OSU Oregon State University

Open Access Articles

Differential DNA methylation in umbilical cord blood of infants exposed to mercury and arsenic in utero

The Faculty of Oregon State University has made this article openly available.
Please share how this access benefits you. Your story matters.

Citation	Cardenas, A., Koestler, D. C., Houseman, E. A., Jackson, B. P., Kile, M. L., Karagas, M. R., & Marsit, C. J. (2015). Differential DNA Methylation in Umbilical Cord Blood of Infants Exposed to Mercury and Arsenic in utero. Epigenetics, 10 (6), 508-515. doi:10.1080/15592294.2015.104602
DOI	10.1080/15592294.2015.1046026
Publisher	Taylor & Francis
Version	Accepted Manuscript
Terms of Use	http://cdss.library.oregonstate.edu/sa-termsofuse

Oregon State University OSU

Title: In Utero Arsenic Exposure and Epigenome-Wide Associations in Placenta, Umbilical Artery and Human Umbilical Vein Endothelial Cells

Authors: Andres Cardenas[1], E. Andres Houseman[1], Andrea A. Baccarelli[2], Quazi Quamruzzaman[3], Mahmuder Rahman[3], Golam Mostofa[3], Robert O. Wright[4], David C. Christiani[2] and Molly L. Kile[1]

Affiliations:

[1] School of Biological and Population Health Sciences, College of Public Health and Human Sciences, Oregon State University, Corvallis, OR USA

[2] Harvard T.H. Chan School of Public Health; Boston, MA USA

[3] Dhaka Community Hospital; Dhaka, Bangladesh

[4] Preventative Medicine and Pediatrics; Mt Sinai School of Medicine; New York, NY USA

Corresponding Author:

Molly L. Kile, ScD, College of Public Health and Human Sciences

Oregon State University, 15 Milam Hall, Corvallis, OR 97331

Telephone: (514) 737-1443 Fax: 541-737-6914

Email: Molly.Kile@OregonState.edu

Running title: Epigenetic Effect of Arsenic in Placenta, Artery and HUVEC

Keywords: Arsenic, DNA methylation, Epigenetics, Illumina 450K, *In-utero* exposure, environmental epigenetics, fetal programming

Competing financial interest: The authors declare no competing financial interests.

ABSTRACT

Exposure to arsenic early in life has been associated with increased risk of several chronic diseases and is believed to alter epigenetic programming in utero. In the present study, we evaluate the epigenome-wide association of arsenic exposure in utero and DNA methylation in placenta (n=37), umbilical artery (n=45) and human umbilical vein endothelial cells (HUVEC) (n=52) in a birth cohort using the Infinium HumanMethylation450 BeadChip array. Unadjusted and cell mixture adjusted associations for each tissue were examined along with enrichment analyses relative to CpG island location and omnibus permutation tests of association among biological pathways. One CpG in artery (cg26587014) and four CpGs in placenta (cg12825509; cg20554753; cg23439277; cg21055948) reached a Bonferroni adjusted level of significance. Several CpGs were differentially methylated in artery and placenta when controlling the false discovery rate (q-value<0.05), but none in HUVEC. Enrichment of hypomethylated CpG islands was observed for artery while hypermethylation of open sea regions were present in placenta relative to prenatal arsenic exposure. The melanogenesis pathway was differentially methylated in artery (Max F $P<0.001$), placenta (Max F $P<0.001$) and HUVEC (Max F $P=0.002$). Similarly, the insulin signaling pathway was differentially methylated in artery (Max F $P=0.02$), placenta (Max F $P=0.03$) and HUVEC (Max F $P=0.002$). Our results show that prenatal arsenic exposure can alter DNA methylation in artery and placenta but not in HUVEC. Further studies are needed to determine if these alterations in DNA methylation mediate the effect of prenatal arsenic exposure and health outcomes later in life.

INTRODUCTION

Over 200 million individuals worldwide are exposed to elevated levels of inorganic arsenic. This is a public health concern because arsenic is a known human carcinogen and chronic exposure is associated with the development of skin, lung, bladder, kidney, liver and potentially prostate cancer.[1] Particularly, early life exposure to arsenic has been associated with the development of many latent health effects including carcinogenesis.[2] Human ecological studies from the Antofagasta region of Chile have associated prenatal and early childhood exposure to arsenic from contaminated municipal water with increased risk of lung and bladder cancer later in life.[3] Increased mortality from acute myocardial infarction and cancers of the bladder, kidney, lung, and liver have also been reported from this population decades after the exposure declined.[4, 5]

Of public health interest is the ability of early life arsenic exposure, particularly exposures occurring *in utero*, to increase disease risk and susceptibility to adverse health conditions later in life. For example, animal models support the involvement of transplacental arsenic exposure in the development and progression of atherosclerosis, consistent with human studies linking early life exposure and cardiovascular disease.[6, 7] Emerging evidence also indicates that exposure to arsenic can disrupt normal immune function and *in utero* exposure can increase the susceptibility and severity of infections later in life.[8-10] Furthermore, arsenic exposure during fetal development has been associated with growth restrictions and adverse perinatal health outcomes such as low birth weight, still births, infant mortality, and preterm births.[11] Lastly, latent adverse neurological health outcomes have also been documented with maternal exposure to arsenic during pregnancy.[12, 13]

~~The exact molecular mechanisms of the toxicological effects attributed to arsenic exposure~~

~~remains elusive and no single mechanism has been identified in the development of arsenic associated diseases and the observed latency of health effects.~~[14] ~~However, T~~The latency of health effects documented in epidemiological studies and animal models along with the observed susceptibility of prenatal exposures are suggestive of an epigenetic mode of action. Fetal programming events involving DNA methylation occur at critical windows of fetal development in a cell-specific manner shown to be sensitive to environmental exposures.[15] Experimental evidence from animal models demonstrate that transplacental exposure to arsenic leads to epigenetic alterations, changes in gene expression and increased incidence of tumors in the offspring.[16, 17] Therefore, it is postulated that epigenomic regulation including, but not limited to, DNA methylation is a potential mechanism of arsenic induced carcinogenesis and latent disease risk.[2, 18, 19] Other likely interacting mechanisms of early life exposure to arsenic and latent disease risk include the development of cancer stem cells and perturbations of immune function.[2]

Several human studies have evaluated the impact of prenatal arsenic exposure on the cord blood and whole blood epigenome..[1] Among these epidemiological studies evaluating cord blood or whole blood DNA methylation no common loci has been identified to be differentially methylated across studies.[20-25] However, significant DNA methylation disruption of unique loci along with enrichment of key regulatory CpG regions has been documented across different study populations.[21-27] Besides studies that examined cord and whole blood epigenome, only two studies to date have evaluated the association between arsenic exposure and CpG methylation of target tissue by evaluating DNA methylation in urothelial carcinoma samples and CpG methylation of exfoliated urothelial cells, respectively.[28, 29] These studies found differentially methylated loci associated with arsenic exposure in key regulatory genes potentially involved in development of arsenic induced urothelial carcinoma.

