Published in final edited form as:
J Expo Sci Environ Epidemiol. 2014 ; 24(2): 145–149. doi:10.1038/jes.2013.89.

BLOOLD METHYLMICS IN RESPONSE TO ARSENIC EXPOSURE IN A LOW–EXPOSED US POPULATION

Xin Liu, MD, PhD[1,2], **Yinan Zheng, BS**[3], **Wei Zhang, PhD**[4], **Xiao Zhang, PhD**[2], **Hongyan Ning, MD**[2], **Kiang Liu, PhD**[2], **Donald M Lloyd-Jones, MD**[2], **Myriam Fornage, PhD**[5], **Ka He, MD, MPH, ScD**[6], and **Lifang Hou, MD, PhD**[2,7,*]

[1]Mary Ann and J. Milburn Smith Child Health Research Program, Ann & Robert H. Lurie Children's Hospital of Chicago Research Center, Department of Pediatrics, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

[2]Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

[3]Driskill Graduate Program (DGP) in Life Sciences, Northwestern University Biomedical Informatics Center (NUBIC), NUCATS, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

[4]Department of Pediatrics, University of Illinois, Chicago, IL, USA

[5]Institute of Molecular Medicine and Division of Epidemiology, School of Public Health, University of Texas Sciences Center, Houston, TX, USA

[6]Department of Epidemiology and Biostatistics, Indiana University School of Public Health, Bloomington, IN, USA

[7]The Robert H. Lurie Comprehensive Cancer Center, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

Abstract

Exposure to arsenic (As) has been associated with cancers, CVD, and neurological disorder. To explore the possible underlying epigenetic mechanisms, a genome-wide study was conducted in low exposed healthy individuals. This study was nested within a prospective study of Coronary Artery Risk Development in Young Adults (CARDIA) by randomly selecting 46 non-smoker and non-diabetic White participants with low (N=23) and high (N=23) As exposure. based on toenail total As measures at examination year 2. We conducted methylomic profiling of white blood cell DNA collected at examination year 15 using the Illumina HumanMethylation450 BeadChip. Multivariate linear regression models were fitted to evaluate the associations between As exposure status and DNA methylation levels at each CpG site. We identified 29 CpG sites with methylation levels associated with As exposure status at a nominal p-value less than 0.0001. Some genes are

[*]**Correspondence and reprint requests should be addressed to:** Lifang Hou MD, PhD, Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, 680 N Lake Shore Drive, Suite 1400, Chicago, IL 60611 USA, Telephone: 312-503-4798, Fax: (312) 908-9588, l-hou@northwestern.edu.

Conflict of interest: The authors declare no conflict of interest

Supplementary information is available at *Journal of Exposure Science and Environmental Epidemiology*' website.

known to be involved in cancers, CVD, and neurological disorder. Pathway analyses further revealed several canonical pathways relevant to the etiology of As-associated diseases. We demonstrated that As exposure is prospectively associated with DNA methylation levels in a number of genes implicated in As-associated diseases. Further studies are required for elucidating the role of epigenetic alterations in the pathogenesis of these diseases.

Keywords

arsenic exposure; methylomic profiling; prospective association

INTRODUCTION

Arsenic (As) is a naturally occurring ubiquitous element found in foods and the environment, such as water, soil, and air. Chronic As exposure is a worldwide health problem. [1] It was classified as a Group 1 carcinogen in 1987 by the International Agency for Research on Cancer (IARC). [2] Accumulative experimental data and epidemiological evidence indicate that As exposure is associated with various cancers [3] such as lung, bladder, kidney, liver and skin cancers, [4–6] and other chronic diseases such as cardiovascular disease (CVD) [7–9] and neurological disorders. [10, 11]

The mechanisms linking As with disease still remain largely unknown. Recent evidence suggests that environmental chemicals may cause diseases via epigenetic mechanisms that regulate gene expression without changing DNA sequencing, such as DNA methylation (i.e., cytosine modification at CpG dinucleotides), the most well-studied epigenetic mechanism in the etiology of disease. [12, 13] Although most studies have been conducted using tissue samples, [14, 15] evidence for As exposure and DNA methylation in blood DNA has also begun to accumulate. [16–19] A global dose-dependent hypermethylation of blood DNA was observed in a Bangladeshi population with chronic As exposure to contaminated drinking water. [16] In a cross-sectional study of Indian individuals, As levels in contaminated water was also associated with global DNA hypermethylation in blood mononuclear cells. [17] Chanda et al. observed blood DNA hypermethylation in the promoter of *p53* and *p16* in As-exposed Indian individuals. [18] Consistently, in a Chinese population, increased DNA methylation in the *p16* promoter was observed in arseniasis patients compared with people without a history of As exposure. [20] However, these previous studies have been limited to the evaluation of global methylation markers or methylation alterations in a small group of genes. A recent genome-wide DNA methylation analysis reported that 183 genes were epigenetically-modified in blood DNA in 16 female Mexicans aged 12–59 years, with half showing signs of arsenicosis. [19] To date, there has been no prospective genome-wide study in a U.S. population to evaluate the effect of As exposure on DNA methylation alterations.

