Title

Linkage Analysis of Urine Arsenic Species Patterns in the Strong Heart Family Study.

Journal

Toxicological Sciences, 148(1)

Authors

Voruganti, Venkata
Cole, Shelley
Haack, Karin
et al.

Publication Date

2015-11-01

DOI

10.1093/toxsci/kfv164

Peer reviewed

OXFORD

SOT | Society of Toxicology
www.toxsci.oxfordjournals.org

TOXICOLOGICAL SCIENCES, 148(1), 2015, 89–100

doi: 10.1093/toxsci/kfv164
Advance Access Publication Date: July 23, 2015
Research Article

Linkage Analysis of Urine Arsenic Species Patterns in the Strong Heart Family Study

Matthew O. Gribble,[*,1] Venkata Saroja Voruganti,[†,‡] Shelley A. Cole,[§] Karin Haack,[§] Poojitha Balakrishnan,[¶,||] Sandra L. Laston,[lll] Maria Tellez-Plaza,[¶,llll] Kevin A. Francesconi,[#] Walter Goessler,[#] Jason G. Umans,[**,††] Duncan C. Thomas,[*] Frank Gilliland,[*] Kari E. North,[a] Nora Franceschini,[a] and Ana Navas-Acien[¶,||,b,c]

[*]Department of Preventive Medicine, University of Southern California, Los Angeles, California; [†]Department of Nutrition, University of North Carolina, Chapel Hill, North Carolina; [‡]UNC Nutrition Research Institute, University of North Carolina at Chapel Hill, Kannapolis, North Carolina; [§]Department of Genetics, Texas Biomedical Research Institute, San Antonio, Texas; [¶]Department of Environmental Health Sciences, Johns Hopkins University, Baltimore, Maryland; [||]Department of Epidemiology, Johns Hopkins Medical Institutions, Baltimore, Maryland; [lll]South Texas Diabetes and Obesity Institute, University of Texas Health Science Center, San Antonio—Regional Academic Health Center, Brownsville, Texas; [llll]Biomedical Research Institute, Hospital Clinic de Valencia—INCLIVA, Valencia, Spain; [#]Institute of Chemistry—Analytical Chemistry, University of Graz, Graz, Austria; [**]Georgetown-Howard Universities Center for Clinical and Translational Science, Washington, District of Columbia; [††]MedStar Health Research Institute, Hyattsville, Maryland; [a]Department of Epidemiology, University of North Carolina, Chapel Hill, North Carolina; [b]Welch Center for Prevention, Epidemiology and Clinical Research, Johns Hopkins Medical Institutions, Baltimore, Maryland; [c]Department of Oncology, Johns Hopkins Medical Institutions, Baltimore, Maryland

[1]To whom correspondence should be addressed at Department of Preventive Medicine, University of Southern California, 2001 N. Soto Street, Mail Code 9234, Los Angeles, California 90032. E-mail: mgribble@usc.edu.

ABSTRACT

Arsenic toxicokinetics are important for disease risks in exposed populations, but genetic determinants are not fully understood. We examined urine arsenic species patterns measured by HPLC-ICPMS among 2189 Strong Heart Study participants 18 years of age and older with data on ~400 genome-wide microsatellite markers spaced ~10 cM and arsenic speciation (683 participants from Arizona, 684 from Oklahoma, and 822 from North and South Dakota). We logit-transformed % arsenic species (% inorganic arsenic, %MMA, and %DMA) and also conducted principal component analyses of the logit % arsenic species. We used inverse-normalized residuals from multivariable-adjusted polygenic heritability analysis for multipoint variance components linkage analysis. We also examined the contribution of polymorphisms in the arsenic metabolism gene AS3MT via conditional linkage analysis. We localized a quantitative trait locus (QTL) on chromosome 10 (LOD 4.12 for %MMA, 4.65 for %DMA, and 4.84 for the first principal component of logit % arsenic species). This peak was partially but not fully explained by measured AS3MT variants. We also localized a QTL for the second principal component of logit % arsenic species on chromosome 5 (LOD 4.21) that was not evident from considering % arsenic species individually. Some other loci were suggestive or significant for 1 geographical area but not overall across all areas, indicating possible locus heterogeneity. This genome-wide linkage scan suggests genetic determinants of arsenic

toxicokinetics to be identified by future fine-mapping, and illustrates the utility of principal component analysis as a novel approach that considers % arsenic species jointly.

Key words: toxicogenetics; toxicokinetics; arsenic metabolism; arsenic species; linkage analysis; Strong Heart Family Study

Inorganic arsenic (iAs) exposure is carcinogenic (IARC, 2004) and increasing evidence also supports its association with cardiovascular disease (Moon *et al.*, 2012; Wu *et al.*, 2014) and diabetes (Maull *et al.*, 2012), although the importance of low-level exposure remains unclear. The iAs species (arsenite, arsenate) are metabolized to form methylated (MMA) and dimethylated (DMA) arsenic species (Thomas, 2009; Vahter, 2002). The enzyme arsenic (III) methyltransferase coded by *AS3MT* on 10q24.32 is believed to be centrally important in arsenic methylation reactions (Wood *et al.*, 2006), although other enzymes might also be involved (Harari *et al.*, 2013; Ren *et al.*, 2010).

There is growing evidence from human samples that arsenic species differ in their toxicities. Proliferation of T cells in fresh blood collected from human participants recruited at University of New Mexico was more strongly inhibited by exposure to trivalent MMA than exposure to trivalent arsenite (Burchiel *et al.*, 2014). In a study using blood collected from 2 male medical students at the National Autonomous University of Mexico, treating lymphocytes with either trivalent MMA or trivalent DMA induced DNA damage via oxidative-stress, but MMA was effective at lower doses (Soto-Reyes *et al.*, 2005). Human lymphoblastoid cell culture experiments suggested that trivalent MMA had greater impact on mitotic indices than arsenate, arsenite, pentavalent MMA, or trivalent or pentavalent DMA; and trivalent MMA and DMA were able to inhibit tubulin polymerization at lower concentrations than arsenite (Kligerman *et al.*, 2005). If arsenic species are toxicologically distinct, then their biotransformation and elimination may be relevant for human disease risks from arsenic exposure.

Inter-individual variation in arsenic kinetic processes has been studied indirectly by comparing the percentages of iAs, MMA, and DMA species (relative to their sum) in urine (Chiou *et al.*, 1997; Gamble and Liu, 2005; Gomez-Rubio *et al.*, 2011; Gribble *et al.*, 2013a; Hopenhayn-Rich *et al.*, 1996; Navas-Acien *et al.*, 2009; Vahter, 2000). The pattern of urine arsenic species is a possible risk factor for skin lesions (Ahsan *et al.*, 2007; Chen *et al.*, 2009; Valenzuela et al., 2009); bladder, urothelial, lung, and skin cancers (Agusa *et al.*, 2010; Chen *et al.*, 2003; Huang *et al.*, 2008a, b; Steinmaus *et al.*, 2006, 2010; Yu *et al.*, 2000) and cardiovascular diseases (Huang *et al.*, 2007, 2009). Endogenous and environmental factors have been associated with urine arsenic species patterns, including sex, age, pregnancy status, nutrition, body mass, smoking, alcohol, diabetes status, and insulin level (Gamble and Hall, 2012; Gomez-Rubio *et al.*, 2011; Gribble *et al.*, 2013a; Su *et al.*, 2012; Tseng, 2009). Arsenic species patterns are also influenced by genetics (Antonelli *et al.*, 2014; Hernandez and Marcos, 2008). In the Strong Heart Study, a large population-based cohort of American Indians with low-moderate arsenic exposure, heritability estimates for each % arsenic species (%iAs, %MMA, and %DMA) were ≥50% (Tellez-Plaza *et al.*, 2013). The causal role of *AS3MT* in arsenic methylation biology has been studied in both functional experiments (Wood *et al.*, 2006) and Mendelian randomization analysis of *AS3MT* variants for urine arsenic species (Pierce *et al.*, 2013). This gene has attracted such interest that there have been several comparative studies of *AS3MT* genetic variation across populations (Fujihara *et al.*, 2007, 2008, 2009, 2011); in study populations from Bangladesh and Northern Mexico, there was a large linkage disequilibrium region around this gene (Engstrom *et al.*, 2013; Gomez-Rubio *et al.*, 2010).

The predominant approach to genetic epidemiology of arsenic kinetics (Hernandez and Marcos, 2008) has been candidate gene or genome-wide association studies (Engstrom *et al.*, 2011; Gomez-Rubio *et al.*, 2010; Gribble *et al.*, 2013b; Hernandez *et al.*, 2008; Lindberg *et al.*, 2007; Meza *et al.*, 2005; Pierce *et al.*, 2012; Schlawicke Engstrom *et al.*, 2009). Although often informative, association studies can have limited power to detect genomic relationships in the presence of allelic heterogeneity (Ott *et al.*, 2011), or for rare variants unless they have a large effect. Association studies are also often performed using genotyping arrays, which may have limited coverage in populations such as American Indians. Family-based linkage studies can complement association studies to find chromosomal regions that co-segregate across generations with urine arsenic species patterns (Ott *et al.*, 2011). We previously reported a preliminary linkage analysis of urine arsenic species patterns in a pilot sample of 487 participants from the Strong Heart Family Study, an ancillary study to the Strong Heart Study, finding suggestive linkage evidence on chromosomes 5 (%iAs), 10 (%DMA), and 11 (%iAs and %MMA) (Tellez-Plaza *et al.*, 2013).

In this study, we aimed to investigate the co-segregation of loci with urine arsenic species in the full sample of the Strong Heart Family Study cohort. In addition to considering each % arsenic species biomarker separately, we also used principal components as a novel approach to describe overall patterns in the 3 % arsenic species measures.

MATERIALS AND METHODS

Study population. The Strong Heart Family Study is composed of 94 extended families from Arizona, Oklahoma, North Dakota, and South Dakota (North *et al.*, 2003). The rich family relationships included 3975 relative pairs in Arizona; 3248 relative pairs in Oklahoma; and 5571 relative pairs in North and South Dakota. Trained nurses and medical assistants collected anthropometric and interview data following standardized protocols (Lee *et al.*, 1990). This study was approved by relevant institutional and tribal ethics review boards. Informed consent was obtained from all participants. For this study, we included Strong Heart Family Study participants with both genotype and arsenic species data measured as part of an ancillary study to evaluate gene-environment interactions for diabetes incidence. Arsenic species data were not available among persons with baseline diabetes. We also excluded 245 participants with one or more urine arsenic species <0.1 μg/l limit of detection, as it is difficult to evaluate arsenic metabolism if arsenic species in urine are undetectable. We also excluded participants with missing data on smoking history ($n = 6$), current alcohol consumption ($n = 3$), body mass index ($n = 8$), educational attainment ($n = 5$), or with a sample error ($n = 2$) for an analysis sample size of $n = 2189$.

