Environment International 191 (2024) 108955

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

Full length article

An epigenome-wide study of selenium status and DNA methylation in the Strong Heart Study

Wil Lieberman-Cribbin [a,*], Arce Domingo-Relloso [a,b], Ronald A. Glabonjat [a], Kathrin Schilling [a], Shelley A. Cole [c], Marcia O'Leary [d], Lyle G. Best [d], Ying Zhang [e], Amanda M. Fretts [f], Jason G. Umans [g,h], Walter Goessler [i], Ana Navas-Acien [a], Maria Tellez-Plaza [j], Allison Kupsco [a]

[a] Department of Environmental Health Sciences, Columbia University Mailman School of Public Health, New York, NY, USA
[b] Department of Biostatistics, Columbia University Mailman School of Public Health, New York, NY, USA
[c] Population Health Program, Texas Biomedical Research Institute, San Antonio, TX, USA
[d] Missouri Breaks Industries Research, Cheyenne River Sioux Tribe, Eagle Butte, SD 57625, USA
[e] Department of Biostatistics and Epidemiology, The University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA
[f] Department of Epidemiology, School of Public Health, University of Washington, Seattle, WA, USA
[g] MedStar Health Research Institute, Hyattsville, MD, USA
[h] Georgetown-Howard Universities Center for Clinical and Translational Science, Washington, DC, USA
[i] Institute of Chemistry, University of Graz, Graz, Austria
[j] Department of Chronic Diseases Epidemiology, National Center for Epidemiology, Instituto de Salud Carlos III, 28029, Madrid, Spain

ARTICLE INFO

Handling Editor: Adrian Covaci

Keywords:
Selenium
Epigenetics
DNA methylation
American Indian populations
Strong Heart Study

ABSTRACT

Background: Selenium (Se) is an essential nutrient linked to adverse health endpoints at low and high levels. The mechanisms behind these relationships remain unclear and there is a need to further understand the epigenetic impacts of Se and their relationship to disease. We investigated the association between urinary Se levels and DNA methylation (DNAm) in the Strong Heart Study (SHS), a prospective study of cardiovascular disease (CVD) among American Indians adults.

Methods: Selenium concentrations were measured in urine (collected in 1989–1991) using inductively coupled plasma mass spectrometry among 1,357 participants free of CVD and diabetes. DNAm in whole blood was measured cross-sectionally using the Illumina MethylationEPIC BeadChip (850 K) Array. We used epigenome-wide robust linear regressions and elastic net to identify differentially methylated cytosine-guanine dinucleotide (CpG) sites associated with urinary Se levels.

Results: The mean (standard deviation) urinary Se concentration was 51.8 (25.1) μg/g creatinine. Across 788,368 CpG sites, five differentially methylated positions (DMP) (hypermethylated: cg00163554, cg18212762, cg11270656, and hypomethylated: cg25194720, cg00886293) were significantly associated with Se in linear regressions after accounting for multiple comparisons (false discovery rate p-value: 0.10). The top hyper-methylated DMP (cg00163554) was annotated to the Disco Interacting Protein 2 Homolog C (*DIP2C*) gene, which relates to transcription factor binding. Elastic net models selected 425 hypo- and hyper-methylated DMPs associated with urinary Se, including three sites (cg00163554 [DIP2C], cg18212762 [*MAP4K2*], cg11270656 [*GPIHBP1*]) identified in linear regressions.

Conclusions: Urinary Se was associated with minimal changes in DNAm in adults from American Indian communities across the Southwest and the Great Plains in the United States, suggesting that other mechanisms may be driving health impacts. Future analyses should explore other mechanistic biomarkers in human populations, determine these relationships prospectively, and investigate the potential role of differentially methylated sites with disease endpoints.

Abbreviations: CVD, Cardiovascular disease; DNAm, DNA methylation; DMP, Differentially methylated position; eGFR, Estimated glomerular filtration rate (eGFR); EWAS, Epigenome-wide association study; FDR, False discovery rate; IRB, Institutional review board; NHANES, National Health and Nutrition Examination Survey; SHFS, Strong Heart Family Study; SHS, Strong Heart Study.

* Corresponding author at: 722 W 168th St, 11th Floor Rm 1107, New York, NY 10032.
E-mail address: wfl2112@cumc.columbia.edu (W. Lieberman-Cribbin).

https://doi.org/10.1016/j.envint.2024.108955
Received 26 January 2024; Received in revised form 19 June 2024; Accepted 13 August 2024
Available online 14 August 2024

1. Introduction

Selenium (Se) is an essential nutrient and naturally occurring element commonly found in rocks and soils (US Department of Health Human Services, 2003). Selenium present in air and water mainly originates from the burning of coal and fossil fuels, while Se intake primarily comes from diet (US Department of Health Human Services, 2003). The physiological functions of Se are achieved by proteins incorporating Se, termed selenoproteins, which play key roles as antioxidants and transporters of Se (Rayman, 2012; Speckmann and Grune, 2015). Normal plasma Se levels among adults have been cited as 60 to 150 μg/L (Smith et al., 2017), with a recommended dietary allowance of 55 μg Se/day in the USA (Fairweather-Tait et al., 2011). Notably, Se has a very narrow physiological range, with both deficient and excess Se associated with adverse health outcomes. Deficiencies in Se have been associated with cancer, immune disorders, thyroid disease, and cardiovascular disease (Rayman, 2012), while excessive Se and selenoproteins have been linked to elevated risk of hyperglycemia (Oo et al., 2018) and diabetes (Ogawa-Wong et al., 2016; Bleys and Navas-Acien, 2007).

American Indian communities have an increased burden of chronic metal exposures and cardiovascular disease (CVD) compared to the general U.S. population (Lee et al., 1990; Pang et al., 2016). In the Strong Heart Study (SHS) and Strong Heart Family Study (SHFS), exposure to environmental metals such as Se, arsenic and cadmium has been associated with an increased risk for CVD (Moon et al., 2013; Moon et al., 2018; Pichler et al., 2019; Oliver-Williams et al., 2018; Tellez-Plaza et al., 2013; Zhao et al., 2022). In the SHS, dietary intake is the primary source of Se, and some communities reside in areas with high concentrations of Se in soil, resulting in elevated intake (68–724 μg/day) compared to the general population (Zhao et al., 2022). Environmental levels of Se and toxic metals are also influenced by the legacy of mining operations in American Indian communities (Lewis et al., 2017), where elevated Se concentrations have been found in plants and animals downstream of uranium mines (Khamkhash et al., 2017; Muscatello and Belknap, 2008; Sharmasarkar, 2002), highlighting the need to study the health impacts of Se in American Indian communities.

