. *Cancer, Causes Control, Pap. Symp. 14*, 303−310.
(20) Currier, J. M., Ishida, M. C., González-Horta, C., Sánchez-Ramírez, B., Ballinas-Casarrubias, L., Gutiérrez-Torres, D. S., Hernández Cerón, R., Viniegra Morales, D., Baeza Terrazas, F. A., Del Razo, L. M., García-Vargas, G. G., Saunders, R. J., Drobná, Z., Fry, R. C., Matoušek, T., Buse, J. B., Mendez, M. A., Loomis, D., and Stýblo, M. (2014) Associations between arsenic species in exfoliated urothelial cells and prevalence of diabetes among residents of Chihuahua, Mexico. *Environ. Health Perspect. 122*, 1088−1094.
(21) Bundschuh, J., Litter, M. I., Parvez, F., Roman-Ross, G., Nicolli, H. B., Jean, J. S., Liu, C. W., Lopez, D., Armienta, M. A., Guilherme, L. R., Cuevas, A. G., Cornejo, L., Cumbal, L., and Toujaguez, R. (2012) One century of arsenic exposure in Latin America: a review of history and occurrence from 14 countries. *Sci. Total Environ. 429*, 2−35.
(22) Del Razo, L. M., García-Vargas, G. G., Valenzuela, O. L., Castellanos, E. H., Sánchez-Peña, L. C., Currier, J. M., Drobná, Z., Loomis, D., and Stýblo, M. (2011) Exposure to arsenic in drinking water is associated with increased prevalence of diabetes: a cross-sectional study in the Zimapán and Lagunera regions in Mexico. *Environ. Health 10*, 73.
(23) Hernández-Zavala, A., Matoušek, T., Drobná, Z., Paul, D. S., Walton, F., Adair, B. M., Jiří, D., Thomas, D. J., and Stýblo, M. (2008) Speciation analysis of arsenic in biological matrices by automated hydride generation-cryotrapping-atomic absorption spectrometry with multiple microflame quartz tube atomizer (multiatomizer). *J. Anal. At. Spectrom. 23*, 342−351.
(24) Fortin, F., Anghel, T., Brochu, P., and Lemieux, N. (2010) Optimizing urothelial cell preparation for the human urinary micronucleus assay. *Toxicol. In Vitro 24*, 1821−1827.
(25) Tyler, D. E. (1962) Stratified squamous epithelium in vesical trigone and urethra—findings correlated with menstrual cycle and age. *Am. J. Anat. 111*, 319−335.
(26) Landman, J., Chang, Y., Kavaler, E., Droller, M. J., and Liu, B. C. (1998) Sensitivity and specificity of NMP-22, telomerase, and BTA in the detection of human bladder cancer. *Urology 52*, 398−402.
(27) Hoque, M. O., Begum, S., Topaloglu, O., Chatterjee, A., Rosenbaum, E., Van Criekinge, W., Westra, W. H., Schoenberg, M., Zahurak, M., Goodman, S. N., and Sidransky, D. (2006) Quantitation of promoter methylation of multiple genes in urine DNA and bladder cancer detection. *J. Natl. Cancer Inst. 98*, 996−1004.
(28) Matoušek, T., Currier, J. M., Trojánková, N., Saunders, R. J., Ishida, M. C., González-Horta, C., Musil, S., Mester, Z., Stýblo, M., and Dědina, J. (2013) Selective hydride generation-cryotrapping-ICP-MS for arsenic speciation analysis at picogram levels: analysis of river and sea water reference materials and human bladder epithelial cells. *J. Anal. At. Spectrom. 28*, 1456−1465.
(29) Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., and Speed, T. P. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics 4*, 249−264.
(30) Edgar, R., Domrachev, M., and Lash, A. E. (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res. 30*, 207−210.
(31) Fry, R. C., Navasumrit, P., Valiathan, C., Svensson, J. P., Hogan, B. J., Luo, M., Bhattacharya, S., Kandjanapa, K., Soontararuks, S., Nookabkaew, S., Mahidol, C., Ruchirawat, M., and Samson, L. D. (2007) Activation of inflammation/NF-kappaB signaling in infants born to arsenic-exposed mothers. *PLoS Genet. 3*, e207.
(32) Davis, A. P., Murphy, C. G., Johnson, R., Lay, J. M., Lennon-Hopkins, K., Saraceni-Richards, C., Sciaky, D., King, B. L., Rosenstein, M. C., Wiegers, T. C., and Mattingly, C. J. (2013) The Comparative Toxicogenomics Database: update 2013. *Nucleic Acids Res. 41*, D1104−D1114.
(33) Rojas, D., Rager, J. E., Smeester, L., Bailey, K. A., Drobná, Z., Rubio-Andrade, M., Stýblo, M., García-Vargas, G., and Fry, R. C. (2014) Prenatal arsenic exposure and the epigenome: Identifying sites of 5-methyl cytosine alterations that predict functional changes in gene expression in newborn cord blood and subsequent birth outcomes. *Toxicol. Sci. 143*, 97−106.
(34) Du, P., Zhang, X., Huang, C. C., Jafari, N., Kibbe, W. A., Hou, L., and Lin, S. M. (2010) Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinf. 11*, 587.
(35) Ehrlich, M. (2002) DNA methylation in cancer: too much, but also too little. *Oncogene 21*, 5400−5413.
(36) Bailey, K. A., Wu, M. C., Ward, W. O., Smeester, L., Rager, J. E., García-Vargas, G., Del Razo, L. M., Drobná, Z., Stýblo, M., and Fry, R. C. (2013) Arsenic and the epigenome: interindividual differences in arsenic metabolism related to distinct patterns of DNA methylation. *J. Biochem. Mol. Toxicol. 27*, 106−115.
(37) Huang, Y. K., Huang, Y. L., Hsueh, Y. M., Yang, M. H., Wu, M. M., Chen, S. Y., Hsu, L. I., and Chen, C. J. (2008) Arsenic exposure, urinary arsenic speciation, and the incidence of urothelial carcinoma: a twelve-year follow-up study. *Cancer, Causes Control, Pap. Symp. 19*, 829−839.
(38) van der Horst, A., de Vries-Smits, A. M., Brenkman, A. B., van Triest, M. H., van den Broek, N., Colland, F., Maurice, M. M., and Burgering, B. M. (2006) FOXO4 transcriptional activity is regulated by monoubiquitination and USP7/HAUSP. *Nat. Cell Biol. 8*, 1064−1073.
(39) James, K. A., Marshall, J. A., Hokanson, J. E., Meliker, J. R., Zerbe, G. O., and Byers, T. E. (2013) A case-cohort study examining lifetime exposure to inorganic arsenic in drinking water and diabetes mellitus. *Environ. Res. 123*, 33−38.
(40) Hall, J. A., Tabata, M., Rodgers, J. T., and Puigserver, P. (2014) USP7 attenuates hepatic gluconeogenesis through modulation of FoxO1 gene promoter occupancy. *Mol. Endocrinol. 28*, 912−924.
(41) Nicholson, B., and Suresh Kumar, K. G. (2011) The multifaceted roles of USP7: new therapeutic opportunities. *Cell Biochem. Biophys. 60*, 61−68.
(42) Bailey, K. A., Laine, J., Rager, J. E., Sebastian, E., Olshan, A., Smeester, L., Drobná, Z., Styblo, M., Rubio-Andrade, M., García-Vargas, G., and Fry, R. C. (2014) Prenatal arsenic exposure and shifts in the newborn proteome: interindividual differences in tumor necrosis factor (TNF)-responsive signaling. *Toxicol. Sci. 139*, 328−337.
(43) Benton, M. A., Rager, J. E., Smeester, L., and Fry, R. C. (2011) Comparative genomic analyses identify common molecular pathways modulated upon exposure to low doses of arsenic and cadmium. *BMC Genomics 12*, 173.
(44) Smith, A. H., Goycolea, M., Haque, R., and Biggs, M. L. (1998) Marked increase in bladder and lung cancer mortality in a region of Northern Chile due to arsenic in drinking water. *Am. J. Epidemiol. 147*, 660−669.
(45) Karagas, M. R., Tosteson, T. D., Morris, J. S., Demidenko, E., Mott, L. A., Heaney, J., and Schned, A. (2004) Incidence of transitional cell carcinoma of the bladder and arsenic exposure in New Hampshire. *Cancer, Causes Control, Pap. Symp. 15*, 465−472.