Urine arsenic measures. Total urine arsenic levels and arsenic species were measured by the Trace Element Laboratory of the

Institute of Chemistry—Analytical Chemistry, University of Graz in Graz, Austria, using high performance liquid chromatography inductively-coupled plasma mass spectrometry (HPLC-ICPMS) (Scheer *et al.*, 2012). Samples had been occasionally thawed and re-frozen in the years between sample collection and chemical analysis, precluding accurate measurement of reduction state of arsenicals in these samples (Scheer *et al.*, 2012). The inter-batch variability was monitored by replicate measurements of 3 urine reference materials with certified arsenic levels of 20.3 μg/l (NIST 2669 I), 50.2 μg/l (NIST 2669 II), or 119 μg/l (NIES 18); the coefficients of variation (CV) ranged from 3.8 to 14.4%, with an overall mean CV of 7.9% ($n = 46$). Percent arsenic species (%iAs, %MMA, and %DMA) were defined as the contributions of inorganic, methylated, and dimethylated arsenical species in urine to the sum of those species. For example, %iAs = iAs/(iAs + MMA + DMA).

Genetic measures. Methods used for microsatellite marker genotyping have been described previously (Tellez-Plaza *et al.*, 2013). In brief, microsatellite markers from the ABI PRISM Linkage Mapping Set-MD 10 version 2.5 (Applied Biosystems, now Thermo Fisher Scientific, Grand Island, New York) set were genotyped in amplified white blood cell DNA using the ABI PRISM 377 Genetic Analyzer. These ~400 markers were spaced roughly 10 cM apart with density ranging from 2.4 to 24.1 cM. Genotypes were called using the Genotyper software (Thermo Fisher Scientific). SNPs in *AS3MT* were typed as part of the Cardio-Metabo DNA Analysis BeadChip following the manufacturer's protocol (Illumina, San Diego, California). In brief, genomic DNA (200 ng) was whole-genome amplified, enzymatically fragmented, and hybridized to the chips containing locus-specific oligomers covalently linked to beads embedded in the chips. The hybridization was followed by a single-base extension with fluorescently labeled nucleotides. All steps were performed on a Tecan Freedom EVO 150 cm liquid handler (Tecan, Männedorf, Switzerland) with Illumina GTS Robot Control software (Illumina). Fluorescence intensities were detected using an iScan (Illumina) and were consequently analyzed using the Illumina GenomeStudio software. Assay accuracy was assessed using sample-dependent and independent controls. Potential misclassification of pedigree relationships was inspected using PREST (McPeek and Sun, 2000; Sun *et al.*, 2002) and Simwalk2/Preswalk (Sobel *et al.*, 2002) software. Mendelian errors and unlikely double recombinants were blanked, with an overall blanking rate <1%. SNP variants were aligned to the UCSC Genome Browser Human Genome Version 18 assembly.

Statistical analyses. A variance components decomposition method implemented in Sequential Oligogenic Linkage Analysis Routines (SOLAR) software was used to estimate heritability and detect linkage to chromosomal locations that affect the variation in arsenic species (Almasy and Blangero, 1998). Analyses were conducted separately for each logit-transformed measure of arsenic species patterns (%iAs, %MMA, and %DMA) and for each study center (Arizona, Oklahoma, and North and South Dakota). We also considered principal components of the logit-arsenic species as another way of describing inter-individual phenotypic differences in urine arsenic species patterns. The first 2 principal components were of substantive interest, while the third principal component was a negligible (eigenvalue 0.01) artifact of the logit-transformation of the 3 dependent % arsenic species variables. Models adjusted for age, age squared, sex, linear, and quadratic interactions between age and sex, study center (Arizona, Oklahoma, and North and South Dakota), obesity (body mass index ≥30 kg/m^2), high school completion (education ≥ 12 years), smoking status (ever/never), drinking status (current drinking yes/no), and higher than median iAs exposure (measured by the sum of inorganic and methylated arsenic species in urine; not corrected for urine creatinine). Sensitivity analysis indicates that adjustment for urine creatinine prior to dichotomization of the urine sum of species would have little impact on our final results (data not shown). Residuals from preliminary heritability analyses controlling for covariates were inverse-normalized and carried forward for linkage analyses.

We used the LOD score (negative log of the *P* value) thresholds set by Lander and Kruglyak (1995) for suggestive (≥1.9) or significant (≥3.3) evidence of linkage. We also stratified each logit-transformed % arsenic species linkage analyses by study center for secondary analyses to examine possible heterogeneity across study centers.

We examined the contribution of *AS3MT* variants to a linkage signal anticipated on chromosome 10 by including 2 SNPs rs17878846 and rs10509760 from that gene region in a conditional linkage analysis. We summarized the linkage disequilibrium between these SNPs and other 10q24 locus variants measured on the Cardio-Metabo DNA Analysis BeadChip through pair-wise r^2. There were 24 persons in our sample without genotypes available for rs17878846 and rs10509760 who were excluded from the conditional linkage analysis.

RESULTS

Arsenic Metabolism Biomarkers and Participant Characteristics

The demographic characteristics of the study participants are summarized in Table 1. Mean age in the sample was 34.3 (SD 14.8) years, 59% were female, and mean body mass index was 31.5 (SD 7.8) kg/m^2.

In this sample, median (interquartile range) urine arsenic species patterns were 10 (7, 14) %iAs, 14 (11, 18) %MMA, and 75 (68, 81) %DMA (Table 1). These patterns varied by sex, age, study center, body mass index, education, smoking, alcohol consumption, and iAs exposure. The arsenic species pattern measure showing the most variability across groups was %DMA. The sample Spearman correlation of total urine arsenic with %iAs was −0.16, with %MMA was −0.23, and with %DMA was +0.22. The heritability of these urine arsenic species patterns, conditional on demographics and high versus low arsenic, varied slightly by study center, ranging from 0.33 for %iAs in Oklahoma to 0.57 for %DMA in Arizona (Supplementary Table 1).

The first principal component appeared mainly to distinguish more versus less DMA in urine relative to the other species, and explained ~80% of the phenotypic variability in urine arsenic species patterns. The second principal component explained 19% of the phenotypic variability and appeared to distinguish %iAs versus %MMA after accounting for differences in %DMA (Supplementary Table 2). The first principal component score had a negative Spearman correlation with total urine arsenic (−0.22), was positively correlated with logit %iAs (0.87) and logit %MMA (0.80) and negatively correlated with logit %DMA (−0.998). The second principal component score had a weak negative Spearman correlation with total urine arsenic (−0.08), and it was negatively correlated with logit %iAs (−0.42), positively correlated with logit %MMA (+0.56), and essentially uncorrelated with logit %DMA (−0.02, *P* value = 0.48).

TABLE 1. Demographics and Urine Arsenic Species Patterns of $n = 2189$ Study Population

Demographic	N (%)	Median urine total arsenic (IQR) (μg/l)	Median arsenobetaine (IQR) (μg/l)	%iAs median (IQR)	%MMA median (IQR)	%DMA Median (IQR)	PC 1 Median (IQR)	PC 2 Median (IQR)
Overall	2189	10.4 (6.5, 17.9)	0.5 (0.3, 1.1)	10 (7, 14)	14 (11, 18)	75 (68, 81)	0.07 (−0.99, 1.07)	0.06 (−0.38,0.49)
Sex								
Male	901 (41)	10.5 (6.4, 17.9)	0.5 (0.3, 1.0)	12 (8, 16)	16 (12, 19)	72 (65, 78)	0.53 (−0.44, 1.39)	0.09 (−0.36, 0.53)
Female	1,288 (59)	10.3 (6.6, 18.0)	0.5 (0.3, 1.1)	9 (7, 13)	13 (10, 17)	77 (71, 83)	−0.26 (−1.37, 0.67)	0.05 (−0.40, 0.46)
Age								
<35 years	1,217 (56)	10.6 (6.7, 18.3)	0.5 (0.3, 0.8)	11 (8, 16)	15 (12, 18)	73 (66, 79)	0.28 (−0.62, 1.25)	0.04 (−0.42, 0.46)
≥35 years	972 (44)	10.2 (6.3, 17.1)	0.5 (0.3, 1.4)	9 (6, 13)	13 (10, 17)	77 (70, 82)	−0.29 (−1.34, 0.79)	0.11 (−0.33, 0.55)
Study center								
Arizona	683 (31)	15.2 (9.4, 25.6)	0.5 (0.3, 1.2)	11 (8, 15)	13 (10, 16)	75 (69, 81)	0.01 (−1.04, 0.91)	−0.16 (−0.67, 0.25)
Oklahoma	684 (31)	8.9 (5.9, 13.2)	0.6 (0.4, 1.2)	9 (7, 13)	14 (11, 18)	76 (69, 82)	−0.10 (−1.20, 0.92)	0.18 (−0.21, 0.56)
North or South Dakota	822 (38)	9.0 (5.7, 15.3)	0.5 (0.3, 0.9)	10 (7, 15)	15 (12, 19)	74 (67, 80)	0.24 (−0.81, 1.22)	0.15 (−0.27, 0.59)
Body mass index								
<30 kg/m^2	1,027 (47)	9.5 (6.0, 16.0)	0.5 (0.3, 0.9)	11 (8, 16)	16 (12, 19)	72 (65, 78)	0.46 (−0.54, 1.39)	0.17 (−0.26, 0.62)
≥30 kg/m^2	1,162 (53)	11.1 (7.0, 19.9)	0.5 (0.3, 1.2)	10 (7, 13)	13 (10, 16)	77 (71, 82)	−0.30 (−1.27, 0.66)	−0.01 (−0.46, 0.39)
Education attained								
<12 years	1,071 (49)	10.6 (6.4, 19.3)	0.5 (0.3, 0.9)	11 (8, 15)	14 (11, 18)	74 (67, 80)	0.20 (−0.84, 1.18)	0.03 (−0.46, 0.46)
≥12 years	1,118 (51)	10.3 (6.6, 16.5)	0.5 (0.3, 1.2)	10 (7, 14)	14 (11, 18)	76 (69, 82)	−0.05 (−1.17, 0.93)	0.11 (−0.32, 0.52)
Alcohol consumption								
Current abstainer	778 (36)	10.3 (6.4, 16.3)	0.5 (0.3, 1.2)	10 (7, 13)	14 (11, 18)	76 (69, 82)	−0.15 (−1.22, 0.89)	0.12 (−0.31, 0.52)
Current drinker	1,411 (64)	10.4 (6.6, 18.8)	0.5 (0.3, 1.0)	11 (8, 15)	14 (11, 18)	74 (67, 80)	0.15 (−0.85, 1.13)	0.04 (−0.44, 0.47)
Smoking status								
Never smoker	968 (44)	10.5 (6.5, 17.8)	0.5 (0.3, 1.1)	10 (7, 14)	14 (11, 18)	75 (69, 81)	0.01 (−1.05, 1.00)	0.09 (−0.32, 0.51)
Ever smoker	1,221 (56)	10.3 (6.5, 18.2)	0.5 (0.3, 1.0)	10 (7, 15)	14 (11, 18)	74 (68, 80)	0.13 (−0.93, 1.10)	0.04 (−0.44, 0.47)
Inorganic as exposure[a]								
Lower than median	1,092 (50)	6.5 (4.7, 8.3)	0.4 (0.3, 0.8)	11 (8, 15)	15 (12, 18)	74 (67, 79)	0.27 (−0.68, 1.19)	0.14 (−0.28, 0.58)
Higher than median	1,097 (50)	17.3 (12.5, 25.9)	0.6 (0.4, 1.4)	10 (7, 14)	13 (10, 17)	76 (69, 82)	−0.15 (−1.24, 0.90)	0.00 (−0.46, 0.38)