Given the complex role of Se in cellular homeostasis and metabolic disorders, a greater understanding of the molecular impacts of elevated Se is necessary. Research suggests that DNA methylation (DNAm) may serve as a biomarker of Se's impacts on human health. Exposure to environmental metals has been associated with changes in DNAm (Bozack et al., 2020; Tellez-Plaza, 2014), and alterations in DNAm have been associated with several disease outcomes including CVD, diabetes, and obesity (Castillo-Díaz et al., 2010). However, the extent of the epigenetic impacts of Se remains unclear and the majority of research on this topic has been performed in murine and cell-based models (Speckmann and Grune, 2015), conveying the need to study the influences of Se on DNAm in human populations. Mechanistically, Se is proposed to affect DNAm through altering one-carbon metabolism (Speckmann and Grune, 2015). Se has been shown to affect global DNAm, as well as inhibit DNA methyltransferase activity and expression, but the mechanisms through which Se affects DNAm are proposed to differ according to Se deficiency or excess, and warrant further research (Jabłońska and Reszka, 2017; Davis et al., 2000).

In the SHS, DNAm has been linked to a variety of adverse health outcomes including lymphatic–hematopoietic cancers (Domingo-Relloso et al., 2021), liver cancer (Slowly et al., 2024), coronary heart disease (Navas-Acien et al., 2021), lung function (Domingo-Relloso et al., 2022), and insulin resistance (Dye et al., 2023). Previous research has also highlighted the link of environmental metals such as arsenic and cadmium to DNAm signatures (Bozack et al., 2020; Domingo-Relloso et al., 2020) and epigenetic aging (Boyer et al., 2023), although the impacts of Se are unknown. DNAm data available in the SHS provide an opportunity to evaluate the epigenetic impacts of Se in a population with potentially elevated Se status, with the potential to explain the influences of Se on chronic disease.

The objective of this study was to investigate the relationship between urinary Se levels and DNAm of American Indian adults in an epigenome-wide association study (EWAS), to identify biological pathways of genes associated with DNAm sites relevant to Se status, and to identify CpGs that may be useful for the development of future biomarkers to identify Se exposure and nutritional status. We hypothesized that Se would be associated with methylation levels at CpG sites in SHS participants, particularly in genes encoding selenoproteins and pathways related to oxidative stress.

2. Methods

2.1. Study population

The SHS is a prospective cohort of 4,549 American Indian adults from Arizona, Oklahoma, North Dakota, and South Dakota assessing CVD and its risk factors (Lee et al., 1990). The SHS protocol was approved by institutional review boards (IRBs), participating tribes, and the respective area Indian Health Service IRBs. All participants provided written, informed consent. At the Phase 1 baseline exam (1989–1991), adults 45–74 years of age were invited to participate and the participation rate was 62 % (Lee et al., 1990; Welty et al., 1995; Stoddart et al., 2000). Participants were evaluated again in study visits at Phase 2 (1993–1995) and Phase 3 (1997–1999). Subsequently, 1,032 participants from one community were not included in this study upon their request. In the current analysis, participants were eligible if they were free of diabetes at baseline, a known factor altering the epigenome (Nilsson et al., 2014) and related to Se status (Huang et al., 2022), had sufficient urine for metal analyses, and provided sufficient DNA for analyses. This resulted in a total of 1,357 participants. The data were collected, analyzed, and reported under agreements made with the sovereign tribal nations that have partnered in this research, which preclude routine modes of data sharing. Requests to access the dataset from qualified researchers trained in human subject confidentiality protocols may be sent to the Strong Heart Study Coordinating Center at https://strongheartstudy.org/. Requests will be reviewed by tribal research partners before data may be released.

Detailed methods on the anthropometric measurements, biospecimen collection (e.g., urine), and the laboratory measurements of relevant biomarkers have been described extensively (Tellez-Plaza et al., 2013; Bozack et al., 2020; Domingo-Relloso et al., 2022). Sociodemographic information was collected from standardized interviews at each study visit including age, sex, body mass index (BMI; kg/m^2), years of schooling/education, and smoking status (never/former/current). Participants reporting never smoking regularly or smoking $<$ 100 cigarettes during their lifetime were classified as never smokers. Former smokers were defined as smoking $>$ 100 cigarettes and not smoking currently, and current smokers were defined as smoking $\geq$ 100 cigarettes and currently smoking. We calculated estimated glomerular filtration rate (eGFR) using age, sex, and urinary creatinine (mg/dL) via the 2009 Chronic Kidney Disease – Epidemiology Collaboration formula (CKD-Epi) (Levey et al., 2009). Diabetes status was defined as meeting one of the following criteria: fasting glucose greater or equal than 126 mg/dL, 2-hour post-load plasma glucose greater or equal than 200 mg/dL, HbA1c greater or equal than 6.5 %, using insulin or an oral hypoglycemic agent (Diagnosis and classification of diabetes mellitus, 2011).

2.2. Urine selenium measurements

Urine Se measures were available at Phase 1, concurrent to DNAm data. Participants provided a morning spot urine sample, and samples were stored $<$ -70 °C at the MedStar Health Research Institute, the central SHS laboratory, as described previously (Lee et al., 1990). Urine Se was analyzed by inductively coupled plasma mass spectrometry (ICPMS) at the Analytical Chemistry for Health and Environment at the University of Graz, Austria (Scheer et al., 2012). Urinary creatinine was

also measured at the Phase 1 visit, and urinary Se was adjusted for creatinine to account for hydration status (O'Brien et al., 2017). The average limit of detection (LOD) for Se was 3.4 μg/L and all measured concentrations were above the LOD. The linearity of the standard curves for Se was > 0.99 and coefficients of variation for standards was < 6 %.