95 Epigenetic reprogramming during fetal development resulting from transplacental exposure
96 is one of the main hypothesized mechanisms of arsenic's associated-disease.[2] To further our
97 understanding of how prenatal arsenic exposure could alter epigenetic programming it is
98 important to evaluate its effect on different tissues with diverse cellular compositions. Evaluating
99 if exposure to arsenic in utero alters DNA methylation of different tissues could yield insights
100 into the etiology of toxicant-mediated disease and epigenetic modifications of relevant tissues
101 with specific biological functions. Subsequently, we examined the association between maternal
102 drinking water arsenic as a proxy of transplacental exposure during fetal development and the
103 epigenome of placenta, umbilical artery and Human Umbilical Vein Endothelial Cells (HUVEC)
104 from a birth cohort conducted in arsenic affected regions of Bangladesh.

105 **RESULTS**

106 The sample size varied by tissue type with a maximum of 52 samples present for HUVEC
107 followed by 45 samples in umbilical artery and 37 placenta samples. Arsenic concentration in
108 maternal drinking water at study enrollment ranged from below the detection limit of <1μg/L to
109 510 μg/L with a mean exposure concentration of 63.7 μg/L. Selected sample characteristics are
110 shown in Table 1.

111 **Arterial Tissue**

112 **Locus-by-Locus Analysis:** In the analysis that was unadjusted for cellular composition, one
113 CpG loci (cg26587014) located in chromosome 19 and not annotated to any gene was
114 differentially methylated in arterial tissue in relation to arsenic exposure using a Bonferroni
115 threshold for statistical significance ($P<1.33\times10^{-7}$). Controlling for the false discovery rate at 5%
116 (q-value<0.05) revealed 2,105 CpGs that were differentially methylated relative to $\log_2$-
117 transformed maternal drinking water arsenic. However, after adjusting for cellular composition

118 using the Houseman reference-free method, no loci reached a Bonferroni corrected level of
119 significance or a q-value<0.05. Unadjusted and adjusted results are shown in Figure 1A and 1B,
120 respectively. The top 100 differentially methylated loci ranked on lowest *P*-value are
121 summarized in supplementary table S1 and S2 for unadjusted and cell mixture adjusted analyses,
122 respectively. In unadjusted analyses, differentially methylated loci with a q-value<0.05 were
123 disproportionately located in CpG islands (54%) compared to the distribution of CpG island
124 probes in the rest of array (33%) ($P<1x10^{-4}$), supplementary Figure S1A. The majority of
125 unadjusted hypomethylated loci with a q-value<0.05 were located in CpG islands (83%), Figure
126 1C. After adjusting for cellular heterogeneity, a similar enrichment of hypomethylated loci in
127 CpG islands was observed among top loci having a nominal p-value$<1x10^{-4}$, Supplementary
128 Figure S1B.
129 **Biological Pathway Analysis:** Omnibus permutation based tests revealed significant
130 associations between *in utero* exposure to arsenic and epigenetic disruption of KEGG biological
131 pathways in arterial tissue (Mean F-statistics *P*=0.009 and maximum F-statistic *P*=0.006).
132 Pathways that were observed to have the strongest association based on the lowest mean F-static
133 level of significance (*P*=0.004) were: maturity onset of diabetes of the young (hsa04950),
134 primary immunodeficiency (hsa05340), ABC transporters (hsa02010), allograft rejection
135 (hsa05330) and vibrio cholerae infection (hsa05110). Differentially methylated pathways
136 observed to have a strong association using a maximum F-statistic level of significance
137 (*P*<0.001) included: the Hedgehog signaling pathway (hsa04340), Melanogenesis (hsa04916),
138 Wnt signaling pathway (hsa04310), Basal cell carcinoma (hsa05217), DNA replication
139 (hsa03030) and the p53 signaling pathway (hsa04115). The summary for all associations
140 between maternal drinking water arsenic and epigenetic disruption of KEGG biological

141 pathways are shown in supplementary table S7.

142 **Placenta Tissue**

143 **Locus-by-Locus Analysis:** In the analyses that were unadjusted for cellular composition, no
144 single CpG loci reached Bonferroni adjusted significance in placenta ($P<1.37\times10^{-7}$). However,
145 two CpG loci (cg26390526; cg03857453) annotated to the Epidermal Filaggrin gene (*FLG*) and
146 the nuclear receptor subfamily 3, group C, member 1 glucocorticoid receptor gene (*NR3C1*) were
147 hypermethylated relative to maternal drinking water arsenic after controlling for the false-
148 discovery rate (q-value<0.05). In analyses that adjusted for cell mixture in the placenta, four
149 CpGs reached Bonferroni adjusted significance: cg12825509 (*TRA2B* gene), cg20554753,
150 cg23439277 (*PLCE1* gene) and cg21055948 (*CD36* gene). Moreover, analyses adjusted for
151 cellular heterogeneity revealed 518 CpG loci that were differentially methylated after controlling
152 for the false discovery rate (q-value<0.05). Unadjusted and cell mixture adjusted results are
153 shown in Figure 2A and 2B, respectively. The top 100 differentially methylated loci ranked on
154 lowest *p*-value are summarized in supplementary table S3 and S4 for unadjusted and cell mixture
155 adjusted results, respectively. For the top unadjusted differentially methylated loci with a
156 nominal $P<1x10^{-4}$ a disproportionate amount of CpGs were located within open sea regions of
157 CpG islands (76%) compared with the distribution of open sea loci in the rest of array (33%)
158 ($P<1x10^{-4}$), supplementary figure 2A. Among these loci, the great majority of hypermethylated
159 CpGs were located within open sea regions (89%) relative to CpG islands, Figure 2C. For the
160 cell mixture adjusted analyses a similar enrichment of hypermethylated loci in open sea regions
161 was observed for loci with a q-value<0.05, supplementary Figure 2B.

162 **Biological Pathway Analysis:** Omnibus permutation based tests indicated that exposure to
163 arsenic *in utero* disrupts methylation of a small number of CpGs within KEGG biological

164 pathways in the placenta tissue (Omnibus maximum F-statistic P=0.004). However, a marginal
165 association among KEGG biological pathways and arsenic exposure was observed using an
166 omnibus Mean F-statistic test for association (P=0.108). KEGG biological pathways that were
167 differentially methylated in relationship to arsenic exposure with a maximum F-statistic $P<1\text{x}10^{-3}$
168 included: Melanogenesis (hsa04916), Neuroactive ligand-receptor interaction (hsa04080),
169 Calcium signaling pathway (hsa04020), GnRH signaling pathway (hsa04912), Dilated
170 cardiomyopathy (hsa05414), Gap junction (hsa04540), Vasopressin-regulated water reabsorption
171 (hsa04962), Vascular smooth muscle contraction (hsa04270), Oocyte meiosis (hsa04114), Vibrio
172 cholerae infection (hsa05110), Progesterone-mediated oocyte maturation (hsa04914) and the
173 Peroxisome pathway (hsa04146). Several other pathways were significantly associated with
174 arsenic exposure using a maximum F-statistic $P<0.05$ and summarized in supplementary table
175 S7.