[a]iAs exposure measured by the sum of urine iAs, MMA, and DMA.

Linkage Disequilibrium for AS3MT Index SNPs

Considering a pair-wise r^2 between SNPs of ≥0.8 to be indicative of linkage disequilibrium, the linkage disequilibrium block for rs17878846 ranges at least from 104 584 497 to 104 783 894. The region in linkage disequilibrium with rs10509760 ranges at least from 104 624 464 to 104 793 052. As a point of comparison, according to the UCSC Genome Browser hg18 assembly, the *AS3MT* gene extends 104 619 200 to 104 651 645, so each of these 2 index SNPs is tagging a broader region extending a bit beyond the *AS3MT* gene boundaries in our sample. Pair-wise SNP r^2 for the linkage disequilibrium blocks are presented in Supplementary Tables 4 and 5.

Linkage Analysis of Arsenic Metabolism Biomarkers

In our linkage analyses ($n = 2189$), we localized a quantitative trait locus (QTL) on chromosome 10 near the *AS3MT* gene (Figs. 1–3), with LOD scores exceeding 4 for both logit- %MMA and %DMA phenotypes, as well as the first principal component of logit % arsenic species (Fig. 4). After conditioning for rs17878846 and rs10509760 at *AS3MT* ($n = 2165$), the conditional LOD score was attenuated by approximately half (Fig. 5). The %DMA LOD score decreased from 4.61 at 124 cM to 2.49 at 116 cM and the first principal component LOD score decreased from 4.84 at 121 cM to 2.73 at 116 cM.

There was a suggestive linkage peak on chromosome 6 (Fig. 1 and Supplementary Fig. 1) for %iAs largely driven by a signal in Arizona (LOD 3.06, 1-LOD support interval from 0 to 35 cM). A small overall population peak in this region was also visible for %DMA (LOD 1.01) and to a lesser extent %MMA (LOD 0.37); however, those peaks did not meet the threshold for suggestive evidence of linkage (Figs. 2 and 3). There was no visible peak on chromosome 6 for either of the principal components (Fig. 4).

The 2 principal components showed very distinct linkage signals (Fig. 4). The first principal component's main signal on chromosome 10 was not observed for the second principal component. In contrast, the second principal component's largest linkage signal was localized to chromosome 5 (LOD 4.21, 1-LOD support interval 111 to 120 cM), in a region not showing suggestive linkage signal for any individual % arsenic species biomarker nor for the first principal component (Supplementary Fig. 2). Similarly, there appeared to be a signal on chromosome 14 for the second principal component (LOD 3.20, 1-LOD support interval 46 to 67 cM) that was stronger than in the % arsenic species analyses.

Center-stratified analyses revealed potential locus heterogeneity, with some loci appearing suggestive or significant at 1 center but not at the other 2, and not in the combined analysis. In particular, Arizona showed a significant linkage signal (LOD 3.74, 1-LOD support interval from 39 to 57 cM) on chromosome 11 for %MMA but this was absent in the other 2 study centers (Fig. 2).

DISCUSSION

This study identified possible genetic determinants of arsenic toxicokinetics and illustrated a novel approach explicitly considering the inter-relationships among the % arsenic species. The strongest evidence for QTL in our study was with the first principal component on chromosome 10 (LOD 4.71), a component that reflected how %DMA levels were inversely related to %iAs and %MMA levels. Smaller peaks on chromosome 10 were

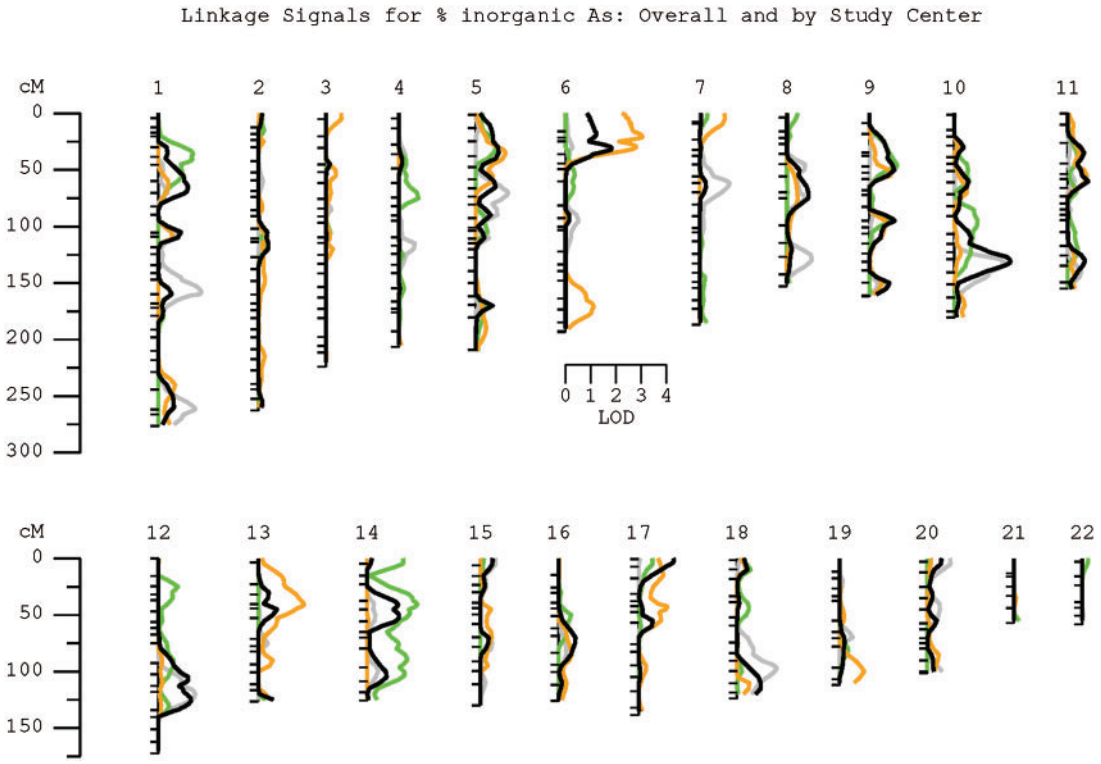

FIG. 1. Multipoint LOD scores for logit-transformed %iAs: overall, and from center-stratified analyses. Black line represents the multipoint LOD score from the combined analysis. Tangerine line is Arizona-restricted analysis, green line is Oklahoma-restricted analysis, and gray line is North or South Dakota-restricted analysis. Adjusted for age, age squared, sex, linear, and quadratic interactions between age and sex, study center (Arizona, Oklahoma, and North or South Dakota), obesity (body mass index $\geq$30 kg/m^2), high-school completion (education $\geq$12 years), smoking status (ever/never), drinking status (current abstainer yes/no), and higher than median iAs exposure (measured by the sum of inorganic and methylated arsenic species in urine).

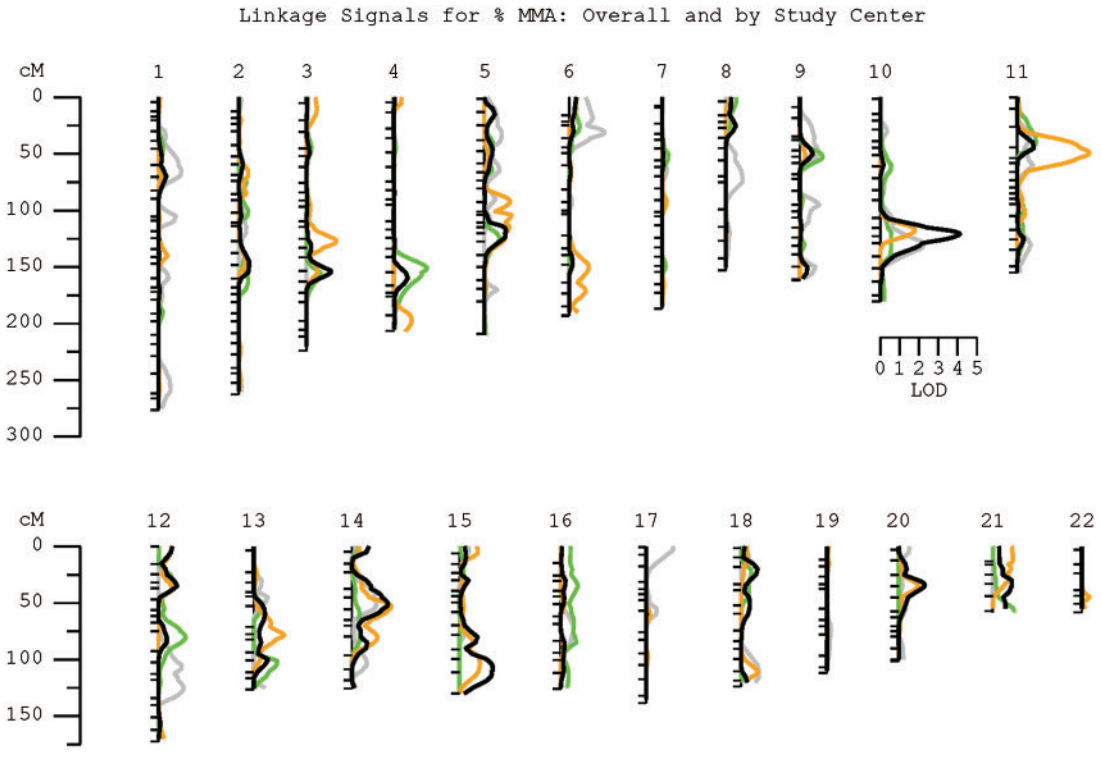

FIG. 2. Multipoint LOD scores for logit-transformed %MMA: overall, and from center-stratified analyses. Black line represents the multipoint LOD score from the combined analysis. Tangerine line is Arizona-restricted analysis, green line is Oklahoma-restricted analysis, and gray line is North or South Dakota-restricted analysis. Adjusted for age, age squared, sex, linear and quadratic interactions between age and sex, study center (Arizona, Oklahoma, and North or South Dakota), obesity (body mass index $\geq$30 kg/m^2), high-school completion (education $\geq$12 years), smoking status (ever/never), drinking status (current abstainer yes/no), and higher than median iAs exposure (measured by the sum of inorganic and methylated arsenic species in urine).