In this study, we performed a prospective genome-wide examination to evaluate whether exposure to As induces DNA methylation alterations in 46 apparently young, middle-aged healthy non-smoker, non-diabetic, White individuals derived from the Coronary Artery Risk Development in Young Adults (CARDIA) study, a large prospective study of young adults.

MATERIAL AND METHODS

Study population

CARDIA is a multi-center perspective study of risk factors for coronary artery disease (CAD) development in young adults free from CVD (N=5,115) and aged 18–30 years at baseline (1985–6). Participants have undergone eight examinations to date, including a baseline examination at year (Y) 0 and follow-up examinations at Y2, 5, 7, 10, 15, 20, and 25, with a 72% examination rate at Y20 (2005–6). A detailed description of the study design, sampling, and response rates was previously published. [21] Institutional Review Boards at each study site reviewed the protocol and procedures, and approved the research. All participants provided written informed consent. The present study included 46 White participants who had available data for Y2 toenail total As level and blood DNA at Y15.

Arsenic Exposure Measurement in CARDIA and Study Subject Selection

At examination at Y2, CARDIA participants were mailed the instructions and materials for collecting toenail samples. Toenail clippings were collected and As level was measured in 4,362 CARDIA participants by Neutron Activation Analysis (NAA) [22] at the University of Missouri Research Reactor. 46 White non-smoker, non-diabetic healthy age- and sex-matched study subjects (23 high- and 23 low- exposed) for the present study were randomly selected from the highest and lowest quartile exposure groups based on the cutoff-points of <0.0649 (Q1: low As exposure) and ≥0.1442 (Q4: high As exposure).

Genome-wide examination of DNA methylation alterations

We performed genome-wide DNA methylation examination in 46 white blood cell (WBC) DNA samples that passed the DNA quality test for our assay using the Illumina Infinium Human Methylation450 BeadChip, which targets ~486,000 CpG sites. A 500ng DNA sample from each selected CARDIA participant was used to perform bisulfite conversion followed by Illumina's protocol for methylation profiling. BeadChips were scanned with an Illumina iScan and then analyzed using the Illumina GenomeStudio software. All experiments were conducted following the manufacturer's protocols in the Genomic Core Facility of the Center for Genetic Medicine at Northwestern University. For the purpose of quality control (QC), in addition to Illumina's build-in QC, we included commercially available known unmethylated (normal B-lymphocytes (NA10923 from Coriell Institute), Camden, NJ) and methylated (colon cancer cells (ATCC: HTB-38), Manassas, VA) control samples in each run, as previously described. [23]

Bioinformatic and statistical analysis

We dropped three samples that were considered to be obvious outliers based on the principle component analysis (PCA) plot (Figure S1) generated from Partek Genomics Suite (Partek GS) (http://www.partek.com/partekgs), and focused on the remaining 43 samples for the statistical analyses. We excluded the CpG probes that are ambiguously mapped to the human genome (hg19). A total of 340,658 probes passed the analysis using Bowtie. [24] To avoid potential bias due to genetic polymorphisms, we also filtered 5,804 CpG probes with the presence of common single nucleotide polymorphisms (SNPs) (i.e., minor allele

frequency (MAF) ≥ 0.01) within the range of 20 base pairs of the CpG sites based on the HapMap European origin populations (CEU: Caucasian residents from Utah, USA) in the dbSNP database (v135). To reduce the effects of differential methylation between males and females on the sex chromosomes, 7,953 CpG sites on chromosome X and Y were excluded. Finally, a total of 326,901CpG sites on autosomes were tested with regards to As exposure (Figure 1). Note that all of the 43 samples had more than 90% CpG sites with detection p-values of less than 10E-05, and none of the 326,901 CpG sites had more than 10% of samples with detection p-values greater than 10E-05.

We then tested the associations between As exposure and DNA methylation at each CpG site. Methylation values of the filtered CpG sites were first transformed into M-values as described previously, [23] and then quantile-normalized. A generalized linear regression model was fitted with adjustment for age and gender, two known factors that influence DNA methylation patterns. [25–27] Moderated t-statistics were computed by empirical Bayes shrinkage of the standard errors. [28] The same procedure was conducted across all tested CpG sites. We then annotated the significant As-associated CpG sites to the corresponding genes based on an Illumina-designed document (http://www.illumina.com/). For the identified differentially methylated genes, we searched for their functions and related disease involvements in the Ingenuity Knowledge Base (http://www.ingenuity.com/products/ pathways_ analysis.html). Heat maps were plotted to visualize the different methylation levels across As exposure status using hierarchical clustering.

Pathway analyses were conducted using GSA-SNP software. [29] Specifically, for each gene, we assigned the 2nd minimum p-value from the association tests for the contained CpG sites to reduce the influence of a few highly significant findings possibly occurring by chance in the genome-wide study, and then performed enrichment analyses using three major canonical pathway (CP) databases (KEGG, BIOCARTA, and REACTOME) enclosed within the Molecular Signatures Database (MSigDB) (http://www.broadinstitute.org/gsea/msigdb/index.jsp). Gene sets with a Benjamini and Hochberg False Discovery Rate (FDR) [30] less than 10E-06 were reported. All of the analyses were carried out using R package and SAS software (version9.3; SAS Institute, Inc., Cary, North Carolina).