176 **Umbilical Vein Endothelial Cells (HUVEC)**

177 **Locus-by-Locus Analysis:** In both unadjusted and cell mixture adjusted analyses no single CpG
178 loci was associated with arsenic exposure at a Bonferroni corrected level of significance
179 ($P<1.44\text{x}10^{-7}$) or a q-value<0.05 after controlling for the false discovery rate, Figure 3A and 3B.
180 Among the top 31 CpG loci with a nominal $P<1\text{x}10^{-4}$ no significant differences were present for
181 the occurrence of top loci relative to CpG island location compared to the rest of the array for
182 unadjusted analyses, Figure 3C. The top 100 differentially methylated loci ranked by lowest p-
183 value are summarized in supplementary table S5 and S6 for unadjusted and cell mixture adjusted
184 results, respectively.

185 **Biological Pathway Analysis:** Omnibus permutation tests for association among KEGG
186 biological pathways indicated that arsenic exposure was not significantly associated with a large

187 number of changes in DNA methylation across pathways in HUVEC (Mean F-statistic P=0.129)
188 and the presence of a small number of strong associations was borderline significant (Max F-
189 statistic P=0.06). Few individual biological pathways reached statistical significance using a
190 maximum F-statistic level of significance. The top differentially methylated biological pathways
191 (maximum F-statistic P=0.002) in HUVEC included: Melanogenesis (hsa04916), Wnt signaling
192 pathway (hsa04310), Basal cell carcinoma (hsa05217) and the Insulin signaling pathway
193 (hsa04910), all KEGG biological pathway based associations are summarized in supplementary
194 table S7.

195 The overlap among CpGs within each tissue for unadjusted and cell mixture adjusted
196 analyses using the top 100 differentially methylated CpGs was 26 loci in artery, 21 loci in
197 placenta and 33 loci for HUVEC, supplementary figure S3. Among the top 100 differentially
198 methylated loci, only one CpG (cg21002651) located within the body of the *CASP1* gene was
199 differentially methylated across two tissues in unadjusted analyses. This loci was
200 hypomethylated in placenta (β=-0.20, $P=5.73\times10^{-6}$) but hypermethylated in HUVEC (β=0.20,
201 $P=1.29\times10^{-4}$) in relationship to maternal drinking water arsenic. No other CpGs overlapped in
202 unadjusted or adjusted analyses.

203 **DISCUSSION**

204 Our study provides evidence that *in utero* exposure to arsenic can disrupt DNA
205 methylation of artery and placenta tissues but the association with umbilical vein endothelial
206 vein cells was marginal. However, the association of prenatal arsenic exposure on the epigenome
207 on artery and placenta depended on the cell mixture adjustment. For instance, the association in
208 artery was attenuated after controlling for cellular heterogeneity but strengthened in placenta. *In*
209 *utero* exposure to arsenic was also associated with DNA methylation levels of key biological

210 pathways across tissues providing new insights into the potential etiology of arsenic-mediated
211 diseases with a plausible epigenetic reprogramming component.

212 In normal tissue, the majority of CpG islands remain unmethylated and methylation of
213 CpG islands located within promoter regions of genes is usually restricted to genes at which
214 there is long-term stabilization of repressed states such as in gene silencing of imprinted genes.[30]
215 However, CpG island methylation is not deterministic of gene expression and further studies are
216 needed to determine if the observed alterations in DNA methylation are associated with
217 biological effects. Conversely, we observed an enrichment of hypomethylated loci in CpG
218 islands relative to prenatal arsenic exposure. This is of particular interest because both animal
219 and human studies have demonstrated that DNA hypomethylation occurs in atherosclerotic
220 lesions and that hypomethylation of CpG islands is observed broadly in human atherosclerotic
221 arteries[31, 32] and in arterial disease pathogenesis.[33] ~~In animal models, *in utero* arsenic exposure~~
222 ~~has been shown to induce the early onset of atherosclerosis along with epidemiological studies~~
223 ~~linking early life exposure with cardiovascular disease.[6, 7] Therefore, we hypothesize that the~~
224 ~~observed hypomethylation of influential genomic regions such as CpG islands could play a role~~
225 ~~in the development of arsenic-associated cardiovascular disease, particularly atherosclerosis of~~
226 ~~arterial tissue. Another early observation from epigenetic cancer studies was the global~~
227 ~~hypomethylation of tumor samples compared to normal tissue mainly at repetitive genomic~~
228 ~~elements and that hypomethylation of these regions can lead to hypermethylation of tumor~~
229 ~~suppressor genes.~~[34] Along with this observation, previous studies of arsenic exposure have
230 characterized hypermethylation of the promoter region of the p53 gene a mechanisms
231 hypothesized to contribute to the carcinogenesis of arsenical compounds.[35, 36] Consistent with
232 these reports, our gene set analysis shows that CpG methylation within the p53 signaling

pathway is associated with arsenic exposure during pregnancy suggesting that artery might be a target tissue for the epigenetic toxicity of arsenic, and potentially involved in carcinogenesis. However, this hypothesis needs to be evaluated.

The placenta is an important regulator of fetal development and intrauterine growth that plays a crucial role mediating the maternal and fetal environment. Furthermore, the placenta is a unique epigenetic target organ as the majority of imprinted genes in animal models are both expressed and imprinted in the placenta and hypothesized to contribute to fetal neurodevelopment.[37, 38]In unadjusted analyses a CpG located in the body of the glucocorticoid receptor gene (*NR3C1*) was significantly hypermethylated in the placenta relative to prenatal arsenic exposure. Previous studies have shown that hypermethylation of the *NR3C1* gene influences cortisol response, infant behavior and self-regulation.[39, 40] Interestingly, a recent experimental study demonstrated that exposure to arsenic *in utero* lowers the activity of the glucocorticoid receptor pathway and these changes were maintained into adolescence of the mouse model.[41] Although further studies are needed to confirm whether these biological findings are connected and related to behavioral outcomes. ~~The placenta has also been characterized as one of the hypomethylated tissues as LINE-1 elements have lower levels of methylation when compared to other tissues. Furthermore, it has been shown that normal human placenta contains partially methylated domains (37%) with the ability to suppress genes and impact tissue-specific functions independent of the tissue of origin.[42] The observed hypermethylation of open sea regions relative to CpG island location could have implications for normal methylation of LINE-1 elements and partially methylated domains, potentially affecting normal biological function and development of the placenta.~~