2.3. DNA methylation

The procedure outlining sample extraction, processing, and analysis has been described in detail previously (Domingo-Relloso et al., 2020). Briefly, buffy coat from fasting blood samples was collected at Phase 1 (1989–1991), and biological specimens were stored at −70 °C. DNA from white blood cells was extracted and stored at the MedStar Health Research Institute. In 2015, blood DNA was shipped to the Texas Biomedical Research Institute where DNA was bisulfite-converted with the EZ DNAm kit (Zymo Research, Irvine CA). DNA was then measured with the Illumina MethylationEPIC BeadChip (850 K) (Illumina, San Diego CA), providing DNAm information across > 850,000 CpGs. To account for potential batch artifacts and confounding effects, samples were randomized across and within plates and replicate and across-plate control samples were included on every plate, as described previously (Domingo-Relloso et al., 2020). Preprocessing of DNAm data was performed in R version 3.6.1 (Team RC, 2020). In total, data were analyzed in six different batches, with each batch consisting of roughly 400 individuals, and combined using the R package minfi (version 1.18.4) (Aryee et al., 2014). 6,159 CpGs were removed that had a p-detection value of > 0.01 in more than 5 % of participants. Single sample normalization was then performed using the preprocessNoob function in minfi (Fortin et al., 2017; Triche et al., 2013) to provide a background correction with dye-bias normalization.

After preprocessing, we had data from 1,357 individuals and 860,079 CpGs. As described previously, cross-hybridizing probes, sex chromosomes, and SNP probes with a minor allele frequency > 0.05 (McCartney et al., 2016) were removed. This resulted in 788,368 CpGs used in this analysis. Data then underwent quality checks, normalization, statistical preprocessing, and beta-value calculations, representing the proportion of unconverted cytosines at specific locations and ranging from 0 to 1 at specific locations, using minfi (Fortin et al., 2017). We estimated Houseman cell proportions (CD8T, CD4T, NK, B cells, monocytes, and neutrophils) (Houseman et al., 2012) using the R package *FlowSorted.Blood.EPIC* as adjustment factors in regression models. Potential batch effects by sample plate, sample row, and DNA isolation time were investigated and accounted for using the combat function in the R package sva (Leek et al., 2019). We annotated CpGs to the nearest gene according to UCSC genome browser (Nassar et al., 2023).

2.4. Statistical analysis

All analyses were conducted in R version 4.1.1. Urine Se concentrations were skewed and modeled in the $\log_2$ scale to reduce the influence of extreme values (**Figure S1**). We performed an EWAS to evaluate the association between urinary Se and each CpG individually using the *limma* package (Ritchie et al., 2015). This utilized the 'eBayes' function, which conducts empirical Bayes shrinkage of standard errors towards a common value to borrow information across genes. Genomic inflation factor and Q-Q plots were used to estimate the extent of unmeasured confounding and genomic inflation. Due to high genomic inflation ($\lambda > 1.50$), surrogate variables were included as additional adjustment factors using *SmartSVA* and *sva*. Ultimately, 12 surrogate variables were selected for inclusion resulting in a λ of 1.11. Models were adjusted for biologically relevant variables that may influence both urinary Se levels and DNAm including age, sex, study center of recruitment, BMI, education status (no high school/some high school/completed high school), smoking status (never/former/current), eGFR, Houseman blood cell composition, and five genetic principal components (to account for population stratification) (Barfield et al., 2014). Multiple comparisons were accounted for using the Benjamini-Hochberg method for false discovery rate (FDR) at a p-value of 0.10. We performed gene ontology enrichment analyses using the gometh function in *missMethyl* (Phipson et al., 2016) to identify Gene Ontology (GO) terms (The Gene Ontology resource, 2021) that contain over-representation of genes with DMPs, and for biological pathways in the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Goto, 2000). A separate EWAS was performed as a sensitivity analysis without the inclusion of the 12 surrogate variables. An additional EWAS was performed as a sensitivity analysis restricted to CpGs located in selenoprotein genes. A further EWAS was performed as a sensitivity analysis according to urinary selenium ≥ 60 μg/g creatinine compared to < 60 μg/g creatinine to explore the effect of Se at a level previously associated with increased incident CVD risk (Zhao et al., 2022). We also tested the interaction between urinary Se and urinary cadmium, as well as between urinary Se and urinary arsenic, among our top CpG sites in sensitivity analyses.

In addition to individual regressions, we performed GLMnet penalized regression (elastic net, R package *glmnet*) to identify relevant DMPs (Benton et al., 2017). The elastic net framework allows us to include all CpGs in the same model, accounting for the complex correlation structure between CpG methylation and performing variable selection for the most relevant sites. The parameters of the elastic net model included 10 cross validation folds and an alpha of 0.05.

3. Results

Among 1,357 participants, the mean (SD) age of participants was 55.4 (8.0) years, 55.8 % were female, and the mean (SD) Se level was 51.8 (25.1) μg/g creatinine (Table 1). The median (IQR) of Se was 45.8 (28.1) μg/g creatinine. Linear regressions investigating each site separately resulted in five DMPs that were significantly associated with urinary Se after multiple comparisons (Table 2; Fig. 1). One DMP was statistically significant at an FDR p value of < 0.05, the other four top sites were marginal, resulting in the use of an FDR p value of 0.10. These sites were cg00163554 (annotated to the Disco Interacting Protein 2 Homolog C [*DIP2C*] gene), cg18212762 (Mitogen-Activated Protein Kinase Kinase Kinase Kinase 2 [*MAP4K2*]), cg25194720 (Potassium

Table 1
Urinary selenium concentrations and participant characteristics.

	Analytical sample (n = 1,357)	SHS* (n = 2,325)
Urinary selenium (μg/g creatinine) mean (SD)	51.8 (25.1)	56.6 (29.2)
Center, n (%):		
Arizona	98 (7.2)	312 (13.4)
Oklahoma	601 (44.3)	981 (42.2)
North & South Dakota	658 (48.5)	1,032 (44.4)
Female, n (%):	757 (55.8)	1,361 (58.5)
Age (years), mean (SD)	55.4 (8.0)	56.2 (8.1)
Education, n (%):		
No High School	212 (15.6)	407 (17.5)
Some High School	299 (22.0)	556 (23.9)
Completed High School	846 (62.4)	1,362 (58.6)
Smoking status, n (%):		
Former	407 (30.0)	748 (32.2)
Never	376 (27.7)	684 (29.4)
Current	574 (42.3)	893 (38.4)
Diabetes at baseline (%)	0 (0)	968 (41.6)
Body mass index (kg/m^2), mean (SD)	29.1 (5.6)	30.3 (6.1)
eGFR (ml/min/1.73 m^2; CKD-Epi), mean (SD)	98.1 (14.4)	97.4 (16.8)

Note: eGFR=estimated glomerular filtration rate; CDK-Epi = Chronic Kidney Disease −Epidemiology Collaboration. The analytical sample in the current analysis was free of diabetes at baseline. *SHS refers to the Strong Heart Study sample that had urine selenium and DNAm data available.