RESULTS

Table 1 shows the characteristics of the study subjects with low (N=21) and high (N=22) As exposure status. The As high-exposure group consisted of more alcohol drinkers (81.8%) than the low-exposure group (57.1%) (p=0.10). There was no significant difference in age, education, marriage status or body mass index (BMI) between the two groups.

We identified 29 CpG sites with methylation levels associated with As exposure status at a nominal p-value less than 0.0001. The percentage difference of mean methylation levels between the As high and low exposure groups ranged from 1% to 31% (Table 2). An overview of the sample relations based on a heatmap of these top CpG sites shows distinct DNA methylation patterns between subjects with high and low As exposure (Figure 2), with over half of the top CpG sites showing elevated methylation levels in high As exposure groups compared to low exposure groups. Genes annotated by these 29 top hits are involved

in three major classes of diseases, including cancers (*HAPLN1BSGFANCGCXCR5AKAP8LADARB2EVLORAOV1SKI*), neurological disorders (*SREBF2NR1H2MS4A4AEGLN1FAAHMAG*), and cardiovascular diseases (*HAPLN1LMF1ADARB2, SREBF2*).

Table 3 lists the significant As exposure-associated canonical pathways (i.e., Benjamini and Hochberg FDR <10E-06). The gene-set based analysis identified As exposure-associated pathways that are involved in neurological disease (i.e., formation of neuronal network and nerve growth factor signaling), cancer (i.e., Wnt signaling pathway), both cancer and neurological disease (i.e., focal adhesion and cell adhesion molecules), and CVD (i.e., viral myocarditis), which are consistent with findings from the CpG site-based analysis listed in Table 2. Detailed information on these significant canonical pathways is provided in Table S1.

DISCUSSION

This is the first prospective genome-wide methylomic study linking DNA methylation markers with As exposure status assessed in toenails collected 13 years ago. We found that As-exposure is prospectively associated with DNA methylation levels in a number of genes and canonical pathways known to be involved in As-associated diseases. These findings, once confirmed by a large independent human sample, should stimulate future investigations to better understand the underlying molecular mechanisms of how As exposure may contribute to the pathogenesis of As-associated diseases via DNA methylation changes.

As exposure has repeatedly been associated with a variety of common diseases, including lung, bladder, kidney, liver and skin cancers, [4–6] CVD, [7–9] and neurological disease. [10, 11] The genes that we identified to be associated with As exposure also have been previously implicated in such diseases. For example, *CXCR5*, a G protein-coupled seven transmembrane receptor for chemokine C-X-C motif chemokine 13, [31] has been demonstrated to be involved in tumor cell adhesion and migration. [32] *SREBF2* encodes a ubiquitously expressed transcription factor that controls cellular cholesterol metabolism. [33] In human atherosclerotic tissues, Fan et al. observed a significant down-regulation of *SREBF2* in atherosclerotic carotid plaques, suggesting that it might be implicated in the progression of atherosclerosis. Recently, Guo et al. identified *SREBF2* to be associated with BMI, an established risk factor for CVD and cancers. [34] In addition, the significant association between *SREBF2* gene polymorphism and schizophrenia indicates that SREBF2-controlled cholesterol biosynthesis is one of the etiological mechanisms for the development of this psychiatry disorder. [35] As such, varied As-associated DNA methylation levels in *SREBF2* might be one of the potential molecular mechanisms shared by several common diseases.

The exact mechanisms via which As may cause methylomic changes are largely unknown. Oxidative stress has been proposed as a link between As exposure and chronic disease. This cellular process has also been shown to induce altered DNA methylation patterns. A recent study demonstrated that reactive oxygen species (ROS) production can alter the expression of genes belonging to DNA methylation machinery. [36] In addition, inorganic As is

enzymatically methylated for detoxification using S-adenosyl methionine (SAM) in the process. [37] Therefore, SAM insufficiency might be another possible mechanism underlying As-induced DNA methylation, given that both As metabolism and DNA methylation need SAM as the methyl donor. Furthermore, As exposure often occurs in relatively resource-poor populations with low dietary intake of methionine, an essential amino acid required for SAM synthesis. [38] In addition, As has also been shown to decrease DNA methyltransferase (*DNMT*) gene expression [14] and enzyme activities. [3] All of these As-induced cellular processes may independently or cooperatively interact with each other to contribute to related DNA methylation changes. Different genes may behave differently with respect to As exposure-associated diseases, and As may cause hypo- or hyper-methylation in each individual gene depending on the role of the gene in cancer and other disease development.

Nevertheless, our findings should be interpreted with caution because of our small sample size and multiple testing issues. With a total of 43 study subjects, none of the top 29 CpG sites remained statistically significant after correction for multiple testing. However, the conclusion that the development of multiple diseases might be influenced by exposure to high As levels in general is less likely to be biased, because, first, at least four genes were indicated to be associated with cancers, CVD, or neurological disorders based on the site-based analyses (Table 2); and second, these inferences were strengthened by substantially significant findings from the pathway-based analyses (Table 3). Taken together, the DNA methylation alterations that we observed in this study may play a role in As-related diseases.

