OXFORD

Occupational Health

Somatic loss of the Y chromosome is associated with arsenic exposure among Bangladeshi men

Kathryn Demanelis,[1,2] **Dayana A Delgado,**[2] **Lin Tong,**[2] **Farzana Jasmine,**[2] **Alauddin Ahmed,**[5] **Tariqul Islam,**[5] **Faruque Parvez,**[6] **Muhammad G Kibriya,**[2] **Joseph H Graziano,**[6] **Habibul Ahsan,**[2,3,4,7] **and Brandon L Pierce**[2,3,4]*

[1]Division of Hematology/Oncology, Department of Medicine, University of Pittsburgh, Pittsburgh, PA, USA, [2]Department of Public Health Sciences, The University of Chicago, Chicago, IL, USA, [3]Department of Human Genetics, The University of Chicago, Chicago, IL, USA, [4]Comprehensive Cancer Center, The University of Chicago, Chicago, IL, USA, [5]UChicago Research Bangladesh, Dhaka, Bangladesh, [6]Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY, USA, [7]Departments of Medicine and Human Genetics, The University of Chicago, Chicago, IL, USA

*Corresponding author. Department of Public Health Sciences, The University of Chicago, 5841 South Maryland Avenue, W264, MC2000, Chicago, IL 60637, USA. E-mail: brandonpierce@uchicago.edu

Abstract

Background: Arsenic exposure increases the risk of several cancers in humans and contributes to genomic instability. Somatic loss of the Y chromosome (LoY) is a potential biomarker of genomic instability and cancer risk. Smoking is associated with LoY, but few other carcinogens have been investigated. We tested the cross-sectional association between arsenic exposure and LoY in leukocytes among genotyped Bangladeshi men (age 20–70 years) from the Health Effects of Arsenic Longitudinal Study.

Methods: We extracted the median of logR-ratios from probes on the Y chromosome (mLRR-chrY) from genotyping arrays ($n = 1364$) and estimated the percentage of cells with LoY (% LoY) from mLRR-chrY. We evaluated the association between arsenic exposure (measured in drinking water and urine) and LoY using multivariable linear and logistic regression models. The association between LoY and incident arsenic-induced skin lesions was also examined.

Results: Ten percent of genotyped men had LoY in at least 5% of cells and % LoY increased with age. Among men randomly selected for genotyping ($n = 778$), higher arsenic in drinking water, arsenic consumed and urinary arsenic were associated with increased % LoY ($P = 0.006$, $P = 0.06$ and $P = 0.13$, respectively). LoY was associated with increased risk of incident skin lesions ($P = 0.008$).

Conclusion: Arsenic exposure was associated with increased LoY, providing additional evidence that arsenic contributes to genomic instability. LoY was associated with developing skin lesions, a risk factor for cancer, suggesting that LoY may be a biomarker of

susceptibility in arsenic-exposed populations. The effect of arsenic on somatic events should be further explored in cancer-prone tissue types.

Key words: Arsenic, Y chromosome loss, skin lesions, copy number variation

---

Key Messages

- Somatic loss of the Y chromosome (LoY) is one of the most common acquired genetic abnormalities in human blood cells and a potential biomarker of genomic instability and ageing in men.
- Among arsenic-exposed Bangladeshi men, we observe an association between arsenic exposure (measured in drinking water and urine) and increased LoY in peripheral blood.
- In this cohort, LoY is also associated with increased risk of incident (subsequent) arsenic-induced skin lesions, an indicator of arsenic toxicity and risk factor for cancer.
- This is the first study linking arsenic, a modifiable and carcinogenic environmental exposure, to an acquired genetic abnormality in a non-cancerous human tissue.

Introduction

Consumption of drinking water contaminated with inorganic arsenic, a known carcinogen, is a serious public health issue in many countries around the world, including Bangladesh, Argentina and the USA.[1] Arsenic exposure increases the risk of cancer, specifically cancers of the skin,[2] prostate,[3] liver,[3,4] lung[3,5,6] and bladder.[5,7,8] In addition, arsenic exposure is associated with increased risk of age-related chronic diseases, including cardiovascular diseases (CVD),[9,10] non-malignant respiratory diseases,[11] diabetes mellitus,[12–14] impaired cognitive function,[15] as well as increased overall mortality.[16]

Although arsenic is a known environmental carcinogen, its mechanisms of carcinogenicity remain to be fully elucidated. Arsenic may exert its carcinogenic effects via epigenetic dysregulation, increased cellular inflammation and oxidative stress, and inhibition of DNA repair[17,18]—mechanisms that also likely contribute to genomic instability and ageing.[19] Long-term arsenic exposure can cause the development of arsenic-induced skin lesions.[20] A dose–response relationship between arsenic exposure and these skin lesions is well established.[21–23] The presence of skin lesions is associated with increased risk of subsequent cancers,[24] suggesting that skin-lesion susceptibility can inform our understanding of the effect of arsenic on carcinogenesis and ageing.

Somatic loss of chromosome Y (LoY) in men is one of the most common acquired (i.e. somatic) genetic abnormalities known to occur in humans. The percentage of blood cells affected by LoY increases with age and LoY in blood cells may be a biomarker of ageing, immune surveillance and genomic instability in men.[25] LoY is frequently observed in hematologic cancers[26] and myelodysplastic

syndrome[27] and is a potential risk factor for non-hematologic cancers. Among Swedish men, LoY was associated with a 4-fold increased risk of cancer[28] and a similar association between LoY and cancer was identified among UK Biobank participants, specifically for stomach, lung, skin and prostate cancers.[29] LoY was associated with prevalent CVD and diabetes, as well as early mortality among UK Biobank participants.[30] LoY has been shown to be associated with cigarette smoking in a dose-dependent manner,[31] suggesting that LoY may be a biomarker of carcinogen-induced genomic damage and cancer susceptibility. In addition, LoY was associated with higher exposure to air pollution[32] and polycyclic aromatic hydrocarbons.[33] To date, the relationship between LoY and environmental arsenic exposure has not been assessed.

In this study, we utilize log-R ratios (LRRs), generated from genotyping array intensity data from probes binding to the Y chromosome, to estimate LoY among Bangladeshi men participating in the Health Effects of Arsenic Longitudinal Study (HEALS). Using baseline whole-blood samples and baseline inorganic arsenic exposure measurements (prior to arsenic mitigation efforts), we characterize the association between exposure to arsenic, a known environmental carcinogen, and somatic LoY in blood. We also examine LoY in relation to risk of incident and prevalent skin lesions and all-cause mortality.

Methods

Study participants

HEALS[34] is a population-based prospective cohort study designed to assess the effects of arsenic exposure on health

in Araihazar, Bangladesh. From 2000 to 2002, 11 746 participants between 18 and 75 years of age were recruited, providing both blood and urine samples at baseline. A total of 10 970 wells in the study area were tested for arsenic and individuals were assigned arsenic exposure values based on the arsenic concentration in their reported primary drinking well(s) and in their urine. Follow-up visits are conducted every 2 years for the entire cohort. The study protocol was approved by the Institutional Review Boards of the University of Chicago, Columbia University and the Bangladesh Medical Research Council. Informed consent was obtained from all participants.

Arsenic exposure assessment

Arsenic in drinking water was measured using graphite furnace atomic absorption spectrometry (and inductively coupled plasma–mass spectrometry for samples with concentrations below the limit of detection).[35] At baseline, participants reported their primary drinking well and the concentration of that well was assigned as their water arsenic exposure. Daily arsenic dose (µg/day) was calculated by multiplying the well water arsenic concentration of the primary well (µg/day) by the self-reported volume of water consumed daily from that well (L/day). If participants also regularly drank from a secondary well (or wells), information from that well was included in the estimation of daily arsenic dose.

