Supplementary Information

Text S1. Identifying genes with differential methylation and differential expression associated with prenatal arsenic exposure from the BEAR cohort [36].

1. Identifying Genes with Differential Expression Associated with Prenatal Arsenic Exposure

In our previous investigation (Rager *et al*. 2014) [36], the Biomarkers of Exposure to ARsenic (BEAR) pregnancy cohort in Gómez Palacio, Mexico, was established to better understand the impacts of prenatal exposure to arsenic (As). In this study, concentrations of total arsenic in maternal urine (U-tAs) were related to genome-wide mRNA expression levels in cord blood from 38 newborns. U-tAs was defined as the sum of the levels of iAs and its monomethylated (MMAs) and dimethylated (DMAs) metabolites. A total of 334 genes were identified as differentially expressed and associated with U-tAs, of which, 110 displayed increased expression levels and 224 displayed decreased expression levels (Rager *et al*. 2014) [36].

2. Identifying Genes with Differential Methylation Associated with Prenatal Arsenic Exposure

In order to identify which of the As-associated genes identified by Rager *et al*. (2014) [36] may be regulated by DNA methylation mediation, the same cord blood samples (n = 38) from the BEAR cohort were also assessed for genome-wide DNA methylation profiles and related to maternal U-tAs. The mean concentration of U-tAs in this cohort was 73.87 µg/L (median = 32.57 µg/L), as measured using methods previously described (Rager *et al*. 2014).

Cord blood samples were collected from newborns immediately following delivery, using Paxgene Blood DNA tubes and extracted using the PaxGene Blood DNA kit (Qiagen, Valencia, CA, USA) per manufacturer's specifications. DNA methylation was assessed using the Illumina HumanMethylation450 BeadChip (Illumina, Inc., San Diego, CA, USA). This platform assesses the methylation levels of a total of 486,428 individual probes each measuring the methylation levels at a single CpG site. Isolated DNA was bisulfite-converted using the EZ DNA methylation kit (Zymo Research, Irvine, CA, USA) and then converted DNA was hybridized onto the array. Microarray data was collected at Expression Analysis, Inc. (Durham, NC, USA; www.expressionanalysis.com). Methylation levels were calculated and expressed as β values, defined as the ratio of the methylated probe intensity to the intensity from both methylated and unmethylated probes, where a β value of 1 indicates the highest level of methylation, as in Joubert *et al*. (2012) [78]. Methylation data were normalized using a quantile-based methodology (Bolstad *et al*. 2003) [79]. Probes with high detection p-values ($p > 0.05$) were treated as unreliable and removed from analysis (n = 1761), as per manufacturer recommendation. Probes that represent known single nucleotide polymorphisms (SNPs) were also removed (n = 59,732), leaving a total of 424,935 probes for further analyses.

Sites of differential DNA methylation associated with U-tAs were identified using a multi-variable regression model, where the dependent variable was DNA methylation and the independent variable was U-tAs. The regression model included covariates that were associated with both exposure and outcome using a bivariate analysis ($p < 0.05$) or are plausibly related to cord blood DNA methylation levels, specifically: Newborn gender (binary variable) and birth weight/gestational age (continuous variable). Significant probes were identified based on a false discovery corrected q-value <0.05. A total of

4771 probes, representing 2919 genes, displayed differential methylation associated with U-tAs. Of those probes, 34% (n = 1621) displayed hypo-methylation as U-tAs levels increased and 66% (n = 3150) displayed hyper-methylation as U-tAs levels increased.

3. Identifying Genes with Differential Methylation and Differential Expression Associated with Prenatal Arsenic Exposure

In order to generate a list of genes with differential methylation and differential expression associated with As, the list of genes with differential methylation was compared against the list of genes previously identified with differential expression associated with U-tAs (Rager *et al*. 2014) [36]. Of the 334 genes that showed differential gene expression, 269 were represented in the platform used to assess DNA methylation and could be matched to their corresponding methylation β value. A total of 20% (n = 54 genes, represented by 74 Illumina probesets) of the differentially expressed genes showed at least one site of differential methylation associated with U-tAs (Additional File 1, Table S8).

References

78. Joubert, B.R.; Håberg, S.E.; Nilsen, R.M.; Wang, X.; Vollset, S.E.; Murphy, S.K.; Huang, Z.; Hoyo, C.; Midttun, Ø.; Cupul-Uicab, L.A.; *et al*. 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environ. Health Perspect*. **2012**, *120*, 1425–1431.
79. Bolstad, B.M.; Irizarry, R.A.; Astrand, M.; Speed, T.P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **2003**, *19*, 185–193.