Table 2
Top differentially methylated positions for urinary selenium.

CpG	Effect estimate (95 % CI Lower limit, Upper limit)	P-value	FDR Adj. p-value	Chromosome	UCSC Gene Name	UCSC Gene Group	Elastic Net Coefficient
cg00163554	0.055 (0.036, 0.074)	5.68E-09	0.009	10	DIP2C	Body	0.008
cg18212762	0.006 (0.004, 0.008)	3.57E-07	0.077	11	MAP4K2	3′UTR	0.021
cg25194720	−0.003 (−0.005, −0.001)	3.64E-07	0.077	11	KCNJ1	5′UTR; TSS200;	–
cg11270656	0.002 (0.001, 0.003)	4.55E-07	0.077	8	GPIHBP1	5′UTR; 1stExon	0.010
cg00886293	−0.009 (−0.012, −0.005)	4.90E-07	0.077	20	–	–	–

Note: Urinary selenium was $\log_2$-transformed. Linear regressions were adjusted for sex, age, study center, smoking (never, former, current), body mass index (kg/m^2), education status (no high school/some high school/completed high school), study center of recruitment, estimated glomerular filtration rate (ml/min/1.73 m^2), houseman cell proportions (CD8, CD4, natural killer, B-cell, monocytes), five principle genetic components, and 12 surrogate variables that captured unmeasured confounding. CpG sites were annotated using the University of California Santa Cruz (UCSC) genome browser. Effect estimates, 95 % CI lower limits, and 95 % CI upper limits are derived from regressions using beta-values. Other measures refer to calculations with M−values.

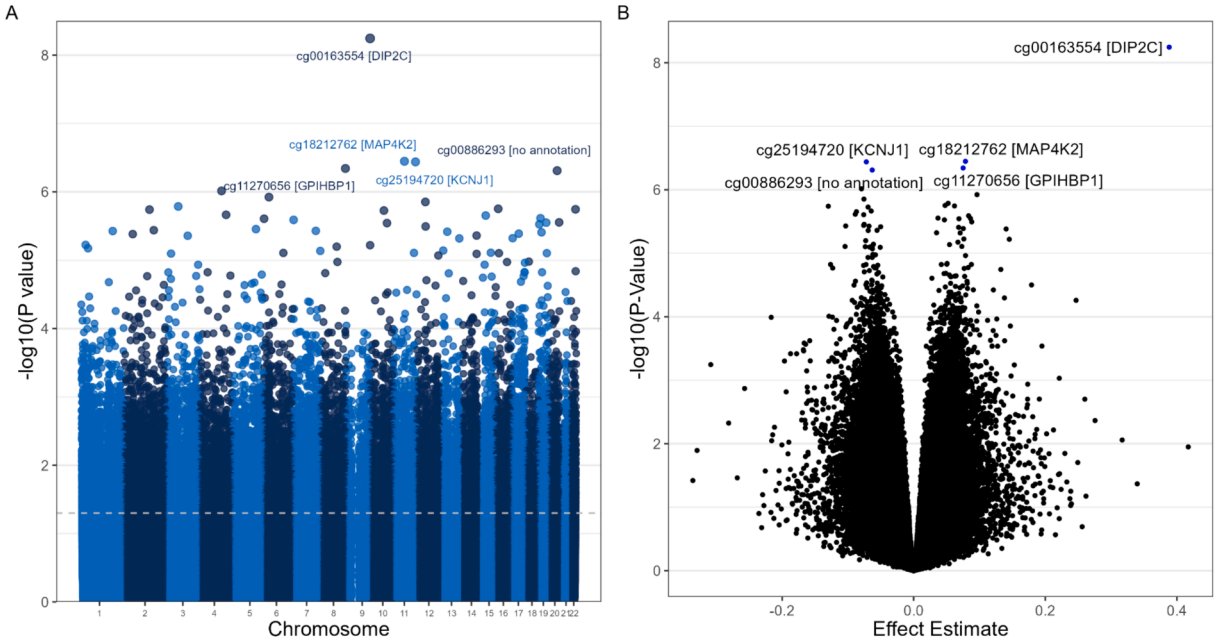

Fig. 1. Manhattan plot (A) and volcano plot (B) of differentially methylated DNA positions associated with urinary selenium. The top five sites significant after multiple comparisons (FDR p < 0.10) from linear regressions are annotated. M−values were used in regressions.

Inwardly Rectifying Channel Subfamily J Member 1 [*KCNJ1*]), cg11270656 (Glycosylphosphatidylinositol Anchored High Density Lipoprotein Binding Protein 1 [*GPIHBP1*]), and cg00886293 (*no annotation found*). cg00163554 was annotated to a south shelf in relationship to a CpG island in the gene body, cg18212762 was annotated to a 3′UTR (untranslated region), cg25194720 was annotated to a 5′UTR and a transcription start site (TSS200), and cg11270656 was annotated to a 5′UTR and the 1st exon. Specifically, a doubling of Se was associated with 5.5 % higher DNAm at cg00163554 (*DIP2C*), 0.6 % higher DNAm at cg18212762 (*MAP4K2*), 0.3 % lower DNAm at cg25194720 (*KCNJ1*), 0.2 % higher DNAm at cg11270656 (*GPIHBP1)*, and 0.9 % lower DNAm at cg00886293 (*no annotation found)*. There were 51,412 CpG's with a significant nominal p-value before multiple testing correction, and the full output from linear regressions can be found in **Table S1**. We did not identify any significantly over-represented GO terms or KEGG pathways after accounting for multiple testing. However, at nominal p < 0.05 the top pathways were granule and spindle formation. Results from the EWAS without the inclusion of surrogate variables are displayed in **Table S2**. This resulted in 49,769 significant CpG sites after FDR correction or 4,037 significant CpG sites after Bonferroni correction.