Total arsenic concentration in urine (collected at baseline) was measured using a graphite furnace atomic absorption spectrometry method with a very low limit of detection (~1 µg/L) in a single laboratory at Columbia University.[36] Creatinine in urine was measured using a colorimetric diagnostics kit (Sigma, St. Louis, MO, USA) and this measure was used in regression models examining urinary arsenic to adjust for hydration status.[37]

Assessment of skin lesions and mortality

At baseline, HEALS study physicians, trained to detect and diagnose skin lesions, recorded the following conditions: melanosis (hyperpigmentation), leucomelanosis (hypopigmentation) or keratosis (hyperkeratotic thickening of skin). Participants were classified as having a baseline lesion if any one of these three conditions was present at baseline and were classified as having an incident lesion if any of the three conditions was present at any follow-up visit. Vital status was determined at each biennial follow-up interview and date of death ascertained primarily from relatives of deceased participants, with essentially no loss to follow-up.[16,38] A verbal autopsy procedure was used to investigate and assign the causes of death (previously

validated in a Bangladeshi population[39–42]). For this study, we use data on incident skin-lesion cases and deaths occurring from 2002 to 2018.

Participant selection and genotyping

Only male HEALS participants with existing genome-wide array data generated using DNA from baseline peripheral blood were included in this study (Figure 1). Men from genotyping Batches 1 and 2 were genotyped using Illumina's HumanCytoSNP-12 v2.1 array. Men from genotyping Batches 3 and 4 were genotyped using Illumina's Multiethnic Global array. DNA extraction and genotyping quality control have been described previously.[43,44] HEALS participants selected for genotyping (in prior studies)[44–48] belonged to one of four groups: (i) randomly selected from a cohort independently of skin-lesion or disease status, (ii) randomly selected from skin-lesion-free participants, (iii) selected based on incident skin lesions or (iv) selected on baseline or incident cardiovascular disease or chronic respiratory disease. All men analysed have been genetically confirmed to be men based on X chromosome single-nucleotide polymorphism (SNP) data.[43,44] Samples that failed to pass genotyping quality control for inclusion in genome-wide studies of arsenic metabolism have also been excluded in our analysis.

Detecting somatic loss of the Y chromosome

The Illumina HumanCytoSNP-12 v2.1 and Multiethnic Global arrays measure a total of 2972 and 1598 SNPs, respectively, in the male-specific region of the Y chromosome (chrY) [2 694 521 to 59 034 049 (hg19/GRCh37) and 2 781 480 to 56 887 902 (GRCh38)]. The LRR is a normalized measure of signal intensity for each SNP included on the array and it is the log-2 of the ratio between the observed and expected signal intensity from two copies of the genome. From these arrays, we obtained the LRRs to calculate the median LRR from chrY (mLRR-chrY), median LRR from chromosome 1 (mLRR-chr1) and SD of LRRs from chromosome 1 (sdLRR-chr1). For men, mLRR-chrYs cluster near 0 (one copy of chrY), indicating that chrY is present in most blood cells, whereas negative mLRR-chrYs correspond to an increasing proportion of cells that lack a chrY (i.e. lower mLRR-chrY indicates more LoY among cells). We removed any samples that had a sdLRR-chr1 > 0.28 (a quality control filter recommended by Illumina).[49] Batch 2 had a high proportion (21%) of samples that failed this criterion compared with the other batches (and higher sdLRR-chr1 across all samples compared with the other batches) and the average mLRR-chrY deviated negatively from zero, indicating the presence of
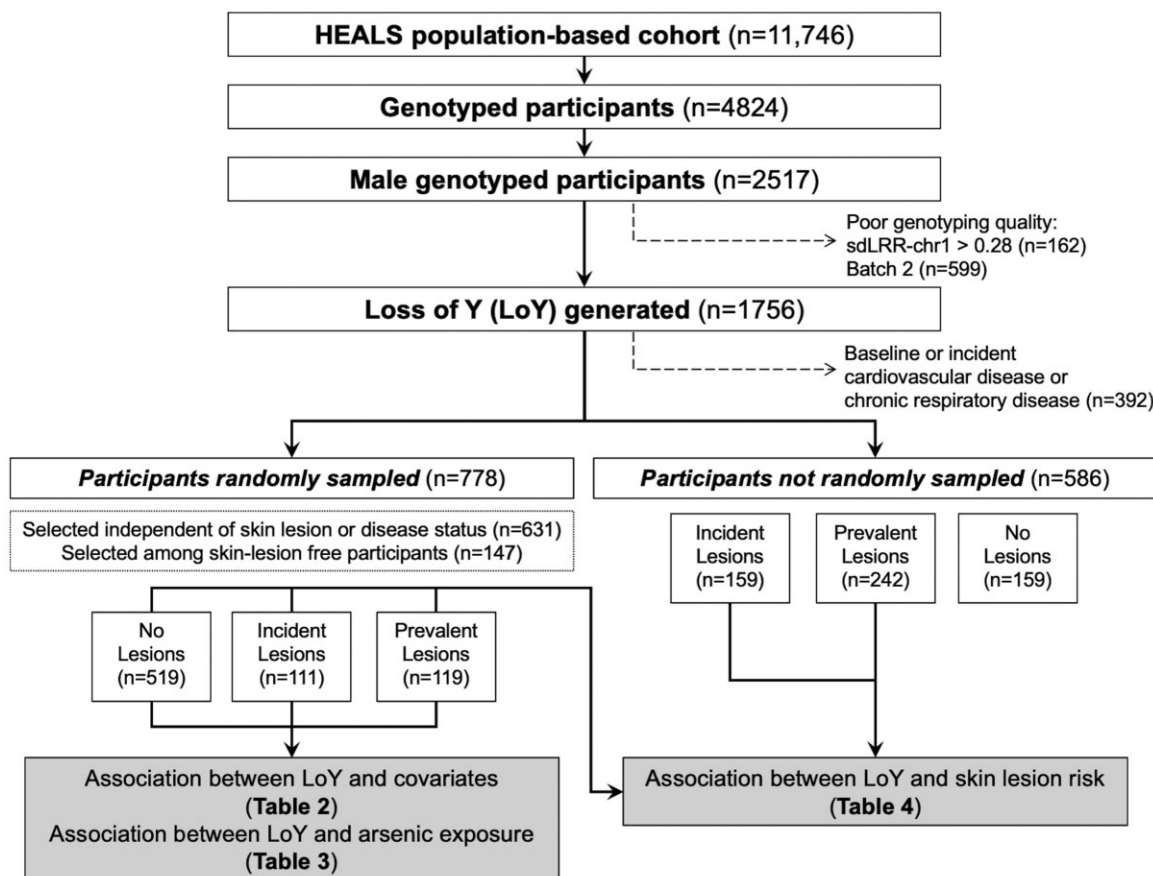
Figure 1 A description of male genotyped HEALS participants in the study and selection for analyses. Frequencies may not sum to the total due to small amounts of missing data. HEALS, Health Effects of Arsenic Longitudinal Study; LoY, loss of Y chromosome; sdLRR-chr1, standard deviation of LRR from chromosome 1

bias in the LRRs (Supplementary Figure S1, available as Supplementary data at *IJE* online). Therefore, all men genotyped in Batch 2 ($n = 758$) were excluded. After visual inspection of chrY LRRs, we observed evidence of an inherited ∼3-kilobase deletion on chrY carried by <2% of genotyped men in our cohort (Supplementary Figure S2, available as Supplementary data at *IJE* online). To reduce the bias of this deletion (and other copy number variable regions on chrY or poorly performing probes) on mLRR-chrY, we excluded chrY probes whose LRRs had an absolute mean >1, absolute skewness >5 or SD >0.4 within each batch (Supplementary Figure S3, available as Supplementary data at *IJE* online). These exclusions eliminated <400 probes within each batch, including the probes affected by the inherited deletion (Supplementary Table S1, available as Supplementary data at *IJE* online).