A few KEGG biological pathways were differentially methylated in relation to maternal

256 drinking water arsenic in all three tissues. Namely, DNA methylation of the melanogenesis
257 pathway was strongly associated with exposure to arsenic in artery, placenta and HUVEC. An
258 early clinical symptom of arsenicosis include the appearance of hyperpigmentation changes of
259 the skin in the trunk, neck and chest regions of the body eventually progressing to the palmar and
260 plantar regions and eventually leading to hyperkeratosis.[43] Consistent with the differential
261 methylation of this pathway in our data, arsenic-associated alterations in DNA methylation of
262 leukocytes has been previously associated with increased risk of developing skin lesions.[44]
263 Lastly, the insulin signaling pathway was observed to be differentially methylated across all
264 three tissues with respect to arsenic exposure. Exposure to arsenic has been consistently
265 associated with Type 2 diabetes and insulin resistance in both animal models and
266 epidemiological studies.[46] Previous studies have documented the epigenetic disruption of several
267 genes involved in the development of diabetes and insulin resistance for individuals chronically
268 exposed to arsenic.[47] Although epigenetic distribution was characterized among these biological
269 pathways relative to arsenic exposure, future studies need to evaluate if these changes are
270 associated with changes in gene expression, metabolism and ultimately pathological phenotypes.
271 It is also important to note that the permutation test used in this analysis evaluates the DNA
272 methylation disruption of the biological pathways at a global level and not on a gene by gene
273 basis. Therefore, it is not possible to determine if individual genes are differentially methylated
274 with regards to arsenic exposure.

275 It is crucial to highlight that HUVEC is a homogenous tissue in terms of cellular
276 composition and was not significantly disrupted in the locus-by-locus analysis and marginally
277 associated among some biological pathways. However, artery and placenta, both representing a
278 diverse mixture of cell types, were observed to be differentially methylated relative to prenatal

279 arsenic exposure. DNA methylation is cell specific playing a key role in tissue differentiation
280 and lineage commitment making this process particularly vulnerable to environmental stimuli
281 and exposures during fetal development. The placenta represents the most diver tissue composed
282 of fetal vascular cells, mesenchymal cells, cytotrophoblast and syncytiotrophoblast that originate
283 from the trophoblast.[48] Furthermore, it has also been observed that the human placenta contains
284 both hematopoietic stem cells and mesenchymal stem cells.[49] All tissues derived from the fetus
285 are an extension of the mesoderm that differentiates during embryonic development to form the
286 umbilical cord and placenta. Therefore, it might be possible to arsenic exposure during fetal
287 development could affect cellular differentiation for placenta and artery but not HUVEC as this
288 is a cellular homogenous tissue. Epidemiologic studies often rely on preserved samples and have
289 limited fresh tissue availability making the sorting or isolation of target cell types not feasible.
290 Therefore, future experimental studies should evaluate the development of cancer stem cells
291 (CSCs) and alterations to the immune function as factors or intermediary mechanisms of the
292 observed epigenetic perturbations, as others have also suggested.[2] Moreover, the interaction
293 between prenatal arsenic exposure and other transplacental contaminants should also be
294 considered, as prenatal exposure to arsenic has been previously shown to interact with other
295 prenatal exposures such as mercury.[50]
296 One of the major strengths for the present study is the epigenome-wide analysis of three
297 different tissues collected from the same maternal-infant pairs yielding insights for the potential
298 biological impact of arsenic exposure during fetal development. Also, the prospective design of
299 this birth cohort along with the exposure assessment early during pregnancy are important
300 qualities that strengthens the temporality of the epigenetic perturbations reported. Although the
301 present study relies on a single water sample during early pregnancy and exposure

302 misclassification cannot be ruled out, previous studies in rural Bangladesh have demonstrated
303 that drinking water arsenic exposure is relatively constant and correlated with biomarkers of
304 internal doses, such as urine and toenails[51, 52] and that arsenic readily crosses the placenta.[53]
305 Additionally, the availability of umbilical samples at birth provides one of the few opportunities
306 for examining epigenetic programming in cardiovascular target tissue in a non-invasive and
307 feasible manner. There are a number of important limitations to our current study including the
308 relatively small sample size and the lack of validation using a complementary DNA methylation
309 method due to sample availability. The lack of reference methylomes for placenta, artery and
310 HUVEC also raise an important challenge when interpreting the observed epigenetic
311 perturbations in tissues that might represent a mixture of cell types such as artery or placenta.
312 However, we implemented a complementary bioinformatics method to adjust for cellular
313 heterogeneity to identify potential perturbations in loci hypothesized to be associated with
314 methylation levels independent of cellular heterogeneity. Furthermore, unique tissue samples
315 were analyzed in separate plates raising the possibility that differences across tissue could be
316 potentially attributed to technical plate effects. Finally, gene expression was not measured and
317 the observed changes in DNA methylation need to be further confirmed and evaluated.
318 Particularly significant association between DNA methylation among KEGG biological
319 pathways might not result in functional gene expression or proteomic alterations within
320 pathways.

321 In conclusion, we show that prenatal arsenic exposure is associated with altered DNA
322 methylation of umbilical artery and placenta tissue but evidence of an association for HUVEC is
323 limited. Furthermore, we present evidence of DNA methylation disruption of key biological
324 pathways across different tissues holding the potential to mediate arsenic-associated diseases

325 previously described from exposures *in utero*.

326 MATERIALS & METHODS

327 Study Population

328 This pilot study was nested within an established birth cohort recruited in Bangladesh
329 (2007-2011) and designed to characterize the potential epigenetic disruption associated with
330 arsenic exposure during pregnancy in different tissues collected at birth. A more detailed
331 explanation of the full birth cohort has been published previously.[21] Briefly, pregnant women
332 with ≤ 16 weeks of gestation confirmed by ultrasound were enrolled in a prospective
333 reproductive birth cohort in Bangladesh. Trained health care workers at community health clinics
334 in Sirajdikhan and Birahimpur recruited pregnant women 18 years of age or older that used a
335 tube-well as their primary drinking water source, planned to live at their current residency during
336 the duration of the pregnancy and received prenatal health care at Dhaka Community Hospital
337 (DCH) or affiliated community clinic. Study participants agreed to deliver at DCH or at home
338 with a DCH trained midwife. Informed consent was obtained from all participants prior to
339 enrollment. All participants were provided with prenatal care and prenatal vitamins offered by
340 DCH. This study was approved by the Human Research Committees at the Harvard School of
341 Public Health, Oregon State University and Dhaka Community Hospital Trust.

342 Three distinct tissues were collected at the time of delivery including: artery from the
343 umbilical cord, placenta, and endothelial cells isolated from the umbilical vein. Since the goal of
344 this pilot study was to examine the potential exposure-response relationship between arsenic and
345 DNA methylation, specimens were selected based on maternal drinking water arsenic
346 concentrations at study enrollment to cover a wide range of exposures (<1-510 μg/L). A total of
347 37 placenta samples, 45 artery samples and 52 HUVEC samples were included in the final

348 analysis.