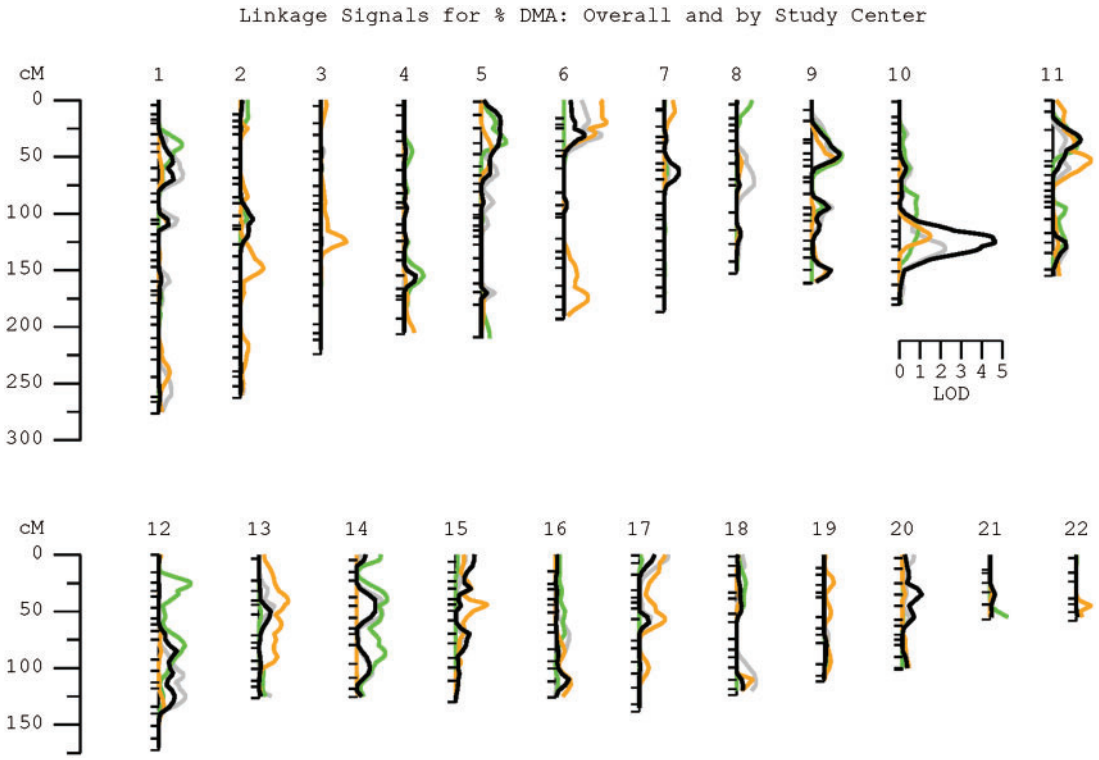

FIG. 3. Multipoint LOD scores for logit-transformed %DMA: overall, and from center-stratified analyses. Black line represents the multipoint LOD score from the combined analysis. Tangerine line is Arizona-restricted analysis, green line is Oklahoma-restricted analysis, and gray line is North or South Dakota-restricted analysis. Adjusted for age, age squared, sex, linear and quadratic interactions between age and sex, study center (Arizona, Oklahoma, and North or South Dakota), obesity (body mass index $\geq$30 kg/m^2), high-school completion (education $\geq$12 years), smoking status (ever/never), drinking status (current abstainer yes/no), and higher than median iAs exposure (measured by the sum of inorganic and methylated arsenic species in urine).

observed for %iAs (LOD 2.21), %MMA (LOD 4.12), and %DMA (LOD 4.61). The second principal component, which distinguished the balance of %iAs and %MMA independent of the level of %DMA, localized a QTL to chromosome 5 (LOD 4.27). This QTL did not pass the suggestive threshold for the % arsenic species when modeled separately.

Principal component scores in the context of arsenic species have several advantages compared with the conventional separate analysis of % arsenic species measures. First, it simplifies the interpretation by consolidating 3 inter-related measures into 2 independent measures. Second, the principal components may be reflecting biological processes more directly. The first component might be reflecting methylation broadly or even the second methylation step specifically, as these would lead to higher DMA relative to the other species. The evidence for a QTL near the *AS3MT* gene seen with the first principal component was stronger than the evidence for a QTL seen with any % arsenic species biomarker, consistent with the principal component score having less measurement error as an indicator of the underlying biological processes. The second component could be reflecting a transport process that would only influence the inter-relationship between iAs and MMA in urine independent of the DMA, or another unknown biological process.

Despite the value of these principal component scores for characterizing inter-individual variability in urine arsenic species patterns, there are limitations to this approach. The orientation of the principal component axes depends on the sample, so principal component scores from different study populations are not directly comparable. Also, the value of an individual's score depends on their compatriots in the study sample. Nevertheless, loci identified by studies using principal component scores can be compared across studies.

This study extends previous arsenic species pattern research in the Strong Heart Family Study (Tellez-Plaza *et al.*, 2013). We confirmed the QTL found in chromosome 10 in our preliminary analysis, which at the time was only suggestive (Tellez-Plaza *et al.*, 2013). Although we failed to replicate earlier findings of suggestive linkage of %iAs on chromosomes 9 and 11 in the overall analyses, on chromosome 11 we found significant evidence for a QTL for %MMA in Arizona. For chromosome 5, our second principal component QTL was in a different part of the chromosome than a suggestive peak for %iAs in the preliminary analysis (1-LOD support interval 29–57 cM). The discrepant findings for chromosomes 5 and 9 might be due to our restriction to diabetes-free participants if the preliminary findings, which included persons with diabetes, were partially informed by genes expressed differently under the conditions of diabetes. We were unable to test for locus-diabetes interactions as arsenic was not measured in participants with diabetes at baseline due to the study design. However, this interaction is possible, as co-morbid disease status can sometimes affect linkage signals (Colilla *et al.*, 2003), diabetes may have epigenetic consequences (Reddy and Natarajan, 2013), and even the earliest arsenic genetic epidemiology research suggested that the genetic determinants of urine arsenic species may be context (ie, age) dependent (Chung *et al.*, 2002).

The field of arsenic genetic epidemiology has heretofore emphasized the role of common genetic variants in explaining population-level variability in arsenic kinetics and

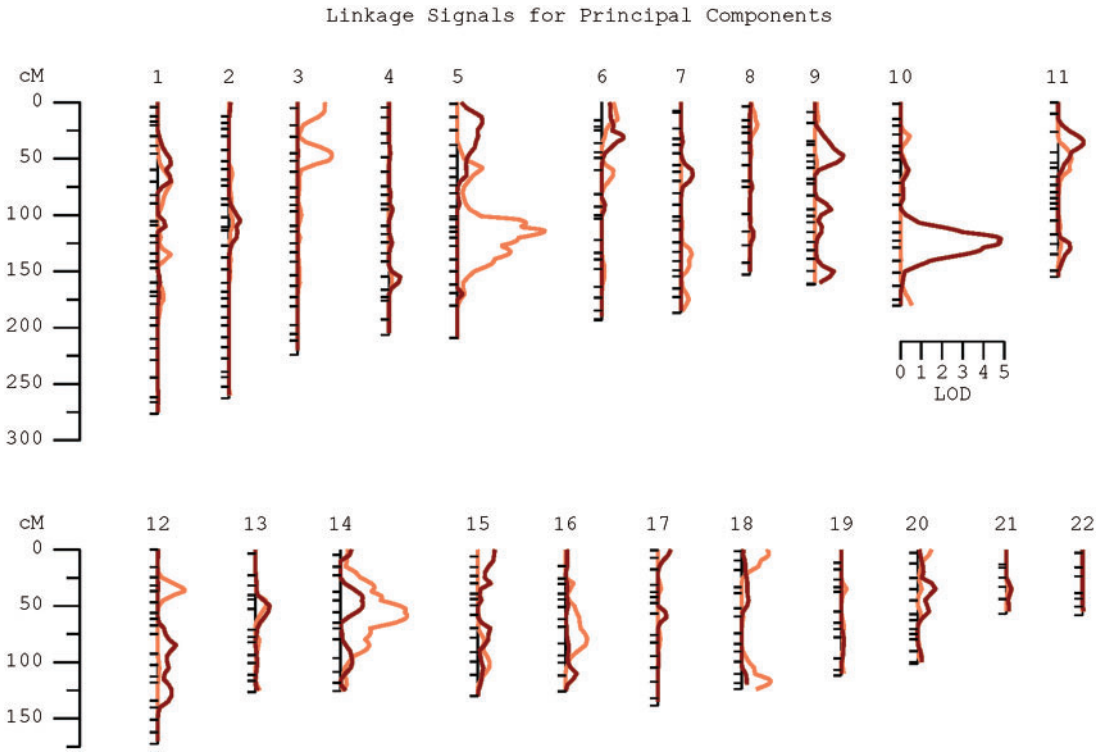

FIG. 4. Multipoint LOD scores signal for principal components of the logit-transformed % arsenic species. Red line is the first principal component and orange line is the second principal component. Adjusted for age, age squared, sex, linear, and quadratic interactions between age and sex, study center (Arizona, Oklahoma, and North or South Dakota), obesity (body mass index $\geq$30 kg/m^2), high-school completion (education $\geq$12 years), smoking status (ever/never), drinking status (current abstainer yes/no), and higher than median iAs exposure (measured by the sum of inorganic and methylated arsenic species in urine).