Overall, we demonstrated that As exposure is prospectively associated with DNA methylation levels in a number of genes and canonical pathways that have been implicated in As-associated diseases. Further studies in larger human samples and aiming at determining the impact of these methylation changes on gene expression are required for elucidating the role of these epigenetic changes in the pathogenesis of As-associated diseases.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The Coronary Artery Risk Development in Young Adults Study (CARDIA) is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with the University of Alabama at Birmingham (N01-HC95095 & N01-HC48047), University of Minnesota (N01-HC48048), Northwestern University (N01-HC48049), Kaiser Foundation Research Institute (N01-HC48050), and George M Eisenberg Foundation. This manuscript has been reviewed by CARDIA for scientific content and consistency of data interpretation with previous CARDIA publications.

References

1. Nuntharatanapong N, Chen K, Sinhaseni P, Keaney JF Jr. EGF receptor-dependent JNK activation is involved in arsenite-induced p21Cip1/Waf1 upregulation and endothelial apoptosis. Am J Physiol Heart Circ Physiol. 2005 Jul; 289(1):H99–H107. [PubMed: 15734884]

2. Overall Evaluations of Carcinogenicity: An Updating of IARC Monographs. IARC Monogr Eval Carcinog Risks Hum. 1987; 1–42(Suppl 7):1–440.
3. Ren X, McHale CM, Skibola CF, Smith AH, Smith MT, Zhang L. An emerging role for epigenetic dysregulation in arsenic toxicity and carcinogenesis. Environ Health Perspect. 2011 Jan; 119(1):11–19. [PubMed: 20682481]
4. Smith AH, Steinmaus CM. Health effects of arsenic and chromium in drinking water: recent human findings. Annu Rev Public Health. 2009 Apr 29.30:107–122. [PubMed: 19012537]
5. Celik I, Gallicchio L, Boyd K, Lam TK, Matanoski G, Tao X, et al. Arsenic in drinking water and lung cancer: a systematic review. Environ Res. 2008 Sep; 108(1):48–55. [PubMed: 18511031]
6. Baastrup R, Sorensen M, Balstrom T, Frederiksen K, Larsen CL, Tjonneland A, et al. Arsenic in drinking-water and risk for cancer in Denmark. Environ Health Perspect. 2008 Feb; 116(2):231–237. [PubMed: 18288323]
7. Zierold KM, Knobeloch L, Anderson H. Prevalence of chronic diseases in adults exposed to arsenic-contaminated drinking water. Am J Public Health. 2004 Nov; 94(11):1936–1937. [PubMed: 15514231]
8. Meliker JR, Wahl RL, Cameron LL, Nriagu JO. Arsenic in drinking water and cerebrovascular disease, diabetes mellitus, and kidney disease in Michigan: a standardized mortality ratio analysis. Environ Health. 2007; 6:4. [PubMed: 17274811]
9. Lewis DR, Southwick JW, Ouellet-Hellstrom R, Rench J, Calderon RL. Drinking water arsenic in Utah: A cohort mortality study. Environ Health Perspect. 1999 May; 107(5):359–365. [PubMed: 10210691]
10. Rahman MM, Chowdhury UK, Mukherjee SC, Mondal BK, Paul K, Lodh D, et al. Chronic arsenic toxicity in Bangladesh and West Bengal, India--a review and commentary. J Toxicol Clin Toxicol. 2001; 39(7):683–700. [PubMed: 11778666]
11. Gong G, O'Bryant SE. The Arsenic Exposure Hypothesis for Alzheimer Disease. Alzheimer Dis Assoc Disord. 2010 May 13.
12. Hou LF, Zhang X, Wang D, Baccarelli A. Environmental chemical exposures and human epigenetics. International Journal of Epidemiology. 2012 Feb; 41(1):79–105. [PubMed: 22253299]
13. Hou LF, Wang D, Baccarelli A. Environmental chemicals and microRNAs. Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis. 2011 Sep 1; 714(1–2):105–112. [PubMed: 21609724]
14. Reichard JF, Schnekenburger M, Puga A. Long term low-dose arsenic exposure induces loss of DNA methylation. Biochemical and biophysical research communications. 2007 Jan 5; 352(1): 188–192. [PubMed: 17107663]
15. Sciandrello G, Caradonna F, Mauro M, Barbata G. Arsenic-induced DNA hypomethylation affects chromosomal instability in mammalian cells. Carcinogenesis. 2004 Mar; 25(3):413–417. [PubMed: 14633664]
16. Pilsner JR, Liu X, Ahsan H, Ilievski V, Slavkovich V, Levy D, et al. Genomic methylation of peripheral blood leukocyte DNA: influences of arsenic and folate in Bangladeshi adults. Am J Clin Nutr. 2007 Oct; 86(4):1179–1186. [PubMed: 17921400]
17. Majumdar S, Chanda S, Ganguli B, Mazumder DN, Lahiri S, Dasgupta UB. Arsenic exposure induces genomic hypermethylation. Environ Toxicol. 2010 Jun; 25(3):315–318. [PubMed: 19437452]
18. Chanda S, Dasgupta UB, Guhamazumder D, Gupta M, Chaudhuri U, Lahiri S, et al. DNA hypermethylation of promoter of gene p53 and p16 in arsenic-exposed people with and without malignancy. Toxicol Sci. 2006 Feb; 89(2):431–437. [PubMed: 16251483]
19. Smeester L, Rager JE, Bailey KA, Guan X, Smith N, Garcia-Vargas G, et al. Epigenetic changes in individuals with arsenicosis. Chem Res Toxicol. 2011 Feb 18; 24(2):165–167. [PubMed: 21291286]
20. Zhang AH, Bin HH, Pan XL, Xi XG. Analysis of p16 gene mutation, deletion and methylation in patients with arseniasis produced by indoor unventilated-stove coal usage in Guizhou, China. J Toxicol Environ Health A. 2007 Jun; 70(11):970–975. [PubMed: 17479413]