349 **Drinking Water Arsenic**

350 Water samples were collected from the tube-well identified by participants as their main
351 source of drinking water at the time of their enrollment into the study as previously described.[21]
352 Briefly, water samples were collected in a 50-mL polypropylene tubes (BD Falcon, BD
353 Bioscience, Bedford, MA), preserved with Reagent Grade nitric acid (Merck, Germany) to a
354 pH<2 and stored at room temperature. Arsenic concentrations were measured by inductively
355 coupled plasma-mass spectrometry (ICP-MS) using the US EPA method 200.8 to determined
356 metals in water (Environmental Laboratory Services, North Syracuse, New York).[54] Average
357 percent recovery for Arsenic from plasmaCal multi-element QC standard #1 solution (SCP
358 Science) was 102% ± 7%. The limit of detection (LOD) for arsenic in drinking water was 1
359 µg/L.

360 **Tissue Collection: Umbilical Artery, Placenta & HUVEC**

361 Trained medical workers present at delivery collected a sample of the umbilical cord and
362 placenta immediately after the delivery was completed. Using sterile techniques, approximately
363 5-7 cm of umbilical vein was dissected out of fresh umbilical cord and rinsed with phosphate
364 buffered saline solution to remove external contamination. The vein lumen was then bisected and
365 the interior cavity was flushed with approximately 100 mL of phosphate buffered solution to
366 remove blood. The interior lumen wall was gently rubbed using a sterile cytology brush to
367 collected endothelial cells. The cytology brush was then vortexed in 1 mL of cell lysis solution
368 (Qiagen) to transfer the cells. The cell lysis solution was then stored at 4 ^{0}C. Samples were
369 shipped to Harvard School of Public Health where the DNA was extracted using DNeasy Blood
370 & Tissue Kit (Qiagen) following manufacturer's instructions.

371 Approximately 1 cm of umbilical cord artery was dissected out of fresh umbilical cord,
372 the exterior of the artery was scraped to remove Wharton's Jelly, and rinsed with phosphate
373 buffered saline solution to remove blood. The arterial cross section was placed in 2 mL of
374 RNAse later and stored at -20 ^{0}C. Samples were shipped to Harvard School of Public Health on
375 dry ice. The artery sample was then minced using a sterile scalpel and added to Maxwell Cell
376 DNA Purification kits (Promega) with an additional 20 μL of Proteinase K (Qiagen). Samples
377 were allowed to sit for 30 minutes before being extracted using the Maxwell 16 Research
378 instrument following manufacturer's instructions.

379 For placenta samples, a one centimeter tissue plug was excised from fresh placenta. The
380 tissue plug was placed into a sterile vial and covered with Tissue-Tek O.C.T. gel (Electron
381 Microscopy Sciences) and frozen at -20 ^{0}C. Samples were then shipped to Harvard School of
382 Public Health on dry ice. Next, approximately 10 grams of placenta tissue was removed from the
383 plug and minced using a sterile scalpel and added to Maxwell Cell DNA Purification kits
384 (Promega) with an additional 20 μL of Proteinase K (Qiagen). Samples were allowed to sit for
385 30 minutes before being extracted using the Maxwell 16 Research instrument following
386 manufacturer's instructions.

387 **DNA Methylation Assessment and Quality Control**

388 DNA was shipped to the University of Minnesota's Biomedical Genomic Center that
389 quantified DNA methylation using the Illumina Infinium HumanMethylation450 BeadChip
390 (Illumina, San Diego, CA) following standard manufacturer's protocols. The
391 HumanMethylation450 BeadChip measures DNA methylation at > 485,000 CpG sites at single
392 nucleotide resolution, covering 99% of the RefSeq genes.

393 Tissues were analyzed in separate plates and randomly allocated to different chips. Data
394 were obtained and processed from raw methylation image files and normalized using internal
395 control probes via the functional normalization method with two principal components to
396 account for technical variation between samples using the *minfi* package of R.[55] DNA
397 methylation was estimated at each CpG as the fraction of DNA molecules whose target CpG loci
398 is methylated and referred to as β-values. Measurements at CpG loci on X and Y chromosomes
399 were excluded from the analysis to avoid gender-specific methylation bias. Previously identified
400 non-specific and cross-reactive probes within the array along with polymorphic CpG loci ($\geq$ 5%
401 of the minor allele frequency) were removed for the analysis.[56] Furthermore, a detection *P-value*
402 was computed for all CpGs and probes with non-significant detection (P>0.01) in greater than
403 10% of the samples were removed from the analysis. After quality control, the total number of
404 autosomal CpGs left in the analysis were 374,320 loci for artery, 365,994 loci for placenta and
405 347,650 loci in HUVEC samples. Finally, a beta-mixture quantile intra sample normalization
406 procedure (BMIQ) was further applied to the data to reduce the potential bias that can arise from
407 type 2 probes as previously described.[57] Strip plots and signal intensities of control probes were
408 visually examined for bisulfite conversion, probe hybridization and single base extension.
409 Density plots for the β-values were examined for all samples at each normalization step
410 described above.

411 **Statistical Analysis**

412 **Unadjusted and Cell-adjusted Locus-by-Locus Analysis:** We first aimed to identify
413 differentially methylated CpG loci in relationship to prenatal arsenic exposure from maternal
414 drinking water. Maternal arsenic concentration in water was right skewed and subsequently $\log_2$-
415 tranformed. In order to evaluate linear associations between prenatal exposure to arsenic and

416 differentially methylated CpG loci, β-values were logit-transformed to M-values previously
417 described to be more appropriate for differential analysis of DNA methylation.[58] In the locus-by-
418 locus approach, two different but complementary methodologies were implemented. First, the
419 linear association between individual CpG methylation on the M-value scale and $\log_2$-
420 transformed arsenic was evaluated adjusting for infant sex using the *limma* function found in the
421 *minfi* package of R. Second, due to the lack of reference methylomes of isolated cell types in
422 placenta, artery or HUVEC, a novel reference-free method of adjusting for cellular heterogeneity
423 was implemented using the *RefFreeEWAS* package of R. The reference-free method is an
424 extension of the original Houseman method that utilizes a deconvolution approach similar to
425 surrogate variable analysis (SVA) that is data driven to identify latent variables or dimensions as
426 surrogates of cellular composition.[59] Using this method, the sex adjusted linear association
427 between individual CpG methylation on the β-value scale and $\log_2$-transformed maternal
428 drinking water arsenic was evaluated using 1000 bootstrap samples for estimating the standard
429 errors of association in placenta, umbilical artery and HUVEC. Results from the unadjusted
430 *limma* models and the reference-free cell mixture adjusted analyses were compared within
431 tissues and across tissues. Enrichment analyses for the distribution of CpGs relative to CpG
432 island location of the top differentially methylated loci based on a q-value<0.05 or a nominal
433 $P<1 \times 10^{-4}$, were compared to the distribution of probes on the rest of the array.
434 **Biological Pathway Analysis:** Omnibus permutation based tests and *p*-values were obtained by
435 mapping subsets of CpGs to their associated genes in specific KEGG biological pathways. Gene
436 sets were compiled from the Kyoto Encyclopedia of Genes and Genomes (KEGG) corresponding
437 to specific biological pathways using the Entrez IDs matched to KEGG biological pathways
438 using the Bioconductor library *org.Hs.eg.db*. The permutation distribution was obtained from