susceptibility. The heterogeneity across study centers, and residual heritability unexplained by the *AS3MT* SNPs included in the conditional linkage model, may suggest an important role for private or population-specific mutations as well. This would have implications for regulatory policies designed to protect the public from arsenic health effects, as "uncertainty factors" for non-cancer risk assessments (Alexeeff *et al.*, 2002; Gaylor and Kodell, 2002) that consider only common modifiers such as common *AS3MT* SNPs, sex, body mass index, or smoking status would underestimate the extent of inter-individual variability, and could result in regulatory standards inadequate to protect sensitive subpopulations. Quantifying the frequency of rare mutations conferring an impact on arsenic kinetics and susceptibility, and their relative importance for arsenic kinetics and for clinical outcomes, is an important and novel direction for future arsenic susceptibility genetic research. Exploring the role of rare variants in arsenic susceptibility is consistent with the broader push in modern genetic epidemiology to understand the role of low frequency and rare genetic variation in explaining phenotypic variance (Agarwala *et al.*, 2013; Cooper *et al.*, 2010; Kiezun *et al.*, 2012; Ng *et al.*, 2010; Sobreira *et al.*, 2010). The family design of the Strong Heart Family Study may be an asset in further exploring the role of rare genetic variation in explaining inter-individual variability in arsenic kinetics (Cheung *et al.*, 2013; Saad and Wijsman, 2014). Lastly, we did not detect any suggestive peaks on chromosome 12, where we recently identified SNPs in the liver transporter *SLCO1B1* explaining a substantial proportion of the phenotypic variance in %MMA and %DMA in a small ($n = 157$) convenience sample (Gribble *et al.*, 2013b). The discrepancy in findings for chromosomes 11 and 12 may be due to the complementary strengths of linkage and association designs to detect different kinds of genetic signals (Ott *et al.*, 2011). It may also be in part due to a spurious finding in our earlier study related to the small sample size or to locus heterogeneity across study center. Apparent locus heterogeneity across study centers might also be due to differential power to detect linkage signals across study centers if measurement error is different across study centers. We did estimate higher heritabilities for %MMA and %DMA (residuals) in Arizona than the other study centers (Supplementary Table 1) which could be consistent with less phenotypic measurement error in that study center (Almasy and Blangero, 2010); or it could reflect a true difference in heritability across study centers, for example if there is an exposure-by-gene interaction observable in the study center with highest arsenic exposure levels. An arsenic exposure-by-locus interaction is plausible as arsenic exposure has epigenetic effects across the genome (Bailey and Fry, 2014) and many genes may have differential expression by arsenic exposure (Chavan *et al.*, 2011; Hou *et al.*, 2014; Jutooru *et al.*, 2010; Klei and Barchowsky, 2008; Tchounwou *et al.*, 2003).

Localization of these broad linkage peaks is challenging due to the presence of multiple potentially relevant genes in these regions, and limited understanding of the biology of arsenic absorption, distribution, metabolism, and elimination in humans. The chromosome 6 linkage signal, suggestive for %iAs but not the other % arsenic species, is near several methyltransferase genes including the euchromatic histone-lysine *N*-methyltransferase *EHMT2* (Kim *et al.*, 2013), threonylcarbamoyladenosine tRNA methylthiotransferase (*CDKAL1*) (Okada *et al.*, 2012), and cap-specific mRNA (nucleoside-2'-*O*-)-methyltransferase 1 (*CMTR1*) (Smietanski *et al.*, 2014). However, it is

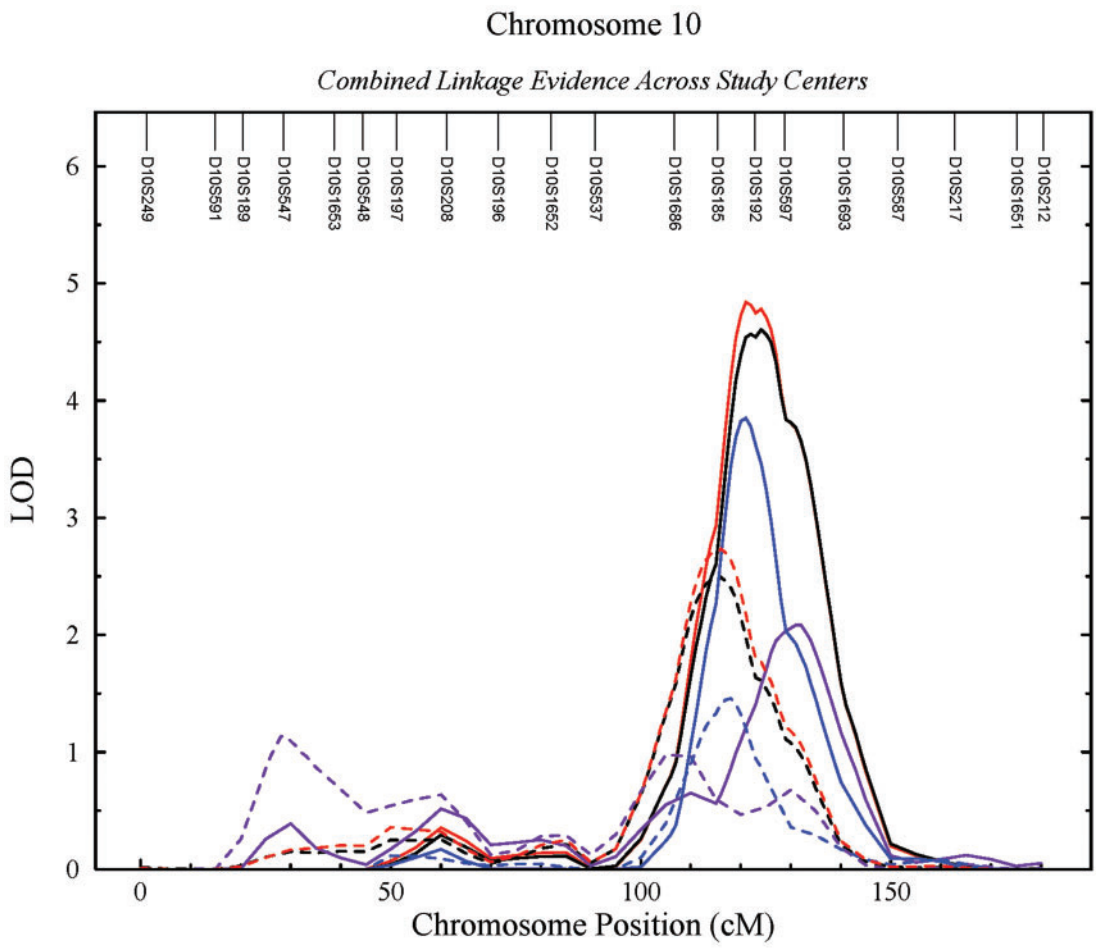

FIG. 5. Evidence from conditional linkage analysis for QTL on chromosome 10. Outcomes: purple line is logit-transformed %iAs, blue line is logit-transformed %MMA, black line is logit-transformed %DMA, and red line is principal component #1. Covariates: solid lines show results from models adjusted for age, age squared, sex, linear, and quadratic interactions between age and sex, study center (Arizona, Oklahoma, and North or South Dakota), obesity (body mass index $\geq$30 kg/m^2), high-school completion (education $\geq$12 years), smoking status (ever/never), drinking status (current abstainer yes/no), and higher than median iAs exposure (measured by the sum of inorganic and methylated arsenic species in urine). Dashed lines show further adjustment for rs17878846 and rs10509760 genotypes.

also possible that the linkage peak is not due to a methyltransferase but to some other part of the arsenic kinetic machinery, known to include reduction enzymes and transporters as well (Hernandez and Marcos, 2008). Transporters near the linkage peak include *HFE*, which has been previously associated with kinetics of metals including iron (Hanson *et al.*, 2001), lead (Hopkins *et al.*, 2008), and manganese (Claus Henn *et al.*, 2011); and *ABCB2*, which is important for antigen processing (Lankat-Buttgereit and Tampe, 2002) and may be relevant to the emerging research on arsenic immunotoxicity (Ahmed *et al.*, 2011; Farzan *et al.*, 2013). It is thought that arsenic metabolism may be related to creatinine pathways (Gamble and Hall, 2012; Thomas, 2009), and there is some toxicological evidence for an arsenic association with renal elimination of uric acid (Jauge and Del-Razo, 1985) although the epidemiological evidence is scarce (Kuo *et al.*, 2015). We cannot exclude the possibility that some of the chromosome 10 evidence for linkage is explained by rare variants within *AS3MT* not tagged by the haplotype index SNPs. The principal components' linkage analyses suggested that there may be distinct genetic determinants for %DMA versus other species, than for the balance of %iAs and %MMA after accounting for %DMA. The chromosome 5 region associated with principal component 2 includes many genes possibly relevant for arsenic kinetics, including solute carrier anion transporter family 4, member C1 (*SLCO4C1*) (Toyohara *et al.*, 2009), and solute carrier anion transporter family 6, member A1 (*SLCO6A1*) (Lee *et al.*, 2004) which both are in the same transporter super-family as the *SLCO1B1* gene preliminarily associated with % arsenic species patterns in a sample from the Strong Heart Family Study (Gribble *et al.*, 2013b; Roth *et al.*, 2012). The chromosome 14 peak for principal component 2 is near potential candidate genes including tRNA methyltransferase 5 (*TRMT5*) (Christian *et al.*, 2013), valosin-containing protein lysine (K) methyltransferase (*VCPKMT*) (Kernstock *et al.*, 2012), and the transporter solute carrier family 38, member 6 (*SLC38A6*) (Sundberg *et al.*, 2008). Linkage fine-mapping studies are needed to further evaluate the signals from this study.

This study has numerous strengths but also several limitations. Information on arsenic in drinking water for the study participants was not available, although arsenic in drinking water has been historically >10 μg/l in the communities from Arizona and in some of the communities from North and South Dakota. In the communities from Oklahoma arsenic levels in drinking water is <10 μg/l, and food is suspected as the main source of arsenic. The interpretation of % arsenic species as a biomarker of iAs metabolism and elimination is strengthened

by the result that arsenobetaine was a small proportion of the total urine arsenic, consistent with little seafood intake in the population. There was a large sample size with rich family relationships allowing stable estimation of the co-segregation of genetic loci with arsenic species phenotypes. This study was restricted to participants without diabetes at baseline as it was nested within an incident diabetes risk factor study; future work should characterize these genetic relationships among persons with diabetes. Findings from this population may or may not generalize to other populations with different distributions of variables, such as body mass index, predictive of urine arsenic species patterns.