We calculated mLRR-chrY from the LRRs of the remaining chrY probes. Within each genotyping batch, we corrected for batch effects by extracting the local regression median obtained from a kernel density estimation of the unadjusted mLRR-chrY distribution (using *density* function in R) and adjusting each mLRR-chrY by this

median, as previously described (Figure 2).[28,31,50,51] mLRR-chrY was inversely correlated with sdLRR-1 (Supplementary Figure S4, available as Supplementary data at *IJE* online) and positively correlated with mLRR-chr1 in all batches (Supplementary Figure S5, available as Supplementary data at *IJE* online), indicating that some of the variance in mLRR-chrY is due to the effect of genome-wide genotyping quality and variation in LRRs, respectively. Both mLRR-chr1 and sdLRR-chr1 were adjusted for as covariates in all statistical analyses of LoY. We converted mLRR-chrYs to estimated percentage of cells with LoY using the formula derived by Danielsson *et al.*: LoY $(\%) = 100 \times [1 - 2^{(2 \times \text{mLRR-chrY})}]$.[51] Based on the percentage of cells with LoY, we constructed binary LoY variables based on the percentage of cells with LoY above the following thresholds: 5% and 10%.

Statistical analysis

Prior to analysis, we removed participants selected for genotyping based on baseline or incident cardiovascular disease or chronic respiratory disease to avoid potential bias
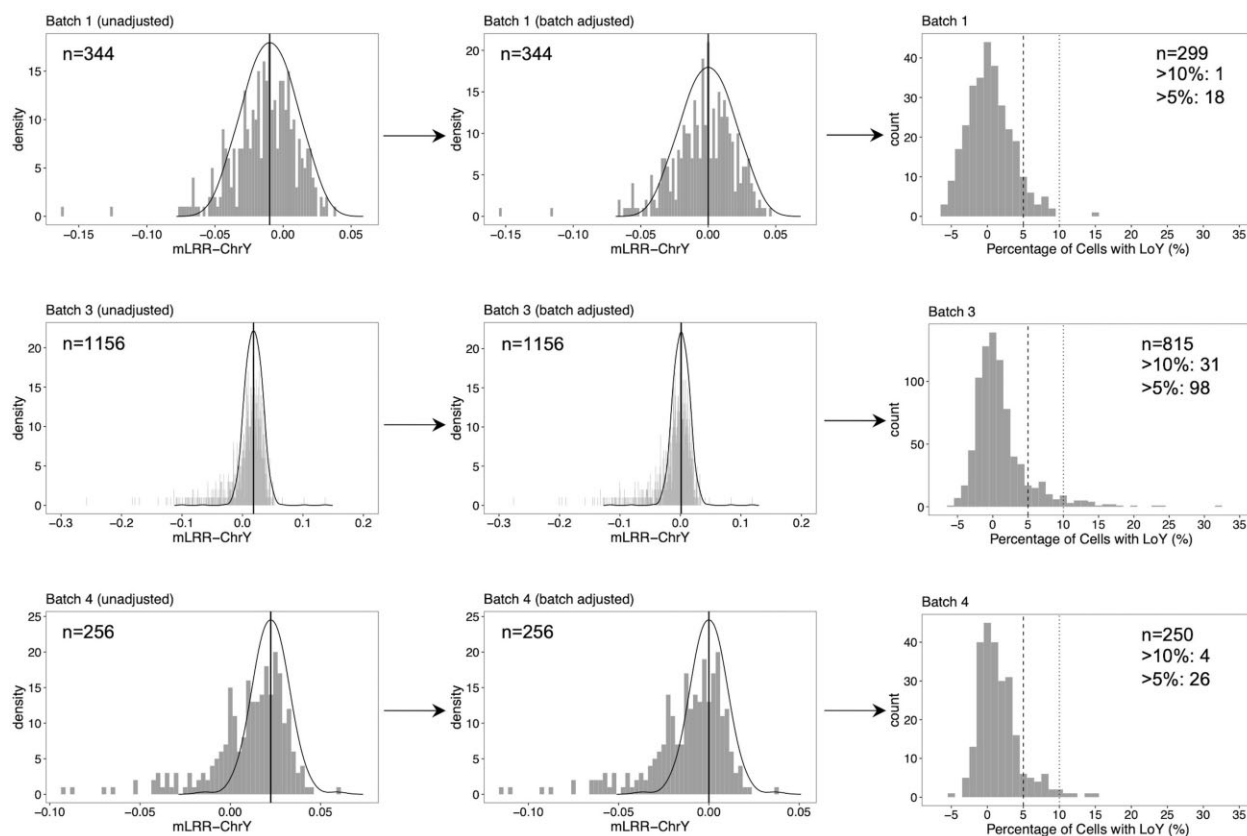
Figure 2 Distribution of unadjusted mLRR-chrY (left), batch-adjusted mLRR-chrY (middle) and estimated percentage of cells with LoY (right) within each genotyping batch. The observed distributions of unadjusted mLRR-chrY (left) and batch-adjusted mLRR-chrY (middle) are shown as grey bars. The distribution of experimental variation in mLRR-chrY in the absence of LoY is modelled (solid line). Batch-adjusted (i.e. centred) mLRR-chrY values (in middle) are obtained by adjusting the unadjusted mLRR-chrY by the local regression median (see 'Methods' section). The distribution of the estimated percentage of cells with LoY (right) are shown and lines correspond to the boundaries for 5% and 10% thresholds. These estimated percentages are derived from the batch-adjusted mLRR-chrY values using the formula: $(1 - 2^{2*mLRRchrY}) \times 100$ (see 'Methods' section). mLRR-chrY, median of LRR from Y chromosome; LoY, loss of Y chromosome

that might occur due to this selection on disease (i.e. selection bias, collider bias).[52] Among the randomly selected genotyped participants, we examined the association between categorical and ordinal quartiles of arsenic exposure (water, daily dose and urinary arsenic) and LoY using linear (for percentage of cells with LoY) and logistic (for thresholds of 5% or 10% or more cells with LoY) regression. Among randomly selected genotyped participants, we performed covariate-adjustment standardization on urinary arsenic[53] in which we (i) adjusted the log-transformed urinary creatinine by age and categorical body mass index (BMI) status using linear regression, (ii) calculated the ratio of observed creatinine to adjusted creatinine (exponentiated) and (iii) divided urinary arsenic by this ratio (expressed as µg/L).To examine LoY and skin-lesion risk, we defined two case groups: (i) incident skin-lesion cases diagnosed during follow-up and (ii) prevalent lesions diagnosed at baseline, both including those occurring within the random sample. These two case groups were compared with a skin-lesion-free group among the randomly selected

genotyped men who did not have or develop skin lesions (i.e. control group). We tested the association between LoY and skin-lesion status using logistic regression. To test the association between LoY and mortality, we compared individuals who died during follow-up to individuals who were alive at the most recent follow-up visit using logistic regression. For all models, we examined two sets of covariate adjustments: (i) age, sdLRR-chr1, mLRR-chr1, deletion status and genotyping batch; and (ii) age, sdLRR-chr1, mLRR-chr1, deletion status, genotyping batch, BMI quartiles (categorical), smoking (categorical: never, former, current) and education status (categorical: 0, 1–5, 6–16 years). Models examining urinary arsenic were adjusted for urinary creatinine. All analyses were conducted in R 4.1.2.

Results

LoY determination and quality control

After performing quality control on the array data, we determined mLRR-chrY in 1756 male HEALS participants.

After removing samples from men selected for genotyping based on cardiovascular or respiratory disease, there were 1364 (78%) men eligible for inclusion in our analysis and 778 (57%) of these men were randomly selected from HEALS for genotyping. mLRR-chrY was inversely correlated with age in all genotyping batches ($P < 0.05$). Age was not associated with mLRR obtained from other chromosomes for any batch (e.g. $P > 0.05$ for chromosome 1) (Supplementary Figure S6, available as Supplementary data at *IJE* online). Among the participants included in this analysis ($n = 1364$), LoY was detected in at least 5% and 10% of cells for 10% ($n = 142$) and 3% ($n = 36$) of males (Figure 2).