21. Friedman GD, Cutter GR, Donahue RP, Hughes GH, Hulley SB, Jacobs DR Jr, et al. CARDIA: study design, recruitment, and some characteristics of the examined subjects. Journal of clinical epidemiology. 1988; 41(11):1105–1116. [PubMed: 3204420]
22. Nichols TA, Morris JS, Mason MM, Spate VL, Baskett CK, Cheng TP, et al. The study of human nails as an intake monitor for arsenic using neutron activation analysis. Journal of Radioanalytical and Nuclear Chemistry. 1998 Oct; 236(1–2):51–56.
23. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics. 2010 Nov 30.11(1):587. [PubMed: 21118553]
24. Zhang X, Mu W, Zhang W. On the analysis of the illumina 450k array data: probes ambiguously mapped to the human genome. Front Genet. 2012; 3:73. [PubMed: 22586432]
25. Boks MP, Derks EM, Weisenberger DJ, Strengman E, Janson E, Sommer IE, et al. The relationship of DNA methylation with age, gender and genotype in twins and healthy controls. PloS one. 2009; 4(8):e6767. [PubMed: 19774229]
26. El-Maarri O, Becker T, Junen J, Manzoor SS, Diaz-Lacava A, Schwaab R, et al. Gender specific differences in levels of DNA methylation at selected loci from human total blood: a tendency toward higher methylation levels in males. Human genetics. 2007 Dec; 122(5):505–514. [PubMed: 17851693]
27. Liu J, Morgan M, Hutchison K, Calhoun VD. A study of the influence of sex on genome wide methylation. PloS one. 2010; 5(4):e10028. [PubMed: 20386599]
28. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol. 2004; 3 Article3.
29. Abdallah HE, Gouider E, Stambouli N, Ben Amor M, Jlizi A, Belhedi N, et al. Structural analysis of two novel mutations in MCFD2 gene causing combined coagulation factors V and VIII deficiency. Blood Cells Mol Dis. 2010 Mar-Apr;44(2):120–123. [PubMed: 20004600]
30. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society B (Methodological). 1995; 57:289–300.
31. Dobner T, Wolf I, Emrich T, Lipp M. Differentiation-specific expression of a novel G protein-coupled receptor from Burkitt's lymphoma. Eur J Immunol. 1992 Nov; 22(11):2795–2799. [PubMed: 1425907]
32. Razis E, Kalogeras KT, Kotoula V, Eleftheraki AG, Nikitas N, Kronenwett R, et al. Improved outcome of high-risk early HER2 positive breast cancer with high CXCL13-CXCR5 messenger RNA expression. Clin Breast Cancer. 2012 Jun; 12(3):183–193. [PubMed: 22607768]
33. Horton JD, Shimomura I, Brown MS, Hammer RE, Goldstein JL, Shimano H. Activation of cholesterol synthesis in preference to fatty acid synthesis in liver and adipose tissue of transgenic mice overproducing sterol regulatory element-binding protein-2. J Clin Invest. 1998 Jun 1; 101(11):2331–2339. [PubMed: 9616204]
34. Guo Y, Lanktree MB, Taylor KC, Hakonarson H, Lange LA, Keating BJ. Gene-centric meta-analyses of 108 912 individuals confirm known body mass index loci and reveal three novel signals. Hum Mol Genet. 2013 Jan 1; 22(1):184–201. [PubMed: 23001569]
35. Le Hellard S, Muhleisen TW, Djurovic S, Ferno J, Ouriaghi Z, Mattheisen M, et al. Polymorphisms in SREBF1 and SREBF2, two antipsychotic-activated transcription factors controlling cellular lipogenesis, are associated with schizophrenia in German and Scandinavian samples. Molecular psychiatry. 2010 May; 15(5):463–472. [PubMed: 18936756]
36. Fratelli M, Goodwin LO, Orom UA, Lombardi S, Tonelli R, Mengozzi M, et al. Gene expression profiling reveals a signaling role of glutathione in redox regulation. Proc Natl Acad Sci U S A. 2005 Sep ;102(39):13998–14003. [PubMed: 16172407]
37. Aposhian HV. Enzymatic methylation of arsenic species and other new approaches to arsenic toxicity. Annu Rev Pharmacol Toxicol. 1997; 37:397–419. [PubMed: 9131259]
38. Anetor JI, Wanibuchi H, Fukushima S. Arsenic exposure and its health effects and risk of cancer in developing countries: micronutrients as host defence. Asian Pac J Cancer Prev. 2007 Jan-Mar; 8(1):13–23. [PubMed: 17477765]