Elastic net regression retained 425 sites associated with urinary Se levels, including three overlapping DMPs identified in linear regressions (**Table S3**). While the interpretability of elastic net coefficients should be approached with caution, we observed hypermethylation in sites annotated to *MAPT*, *MAP4K2*, *TELO2*, *HS3ST3B1*, *CRTC1*, *TFRC*, and *AUTS2*, as well as hypomethylation in *ABLIM2*, *PDK4*, *CLEC1A*, and *ALPK3*. Annotation performed on the top five sites found pathways for chylomicrons and basal membranes but these were not significant after accounting for multiple testing.

We additionally performed a sensitivity analysis restricted to CpGs located in selenoprotein genes (Kryukov et al., 2003; Stelzer et al., 2016) (**Table S4**). While there were 662 CpG sites annotated to genes encoding selenoproteins, none of these were significantly associated with urinary Se levels in linear regressions after multiple testing correction. The EWAS exploring selenium ≥ 60 μg/g creatinine compared to < 60 μg/g creatinine reported consistent results as the main analysis, with only cg00163554 statistically significant after FDR correction (M−value, 95 % confidence interval: 0.471, 0.303, 0.640; **Table S5**). No significant

findings were observed in the interaction between urinary Se and urinary cadmium (**Table S6**) or between urinary Se and urinary arsenic (**Table S7**).

4. Discussion

This study evaluated the association of urinary Se with DNAm and is the first to study the epigenetic impacts of Se in a population of American Indians. In this sample of adults from the SHS, urinary Se was distinctly associated with five DMPs (cg00163554 [*DIP2C*], cg18212762 [*MAP4K2*], cg25194720 [*KCNJ1*], cg11270656 [*GPIHBP1*], cg00886293 [*no annotation found*]) in linear regressions, and elastic net regression retained 425 DMPs relevant to urinary Se, including three identified in linear regressions (cg00163554, cg18212762, cg11270656). Significant DMPs were not enriched for biological pathways. Although we hypothesized that DMPs associated with Se would be localized in selenoproteins and genes related to oxidative stress, this was not confirmed by our analysis.

In the literature, Se and DNAm have been investigated in a cancer context, where high Se exposure is proposed to inhibit DNA methyltransferase expression and activity (Speckmann and Grune, 2015; Ewa, 2017). Selenium has also been described to affect methylation of specific tumor suppressor genes, while many cancer phenotypes including head and neck, urogenital, gynecological, gastrointestinal, skin, and blood have been characterized by altered methylation of selenoprotein-encoding genes (Ewa, 2017; Tian et al., 2020). In the general population, it is possible that Se-linked DNAm changes may result from changing requirements for physiological Se, or potential increased oxidative stress from excess Se. While Se has been characterized as having antioxidant properties (Tinggi, 2008), research has also identified that certain Se compounds such as selenite and methylselenol can induce oxidative stress (Drake, 2006; Spallholz et al., 2004).

Few DMPs were detected in our primary EWAS. However, genes annotated to the top DMPs in this EWAS have been previously investigated in the literature in relation to Se and health. *DIP2C* is a protein coding gene that is expressed in a variety of tissues and is related to transcription factor binding (Uhlén et al., 2015). *DIP2C* is relevant to cancer and was upregulated in a Phase I trial of Se plus chemotherapy in gynecologic cancers (Song et al., 2018). In a separate study of *DIP2C* knockout in RKO colorectal cancer cells, a loss of *DIP2C* was related to various changes in DNAm and gene expression in the cells (Larsson et al., 2017) as well as somatic mutations in breast cancer (Jiao et al., 2012). Differential methylation of *DIP2C* has also been linked to metalloids, with *DIP2C* found to be differentially methylated relative to arsenic in two populations from Bangladesh and Spain (Domingo-Relloso et al., 2022), as well as amongst a pregnancy cohort in Mexico (Rojas et al., 2015); underscoring the need to study how *DIP2C* is related to both Se and other metalloids. In this study we report a doubling of Se was associated with a 5.5 % increase in DNAm in cg00163554 (*DIP2C*). Previous findings of methylation of *DIP2C* with respect to arsenic reported differences in 5-mC proportions (Domingo-Relloso et al., 2022), or correlations between maternal urinary arsenic and differentially methylated probes (Rojas et al., 2015), that were comparable in magnitude and direction to our study.

Although the effect estimates for additional DMPs were small in magnitude, the second top site of the EWAS (*MAP4K2*) is part of the mitogen-activated protein kinases (MAPK) signal transduction and pro-inflammatory response pathways that have been implicated in a variety of diseases including cancers and diabetes (Lawrence et al., 2008). Due to the role of the selenoprotein glutathione peroxidase in ameliorating oxidative damage, research has explored the ability of Se to inhibit inflammatory events and its role in the MAPK signaling cascade (Rashtchizadeh et al., 2015). While comprehensive data is largely missing from human populations, the literature supports the links between Se and activation of MAP4K pathways (Rashtchizadeh et al., 2015; Stapleton et al., 1997; Zhao et al., 2022; Hei et al., 1998), encouraging further exploration of the epigenetic impacts of Se on MAP4K signaling.

Elastic net regression analysis has the advantage of identifying DMPs that may be predictive of Se levels and has previously been successfully implemented in biomarker development (Colicino et al., 2021). The genes annotated to the top DMPs in elastic net regression have relevance to several Se-related diseases. Specifically, cg11018313 was annotated to *ALPK3* which is proposed to be involved in cardiac muscle cell development and is associated with cardiomyopathy (Walsh and Bezzina, 2021; Agarwal et al., 2022). *ALPK3* has been identified in relation to Keshan's disease (Huang et al., 2022), which is associated with low Se intake, suggesting that further research is needed on how Se is potentially associated with *ALPK3* at both low and high Se concentrations. Cg17075888 was annotated to the *PDK4* gene, which codes for a protein that regulates glucose metabolism and is regulated by insulin (Kaur et al., 2022). This warrants further study in relation to Se, which has been identified as an insulin mimetic, and excesses of Se have been associated with insulin resistance (Fontenelle et al., 2018; Casanova and Monleon, 2023). Additional sites may have relevance to cancer and neurogenerative diseases. Cg19589194 was annotated to *TELO2*, which encodes for an S-phase checkpoint protein and may be relevant to DNA repair (Feng et al., 2016). *TELO2* is upregulated in colorectal cancers (Guo et al., 2021), for which Se may play a preventative role (Peters and Takata, 2008). Cg26319641 was annotated to *CRTC1*, a transcription coactivator, which has been associated with both mucoepidermoid carcinoma and salivary gland carcinoma (Chen et al., 2021). Lastly, cg21953146 was annotated to the *MAPT* gene that encodes for a microtubule-associated protein tau, and mutations in this gene have been associated with Alzheimer's disease and dementia (Strang et al., 2019). Thus, these sites may serve has markers of health and disease for several pathologies.