Relationship between LoY and participant characteristics

The median age of our randomly selected sample of men was 40 years (range: 20–70 years). Most of these men were current smokers (64%). Among all samples, 270 (20%) and 361 (26%) men had incident or prevalent lesions, respectively (Table 1). Among randomly selected genotyped men, the percentage of cells with LoY increased with age across all genotyping batches (Supplementary Figure S7, available as Supplementary data at *IJE* online). Consistently with prior studies, we observed that non-smokers had a lower percentage of cells with LoY compared with former ($P = 0.043$) smokers, but this

Table 1 Characteristics of genotyped male HEALS participants included in analyses of LoY

	Randomly selected[a]	Randomly selected and skin-lesion-free[b]	Incident skin lesions	Prevalent skin lesions
N	778	519	270	361
Age (years)				
<30	86 (11.1%)	70 (13.5%)	7 (2.6%)	16 (4.4%)
30–39	282 (36.2%)	204 (39.3%)	55 (20.4%)	92 (25.5%)
40–49	241 (31.0%)	147 (28.3%)	106 (39.3%)	136 (37.7%)
50–59	135 (17.4%)	83 (16.0%)	78 (28.9%)	93 (25.8%)
>59	34 (4.4%)	15 (2.9%)	24 (8.9%)	24 (6.6%)
Body mass index (kg/m^2)				
<17.5	195 (25.1%)	120 (23.1%)	72 (26.9%)	111 (31.0%)
17.5–18.8	193 (24.8%)	116 (22.4%)	77 (28.7%)	92 (25.7%)
18.9–21.1	195 (25.1%)	141 (27.2%)	62 (23.1%)	94 (26.3%)
>21.1	194 (25.0%)	142 (27.4%)	57 (21.3%)	61 (17.0%)
Cigarette smoking				
Never	207 (26.6%)	152 (29.3%)	54 (20.0%)	63 (17.5%)
Former	76 (9.8%)	43 (8.3%)	47 (17.4%)	59 (16.3%)
Current	495 (63.6%)	324 (62.4%)	169 (62.6%)	239 (66.2%)
Education (years)				
0	281 (36.1%)	187 (36.0%)	102 (37.8%)	175 (48.5%)
1–5	258 (33.2%)	164 (31.6%)	95 (35.2%)	109 (30.2%)
6–16	239 (30.7%)	168 (32.4%)	73 (27.0%)	77 (21.3%)
Water arsenic (µg/L)				
<17	195 (25.1%)	150 (28.9%)	59 (21.9%)	51 (14.1%)
17–71	191 (24.6%)	142 (27.4%)	40 (14.8%)	80 (22.2%)
72–153	196 (25.2%)	132 (25.4%)	64 (23.7%)	98 (27.1%)
>153	196 (25.2%)	95 (18.3%)	107 (39.6%)	132 (36.6%)
Arsenic dose (µg/day)				
<37	195 (25.1%)	150 (28.9%)	58 (21.5%)	46 (12.7%)
37–169	194 (24.9%)	151 (29.1%)	46 (17.0%)	73 (20.2%)
170–419	194 (24.9%)	118 (22.7%)	70 (25.9%)	110 (30.5%)
>419	195 (25.1%)	100 (19.3%)	96 (35.6%)	132 (36.6%)
Urinary arsenic (µg/L)				
<50	190 (24.8%)	136 (26.4%)	47 (17.9%)	67 (18.7%)
50–103	192 (25.1%)	145 (28.2%)	55 (20.9%)	67 (18.7%)
104–196	191 (25.0%)	129 (25.0%)	73 (27.8%)	114 (31.8%)
>196	192 (25.1%)	105 (20.4%)	88 (33.5%)	111 (30.9%)

[a]Frequencies may not sum to the total due to small amounts of missing data. Randomly selected genotyped men may include participants who have or develop skin lesions (see 'Methods' section).

[b]Randomly selected and skin-lesion-free group were considered our control group and includes randomly selected genotyped men who did not have or develop skin lesions (i.e. free of skin lesions). HEALS, Health Effects of Arsenic Longitudinal Study; LoY, loss of Y chromosome.

association was attenuated after adjustment for age, genotyping batch and other covariates (Table 2). Normal BMI (third quartile) was associated with a lower percentage of cells with LoY compared with underweight BMI (first quartile) before and after adjustment ($P = 0.0094$ and $P = 0.050$, respectively). The risk of having $\geq 5\%$ cells with LoY increased with age and with former smoking ($P = 0.023$), but associations with smoking were attenuated after covariate adjustment.

Arsenic exposure is associated with increased percentage of cells with LoY

Over 80% of men included in our analysis ($n = 1109$) consumed arsenic-contaminated water with arsenic concentrations above the WHO action level of 10 μg/L. Among randomly selected genotyped men, increasing arsenic exposure was associated with increased percentage of cells with LoY (Table 3). Compared with the lowest quartile, the highest quartile of water arsenic and arsenic dose were associated with an increased percentage of cells with LoY ($P = 0.0054$ and $P = 0.010$, respectively). The third quartile of urinary arsenic was associated with an increased percentage of cells with LoY compared with the lowest quartile ($P = 0.011$). The percentage of cells with LoY increased by 0.24 (95% CI: 0.06, 0.42) and 0.16 (95% CI: –0.02, 0.34) across water arsenic quartiles ($P = 0.0056$) and arsenic dose quartiles ($P = 0.061$). When we stratified by genotyping batch, we confirmed that the relationship between arsenic exposure and the percentage of cells with LoY was not strongly driven by any single genotyping batch (Supplementary Table S2, available as Supplementary data at *IJE* online). Risk of having >10% of cells with LoY increased across quartiles of water arsenic by 1.93-fold (95% CI: 0.96, 3.88).

LoY and risk of arsenic-induced skin lesions

A one-unit increase in the percentage of cells with LoY was associated with increased risk of incident skin lesions [odds ratio (OR) = 1.10; 95% CI: 1.03, 1.18] after adjustment (Table 4). Stratifying by genotyping batch, the percentage

Table 2 Association between LoY and selected covariates among the randomly selected genotyped male HEALS participants ($n = 778$)