39. Song H, Tong D, Cha Z, Bai J. C-X-C chemokine receptor type 5 gene polymorphisms are associated with non-Hodgkin lymphoma. Molecular biology reports. 2012 Sep; 39(9):8629–8635. [PubMed: 22707196]

40. Oguro R, Kamide K, Katsuya T, Akasaka H, Sugimoto K, Congrains A, et al. A single nucleotide polymorphism of the adenosine deaminase, RNA-specific gene is associated with the serum triglyceride level, abdominal circumference, and serum adiponectin concentration. Experimental gerontology. 2012 Feb; 47(2):183–187. [PubMed: 22210125]

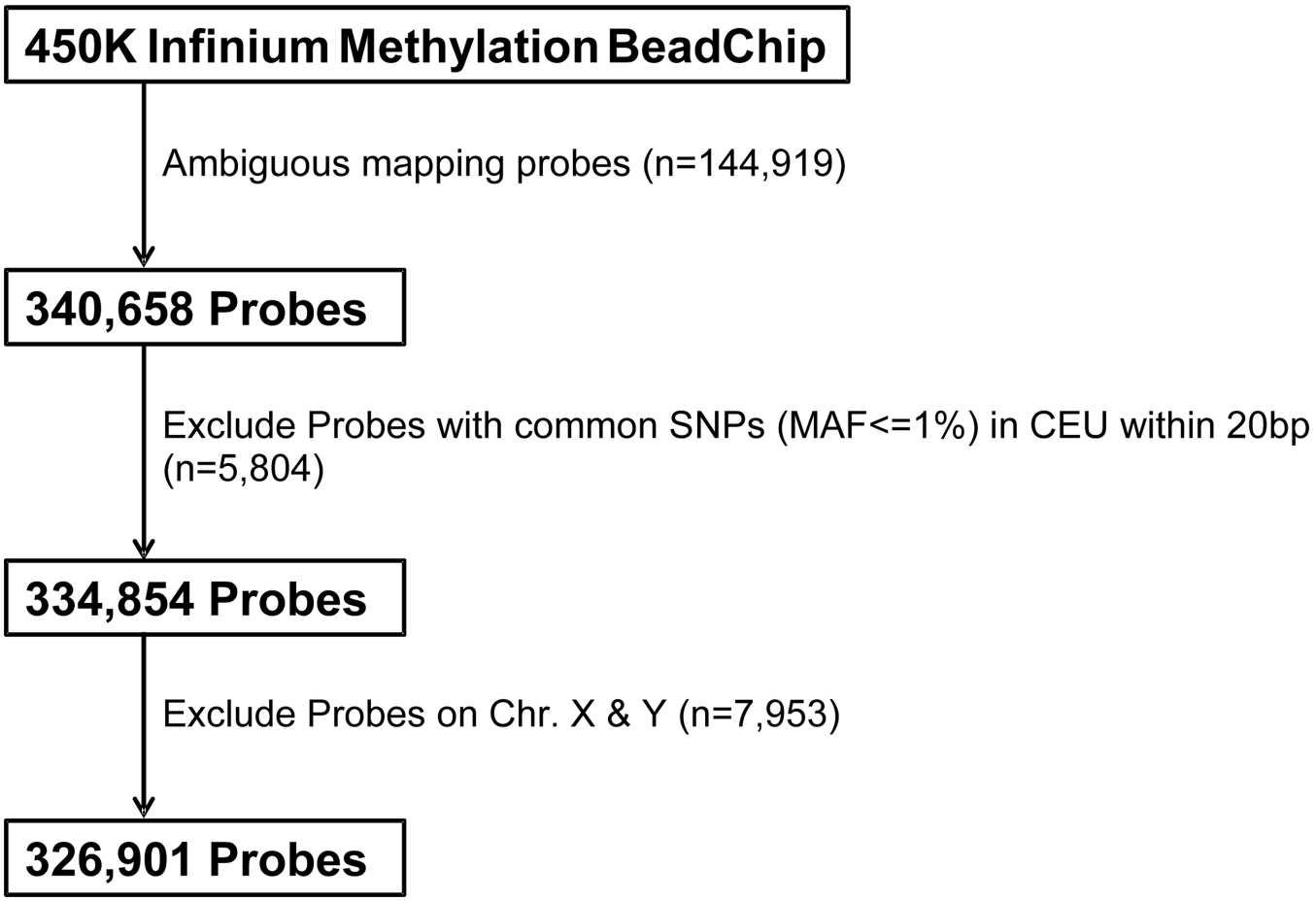

Figure 1. Filter process of Illumina 450k Infinium Methylation BeadChip

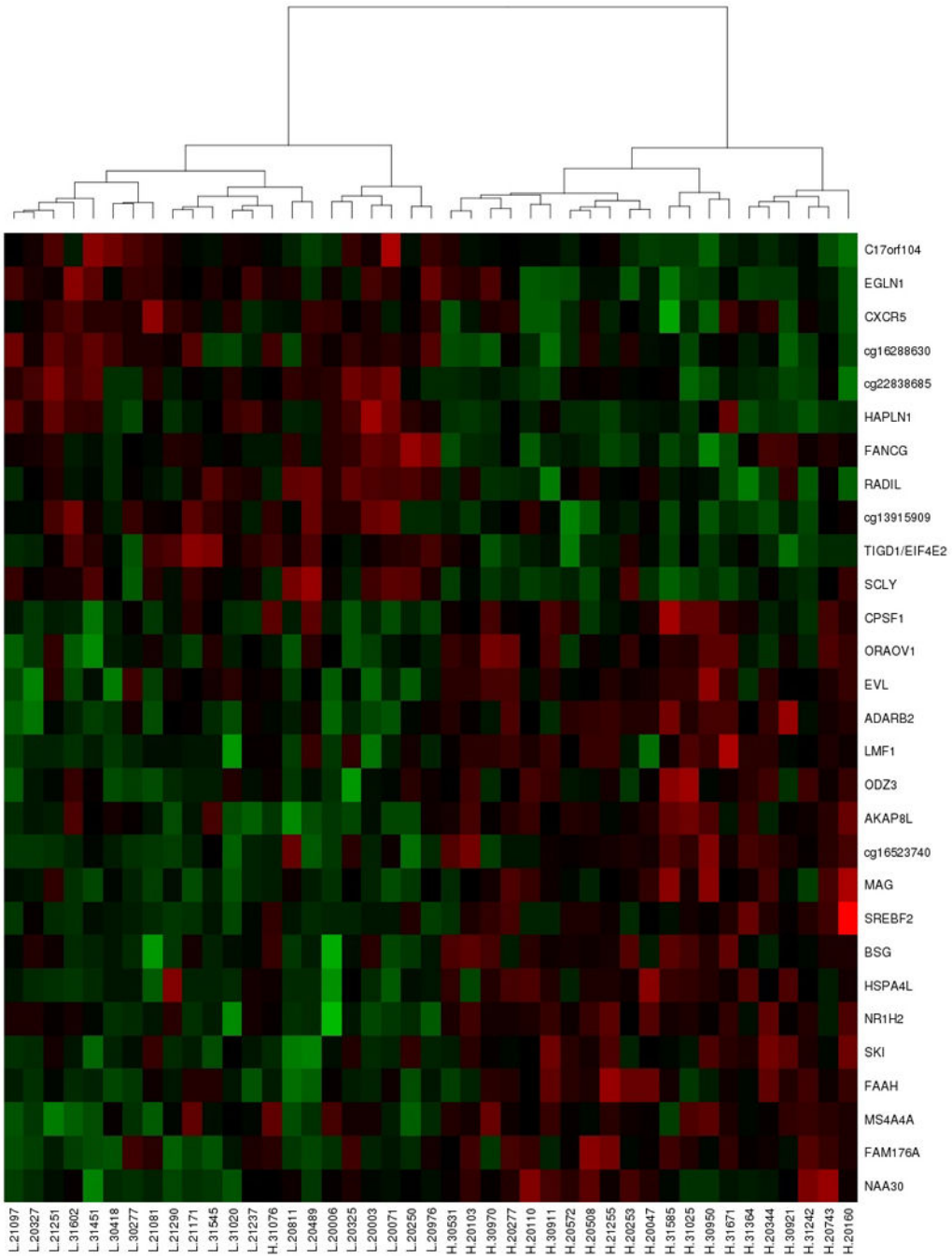

Figure 2. Heat map of 29 top CpG sites with methylation levels associated with As exposure

Note. DNA methylation heatmap of methylated genes (nominal p<0.0001) in WBC DNA from As highly exposed subjects compared with that from low exposed subjects. Each row represents a CpG site with columns corresponding to each sample. Higher methylation levels are shaded in red and lower levels are in green. The dendrogram shows the results of unsupervised hierarchical clustering of 43 samples based on 29 CpG sites, which separates

As highly exposed subjects (labeled as H at the bottom) from As low exposed subjects (labeled as L).

Table 1

Characteristics of 43 Study Subjects in CARDIA by Arsenic Exposure Status

Variables	As high-exposed (N=22)	As low-exposed (N=21)	P-value[a]
Male, n (%)	12 (54.5)	10 (47.6)	0.76
Age [years], mean (SD)	26.0±2.6	27.0±2.1	0.19
Education [years], mean (SD)	15.6±2.2	16.5±2.1	0.21
Married, n (%)	8 (36.4)	11 (52.4)	0.36
Alcohol drinker, n (%)	18 (81.8)	12 (57.1)	0.10
Body mass index [kg/m^2], mean (SD)	23.9±3.4	23.3±3.7	0.57

[a] p-values were calculated using Student's t-test and Fisher's exact test for continuous and categorical variables, respectively.