DOI: 10.1021/tx500393y
Chem. Res. Toxicol. XXXX, XXX, XXX−XXX

(46) Sanders, A. P., Smeester, L., Rojas, D., DeBussycher, T., Wu, M. C., Wright, F. A., Zhou, Y. H., Laine, J. E., Rager, J. E., Swamy, G. K., Ashley-Koch, A., Lynn Miranda, M., and Fry, R. C. (2014) Cadmium exposure and the epigenome: Exposure-associated patterns of DNA methylation in leukocytes from mother−baby pairs. *Epigenetics 9*, 212−221.

(47) Smith, Z. D., and Meissner, A. (2013) DNA methylation: roles in mammalian development. *Nat. Rev. Genet. 14*, 204−220.

(48) Mo, M. L., Okamoto, J., Chen, Z., Hirata, T., Mikami, I., Bosco-Clement, G., Li, H., Zhou, H. M., Jablons, D. M., and He, B. (2013) Down-regulation of SIX3 is associated with clinical outcome in lung adenocarcinoma. *PLoS One 8*, e71816.

(49) Li, L., Zhang, A., Cao, X., Chen, J., Xia, Y., Zhao, H., and Shen, A. (2013) General transcription factor IIb overexpression and a potential link to proliferation in human hepatocellular carcinoma. *Pathol. Oncol. Res. 19*, 195−203.

(50) Steffen, J., Seeger, M., Koch, A., and Krüger, E. (2010) Proteasomal degradation is transcriptionally controlled by TCF11 via an ERAD-dependent feedback loop. *Mol. Cell 40*, 147−158.

(51) Lee, H., Jun, S. Y., Lee, Y. S., Lee, H. J., Lee, W. S., and Park, C. S. (2014) Expression of miRNAs and ZEB1 and ZEB2 correlates with histopathological grade in papillary urothelial tumors of the urinary bladder. *Virchows Arch. 464*, 213−220.

(52) Majid, S., Dar, A. A., Saini, S., Deng, G., Chang, I., Greene, K., Tanaka, Y., Dahiya, R., and Yamamura, S. (2013) MicroRNA-23b functions as a tumor suppressor by regulating Zeb1 in bladder cancer. *PLoS One 8*, e67686.

DOI: 10.1021/tx500393y
Chem. Res. Toxicol. XXXX, XXX, XXX−XXX