In the general U.S. population reflected in the National Health and Nutrition Examination Survey (NHANES), Se values are mainly assessed in serum, preventing direct comparison of the Se concentrations measured in urine here to those measured in serum. However, urine Se is increasingly recognized as a viable population-level biomarker of Se and has been shown to reflect dietary intake (Sanz Alaejos and Díaz, 1993; Phiri et al., 2020). Despite this, it is necessary for studies to minimize the variability of urinary Se concentrations across participants by adjusting for urine creatinine or specific gravity (Ohira et al., 2008), as opposed to reporting unadjusted Se values (Urbano et al., 2021; Li et al., 2022), in order to facilitate more direct comparisons among different populations. In the current analysis, the mean urinary Se concentration was 51.8 μg/g creatinine, which was higher than urinary Se concentrations reported in a Chinese population (geometric mean: 19.8 μg/g) (Wu et al., 2018), and similar to specific-gravity adjusted Se geometric means reported across the Multi-Ethnic Study of Atherosclerosis (49.4 μg/L) and the SHS (58.6 μg/L) (Pang et al., 2016).

Epidemiological studies have supported inverse associations between Se status and cancer risk for several cancer types, although randomized clinical trials of Se and prostate cancer risk have shown conflicting results (Vinceti et al., 2018). Epidemiological studies and clinical trials on Se and CVD risk have also been conflicting, with some studies showing inverse associations between serum Se concentrations and coronary heart disease (Flores-Mateo et al., 2006) and Se supplementation lowering high density lipoprotein and plasma cholesterol (Cold et al., 2015), while other studies have not supported this relationship (Hasani et al., 2018; Rayman et al., 2011; Stranges et al., 2010).

This study builds on prior research on metals, epigenetic changes, and health outcomes in the SHS. Findings from the SHS have supported that elevated Se concentrations as measured in urine is a risk factor for CVD, with the highest risk of CVD occurring when urinary Se levels exceed 60 μg/g creatinine (Zhao et al., 2022). Further, previous research in the SHS has identified specific DNAm changes resulting from cadmium (Domingo-Relloso et al., 2020) and arsenic (Bozack et al., 2020), as well as distinct epigenetic changes in relation to lung functional

outcomes (Domingo-Relloso et al., 2022), diabetes outcomes (Domingo-Relloso et al., 2022), cancer (Domingo-Relloso et al., 2021), and coronary heart disease (Navas-Acien et al., 2021). In a separate EWAS performed in the SHS, there were 36 differentially methylated CpGs annotated to *DIP2C* and nominally-significant with respect to arsenic (Bozack et al., 2020). These effect sizes ranged from −0.062 to 0.103, which were similar in magnitude to the effect size observed with Se here (0.055). In the present study there was no overlap among the five significant DMPs with urinary Se and those sites identified as relevant to cadmium or associated with studied health outcomes in linear regressions. Thus, unfortunately, we were unable to conduct a mediation analysis linking these DMPs to health outcomes. However, a previous EWAS performed in the SHS with respect to urinary cadmium identified 24 nominally-significant differentially associated CpGs annotated to *DIP2C*, although none of these were significant after FDR correction, and effect sizes were smaller (−0.038 to 0.039) than reported here (Domingo-Relloso et al., 2020). Further, there was overlap in sites retained by elastic net in both the current analysis and a separate EWAS in the SHS investigating lung function. For example, there were six overlapping sites (cg27227029 [*no annotation found*], cg10276834 [*CHRM2*], cg13054419 [*no annotation found*], cg12899747 [*no annotation found*], cg02879453 [*ADCY7*], cg00986762 [*PSMB2*]) retained in the Se models and in elastic net models of forced expiratory volume in 1 s (FEV1), and 10 overlapping sites (cg12451436 [*RBCK1*], cg19645861 [*no annotation found*], cg27227029 [*no annotation found*], cg05141217 [*no annotation found*], cg17075888 [*PDK4*], cg03122840 [*no annotation found*], cg10276834 [*CHRM2*], cg18557149 [*LOC650226*], cg00758961 [*SLC2A9]*, cg02879453 [*ADCY7*]) relative to Se and forced vital capacity (FVC) (Domingo-Relloso et al., 2022). Further, they identified significant DMPs after FDR correction annotated to *DIP2C* for FEV1 (cg12531838), FEV1/FVC (cg27598761), airflow limitation (cg15713378), and restrictive pattern (cg24718710). For sites annotated to *DIP2C*, this study reported mean differences (95 % confidence intervals) for FEV1 (0.06 (0.01, 0.1)) and FEV1/FVC (−1.85 (−2.86, −0.83)), as well as odds ratios (95 % confidence intervals) for airflow limitation (1.76 (1.27, 2.43)) and restrictive pattern (0.68 (0.52, 0.9)), which are comparable to effect sies reported here. While the lack of overlap in findings across metals may suggest a unique epigenetic profile for individual trace elements, overlap in sites identified through elastic net approaches may relate to shared biomarkers of metals exposure. Given the relatively few identified DMPs in our primary EWAS, the potential mediating impact of epigenetic modifications between Se and adverse health outcomes may not be relevant.

This study has limitations. Dietary information was missing at the SHS baseline exam, which could help explain differences in urinary Se levels across participants, as well as influence epigenetic modifications. Urinary Se was used as blood Se was not available in the SHS. However, urinary measurements have been used as biomarkers of Se status and can be used to assess the majority of absorbed Se that is not retained (Phiri et al., 2020; Combs, 2015). Due to the presence of unmeasured confounding in linear regressions, surrogate variables were calculated and included in analyses. CpGs were annotated to the nearest gene according to the Illumina manifest, but these annotations carry uncertainty and are not always exact. Smoking status was self-reported, and it is possible that misclassification of smoking status could influence results. An FDR p-value of 0.10 was used rather than a p-value of 0.05, as one site was statistically significant at an FDR p value of 0.05, and the other four top sites were marginally significant. However, the goal of this study was discovery and investigating the relationship between Se and DNAm, and it is important to replicate these findings in additional cohorts. Future studies should explore the role of Se and DNAm in respect to adverse health outcomes. Given limited significant sites after FDR correction, we do not report significant findings from gene ontology and interaction analyses, and these effort were likely underpowered. Nonetheless, this study has several strengths, including the large sample size of participants with both urinary Se and blood DNAm data available, the high-quality exposure assessment, and the availability of confounders for participants. Additionally, we were able to test for DNAm changes at > 700,000 sites on the EPIC array.