| | | | Percentage of cells with LoY (%) | | | | | >5% of cells | | | |
| | | | Unadjusted model | | Full model[a] | | | Unadjusted model | | Full model[a] | |
	n	Mean (SD)	β (SE)	*P*	Beta (SE)	*P*	LoY	OR (95% CI)	*P*	OR (95% CI)	*P*
Age (years)											
<30	86	–0.13 (3.03)	Ref.	–	Ref.	–	5	1.00	–	1.00	–
30–39	282	0.58 (3.30)	0.71 (0.41)	0.083	0.40 (0.33)	0.24	17	1.04 (0.37, 2.90)	0.94	0.53 (0.13, 2.13)	0.37
40–49	241	1.04 (3.18)	1.17 (0.42)	0.0053	0.89 (0.35)	0.011	21	1.55 (0.56, 4.24)	0.40	0.89 (0.23, 3.47)	0.87
50–59	135	1.61 (3.82)	1.74 (0.46)	<0.001	1.76 (0.39)	<0.001	17	2.33 (0.83, 6.58)	0.11	2.32 (0.59, 9.23)	0.23
>59	34	2.16 (3.54)	2.29 (0.68)	<0.001	1.97 (0.55)	<0.001	8	4.98 (1.50, 16.57)	0.0088	4.66 (0.97, 22.32)	0.054
Body mass index (kg/m^2)											
<17.5	195	1.39 (3.78)	Ref.	–	Ref.	–	23	1.00	–	1.00	–
17.5–18.8	193	0.86 (3.18)	–0.53 (0.34)	0.12	–0.29 (0.27)	0.28	16	0.68 (0.35, 1.32)	0.25	0.85 (0.37, 1.94)	0.70
18.9–21.1	195	0.50 (2.81)	–0.89 (0.34)	0.0094	–0.53 (0.27)	0.050	11	0.45 (0.21, 0.94)	0.035	0.51 (0.20, 1.34)	0.17
>21.1	194	0.82 (3.63)	–0.58 (0.34)	0.093	–0.31 (0.28)	0.27	18	0.76 (0.40, 1.47)	0.42	1.02 (0.43, 2.42)	0.97
Cigarette smoking status											
Never	207	0.60 (3.36)	Ref.	–	Ref.	–	14	1.00	–	1.00	–
Former	76	1.52 (3.50)	0.92 (0.45)	0.043	0.10 (0.37)	0.78	12	2.58 (1.14, 5.88)	0.023	2.81 (0.88, 8.93)	0.081
Current	495	0.92 (3.36)	0.32 (0.28)	0.25	0.19 (0.23)	0.41	42	1.28 (0.68, 2.39)	0.44	2.11 (0.82, 5.43)	0.12
Education (years)											
0	281	1.14 (3.58)	Ref.	–	Ref.	–	29	1.00	–	1.00	–
1–5	258	0.72 (3.25)	–0.42 (0.29)	0.15	–0.14 (0.23)	0.55	19	0.69 (0.38, 1.26)	0.23	0.81 (0.38, 1.71)	0.58
6–16	239	0.79 (3.27)	–0.35 (0.30)	0.24	–0.20 (0.24)	0.42	20	0.79 (0.44, 1.44)	0.45	0.84 (0.37, 1.88)	0.67

Results from linear and logistic regression models are presented for participant covariates on percentage of cells with LoY (continuous outcome) and threshold based on percentage of cells with LoY (i.e. 5%, binary outcome), respectively.

[a]All models were adjusted for age, genotyping batch, sdLRR-chr1, mLRR-chr1, deletion status, smoking status, education status and BMI category. For categorical age, model includes categorical age instead of age. HEALS, Health Effects of Arsenic Longitudinal Study; LoY, loss of Y chromosome; OR, odds ratio; sdLRR-chr1, standard deviation of LRR from chromosome 1; mLRR-chr1, median of LRR from chromosome 1; BMI, body mass index.

Table 3 Association between arsenic exposure and LoY within randomly selected genotyped male HEALS participants ($n = 778$)

	n	Percentage of cells with LoY (%)			>5% of cells			>10% of cells		
		Mean (SD)	β (SE)	P	LoY	OR (95% CI)	P	LoY	OR (95% CI)	P
Water arsenic (µg/L)										
<17	195	0.39 (3.18)	Ref.	–	13	1.00	–	3	1.00	–
17–71	191	0.78 (3.25)	0.42 (0.27)	0.11	11	1.14 (0.41, 3.13)	0.80	2	0.72 (0.05, 10.87)	0.81
72–153	196	1.06 (3.35)	0.55 (0.27)	0.043	25	2.90 (1.16, 7.23)	0.022	2	0.36 (0.03, 5.04)	0.45
>153	196	1.35 (3.66)	0.75 (0.27)	0.0054	19	1.54 (0.60, 3.93)	0.37	9	5.38 (0.68, 42.42)	0.11
Ordinal[b]			0.24 (0.09)	0.0056		1.22 (0.92, 1.62)	0.18		1.93 (0.96, 3.88)	0.066
Arsenic dose (µg/day)										
<37	195	0.35 (3.18)	Ref.	–	12	1.00	–	3	1.00	–
37–169	194	1.29 (3.58)	0.92 (0.27)	<0.001	20	2.82 (1.10, 7.19)	0.030	4	1.77 (0.16, 19.08)	0.64
170–419	194	0.74 (3.22)	0.45 (0.27)	0.093	19	2.36 (0.90, 6.21)	0.082	1	0.19 (0.01, 4.34)	0.30
>419	195	1.21 (3.46)	0.70 (0.27)	0.010	17	1.30 (0.47, 3.54)	0.61	8	6.08 (0.68, 54.59)	0.11
Ordinal[b]			0.16 (0.09)	0.061		1.03 (0.77, 1.37)	0.85		1.69 (0.82, 3.48)	0.16
Urinary arsenic (covariate adjusted, µg/L)[a]										
<56	191	0.42 (3.28)	Ref.	–	15	1.00	–	2	1.00	–
56–98	191	0.97 (3.45)	0.57 (0.27)	0.033	16	1.46 (0.59, 3.63)	0.41	4	9.15 (0.75, 111.18)	0.082
99–168	191	1.10 (3.44)	0.69 (0.27)	0.011	19	1.69 (0.68, 4.18)	0.26	4	8.56 (0.69, 106.50)	0.095
>168	191	0.95 (3.30)	0.40 (0.27)	0.14	15	0.77 (0.30, 2.00)	0.59	6	9.87 (0.79, 123.69)	0.076
Ordinal[b]			0.13 (0.09)	0.13		0.95 (0.71, 1.26)	0.70		1.76 (0.88, 3.52)	0.11

Results from linear and logistic regression models are presented for categorical arsenic quartile exposure on percentage of cells with LoY (continuous outcome) and thresholds based on percentage of cells with LoY (i.e. 5% and 10%, binary outcome), respectively. All models were adjusted for age, genotyping batch, sdLRR-chr1, mLRR-chr1, deletion status, smoking status, education status and BMI category.

[a]Urinary arsenic was corrected for urinary creatinine using a covariate-adjustment approach. Creatinine was adjusted for age and BMI category. See 'Methods' section.

[b]Arsenic quartile exposures treated as an ordinal predictor in statistical models. Association estimate corresponds to a unit increase in the arsenic quartile. HEALS, Health Effects of Arsenic Longitudinal Study; LoY, loss of Y chromosome; OR, odds ratio; sdLRR-chr1, standard deviation of LRR from chromosome 1; mLRR-chr1, median of LRR from chromosome 1; BMI, body mass index.

of cells with LoY showed evidence of association with increased risk of incident skin lesions in Batches 1 and 3 (Supplementary Table S3, available as Supplementary data at *IJE* online). We did not observe an association between the percentage of cells with LoY and increased risk of prevalent skin lesions overall or within any genotyping batch (Supplementary Table S4, available as Supplementary data at *IJE* online). Men with >5% of cells with LoY were more than twice as likely to have incident skin lesions (OR = 2.58; 95% CI: 1.29, 5.18) compared with men with <5% of cells with LoY after adjustment for age, daily arsenic dose and other covariates. For both incident and prevalent skin lesions, there was no association between having ≥10% cells with LoY and risk of these lesions, possibly due to the low prevalence within our study population. Given that the presence of skin lesions is associated with mortality among Bangladeshi men, we examined mortality among our samples and observed mortality increased with percentage of cells with LoY (OR = 1.08; 95% CI: 1.03, 1.13) and having ≥5% cells with LoY (OR = 2.07, 95% CI: 1.22, 3.51); however, these associations were no longer present after adjusting for age, genotyping batch and other

covariates (Supplementary Table S5, available as Supplementary data at *IJE* online).