439 unadjusted cell mixture models by permuting the exposure with respect to measured DNA
440 methylation over subgroups of CpGs defined by biological pathways (1000 permutations).
441 Unadjusted cell mixture methylation analyses were used for the KEGG pathways as the
442 reference-free Houseman method is unable to accommodate for the permutation test of CpGs
443 across individual biological pathways. Pathway based associations of DNA methylation with
444 prenatal arsenic exposure as a continuous variable were summarized using a maximum nominal
445 F-statistics *p*-value (akin to a minimum *p*-value) and an average nominal F-statistic *p*-value. The
446 maximum and minimum F-statistic *p*-values are better suited for detecting a small number of
447 strong associations and a large number of more variable associations, respectively, as previously
448 described.[21] This approach allowed us to test for significant DNA methylation disruption across
449 single KEGG pathways and not over individual genes.

450 All statistical analyses were performed using the R statistical package version 3.2.0
451 (http://www.R-project.org).

452
453
454
455
456
457
458

459 **Acknowledgements**: This work was supported by the US National Institute of Environmental
460 Health Sciences (NIEHS) grants R01 ES015533, K01 ES017800, R01 ES016454, P30
461 ES000210, and P30 ES000002.

REFERENCES

1.Argos M. Arsenic Exposure and Epigenetic Alterations: Recent Findings Based on the Illumina 450K DNA Methylation Array. Current Environmental Health Reports 2015; 2:137-44.

485 2.Bailey KA SA, Tokar EJ, Graziano JH, Kim KW, Navasumrit P, Ruchirawat M, Thiantanawat
486 A, Suk WA, Fry RC. Mechanisms Underlying Latent Disease Risk Associated with Early-Life
487 Arsenic Exposure: Current Research Trends and Scientific Gaps. Environ Health Perspect 2015;
488 [Epub] DOI:10.1289/ehp.1409360.
489 3.Steinmaus C, Ferreccio C, Acevedo J, Yuan Y, Liaw J, Durán V, Cuevas S, García J, Meza R,
490 Valdés R. Increased lung and bladder cancer incidence in adults after in utero and early-life
491 arsenic exposure. Cancer Epidemiology Biomarkers & Prevention 2014; 23:1529-38.
492 4.Yuan Y, Marshall G, Ferreccio C, Steinmaus C, Selvin S, Liaw J, Bates MN, Smith AH. Acute
493 myocardial infarction mortality in comparison with lung and bladder cancer mortality in arsenic-
494 exposed region II of Chile from 1950 to 2000. American journal of epidemiology 2007;
495 166:1381-91.
496 5.Yuan Y, Marshall G, Ferreccio C, Steinmaus C, Liaw J, Bates M, Smith AH. Kidney cancer
497 mortality: fifty-year latency patterns related to arsenic exposure. Epidemiology 2010; 21:103-8.
498 6.Farzan SF, Karagas MR, Chen Y. In utero and early life arsenic exposure in relation to long-
499 term health and disease. Toxicology and applied pharmacology 2013; 272:384-90.
500 7.Srivastava S, D'Souza SE, Sen U, States JC. In utero arsenic exposure induces early onset of
501 atherosclerosis in ApoE−/− mice. Reproductive Toxicology 2007; 23:449-56.
502 8.Cardenas A, Smit E, Houseman EA, Kerkvliet NI, Bethel JW, Kile ML. Arsenic Exposure and
503 Prevalence of the Varicella Zoster Virus in the United States: NHANES (2003-2004 and 2009-
504 2010). Environ Health Perspect 2015; 123:590–6.
505 9.Farzan SF, Korrick S, Li Z, Enelow R, Gandolfi AJ, Madan J, Nadeau K, Karagas MR. In
506 utero arsenic exposure and infant infection in a United States cohort: A prospective study.
507 Environmental research 2013; 126:24-30.

508 10.Rahman A, Vahter M, Ekstrom E-C, Persson L-Å. Arsenic exposure in pregnancy increases
509 the risk of lower respiratory tract infection and diarrhea during infancy in Bangladesh.
510 Environmental health perspectives 2010; 119:719-24.
511 11.Quansah R, Armah FA, Essumang DK, Luginaah I, Clarke E, Marfoh K, Cobbina SJ,
512 Nketiah-Amponsah E, Namujju PB, Obiri S. Association of Arsenic with Adverse Pregnancy
513 Outcomes/Infant Mortality: A Systematic Review and Meta-Analysis. Environmental health
514 perspectives 2015; 123:412.
515 12.Hamadani J, Tofail F, Nermell B, Gardner R, Shiraji S, Bottai M, Arifeen S, Huda SN, Vahter
516 M. Critical windows of exposure for arsenic-associated impairment of cognitive function in pre-
517 school girls and boys: a population-based cohort study. International journal of epidemiology
518 2011; 40:1593-604.
519 13.Tanaka H, Tsukuma H, Oshima A. Long-term prospective study of 6104 survivors of arsenic
520 poisoning during infancy due to contaminated milk powder in 1955. Journal of Epidemiology
521 2010; 20:439.
522 14.Bailey KA, Fry RC. Arsenic-associated changes to the epigenome: what are the functional
523 consequences? Current Environmental Health Reports 2014; 1:22-34.
524 15.Marsit CJ. Influence of environmental exposure on human epigenetic regulation. The Journal
525 of experimental biology 2015; 218:71-9.
526 16.Waalkes MP, Liu J, Diwan BA. Transplacental arsenic carcinogenesis in mice. Toxicology
527 and applied pharmacology 2007; 222:271-80.
528 17.Waalkes MP, Qu W, Tokar EJ, Kissling GE, Dixon D. Lung tumors in mice induced by
529 "whole-life" inorganic arsenic exposure at human-relevant doses. Archives of toxicology 2014;
530 88:1619-29.