In conclusion, this study identified major loci likely involved in arsenic kinetic processes, located on chromosomes 5, 6, and 10. These loci, together with a few loci that might be relevant for Arizona communities, should be followed up in fine-mapping studies. The evidence for linkage differed by study region, and was only partly explained by common variation in *AS3MT*, indicating that rare variation may be an important part of inter-individual variability in arsenic susceptibility, or perhaps that additional genes on chromosome 10 are important for arsenic processing. The novel approach of considering the inter-relationship of the % arsenic species through principal components analysis provided additional insights by detecting a strong linkage signal on chromosome 5, and may be a fruitful approach for future arsenic susceptibility genetics research. Arsenic exposure has been associated with incident cardiovascular disease (Moon *et al.*, 2013), cancer mortality (Garcia-Esquinas *et al.*, 2013), and diabetes prevalence (Gribble *et al.*, 2012) in the Strong Heart Study. Further evaluating the role of genetic determinants of arsenic dose and toxicity and of gene-environment interactions may contribute to the prevention and control of arsenic-related disease (Flanders, 2006; Rothman and Greenland, 2005).

SUPPLEMENTARY DATA

Supplementary data are available online at http://toxsci.oxfordjournals.org/.

FUNDING

National Heart Lung and Blood Institute (R01HL090863 and by SHS grants HL41642, HL41652, HL41654 and HL65521) and the National Institute of Environmental Health Sciences (R01ES021367); National Institute for Environmental Health Sciences (T32ES013678-07 to M.O.G.). M. Tellez-Plaza was supported by the Strategic Action for Research in Health Sciences (CP12/03080) from the Spanish Ministry of Economy and Competitiveness and the European Funds for Regional Development (FEDER). This investigation was conducted in facilities constructed with support from the National Center for Research Resources, National Institutes of Health (C06 RR013556 and C06 RR017515).

ACKNOWLEDGMENTS

The authors declare no conflicts of interest.

REFERENCES

Agarwala, V., Flannick, J., Sunyaev, S., and Altshuler, D. (2013). Evaluating empirical bounds on complex disease genetic architecture. *Nat. Genet.* **45**, 1418–1427.

Agusa, T., Kunito, T., Kubota, R., Inoue, S., Fujihara, J., Minh, T. B., Ha, N. N., Tu, N. P., Trang, P. T., Chamnan, C., *et al.* (2010). Exposure, metabolism, and health effects of arsenic in residents from arsenic-contaminated groundwater areas of Vietnam and Cambodia: a review. *Rev. Environ. Health* **25**, 193–220.

Ahmed, S., Mahabbat-e Khoda, S., Rekha, R. S., Gardner, R. M., Ameer, S. S., Moore, S., Ekstrom, E. C., Vahter, M., and Raqib, R. (2011). Arsenic-associated oxidative stress, inflammation, and immune disruption in human placenta and cord blood. *Environ. Health Perspect.* **119**, 258–264.

Ahsan, H., Chen, Y., Kibriya, M. G., Slavkovich, V., Parvez, F., Jasmine, F., Gamble, M. V., and Graziano, J. H. (2007). Arsenic metabolism, genetic susceptibility, and risk of premalignant skin lesions in Bangladesh. *Cancer Epidemiol. Biomark. Prev.* **16**, 1270–1278.

Alexeeff, G. V., Broadwin, R., Liaw, J., and Dawson, S. V. (2002). Characterization of the LOAEL-to-NOAEL uncertainty factor for mild adverse effects from acute inhalation exposures. *Regul. Toxicol. Pharmacol.* **36**, 96–105.

Almasy, L., and Blangero, J. (1998). Multipoint quantitative-trait linkage analysis in general pedigrees. *Am. J. Hum. Genet.* **62**, 1198–1211.

Almasy, L., and Blangero, J. (2010). Variance component methods for analysis of complex phenotypes. *Cold Spring Harbor Protoc.* **2010**, pdb.top77. doi: 10.1101/pdb.top77.

Antonelli, R., Shao, K., Thomas, D. J., Sams, R., 2nd, and Cowden, J. (2014). AS3MT, GSTO, and PNP polymorphisms: impact on arsenic methylation and implications for disease susceptibility. *Environ. Res.* **132**, 156–167.

Bailey, K., and Fry, R. (2014). Arsenic-associated changes to the epigenome: what are the functional consequences? *Curr. Environ. Health Rep.* **1**, 22–34.

Burchiel, S. W., Lauer, F. T., Beswick, E. J., Gandolfi, A. J., Parvez, F., Liu, K. J., and Hudson, L. G. (2014). Differential susceptibility of human peripheral blood T cells to suppression by environmental levels of sodium arsenite and monomethylarsonous acid. *PLoS One* **9**, e109192.

Chavan, H., Oruganti, M., and Krishnamurthy, P. (2011). The ATP-binding cassette transporter ABCB6 is induced by arsenic and protects against arsenic cytotoxicity. *Toxicol. Sci.* **120**, 519–528.

Chen, Y. C., Guo, Y. L., Su, H. J., Hsueh, Y. M., Smith, T. J., Ryan, L. M., Lee, M. S., Chao, S. C., Lee, J. Y., and Christiani, D. C. (2003). Arsenic methylation and skin cancer risk in southwestern Taiwan. *J. Occup. Environ. Med.* **45**, 241–248.

Chen, Y., Parvez, F., Gamble, M., Islam, T., Ahmed, A., Argos, M., Graziano, J. H., and Ahsan, H. (2009). Arsenic exposure at low-to-moderate levels and skin lesions, arsenic metabolism, neurological functions, and biomarkers for respiratory and cardiovascular diseases: review of recent findings from the Health Effects of Arsenic Longitudinal Study (HEALS) in Bangladesh. *Toxicol. Appl. Pharmacol.* **239**, 184–192.

Cheung, C. Y., Thompson, E. A., and Wijsman, E. M. (2013). GIGI: an approach to effective imputation of dense genotypes on large pedigrees. *Am. J. Hum. Genet.* **92**, 504–516.

Chiou, H. Y., Hsueh, Y. M., Hsieh, L. L., Hsu, L. I., Hsu, Y. H., Hsieh, F. I., Wei, M. L., Chen, H. C., Yang, H. T., Leu, L. C., *et al.* (1997). Arsenic methylation capacity, body retention, and null genotypes of glutathione S-transferase M1 and T1 among current arsenic-exposed residents in Taiwan. *Mutat. Res.* **386**, 197–207.

Christian, T., Gamper, H., and Hou, Y. M. (2013). Conservation of structure and mechanism by Trm5 enzymes. *RNA* **19**, 1192–1199.

Chung, J. S., Kalman, D. A., Moore, L. E., Kosnett, M. J., Arroyo, A. P., Beeris, M., Mazumder, D. N., Hernandez, A. L., and Smith, A. H. (2002). Family correlations of arsenic methylation patterns in children and parents exposed to high concentrations of arsenic in drinking water. *Environ. Health Perspect.* **110**, 729–733.

Claus Henn, B., Kim, J., Wessling-Resnick, M., Tellez-Rojo, M. M., Jayawardene, I., Ettinger, A. S., Hernandez-Avila, M., Schwartz, J., Christiani, D. C., Hu, H., *et al.* (2011). Associations of iron metabolism genes with blood manganese levels: a population-based study with validation data from animal models. *Environ. Health* **10**, 97.

Colilla, S., Nicolae, D., Pluzhnikov, A., Blumenthal, M. N., Beaty, T. H., Bleecker, E. R., Lange, E. M., Rich, S. S., Meyers, D. A., Ober, C., *et al.* (2003). Evidence for gene-environment interactions in a linkage study of asthma and smoking exposure. *J. Allergy Clin. Immunol.* **111**, 840–846.

Cooper, D. N., Chen, J. M., Ball, E. V., Howells, K., Mort, M., Phillips, A. D., Chuzhanova, N., Krawczak, M., Kehrer-Sawatzki, H., and Stenson, P. D. (2010). Genes, mutations, and human inherited disease at the dawn of the age of personalized genomics. *Hum. Mutat.* **31**, 631–655.

Engstrom, K., Vahter, M., Mlakar, S. J., Concha, G., Nermell, B., Raqib, R., Cardozo, A., and Broberg, K. (2011). Polymorphisms in arsenic(+III oxidation state) methyltransferase (AS3MT) predict gene expression of AS3MT as well as arsenic metabolism. *Environ. Health Perspect.* **119**, 182–188.

Engstrom, K. S., Hossain, M. B., Lauss, M., Ahmed, S., Raqib, R., Vahter, M., and Broberg, K. (2013). Efficient arsenic metabolism—the AS3MT haplotype is associated with DNA methylation and expression of multiple genes around AS3MT. *PLoS One* **8**, e53732.

Farzan, S. F., Korrick, S., Li, Z., Enelow, R., Gandolfi, A. J., Madan, J., Nadeau, K., and Karagas, M. R. (2013). In utero arsenic exposure and infant infection in a United States cohort: a prospective study. *Environ. Res.* **126**, 24–30.

Flanders, W. D. (2006). On the relationship of sufficient component cause models with potential outcome (counterfactual) models. *Eur. J. Epidemiol.* **21**, 847–853.

Fujihara, J., Fujii, Y., Agusa, T., Kunito, T., Yasuda, T., Moritani, T., and Takeshita, H. (2009). Ethnic differences in five intronic polymorphisms associated with arsenic metabolism within human arsenic (+3 oxidation state) methyltransferase (AS3MT) gene. *Toxicol. Appl. Pharmacol.* **234**, 41–46.

Fujihara, J., Kunito, T., Agusa, T., Yasuda, T., Iida, R., Fujii, Y., and Takeshita, H. (2007). Population differences in the human arsenic (+3 oxidation state) methyltransferase (AS3MT) gene polymorphism detected by using genotyping method. *Toxicol. Appl. Pharmacol.* **225**, 251–254.

Fujihara, J., Soejima, M., Koda, Y., Kunito, T., and Takeshita, H. (2008). Asian specific low mutation frequencies of the M287T polymorphism in the human arsenic (+3 oxidation state) methyltransferase (AS3MT) gene. *Mutat. Res.* **654**, 158–161.