Discussion

In this study of arsenic-exposed Bangladeshi men, we observed an association between arsenic exposure assessed in both urine and drinking water and increased risk of somatic LoY in peripheral blood. Although arsenic is known to affect processes that promote carcinogenesis and contribute to genomic instability (e.g. DNA repair, methylation), this is the first study to link arsenic exposure to LoY, a potential biomarker of genomic instability, in an epidemiological study. LoY is the most common acquired genetic alteration and frequently observed form of clonal mosaicism known to occur in humans, and this study provides further evidence that LoY can be induced by modifiable carcinogenic environmental exposures. We also found suggestive evidence that LoY was associated with increased risk of developing arsenical skin lesions, an indicator of arsenic toxicity and a risk factor for cancer.[24]

Table 4 Association between LoY and skin-lesion risk among male genotyped HEALS participants

	Controls	Cases	Unadjusted model		Age + batch model[a]		Full model[b]		Full + arsenic model[c]	
			OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P
Incident lesions										
n	519	270								
% cells with LoY	0.62 (3.40)	1.33 (3.64)	1.06 (1.01, 1.10)	0.0080	1.09 (1.02, 1.17)	0.011	1.09 (1.02, 1.16)	0.015	1.10 (1.03, 1.18)	0.0077
>5% of cells	38 (7.3%)	39 (14.4%)	2.14 (1.33, 3.43)	0.0017	2.32 (1.19, 4.53)	0.014	2.31 (1.17, 4.56)	0.016	2.58 (1.29, 5.18)	0.0075
>10% of cells	10 (1.9%)	7 (2.6%)	1.35 (0.51, 3.60)	0.54	0.60 (0.18, 1.99)	0.41	0.65 (0.19, 2.15)	0.48	0.76 (0.23, 2.56)	0.66
Prevalent lesions										
n	519	361								
% cells with LoY	0.62 (3.40)	2.08 (5.12)	1.11 (1.07, 1.16)	<0.001	1.04 (0.97, 1.10)	0.27	1.03 (0.97, 1.09)	0.36	1.03 (0.97, 1.10)	0.38
>5% of cells	38 (7.3%)	46 (12.7%)	1.85 (1.18, 2.91)	0.0078	0.77 (0.40, 1.49)	0.44	0.72 (0.37, 1.42)	0.35	0.68 (0.34, 1.34)	0.26
>10% of cells	10 (1.9%)	14 (3.9%)	2.05 (0.90, 4.68)	0.087	0.53 (0.18, 1.55)	0.24	0.55 (0.18, 1.63)	0.28	0.51 (0.17, 1.53)	0.23

Results from logistic regression models are presented for the percentage of cells with LoY (% cells with LoY, continuous predictor) and thresholds based on the percentage of cells with LoY (i.e. 5% and 10%, binary predictor) on incident and prevalent lesion risk. Odds ratios (ORs) and 95% CIs are reported.

[a]Model adjusted for age, genotyping batch, sdLRR-chr1, mLRR-chr1 and deletion status.

[b]Model adjusted for age, genotyping batch, sdLRR-chr1, mLRR-chr1, deletion status, smoking status, education status and BMI category.

[c]Model adjusted for age, genotyping batch, sdLRR-chr1, mLRR-chr1, deletion status, smoking status, education status, BMI category and daily arsenic dose (quartiles). HEALS, Health Effects of Arsenic Longitudinal Study; LoY, loss of Y chromosome; sdLRR-chr1, standard deviation of LRR from chromosome 1; mLRR-chr1, median of LRR from chromosome 1; BMI, body mass index.

Although this is the first human study to report an association between arsenic exposure and LoY, arsenic compounds have been shown in vitro to induce genomic instability and aneuploidy in human fibroblasts,[54] Chinese hamster cells[55] and human lung epithelial and keratinocytes.[56] The mechanisms of arsenic-induced carcinogenesis are not well understood but oxidative stress, inflammation and epigenetic alterations have been implicated as potentially important mechanisms.[57] Prior studies of tissues from arsenic-exposed individuals from Argentina and Chile point to genome instability as an observable component of arsenic-induced carcinogenesis. Among arsenic-exposed Andean women, urinary arsenic was associated with increased copy number alterations in blood.[58] Higher arsenic exposure was also associated with higher levels of chromosomal instability in bladder tumours[59] and numerous copy number alterations in squamous cell lung tumors[60] diagnosed in Argentinian and Chilean patients. Arsenic-exposed individuals with skin lesions from West Bengal had higher levels of chromosomal aberrations compared with exposed individuals without lesions.[61]

LoY in blood may be an important biomarker to utilize to better understand the potential carcinogenicity of environmental exposures. Although the link between LoY in blood and cancer risk remains to be elucidated, LoY in blood may be a biomarker of prior carcinogenic exposure or indicate impaired immune cells that cannot properly remove precancerous cells from tissues.[29] Cigarette smoking—a mixture of numerous carcinogens, including arsenic—increases risk of LoY.[31] Among older US men (>65 years), PM10, a Group I carcinogen and air pollutant, was associated with increased leukocyte LoY.[32] Higher exposure to carcinogenic polycyclic aromatic hydrocarbons (PAHs) was also associated with increased LoY in middle-aged male coke oven workers.[33] Although PAHs are directly genotoxic and arsenic is not, their associations with LoY indicate that these exposures may contribute to genomic instability in blood.

The association of arsenic exposure with both LoY and cancer risk[2,5–8,24] could be attributable to one or more proposed hypotheses.[28,31] The first is that exposure may cause multiple chromosomal abnormalities, including LoY, and LoY serves as a biomarker of exposure-induced chromosomal damage and/or instability. These genetic abnormalities (including LoY in itself) could contribute to carcinogenesis and increase cancer risk. In this scenario, associations between LoY and subsequent health outcomes are not necessarily causal, as LoY may be a proxy for other chromosomal abnormalities that promote carcinogenesis. Furthermore, LoY in blood may be a proxy for LoY (and other abnormalities) in other tissue types if exposures have common effects across multiple tissues. Future studies need

to examine the extent to which LoY is shared across human tissues and whether environmental exposures contribute to that sharing. The second hypothesis is that LoY in blood is an indicator (and potentially a cause) of impaired immune function.[28] Interestingly, there is ample evidence indicating that arsenic exposure can lead to impaired immune responses.[62] Thus, exposures that increase LoY in leukocytes and reduce immune surveillance could increase cancer risk by reducing the tumour suppressor activities of the immune system. This reduced immune surveillance could likely have implications for a wide array of health conditions. Additional research is needed to characterize the relationship between LoY and immune function and the connection to health outcomes.

Our study has several limitations. Many of our participants were aged <60 years and the frequency of LoY in cells is expected to be lower in younger men. In addition, there may be residual and unmeasured confounding related to unmeasured exposures, lifestyle factors, technical genotyping artefacts and other factors. Although the water from many of the wells with arsenic-contaminated groundwater in Bangladesh may also have high levels of other toxic metals,[63] it did not appear that the water from wells in the Araihazar area also contained high levels of these metals that could modestly contribute to the observed association between arsenic and LoY. Within HEALS, a small study found that urinary cadmium, lead and other metal levels (besides arsenic) were low in HEALS and within US reference ranges, indicating that this population is not likely co-exposed to other toxic metals at high levels.[64] It is important to acknowledge that this is a single measurement of LoY taken at baseline and technical artefacts related to genotyping and cell-type distribution could contribute to measurement error in LoY estimates. Our study cannot evaluate the impact of arsenic on the longitudinal change in LoY. We analysed the association between arsenic exposure and LoY among randomly selected genotyped men to minimize selection biases. Analyses of skin-lesion risk and LoY are also not likely to be influenced by selection bias given the case–cohort design of our genotyping study in which we attempted to genotype all skin-lesion cases. Although our LoY estimates were obtained from genotyping array data, prior studies have shown that LoY estimated from genotyping arrays is highly concordant with estimates from next-generation sequencing[28] and qPCR.[65] Additionally, the estimated proportion (3%) of men included in our analysis that had ≥10% cells with LoY in their blood is consistent with a prior estimate that 5% of men at age 50 years have >10% of cells with LoY.[66] Our study's strengths included a broad range of arsenic exposure taken at baseline that also reflects a relatively constant exposure, especially when measured in water, multiple approaches for assessing arsenic exposure and a large sample of arsenic-exposed individuals with genotyping data.