5. Conclusions

In a sample of SHS participants, we observed cross-sectional associations between urinary Se and DNAm at select sites analyzed in 1989–1991 among American Indian adults. However, our hypothesis that DMPs associated with Se would be enriched in selenoproteins and genes related to oxidative stress was not confirmed. We did not observe robust epigenome-wide associations between Se and DNAm in a relatively well-powered analysis. This suggests that the impacts of Se on health may not proceed mechanistically via DNAm. However, identified sites may still be relevant for future work on DNAm exposure biomarkers. Future studies should examine additional mechanisms of Se impacts on health in American Indian populations, including exploration of these relationships prospectively and the potential role of differentially methylated sites with disease.

CRediT authorship contribution statement

Wil Lieberman-Cribbin: Writing – review & editing, Writing – original draft, Formal analysis. **Arce Domingo-Relloso:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Ronald A. Glabonjat:** Writing – review & editing, Investigation, Data curation. **Kathrin Schilling:** Writing – review & editing, Investigation, Data curation. **Shelley A. Cole:** Writing – review & editing, Software, Project administration, Investigation, Funding acquisition, Data curation. **Marcia O'Leary:** Writing – review & editing, Resources, Project administration. **Lyle G. Best:** Writing – review & editing, Project administration, Investigation, Funding acquisition. **Ying Zhang:** Writing – review & editing, Project administration, Data curation. **Amanda M. Fretts:** Writing – review & editing, Project administration, Investigation, Funding acquisition. **Jason G. Umans:** Writing – review & editing, Project administration, Investigation, Funding acquisition. **Walter Goessler:** Writing – review & editing, Investigation. **Ana Navas-Acien:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Maria Tellez-Plaza:** Writing – review & editing, Formal analysis, Data curation. **Allison Kupsco:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data underlying this article cannot be shared publicly in an unrestricted manner due to limitations in the consent forms and the agreements between the SHS tribal communities and investigators

Acknowledgments

We thank the dedication of the SHS participants, investigators and staff, without whom this work would not have been possible.

Sources of funding

This work was supported by grants by the National Heart, Lung, and Blood Institute (NHLBI) (under contract numbers 75N92019D00027, 75N92019D00028, 75N92019D00029, and 75N92019D00030) and previous grants (R01HL090863, R01HL109315, R01HL109301,

R01HL109284, R01HL109282, and R01HL109319 and cooperative agreements U01HL41642, U01HL41652, U01HL41654, U01HL65520, and U01HL65521), by the National Institute of Environmental Health Sciences (R01ES021367, R01ES025216, P42ES033719, P30ES009089, T32 ES007322). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the Indian Health Service (IHS). The funders had no role in study design, data collection and analysis, preparation of the manuscript, or decision to submit the manuscript for publication.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2024.108955.

References

Agarwal, R., Wakimoto, H., Paulo, J.A., Zhang, Q., Reichart, D., Toepfer, C., Sharma, A., Tai, A.C., Lun, M., Gorham, J., et al., 2022. Pathogenesis of cardiomyopathy caused by variants in ALPK3, an essential pseudokinase in the cardiomyocyte nucleus and sarcomere. Circulation 146, 1674–1693. https://doi.org/10.1161/circulationaha.122.059688.

Aryee, M.J., Jaffe, A.E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A.P., Hansen, K.D., Irizarry, R.A., 2014. Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics (Oxford, England) 30, 1363–1369. https://doi.org/10.1093/bioinformatics/btu049.

Barfield, R.T., Almli, L.M., Kilaru, V., Smith, A.K., Mercer, K.B., Duncan, R., Klengel, T., Mehta, D., Binder, E.B., Epstein, M.P., et al., 2014. Accounting for population stratification in DNA methylation studies. Genet. Epidemiol. 38, 231–241. https://doi.org/10.1002/gepi.21789.

Benton MC, Sutherland HG, Macartney-Coxson D, Haupt LM, Lea RA, Griffiths LRJA. Methylome-wide association study of whole blood DNA in the Norfolk Island isolate identifies robust loci associated with age. 2017;9:753.

Bleys J, Navas-Acien A, Guallar EJDc. Serum selenium and diabetes in US adults. 2007; 30:829-834.

Boyer, K., Domingo-Relloso, A., Jiang, E., Haack, K., Goessler, W., Zhang, Y., Umans, J. G., Belsky, D.W., Cole, S.A., Navas-Acien, A., et al., 2023. Metal mixtures and DNA methylation measures of biological aging in American Indian populations. Environ. Int. 178, 108064 https://doi.org/10.1016/j.envint.2023.108064.

Bozack, A.K., Domingo-Relloso, A., Haack, K., Gamble, M.V., Tellez-Plaza, M., Umans, J. G., Best, L.G., Yracheta, J., Gribble, M.O., Cardenas, A., 2020. Locus-specific differential DNA methylation and urinary arsenic: An epigenome-wide association study in blood among adults with low-to-moderate arsenic exposure. Environ. Health Perspect. 128, 067015.

Casanova, P., Monleon, D., 2023. Role of selenium in type 2 diabetes, insulin resistance and insulin secretion. World J. Diabetes 14, 147–158. https://doi.org/10.4239/wjd.v14.i3.147.

Castillo-Díaz SA, Garay-Sevilla ME, Hernández-González MA, Solís-Martínez MO, Zaina SJIjomm. Extensive demethylation of normally hypermethylated CpG islands occurs in human atherosclerotic arteries. 2010;26:691-700.

Chen, Z., Ni, W., Li, J.L., Lin, S., Zhou, X., Sun, Y., Li, J.W., Leon, M.E., Hurtado, M.D., Zolotukhin, S., et al., 2021. The CRTC1-MAML2 fusion is the major oncogenic driver in mucoepidermoid carcinoma. JCI Insight 6. https://doi.org/10.1172/jci.insight.139497.