Fujihara, J., Yasuda, T., Kato, H., Yuasa, I., Panduro, A., Kunito, T., and Takeshita, H. (2011). Genetic variants associated with arsenic metabolism within human arsenic (+3 oxidation state) methyltransferase show wide variation across multiple populations. *Arch. Toxicol.* **85**, 119–125.

Gamble, M. V., and Hall, M. N. (2012). Relationship of creatinine and nutrition with arsenic metabolism. *Environ. Health Perspect.* **120**, a145–a146.

Gamble, M. V., and Liu, X. (2005). Urinary creatinine and arsenic metabolism. *Environ. Health Perspect.* **113**, A442; author reply A442–A443.

Garcia-Esquinas, E., Pollan, M., Umans, J. G., Francesconi, K. A., Goessler, W., Guallar, E., Howard, B., Farley, J., Best, L. G., and Navas-Acien, A. (2013). Arsenic exposure and cancer mortality in a US-based prospective cohort: the Strong Heart Study. *Cancer Epidemiol. Biomarker. Prevent.* **22**, 1944–1953.

Gaylor, D. W., and Kodell, R. L. (2002). A procedure for developing risk-based reference doses. *Regul. Toxicol. Pharmacol.* **35**, 137–141.

Gomez-Rubio, P., Meza-Montenegro, M. M., Cantu-Soto, E., and Klimecki, W. T. (2010). Genetic association between intronic variants in AS3MT and arsenic methylation efficiency is focused on a large linkage disequilibrium cluster in chromosome 10. *J. Appl. Toxicol.* **30**, 260–270.

Gomez-Rubio, P., Roberge, J., Arendell, L., Harris, R. B., O'Rourke, M. K., Chen, Z., Cantu-Soto, E., Meza-Montenegro, M. M., Billheimer, D., Lu, Z., *et al.* (2011). Association between body mass index and arsenic methylation efficiency in adult women from southwest U.S. and northwest Mexico. *Toxicol. Appl. Pharmacol.* **252**, 176–182.

Gribble, M. O., Crainiceanu, C. M., Howard, B. V., Umans, J. G., Francesconi, K. A., Goessler, W., Zhang, Y., Silbergeld, E. K., Guallar, E., and Navas-Acien, A. (2013a). Body composition and arsenic metabolism: a cross-sectional analysis in the Strong Heart Study. *Environ. Health* **12**, 107.

Gribble, M. O., Howard, B. V., Umans, J. G., Shara, N. M., Francesconi, K. A., Goessler, W., Crainiceanu, C. M., Silbergeld, E. K., Guallar, E., and Navas-Acien, A. (2012). Arsenic exposure, diabetes prevalence, and diabetes control in the Strong Heart Study. *Am. J. Epidemiol.* **176**, 865–874.

Gribble, M. O., Voruganti, V. S., Cropp, C. D., Francesconi, K. A., Goessler, W., Umans, J. G., Silbergeld, E. K., Laston, S. L., Haack, K., Kao, W. H. L.,*et al.* (2013b). SLCO1B1 variants and urine arsenic metabolites in the Strong Heart Family Study. *Toxicol. Sci.* **136**, 19–25.

Hanson, E. H., Imperatore, G., and Burke, W. (2001). HFE gene and hereditary hemochromatosis: a HuGE review. Human Genome Epidemiology. *Am. J. Epidemiol.* **154**, 193–206.

Harari, F., Engstrom, K., Concha, G., Colque, G., Vahter, M., and Broberg, K. (2013). N-6-adenine-specific DNA methyltransferase 1 (N6AMT1) polymorphisms and arsenic methylation in Andean women. *Environ. Health Perspect.* **121**, 797–803.

Hernandez, A., and Marcos, R. (2008). Genetic variations associated with interindividual sensitivity in the response to arsenic exposure. *Pharmacogenomics* **9**, 1113–1132.

Hernandez, A., Xamena, N., Sekaran, C., Tokunaga, H., Sampayo-Reyes, A., Quinteros, D., Creus, A., and Marcos, R. (2008). High arsenic metabolic efficiency in AS3MT287Thr allele carriers. *Pharmacogenet. Genomics* **18**, 349–355.

Hopenhayn-Rich, C., Biggs, M. L., Smith, A. H., Kalman, D. A., and Moore, L. E. (1996). Methylation study of a population environmentally exposed to arsenic in drinking water. *Environ. Health Perspect.* **104**, 620–628.

Hopkins, M. R., Ettinger, A. S., Hernandez-Avila, M., Schwartz, J., Tellez-Rojo, M. M., Lamadrid-Figueroa, H., Bellinger, D., Hu, H., and Wright, R. O. (2008). Variants in iron metabolism genes predict higher blood lead levels in young children. *Environ. Health Perspect.* **116**, 1261–1266.

Hou, Y., Wang, Y., Wang, H., and Xu, Y. (2014). Induction of glutathione synthesis in human hepatocytes by acute and chronic arsenic exposure: differential roles of mitogen-activated protein kinases. *Toxicology* **325**, 96–106.

Huang, Y. K., Huang, Y. L., Hsueh, Y. M., Yang, M. H., Wu, M. M., Chen, S. Y., Hsu, L. I., and Chen, C. J. (2008a). Arsenic exposure, urinary arsenic speciation, and the incidence of

urothelial carcinoma: a twelve-year follow-up study. *Cancer Causes Control* **19**, 829–839.

Huang, Y. K., Pu, Y. S., Chung, C. J., Shiue, H. S., Yang, M. H., Chen, C. J., and Hsueh, Y. M. (2008b). Plasma folate level, urinary arsenic methylation profiles, and urothelial carcinoma susceptibility. *Food Chem. Toxicol.* **46**, 929–938.

Huang, Y. K., Tseng, C. H., Huang, Y. L., Yang, M. H., Chen, C. J., and Hsueh, Y. M. (2007). Arsenic methylation capability and hypertension risk in subjects living in arseniasis-hyperendemic areas in southwestern Taiwan. *Toxicol. Appl. Pharmacol.* **218**, 135–142.

Huang, Y. L., Hsueh, Y. M., Huang, Y. K., Yip, P. K., Yang, M. H., and Chen, C. J. (2009). Urinary arsenic methylation capability and carotid atherosclerosis risk in subjects living in arsenicosis-hyperendemic areas in southwestern Taiwan. *Sci. Total Environ.* **407**, 2608–2614.

IARC (2004). Some drinking-water disinfectants and contaminants, including Arsenic. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*.

Jauge, P., and Del-Razo, L. M. (1985). Uric acid levels in plasma and urine in rats chronically exposed to inorganic As (III) and As(V). *Toxicol. Lett.* **26**, 31–35.

Jutooru, I., Chadalapaka, G., Sreevalsan, S., Lei, P., Barhoumi, R., Burghardt, R., and Safe, S. (2010). Arsenic trioxide downregulates specificity protein (Sp) transcription factors and inhibits bladder cancer cell and tumor growth. *Exp. Cell Res.* **316**, 2174–2188.

Kernstock, S., Davydova, E., Jakobsson, M., Moen, A., Pettersen, S., Maelandsmo, G. M., Egge-Jacobsen, W., and Falnes, P. O. (2012). Lysine methylation of VCP by a member of a novel human protein methyltransferase family. *Nat. Commun.* **3**, 1038.

Kiezun, A., Garimella, K., Do, R., Stitziel, N. O., Neale, B. M., McLaren, P. J., Gupta, N., Sklar, P., Sullivan, P. F., Moran, J. L., *et al.* (2012). Exome sequencing and the genetic basis of complex traits. *Nat. Genet.* **44**, 623–630.

Kim, Y. J., Kim, H. Y., Lee, J. H., Yu, S. J., Yoon, J. H., Lee, H. S., Kim, C. Y., Cheong, J. Y., Cho, S. W., Park, N. H., *et al.* (2013). A genome-wide association study identified new variants associated with the risk of chronic hepatitis B. *Hum. Mol. Genet.* **22**, 4233–4238.

Klei, L. R., and Barchowsky, A. (2008). Positive signaling interactions between arsenic and ethanol for angiogenic gene induction in human microvascular endothelial cells. *Toxicol. Sci.* **102**, 319–327.

Kligerman, A. D., Doerr, C. L., and Tennant, A. H. (2005). Oxidation and methylation status determine the effects of arsenic on the mitotic apparatus. *Mol. Cell Biochem.* **279**, 113–121.

Kuo, C. C., Weaver, V., Fadrowski, J. J., Lin, Y. S., Guallar, E., and Navas-Acien, A. (2015). Arsenic exposure, hyperuricemia, and gout in US adults. *Environ. Int.* **76C**, 32–40.

Lander, E., and Kruglyak, L. (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat. Genet.* **11**, 241–247.

Lankat-Buttgereit, B., and Tampe, R. (2002). The transporter associated with antigen processing: function and implications in human diseases. *Physiol. Rev.* **82**, 187–204.

Lee, E. T., Welty, T. K., Fabsitz, R., Cowan, L. D., Le, N. A., Oopik, A. J., Cucchiara, A. J., Savage, P. J., and Howard, B. V. (1990). The Strong Heart Study. A study of cardiovascular disease in American Indians: design and methods. *Am. J. Epidemiol.* **132**, 1141–1155.

Lee, S. Y., Williamson, B., Caballero, O. L., Chen, Y. T., Scanlan, M. J., Ritter, G., Jongeneel, C. V., Simpson, A. J., and Old, L. J. (2004). Identification of the gonad-specific anion transporter SLCO6A1 as a cancer/testis (CT) antigen expressed in human lung cancer. *Cancer Immun.* **4**, 13.

Lindberg, A. L., Kumar, R., Goessler, W., Thirumaran, R., Gurzau, E., Koppova, K., Rudnai, P., Leonardi, G., Fletcher, T., and Vahter, M. (2007). Metabolism of low-dose inorganic arsenic in a central European population: influence of sex and genetic polymorphisms. *Environ. Health Perspect.* **115**, 1081–1086.

Maull, E. A., Ahsan, H., Edwards, J., Longnecker, M. P., Navas-Acien, A., Pi, J., Silbergeld, E. K., Styblo, M., Tseng, C. H., Thayer, K. A., *et al.* (2012). Evaluation of the association between arsenic and diabetes: a National Toxicology Program workshop review. *Environ. Health Perspect.* **120**, 1658–1670.

McPeek, M. S., and Sun, L. (2000). Statistical tests for detection of misspecified relationships by use of genome-screen data. *Am. J. Hum. Genet.* **66**, 1076–1094.