531 18.Xie Y, Liu J, Benbrahim-Tallaa L, Ward JM, Logsdon D, Diwan BA, Waalkes MP. Aberrant
532 DNA methylation and gene expression in livers of newborn mice transplacentally exposed to a
533 hepatocarcinogenic dose of inorganic arsenic. Toxicology 2007; 236:7-15.
534 19.Reichard JF, Puga A. Effects of arsenic exposure on DNA methylation and epigenetic gene
535 regulation. Epigenomics 2010; 2:87-104.
536 20.Koestler DC, Avissar-Whiting M, Houseman EA, Karagas MR, Marsit CJ. Differential DNA
537 methylation in umbilical cord blood of infants exposed to low levels of arsenic in utero.
538 Hnvironmental Health Perspectives 2013; 121:971–7.
539 21.Kile ML, Houseman EA, Baccarelli AA, Quamruzzaman Q, Rahman M, Mostofa G,
540 Cardenas A, Wright RO, Christiani DC. Effect of prenatal arsenic exposure on DNA methylation
541 and leukocyte subpopulations in cord blood. Epigenetics 2014; 9:774-82.
542 22.Rojas D, Rager JE, Smeester L, Bailey KA, Drobná Z, Rubio-Andrade M, Stýblo M, García-
543 Vargas G, Fry RC. Prenatal arsenic exposure and the epigenome: identifying sites of 5-
544 methylcytosine alterations that predict functional changes in gene expression in newborn cord
545 blood and subsequent birth outcomes. Toxicological Sciences 2015; 143:97-106.
546 23.Broberg K, Ahmed S, Engström K, Hossain M, Jurkovic Mlakar S, Bottai M, Grandér M,
547 Raqib R, Vahter M. Arsenic exposure in early pregnancy alters genome-wide DNA methylation
548 in cord blood, particularly in boys. Journal of developmental origins of health and disease 2014;
549 5:288-98.
550 24.Liu X, Zheng Y, Zhang W, Zhang X, LIoyd-Jones DM, Baccarelli AA, Ning H, Fornage M,
551 He K, Liu K. Blood methylomics in response to arsenic exposure in a low-exposed US
552 population. Journal of Exposure Science and Environmental Epidemiology 2014; 24:145-9.

553 25.Argos M, Chen L, Jasmine F, Tong L, Pierce BL, Roy S, Paul-Brutus R, Gamble MV, Harper
554 KN, Parvez F. Gene-specific differential DNA methylation and chronic arsenic exposure in an
555 epigenome-wide association study of adults in Bangladesh. Environ Health Perspect 2015;
556 123:64-71.
557 26.Koestler DC, Avissar-Whiting M, Houseman EA, Karagas MR, Marsit CJ. Differential DNA
558 Methylation in Umbilical Cord Blood of Infants Exposed to Low Levels of Arsenic in Utero.
559 Environmental health perspectives 2013; 121:971.
560 27.Seow WJ, Kile ML, Baccarelli AA, Pan WC, Byun HM, Mostofa G, Quamruzzaman Q,
561 Rahman M, Lin X, Christiani DC. Epigenome-wide DNA methylation changes with
562 development of arsenic-induced skin lesions in Bangladesh: A case–control follow-up study.
563 Environmental and molecular mutagenesis 2014; 55:449-56.
564 28.Yang T-Y, Hsu L-I, Chiu AW, Pu Y-S, Wang S-H, Liao Y-T, Wu M-M, Wang Y-H, Chang
565 C-H, Lee T-C. Comparison of genome-wide DNA methylation in urothelial carcinomas of
566 patients with and without arsenic exposure. Environmental research 2014; 128:57-63.
567 29.Rager JE, Tilley SK, Tulenko SE, Smeester L, Ray PD, Yosim A, Currier JM, Ishida MC,
568 González-Horta MdC, Sánchez-Ramírez B. Identification of Novel Gene Targets and Putative
569 Regulators of Arsenic-Associated DNA Methylation in Human Urothelial Cells and Bladder
570 Cancer. Chemical research in toxicology 2015; 28:1144-55.
571 30.Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nature
572 Reviews Genetics 2012; 13:484-92.
573 31.Castillo-Díaz SA, Garay-Sevilla ME, Hernández-González MA, Solís-Martínez MO, Zaina S.
574 Extensive demethylation of normally hypermethylated CpG islands occurs in human
575 atherosclerotic arteries. International journal of molecular medicine 2010; 26:691-700.

576 32.Hiltunen MO, Ylä-Herttuala S. DNA methylation, smooth muscle cells, and atherogenesis.
577 Arteriosclerosis, thrombosis, and vascular biology 2003; 23:1750-3.
578 33.Coit P, De Lott LB, Nan B, Elner VM, Sawalha AH. DNA methylation analysis of the
579 temporal artery microenvironment in giant cell arteritis. Annals of the rheumatic diseases
580 2015:annrheumdis-2014-207116.
581 34. Ehrlich M. DNA hypomethylation in cancer cells. Epigenomics 2009; 1:239-59.
582 35.Mass MJ, Wang L. Arsenic alters cytosine methylation patterns of the promoter of the tumor
583 suppressor gene p53 in human lung cells: a model for a mechanism of carcinogenesis. Mutation
584 Research/Reviews in Mutation Research 1997; 386:263-77.
585 36.Intarasunanont P, Navasumrit P, Waraprasit S, Chaisatra K, Suk WA, Mahidol C, Ruchirawat
586 M. Effects of arsenic exposure on DNA methylation in cord blood samples from newborn babies
587 and in a human lymphoblast cell line. Environ Health 2012; 11:31.
588 37.Tunster SJ, Jensen AB, John RM. Imprinted genes in mouse placental development and the
589 regulation of fetal energy stores. Reproduction 2013; 145:R117-R37.
590 38.Lesseur C, Paquette AG, Marsit CJ. Epigenetic regulation of infant neurobehavioral
591 outcomes. Medical epigenetics 2014; 2:71-9.
592 39.Oberlander TF, Weinberg J, Papsdorf M, Grunau R, Misri S, Devlin AM. Prenatal exposure
593 to maternal depression, neonatal methylation of human glucocorticoid receptor gene (NR3C1)
594 and infant cortisol stress responses. Epigenetics 2008; 3:97-106.
595 40.Conradt E, Fei M, LaGasse L, Tronick EZ, Guerin D, Gorman D, Marsit CJ, Lester B.
596 Prenatal predictors of infant self-regulation: The contributions of placental DNA methylation of
597 NR3C1 and neuroendocrine activity. Name: Frontiers in Behavioral Neuroscience 2015; 9:130.