In conclusion, we provide evidence that arsenic exposure, assessed in both urine and drinking water, is associated with increased LoY in peripheral blood among Bangladeshi men. LoY showed suggestive evidence of an association with increased risk of incident skin lesions—a hallmark of chronic arsenic exposure and risk factor for cancer in arsenic-exposed individuals. Based on these findings, we consider LoY to be a promising candidate biomarker of the effect of arsenic exposure and susceptibility to arsenic toxicity.

Ethics approval

The study protocol was approved by the Institutional Review Boards of the University of Chicago, Columbia University and the Bangladesh Medical Research Council. Informed consent was obtained from all participants.

Data availability

R scripts used for the analyses in this article are available upon request. Individual-level genomic data cannot currently be made publicly available.

Supplementary data

Supplementary data are available at *IJE* online.

Author contributions

K.D., D.A.D. and B.L.P. conceived and designed the study. A.A., T.I., F.P., M.G.K., J.H.G., H.A. and B.L.P. developed and conducted the genetic association study of arsenic metabolism within the HEALS cohort. F.J. and M.G.K. generated genotyping data for all samples in this analysis. L.T. performed quality control and processing of genotyping data. K.D., D.A.D., L.T. and B.L.P. contributed to the statistical analyses in the study. K.D. and B.L.P. wrote the manuscript. All authors contributed to the revision and review of the manuscript.

Funding

Acknowledgements

Conflict of interest

None declared.

References

1. Naujokas MF, Anderson B, Ahsan H *et al.* The broad scope of health effects from chronic arsenic exposure: update on a worldwide public health problem. *Environ Health Perspect* 2013;**121**: 295–302.

2. Karagas MR, Gossai A, Pierce B, Ahsan H. Drinking water arsenic contamination, skin lesions, and malignancies: a systematic review of the global evidence. *Curr Environ Health Rep* 2015;**2**: 52–68.

3. García-Esquinas E, Pollán M, Umans JG *et al.* Arsenic exposure and cancer mortality in a US-based prospective cohort: the strong heart study. *Cancer Epidemiol Biomarkers Prev* 2013;**22**: 1944–53.

4. Wang W, Cheng S, Zhang D. Association of inorganic arsenic exposure with liver cancer mortality: a meta-analysis. *Environ Res* 2014;**135**:120–25.

5. Lamm SH, Ferdosi H, Dissen EK, Li J, Ahn J. A systematic review and meta-regression analysis of lung cancer risk and inorganic arsenic in drinking water. *Int J Environ Res Public Health* 2015;**12**:15498–515.

6. Celik I, Gallicchio L, Boyd K *et al.* Arsenic in drinking water and lung cancer: a systematic review. *Environ Res* 2008;**108**:48–55.

7. Gamboa-Loira B, Cebrian ME, Franco-Marina F, Lopez-Carrillo L. Arsenic metabolism and cancer risk: a meta-analysis. *Environ Res* 2017;**156**:551–58.

8. Saint-Jacques N, Parker L, Brown P, Dummer TJ. Arsenic in drinking water and urinary tract cancers: a systematic review of 30 years of epidemiological evidence. *Environ Health* 2014;**13**: 44.

9. Moon KA, Oberoi S, Barchowsky A *et al.* A dose-response meta-analysis of chronic arsenic exposure and incident cardiovascular disease. *Int J Epidemiol* 2017;**46**:1924–39.

10. Navas-Acien A, Sharrett AR, Silbergeld EK *et al.* Arsenic exposure and cardiovascular disease: a systematic review of the epidemiologic evidence. *Am J Epidemiol* 2005;**162**:1037–49.

11. Sanchez TR, Powers M, Perzanowski M, George CM, Graziano JH, Navas-Acien A. A Meta-analysis of Arsenic Exposure and Lung Function: Is There Evidence of Restrictive or Obstructive Lung Disease?. *Curr Envir Health Rpt* 2018;**5**:244–54.

12. Maull EA, Ahsan H, Edwards J *et al.* Evaluation of the association between arsenic and diabetes: a national toxicology program workshop review. *Environ Health Perspect* 2012;**120**: 1658–70.

13. Thayer KA, Heindel JJ, Bucher JR, Gallo MA. Role of environmental chemicals in diabetes and obesity: a National Toxicology Program workshop review. *Environ Health Perspect* 2012;**120**: 779–89.

14. Navas-Acien A, Silbergeld EK, Streeter RA, Clark JM, Burke TA, Guallar E. Arsenic exposure and type 2 diabetes: a systematic review of the experimental and epidemiological evidence. *Environ Health Perspect* 2006;**114**:641–48.

15. Rodriguez-Barranco M, Lacasana M, Aguilar-Garduno C *et al.* Association of arsenic, cadmium and manganese exposure with neurodevelopment and behavioural disorders in children: a systematic review and meta-analysis. *Sci Total Environ* 2013; **454-455**:562–77.

16. Argos M, Kalra T, Rathouz PJ *et al.* Arsenic exposure from drinking water, and all-cause and chronic-disease mortalities in Bangladesh (HEALS): a prospective cohort study. *Lancet* 2010; **376**:252–58.

17. Chen QY, DesMarais T, Costa M. Metals and mechanisms of carcinogenesis. *Annu Rev Pharmacol Toxicol* 2019;**59**:537–54.

18. Tam LM, Price NE, Wang Y. Molecular mechanisms of arsenic-induced disruption of DNA repair. *Chem Res Toxicol* 2020;**33**: 709–26.

19. Lopez-Otin C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. *Cell* 2013;**153**:1194–217.

20. Byrd DM, Roegner ML, Griffiths JC *et al.* Carcinogenic risks of inorganic arsenic in perspective. *Int Arch Occup Environ Health* 1996;**68**:484–94.

21. Yoshida T, Yamauchi H, Fan Sun G. Chronic health effects in people exposed to arsenic via the drinking water: dose-response relationships in review. *Toxicol Appl Pharmacol* 2004;**198**: 243–52.

22. Ahsan H, Chen Y, Parvez F *et al.* Arsenic exposure from drinking water and risk of premalignant skin lesions in Bangladesh: baseline results from the Health Effects of Arsenic Longitudinal Study. *Am J Epidemiol* 2006;**163**:1138–48.

23. Argos M, Kalra T, Pierce BL *et al.* A prospective study of arsenic exposure from drinking water and incidence of skin lesions in Bangladesh. *Am J Epidemiol* 2011;**174**:185–94.

24. Hsu LI, Chen GS, Lee CH *et al.* Use of arsenic-induced palmoplantar hyperkeratosis and skin cancers to predict risk of subsequent internal malignancy. *Am J Epidemiol* 2013;**177**:202–12.

25. Forsberg LA. Loss of chromosome Y (LOY) in blood cells is associated with increased risk for disease and mortality in aging men. *Hum Genet* 2017;**136**:657–63.

26. Cantú ES, Moses MD, Nemana LJ, Pierre RV. Sex chromosome loss in adults with haematological neoplasms. *Br J Haematol* 2015;**169**:899–901.

27. Ouseph MM, Hasserjian RP, Dal Cin P *et al.* Genomic alterations in patients with somatic loss of the Y chromosome as the sole cytogenetic finding in bone marrow cells. *Haematologica* 2021;**106**:555–64.

28. Forsberg LA, Rasi C, Malmqvist N *et al.* Mosaic loss of chromosome Y in peripheral blood is associated with shorter survival and higher risk of cancer. *Nat Genet* 2014;**46**:624–28.

29. Loftfield E, Zhou W, Yeager M, Chanock SJ, Freedman ND, Machiela MJ. Mosaic Y loss is moderately associated with solid tumor risk. *Cancer Res* 2019;**79**:461–66.

30. Loftfield E, Zhou W, Graubard BI *et al.* Predictors of mosaic chromosome Y loss and associations with mortality in the UK Biobank. *Sci Rep* 2018;**8**:12316.