Meza, M. M., Yu, L., Rodriguez, Y. Y., Guild, M., Thompson, D., Gandolfi, A. J., and Klimecki, W. T. (2005). Developmentally restricted genetic determinants of human arsenic metabolism: association between urinary methylated arsenic and CYT19 polymorphisms in children. *Environ. Health Perspect.* **113**, 775–781.

Moon, K., Guallar, E., and Navas-Acien, A. (2012). Arsenic exposure and cardiovascular disease: an updated systematic review. *Curr. Atheroscler Rep.* **14**, 542–555.

Moon, K. A., Guallar, E., Umans, J. G., Devereux, R. B., Best, L. G., Francesconi, K. A., Goessler, W., Pollak, J., Silbergeld, E. K., Howard, B. V., *et al.* (2013). Association between exposure to low to moderate arsenic levels and incident cardiovascular disease. A prospective cohort study. *Ann. Intern. Med.* **159**, 649–659.

Navas-Acien, A., Umans, J. G., Howard, B. V., Goessler, W., Francesconi, K. A., Crainiceanu, C. M., Silbergeld, E. K., and Guallar, E. (2009). Urine arsenic concentrations and species excretion patterns in American Indian communities over a 10-year period: the Strong Heart Study. *Environ. Health Perspect.* **117**, 1428–1433.

Ng, S. B., Nickerson, D. A., Bamshad, M. J., and Shendure, J. (2010). Massively parallel sequencing and rare disease. *Hum. Mol. Genet.* **19**, R119–R124.

North, K. E., Howard, B. V., Welty, T. K., Best, L. G., Lee, E. T., Yeh, J. L., Fabsitz, R. R., Roman, M. J., and MacCluer, J. W. (2003). Genetic and environmental contributions to cardiovascular disease risk in American Indians: the Strong Heart Family Study. *Am. J. Epidemiol.* **157**, 303–314.

Okada, Y., Kubo, M., Ohmiya, H., Takahashi, A., Kumasaka, N., Hosono, N., Maeda, S., Wen, W., Dorajoo, R., Go, M. J., *et al.* (2012). Common variants at CDKAL1 and KLF9 are associated with body mass index in east Asian populations. *Nat. Genet.* **44**, 302–306.

Ott, J., Kamatani, Y., and Lathrop, M. (2011). Family-based designs for genome-wide association studies. *Nat. Rev. Genet.* **12**, 465–474.

Pierce, B. L., Kibriya, M. G., Tong, L., Jasmine, F., Argos, M., Roy, S., Paul-Brutus, R., Rahaman, R., Rakibuz-Zaman, M., Parvez, F., *et al.* (2012). Genome-wide association study identifies chromosome 10q24.32 variants associated with arsenic metabolism and toxicity phenotypes in Bangladesh. *PLoS Genet.* **8**, e1002522.

Pierce, B. L., Tong, L., Argos, M., Gao, J., Farzana, J., Roy, S., Paul-Brutus, R., Rahaman, R., Rakibuz-Zaman, M., Parvez, F., *et al.* (2013). Arsenic metabolism efficiency has a causal role in arsenic toxicity: Mendelian randomization and

gene-environment interaction. *Int. J. Epidemiol.* **42**, 1862–1871.

Reddy, M. A., and Natarajan, R. (2013). Role of epigenetic mechanisms in the vascular complications of diabetes. *Subcell Biochem.* **61**, 435–454.

Ren, X., Aleshin, M., Jo, W. J., Dills, R., Kalman, D. A., Vulpe, C. D., Smith, M. T., and Zhang, L. (2010). Involvement of N-6 adenine-specific DNA methyltransferase 1 (N6AMT1) in arsenic biomethylation and its role in arsenic-induced toxicity. *Environ. Health Perspect.* **119**, 771–777.

Roth, M., Obaidat, A., and Hagenbuch, B. (2012). OATPs, OATs and OCTs: the organic anion and cation transporters of the SLCO and SLC22A gene superfamilies. *Br. J. Pharmacol.* **165**, 1260–1287.

Rothman, K. J., and Greenland, S. (2005). Causation and causal inference in epidemiology. *Am. J. Public Health* **95**(S1), S144–S150.

Saad, M., and Wijsman, E. M. (2014). Combining family- and population-based imputation data for association analysis of rare and common variants in large pedigrees. *Genet. Epidemiol.*, **38**, 579–590.

Scheer, J., Findenig, S., Goessler, W., Francesconi, K., Howard, B., Umans, J., Pollak, J., Tellez-Plaza, M., Silbergeld, E., Guallar, E., *et al.* (2012). Arsenic species and selected metals in human urine: validation of HPLC-ICPMS and ICPMS procedures for a long-term population-based epidemiological study. *Anal. Methods* **4**, 406–413.

Schlawicke Engstrom, K., Nermell, B., Concha, G., Stromberg, U., Vahter, M., and Broberg, K. (2009). Arsenic metabolism is influenced by polymorphisms in genes involved in one-carbon metabolism and reduction reactions. *Mutat. Res.* **667**, 4–14.

Smietanski, M., Werner, M., Purta, E., Kaminska, K. H., Stepinski, J., Darzynkiewicz, E., Nowotny, M., and Bujnicki, J. M. (2014). Structural analysis of human 2'-O-ribose methyltransferases involved in mRNA cap structure formation. *Nat. Commun.* **5**, 3004.

Sobel, E., Papp, J. C., and Lange, K. (2002). Detection and integration of genotyping errors in statistical genetics. *Am. J. Hum. Genet.* **70**, 496–508.

Sobreira, N. L., Cirulli, E. T., Avramopoulos, D., Wohler, E., Oswald, G. L., Stevens, E. L., Ge, D., Shianna, K. V., Smith, J. P., Maia, J. M.,*et al.* (2010). Whole-genome sequencing of a single proband together with linkage analysis identifies a Mendelian disease gene. *PLoS Genet.* **6**, e1000991.

Soto-Reyes, E., Del Razo, L. M., Valverde, M., and Rojas, E. (2005). Role of the alkali labile sites, reactive oxygen species and antioxidants in DNA damage induced by methylated trivalent metabolites of inorganic arsenic. *Biometals* **18**, 493–506.

Steinmaus, C., Bates, M. N., Yuan, Y., Kalman, D., Atallah, R., Rey, O. A., Biggs, M. L., Hopenhayn, C., Moore, L. E., Hoang, B. K., *et al.* (2006). Arsenic methylation and bladder cancer risk in case-control studies in Argentina and the United States. *J. Occup. Environ. Med.* **48**, 478–488.

Steinmaus, C., Yuan, Y., Kalman, D., Rey, O. A., Skibola, C. F., Dauphine, D., Basu, A., Porter, K. E., Hubbard, A., Bates, M. N.,*et al.* (2010). Individual differences in arsenic metabolism and lung cancer in a case-control study in Cordoba, Argentina. *Toxicol. Appl. Pharmacol.* **247**, 138–145.

Su, C. T., Lin, H. C., Choy, C. S., Huang, Y. K., Huang, S. R., and Hsueh, Y. M. (2012). The relationship between obesity, insulin and arsenic methylation capability in Taiwan adolescents. *Sci. Total Environ.* **414**, 152–158.

Sun, L., Wilder, K., and McPeek, M. S. (2002). Enhanced pedigree error detection. *Hum. Hered.* **54**, 99–110.

Sundberg, B. E., Waag, E., Jacobsson, J. A., Stephansson, O., Rumaks, J., Svirskis, S., Alsio, J., Roman, E., Ebendal, T., Klusa, V., *et al.* (2008). The evolutionary history and tissue mapping of amino acid transporters belonging to solute carrier families SLC32, SLC36, and SLC38. *J. Mol. Neurosci.* **35**, 179–193.

Tchounwou, P. B., Yedjou, C. G., and Dorsey, W. C. (2003). Arsenic trioxide-induced transcriptional activation of stress genes and expression of related proteins in human liver carcinoma cells (HepG2). *Cell. Mol. Biol.* **49**, 1071–1079.

Tellez-Plaza, M., Gribble, M. O., Voruganti, V. S., Francesconi, K. A., Goessler, W., Umans, J. G., Silbergeld, E. K., Guallar, E., Franceschini, N., North, K. E., *et al.* (2013). Heritability and preliminary genome-wide linkage analysis of arsenic metabolites in urine. *Environ. Health Perspect.* **121**, 345–351.

Thomas, D. J. (2009). Unraveling arsenic-glutathione connections. *Toxicol. Sci.* **107**, 309–311.

Toyohara, T., Suzuki, T., Morimoto, R., Akiyama, Y., Souma, T., Shiwaku, H. O., Takeuchi, Y., Mishima, E., Abe, M., Tanemoto, M., *et al.* (2009). SLCO4C1 transporter eliminates uremic toxins and attenuates hypertension and renal inflammation. *J. Am. Soc. Nephrol.* **20**, 2546–2555.

Tseng, C. H. (2009). A review on environmental factors regulating arsenic methylation in humans. *Toxicol. Appl. Pharmacol.* **235**, 338–350.

Vahter, M. (2000). Genetic polymorphism in the biotransformation of inorganic arsenic and its role in toxicity. *Toxicol. Lett.* **112–113**, 209–217.

Vahter, M. (2002). Mechanisms of arsenic biotransformation. *Toxicology* **181–182**, 211–217.

Valenzuela, O. L., Drobna, Z., Hernandez-Castellanos, E., Sanchez-Pena, L. C., Garcia-Vargas, G. G., Borja-Aburto, V. H., Styblo, M., and Del Razo, L. M. (2009). Association of AS3MT polymorphisms and the risk of premalignant arsenic skin lesions. *Toxicol. Appl. Pharmacol.* **239**, 200–207.

Wood, T. C., Salavagionne, O. E., Mukherjee, B., Wang, L., Klumpp, A. F., Thomae, B. A., Eckloff, B. W., Schaid, D. J., Wieben, E. D., and Weinshilboum, R. M. (2006). Human arsenic methyltransferase (AS3MT) pharmacogenetics: gene resequencing and functional genomics studies. *J. Biol. Chem.* **281**, 7364–7373.

Wu, F., Molinaro, P., and Chen, Y. (2014). Arsenic exposure and subclinical endpoints of cardiovascular diseases. *Curr. Environ. Health Rep.* **1**, 148–162.

Yu, R. C., Hsu, K. H., Chen, C. J., and Froines, J. R. (2000). Arsenic methylation capacity and skin cancer. *Cancer Epidemiol. Biomark. Prev.* **9**, 1259–1262.