Table 2

Top CpG Sites with Methylation Levels Associated with High As Exposure ($p<10^{-4}$)

IllumID	Gene Symbol[a]	CpG Features	%Diff.[b]	P-value	Associated Diseases[c]
cg15083037	HAPLN1	TSS1500	30.9	9.46E-05	Cancer, non-insulin-dependent diabetes
cg13611173	NAA30	5'UTR;1stExon	27.8	5.43E-05	--
cg06696800	FAM176A	5'UTR;1stExon	26.7	5.08E-05	--
cg07746918	SREBF2	1stExon	19.7	6.54E-05	Schizophrenia; Alzheimer's Disease; BMI
cg15225378	NR1H2	TSS200	19.4	9.78E-06	Alzheimer's Disease
cg00469380	RADIL	Body	17.8	4.46E-05	--
cg18025430	MS4A4A	TSS200	17.3	7.34E-05	Flu, Febrile seizure; late-onset Alzheimer's Disease (LOAD)
cg05622915	BSG	TSS200 TSS1500	16.8	3.72E-06	Endometriosis; Hepatocellular Carcinoma; Measles virus infection; Liver cancer; Osteolytic Bone Disease; Primary biliary Cirrhosis, Scleosing cholangitis; Renal Cancer; Renal-cell Carcinoma; Crohn's Disease
cg24565496	LMF1	TSS1500	14.9	6.42E-05	Combined lipase deficiency
cg17301015	HSPA4L	5'UTR;1stExon	12.9	9.98E-05	--
cg07059784[d]	TIGD1	TSS1500	12.2	8.17E-05	--
cg07059784[d]	EIF4E2	Body	12.2	8.17E-05	Survival of advanced non-small cell lung cancer treated with carboplatin and paclitaxel
cg16334840	SCLY	TSS1500	11.7	5.41E-05	--
cg05293216	FANCG	TSS1500	9.6	7.68E-05	Breast Cancer, Young-onset pancreatic cancer, Fanconi's anemia group G, Chromosomal aberration
cg18728264	CXCR5	1stExon;3'UTR	9.2	1.00E-04	Rheumatoid arthritis, Hydronephrosis, B-cell non-Hodgkin's disease, Sjogren's syndrome, Endometriosis, Lymphoid tissue lymphoma, Non-Hodgkin's lymphoma[39]
cg08097581	AKAP8L	Body	7.4	8.16E-05	Leiomyomatosis, Uterine cancer/leiomyoma
cg09514185	ADARB2	Body	7.0	2.84E-06	Metabolic disorders[40]; Coronary artery disease, Acute lymphocytic leukemia
cg19921581	FAAH	TSS1500	5.3	2.34E-05	Alzheimer's disease, Dyslipidemia
cg20682143	EGLN1	TSS1500	5.3	1.91E-05	Parkinson's disease; Familial erythrocytosis
cg18048983	C17orf104	Body	5.2	5.76E-05	--
cg18550847	EVL	3'UTR	4.5	4.34E-05	Serous adenocarcinoma

IllumID	Gene Symbol[a]	CpG Features	%Diff.[b]	P-value	Associated Diseases[c]
cg22838685	_	_	3.2	1.99E-06	--
cg16288630	_	_	2.9	2.10E-05	--
cg17594004	ODZ3	Body	2.8	9.81E-05	--
cg13915909	_	_	2.3	1.83E-05	--
cg04690840	MAG	Body	2.2	7.70E-05	Schizophrenia, Alzheimer's disease
cg16523740	_	_	2.0	2.25E-05	--
cg09671094	CPSF1	Body	2.0	7.47E-05	--
cg09373227	ORAOV1	Body	1.6	7.09E-05	Head and neck squamous cell carcinoma
cg16575461	SKI	Body	1.3	9.14E-05	Gastric cancer, Rhabdomyosarcoma

[a] Genes were sorted by % Difference (%Diff.)

[b] % Diff. = ((absolute mean difference between methylation values in As high- and low-exposed subjects)/(summation of these two means/2))×100%.

[c] From Ingenuity Knowledge Base and PubMed search; gene expression or protein levels or genetic polymorphisms are associated with diseases in humans.

[d] cg07059784 is annotated to two genes.

Table 3

Pathway Analyses of the Associations between As Exposure and Genome-wide CpG Methylation Levels[a]

Gene Set[b]	Brief Description	p-value	Corrected p-value[c]
KEGG_FOCAL_ADHESION	Focal adhesion at the cell-extracellular matrix contact points	1.92E-11	3.45E-09
KEGG_WNT_SIGNALING_PATHWAY	Wnt signaling pathway	4.75E-09	4.25E-07
KEGG_AXON_GUIDANCE	A key stage in the formation of neuronal network	5.87E-09	4.25E-07
KEGG_REGULATION_OF_ACTIN_CYTOSKELETON	Cell motility: regulation of actin cytoskeleton	2.10E-07	9.41E-06
KEGG_VIRAL_MYOCARDITIS	Cardiovascular disease: Viral Myocarditis	2.37E-07	9.41E-06
KEGG_CELL_ADHESION_MOLECULES_CAMS	Signaling molecules and interaction: cell adhesion molecules	2.65E-07	9,41E-06
REACTOME_SIGNALLING_BY_NGF	Signal transduction: nerve growth factor (NGF) signaling	2.64E-11	1.61E-08
REACTOME_TRANSMISSION_ACROSS_CHEMICAL_SYNAPSES	Neuronal system (Transmission across Chemical Synapses)	5.95E-10	1.82E-07

[a] Output from GSA-SNP (http://gsa.muldas.org/).

[b] Gene sets and contained genes are detailed in Table S1.

[c] Benjamini and Hochberg False Discovery Rate.