31. Dumanski JP, Rasi C, Lonn M *et al.* Mutagenesis: smoking is associated with mosaic loss of chromosome Y. *Science* 2015;**347**: 81–83.

32. Wong JYY, Margolis HG, Machiela M *et al.* Outdoor air pollution and mosaic loss of chromosome Y in older men from the Cardiovascular Health Study. *Environ Int* 2018;**116**:239–47.

33. Liu Y, Bai Y, Wu X *et al.* Polycyclic aromatic hydrocarbons exposure and their joint effects with age, smoking, and TCL1A variants on mosaic loss of chromosome Y among coke-oven workers. *Environ Pollut* 2020;**258**:113655.

34. Ahsan H, Chen Y, Parvez F *et al.* Health Effects of Arsenic Longitudinal Study (HEALS): description of a multidisciplinary epidemiologic investigation. *J Expo Sci Environ Epidemiol* 2006;**16**:191–205.

35. Cheng Z, Zheng Y, Mortlock R, Van Geen A. Rapid multi-element analysis of groundwater by high-resolution inductively coupled plasma mass spectrometry. *Anal Bioanal Chem* 2004;**379**:512–18.

36. Nixon DE, Mussmann GV, Eckdahl SJ, Moyer TP. Total arsenic in urine: palladium-persulfate vs nickel as a matrix modifier for graphite furnace atomic absorption spectrophotometry. *Clin Chem* 1991;**37**:1575–79.

37. Nermell B, Lindberg AL, Rahman M *et al.* Urinary arsenic concentration adjustment factors and malnutrition. *Environ Res* 2008;**106**:212–18.

38. Pierce BL, Kalra T, Argos M *et al.* A prospective study of body mass index and mortality in Bangladesh. *Int J Epidemiol* 2010;**39**:1037–45.

39. Chowdhury ME, Botlero R, Koblinsky M, Saha SK, Dieltiens G, Ronsmans C. Determinants of reduction in maternal mortality in Matlab, Bangladesh: a 30-year cohort study. *Lancet* 2007;**370**:1320–28.

40. Hurt LS, Ronsmans C, Saha S. Effects of education and other socioeconomic factors on middle age mortality in rural Bangladesh. *J Epidemiol Community Health* 2004;**58**:315–20.

41. Ronsmans C, Vanneste AM, Chakraborty J, van Ginneken J. Decline in maternal mortality in Matlab, Bangladesh: a cautionary tale. *Lancet* 1997;**350**:1810–14.

42. Ronsmans C, Vanneste AM, Chakraborty J, Van Ginneken J. A comparison of three verbal autopsy methods to ascertain levels and causes of maternal deaths in Matlab, Bangladesh. *Int J Epidemiol* 1998;**27**:660–66.

43. Pierce BL, Tong L, Argos M *et al.* Arsenic metabolism efficiency has a causal role in arsenic toxicity: Mendelian randomization and gene-environment interaction. *Int J Epidemiol* 2013;**42**:1862–71.

44. Pierce BL, Kibriya MG, Tong L *et al.* Genome-wide association study identifies chromosome 10q24.32 variants associated with arsenic metabolism and toxicity phenotypes in Bangladesh. *PLoS Genet* 2012;**8**:e1002522.

45. Howe CG, Liu X, Hall MN *et al.* Associations between blood and urine arsenic concentrations and global levels of post-translational histone modifications in Bangladeshi men and women. *Environ Health Perspect* 2016;**124**:1234–40.

46. Gamble MV, Liu X, Ahsan H *et al.* Folate, homocysteine, and arsenic metabolism in arsenic-exposed individuals in Bangladesh. *Environ Health Perspect* 2005;**113**:1683–88.

47. Chen Y, Wu F, Liu M *et al.* A prospective study of arsenic exposure, arsenic methylation capacity, and risk of cardiovascular disease in Bangladesh. *Environ Health Perspect* 2013;**121**:832–38.

48. Niedzwiecki MM, Liu X, Zhu H *et al.* Serum homocysteine, arsenic methylation, and arsenic-induced skin lesion incidence in Bangladesh: a one-carbon metabolism candidate gene study. *Environ Int* 2018;**113**:133–42.

49. Forsberg LA, Halvardson J, Rychlicka-Buniowska E *et al.* Mosaic loss of chromosome Y in leukocytes matters. *Nat Genet* 2019;**51**:4–7.

50. Dumanski JP, Lambert JC, Rasi C *et al.*; European Alzheimer's Disease Initiative Investigators. Mosaic loss of chromosome Y in blood is associated with Alzheimer Disease. *Am J Hum Genet* 2016;**98**:1208–19.

51. Danielsson M, Halvardson J, Davies H *et al.* Longitudinal changes in the frequency of mosaic chromosome Y loss in peripheral blood cells of aging men varies profoundly between individuals. *Eur J Hum Genet* 2020;**28**:349–57.

52. Pearce N, Richiardi L. Commentary: Three worlds collide: Berkson's bias, selection bias and collider bias. *Int J Epidemiol* 2014;**43**:521–24.

53. O'Brien KM, Upson K, Cook NR, Weinberg CR. Environmental chemicals in urine and blood: improving methods for creatinine and lipid adjustment. *Environ Health Perspect* 2016;**124**:220–27.

54. Yih LH, Ho IC, Lee TC. Sodium arsenite disturbs mitosis and induces chromosome loss in human fibroblasts. *Cancer Res* 1997;**57**:5051–59.

55. Sciandrello G, Barbaro R, Caradonna F, Barbata G. Early induction of genetic instability and apoptosis by arsenic in cultured Chinese hamster cells. *Mutagenesis* 2002;**17**:99–103.

56. Ganapathy S, Liu J, Xiong R, Yu T, Makriyannis A, Chen C. Chronic low dose arsenic exposure preferentially perturbs mitotic phase of the cell cycle. *Genes Cancer* 2019;**10**:39–51.

57. Martinez VD, Vucic EA, Becker-Santos DD, Gil L, Lam WL. Arsenic exposure and the induction of human cancers. *J Toxicol* 2011;**2011**:431287.

58. Wojdacz TK, Bottai M, Vahter M, Broberg K. Exposure to arsenic and intra-chromosomal instability in blood. *Metallomics* 2014;**6**:1387–89.

59. Moore LE, Smith AH, Eng C *et al.* Arsenic-related chromosomal alterations in bladder cancer. *J Natl Cancer Inst* 2002;**94**:1688–96.

60. Martinez VD, Buys TP, Adonis M *et al.* Arsenic-related DNA copy-number alterations in lung squamous cell carcinomas. *Br J Cancer* 2010;**103**:1277–83.

61. Banerjee M, Sarma N, Biswas R, Roy J, Mukherjee A, Giri AK. DNA repair deficiency leads to susceptibility to develop arsenic-induced premalignant skin lesions. *Int J Cancer* 2008;**123**:283–87.

62. Dangleben NL, Skibola CF, Smith MT. Arsenic immunotoxicity: a review. *Environ Health* 2013;**12**:73.

63. Frisbie SH, Ortega R, Maynard DM, Sarkar B. The concentrations of arsenic and other toxic elements in Bangladesh's drinking water. *Environ Health Perspect* 2002;**110**:1147–53.

64. Sanchez TR, Slavkovich V, LoIacono N *et al.* Urinary metals and metal mixtures in Bangladesh: exploring environmental sources in the Health Effects of Arsenic Longitudinal Study (HEALS). *Environ Int* 2018;**121**:852–60.

65. Zhou W, Machiela MJ, Freedman ND *et al.* Mosaic loss of chromosome Y is associated with common variation near TCL1A. *Nat Genet* 2016;**48**:563–68.

66. Baliakas P, Forsberg LA. Chromosome Y loss and drivers of clonal hematopoiesis in myelodysplastic syndrome. *Haematologica* 2021;**106**:329–31.