598 41.Caldwell KE, Labrecque MT, Solomon BR, Ali A, Allan AM. Prenatal arsenic exposure
599 alters the programming of the glucocorticoid signaling system during embryonic development.
600 Neurotoxicology and teratology 2015; 47:66-79.
601 42.Schroeder DI, Blair JD, Lott P, Yu HOK, Hong D, Crary F, Ashwood P, Walker C, Korf I,
602 Robinson WP. The human placenta methylome. Proceedings of the national academy of sciences
603 2013; 110:6037-42.
604 43.Wahed M, Rahman M, Vahter M. Prevalence of arsenic exposure and skin lesions. A
605 population based survey in Matlab, Bangladesh. Journal of Epidemiology and Community health
606 2006:242-8.
607 44.Pilsner RJ, Liu X, Ahsan H, Ilievski V, Slavkovich V, Levy D, Factor-Litvak P, Graziano JH,
608 Gamble MV. Folate deficiency, hyperhomocysteinemia, low urinary creatinine, and
609 hypomethylation of leukocyte DNA are risk factors for arsenic-induced skin lesions.
610 Environmental health perspectives 2009; 117:254-60.
611 45.Yu H-S, Liao W-T, Chai C-Y. Arsenic carcinogenesis in the skin. Journal of biomedical
612 science 2006; 13:657-66.
613 46.Pan W-C, Seow WJ, Kile ML, Hoffman EB, Quamruzzaman Q, Rahman M, Mahiuddin G,
614 Mostofa G, Lu Q, Christiani DC. Association of low to moderate levels of arsenic exposure with
615 risk of type 2 diabetes in Bangladesh. American journal of epidemiology 2013; 178:1563-70.
616 47.Bailey KA, Wu MC, Ward WO, Smeester L, Rager JE, García-Vargas G, Del Razo LM,
617 Drobná Z, Stýblo M, Fry RC. Arsenic and the epigenome: interindividual differences in arsenic
618 metabolism related to distinct patterns of DNA methylation. Journal of biochemical and
619 molecular toxicology 2013; 27:106-15.

620 48.Wang Y. Vascular biology of the placenta. Colloquium Series on Integrated Systems
621 Physiology: from Molecule to Function: Morgan & Claypool Life Sciences, 2010:1-98.
622 49.Fukuchi Y, Nakajima H, Sugiyama D, Hirose I, Kitamura T, Tsuji K. Human placenta-
623 derived cells have mesenchymal stem/progenitor cell potential. Stem cells 2004; 22:649-58.
624 50.Cardenas A, Koestler DC, Houseman EA, Jackson BP, Kile ML, Karagas MR, Marsit CJ.
625 Differential DNA Methylation in Umbilical Cord Blood of Infants Exposed to Mercury and
626 Arsenic in utero. Epigenetics 2015; (10)6:508-15.
627 51.Kile ML, Hoffman E, Hsueh Y-M, Afroz S, Quamruzzaman Q, Rahman M, Mahiuddin G,
628 Ryan L, Christiani DC. Variability in biomarkers of arsenic exposure and metabolism in adults
629 over time. Environ Health Perspect 2009; 117:455-60.
630 52.Kile ML, Houseman EA, Rodrigues E, Smith TJ, Quamruzzaman Q, Rahman M, Mahiuddin
631 G, Su L, Christiani DC. Toenail arsenic concentrations, GSTT1 gene polymorphisms, and
632 arsenic exposure from drinking water. Cancer Epidemiology Biomarkers & Prevention 2005;
633 14:2419-26.
634 53.Concha G, Vogler G, Lezcano D, Nermell B, Vahter M. Exposure to inorganic arsenic
635 metabolites during early human development. Toxicological Sciences 1998; 44:185-90.
636 54.Creed J, Brockhoff C, Martin T. Method 200.8: Determination of trace elements in waters and
637 wastes by inductively-coupled plasma-mass spectrometry. Environmental Monitoring Systems
638 Laboratory, Office of Research and Development, US Environmental Protection Agency,
639 Cincinnati, OH, Rev 1994; 5.
640 55.Fortin J-P, Labbe A, Lemire M, Zanke BW, Hudson TJ, Fertig EJ, Greenwood CM, Hansen
641 KD. Functional normalization of 450k methylation array data improves replication in large
642 cancer studies. Genome biology 2014; 15:503.

643 56.Chen Y-a, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, Gallinger S,
644 Hudson TJ, Weksberg R. Discovery of cross-reactive probes and polymorphic CpGs in the
645 Illumina Infinium HumanMethylation450 microarray. Epigenetics 2013; 8:203-9.
646 57.Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, Beck S. A
647 beta-mixture quantile normalization method for correcting probe design bias in Illumina
648 Infinium 450 k DNA methylation data. Bioinformatics 2013; 29:189-96.
649 58.Du P, Zhang X, Huang C-C, Jafari N, Kibbe WA, Hou L, Lin SM. Comparison of Beta-value
650 and M-value methods for quantifying methylation levels by microarray analysis. BMC
651 bioinformatics 2010; 11:587.
652 59.Houseman EA, Molitor J, Marsit CJ. Reference-free cell mixture adjustments in analysis of
653 DNA methylation data. Bioinformatics 2014; 30:1431-9.
654
655
656
657
658
659
660
661
662
663 **Figure Legends**

664 **Figure 1.** Locus-by-locus epigenome-wide analysis for umbilical artery: volcano plots for the
665 association between log_2-transformed maternal drinking water arsenic (A) unadjusted for cellular

666 heterogeneity and (B) adjusting for cellular heterogeneity using the Houseman reference-free
667 method. (C) Distribution of differentially methylated loci (q-value<0.05) relative to CpG islands
668 for the unadjusted cell mixture analysis.
669
670 **Figure 2.** Locus-by-locus epigenome-wide analysis for placenta: volcano plots for the
671 association between $\log_2$-transformed maternal drinking water arsenic (A) unadjusted for cellular
672 heterogeneity and (B) adjusting for cellular heterogeneity using the Houseman reference-free
673 method. (C) Distribution of differentially methylated loci (nominal $P<1\times10^{-4}$) relative to CpG
674 islands for the unadjusted cell mixture analysis.
675
676 **Figure 3.** Locus-by-locus epigenome-wide analysis for HUVEC: volcano plots for the
677 association between $\log_2$-transformed maternal drinking water arsenic: (A) unadjusted for
678 cellular heterogeneity and (B) adjusting for cellular heterogeneity using the Houseman reference-
679 free method. (C) Distribution of differentially methylated loci (nominal p-value$<1\times10^{-4}$) relative
680 to CpG islands for the unadjusted cell mixture analysis.
681
682
683
684
685
686 **Tables**
687 **Table 1.** Sample characteristics for the 52 mother-infant pairs eligible for the analysis

Sample characteristics	Mean±SD	Range
Drinking water arsenic at recruitment (μg/L)	63.7±116.5	<1 - 510

Gestational age at recruitment (weeks)	12.2±2.5	6 - 16
Gestational age at delivery (weeks)	37.6±2.1	33 - 41
Birth weight (grams)	2923±372	2080 - 4050
Gender	**N (%)**	
Male	33 (63.5 %)	
Female	19 (36.5%)	
Number of samples available by tissue	**n**	**CpG loci analyzed**
HUVEC	52	347,650
Artery	45	374,320
Placenta	37	365,994

688