Cold, F., Winther, K.H., Pastor-Barriuso, R., Rayman, M.P., Guallar, E., Nybo, M., Griffin, B.A., Stranges, S., Cold, S., 2015. Randomised controlled trial of the effect of long-term selenium supplementation on plasma cholesterol in an elderly Danish population. Br. J. Nutr. 114, 1807–1818. https://doi.org/10.1017/s0007114515003499.

Colicino, E., Just, A., Kioumourtzoglou, M.-A., Vokonas, P., Cardenas, A., Sparrow, D., Weisskopf, M., Nie, L.H., Hu, H., Schwartz, J.D., et al., 2021. Blood DNA methylation biomarkers of cumulative lead exposure in adults. J. Eposure Sci. Environ. Epidemiol. 31, 108–116. https://doi.org/10.1038/s41370-019-0183-9.

Combs Jr., G.F., 2015. Biomarkers of selenium status. Nutrients 7, 2209–2236. https://doi.org/10.3390/nu7042209.

Davis, C.D., Uthus, E.O., Finley, J.W., 2000. Dietary selenium and arsenic affect DNA methylation in vitro in Caco-2 cells and in vivo in rat liver and colon. J. Nutr. 130, 2903–2909. https://doi.org/10.1093/jn/130.12.2903.

Diagnosis and classification of diabetes mellitus. *Diabetes care*. 2011;34 Suppl 1:S62-69. doi: 10.2337/dc11-S062.

Domingo-Relloso, A., Riffo-Campos, A.L., Haack, K., Rentero-Garrido, P., Ladd-Acosta, C., Fallin, D.M., Tang, W.Y., Herreros-Martinez, M., Gonzalez, J.R., Bozack, A.K., 2020. Cadmium, smoking, and human blood DNA methylation profiles in adults from the strong heart study. Environ. Health Perspect. 128, 067005.

Domingo-Relloso, A., Huan, T., Haack, K., Riffo-Campos, A.L., Levy, D., Fallin, M.D., Terry, M.B., Zhang, Y., Rhoades, D.A., Herreros-Martinez, M., et al., 2021. DNA methylation and cancer incidence: Lymphatic-hematopoietic versus solid cancers in the Strong Heart Study. Clin. Epigenet. 13, 43. https://doi.org/10.1186/s13148-021-01030-8.

Domingo-Relloso, A., Gribble, M.O., Riffo-Campos, A.L., Haack, K., Cole, S.A., Tellez-Plaza, M., Umans, J.G., Fretts, A.M., Zhang, Y., Fallin, M.D., et al., 2022. Epigenetics of type 2 diabetes and diabetes-related outcomes in the Strong Heart Study. Clin. Epigenet. 14, 177. https://doi.org/10.1186/s13148-022-01392-7.

Domingo-Relloso, A., Bozack, A., Kiihl, S., Rodriguez-Hernandez, Z., Rentero-Garrido, P., Casasnovas, J.A., Leon-Latre, M., Garcia-Barrera, T., Gomez-Ariza, J.L., Moreno, B., et al., 2022. Arsenic exposure and human blood DNA methylation and hydroxymethylation profiles in two diverse populations from Bangladesh and Spain. Environ. Res. 204, 112021 https://doi.org/10.1016/j.envres.2021.112021.

Domingo-Relloso, A., Riffo-Campos, A.L., Powers, M., Tellez-Plaza, M., Haack, K., Brown, R.H., Umans, J.G., Fallin, M.D., Cole, S.A., Navas-Acien, A., et al., 2022. An epigenome-wide study of DNA methylation profiles and lung function among American Indians in the Strong Heart Study. Clin. Epigenet. 14, 75. https://doi.org/10.1186/s13148-022-01294-8.

Drake, E.N., 2006. Cancer chemoprevention: Selenium as a prooxidant, not an antioxidant. Med. Hypotheses 67, 318–322. https://doi.org/10.1016/j.mehy.2006.01.058.

Dye, C.K., Domingo-Relloso, A., Kupsco, A., Tinkelman, N.E., Spratlen, M.J., Bozack, A. K., Tellez-Plaza, M., Goessler, W., Haack, K., Umans, J.G., et al., 2023. Maternal DNA methylation signatures of arsenic exposure is associated with adult offspring insulin resistance in the Strong Heart Study. Environ. Int. 173, 107774 https://doi.org/10.1016/j.envint.2023.107774.

Ewa J, Edyta RJAiCR. Chapter Eight: Selenium and Epigenetics in Cancer: Focus on DNA Methylation. 2017;136:193-234.

Fairweather-Tait, S.J., Bao, Y., Broadley, M.R., Collings, R., Ford, D., Hesketh, J.E., Hurst, R., 2011. Selenium in human health and disease. Antioxid. Redox Signal. 14, 1337–1383. https://doi.org/10.1089/ars.2010.3275.

Feng, S.W., Chen, Y., Tsai, W.C., Chiou, H.C., Wu, S.T., Huang, L.C., Lin, C., Hsieh, C.C., Yang, Y.J., Hueng, D.Y., 2016. Overexpression of TELO2 decreases survival in human high-grade gliomas. Oncotarget 7, 46056–46066. https://doi.org/10.18632/oncotarget.10021.

Flores-Mateo, G., Navas-Acien, A., Pastor-Barriuso, R., Guallar, E., 2006. Selenium and coronary heart disease: a meta-analysis. Am. J. Clin. Nutr. 84, 762–773. https://doi.org/10.1093/ajcn/84.4.762.

Fontenelle LC, Feitosa MM, Morais JBS, Severo JS, Freitas TECd, Beserra JB, Henriques GS, Marreiro DdNJBJoPS. The role of selenium in insulin resistance. 2018;54.

Fortin, J.P., Triche Jr., T.J., Hansen, K.D., 2017. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. Bioinformatics (Oxford, England) 33, 558–560. https://doi.org/10.1093/bioinformatics/btw691.

Guo, Z., Zhang, X., Zhu, H., Zhong, N., Luo, X., Zhang, Y., Tu, F., Zhong, J., Wang, X., He, J., et al., 2021. TELO2 induced progression of colorectal cancer by binding with RICTOR through mTORC2. Oncol. Rep. 45, 523–534. https://doi.org/10.3892/or.2020.7890.

Hasani M, Djalalinia S, Sharifi F, Varmaghani M, Zarei M, Abdar ME, Asayesh H, Noroozi M, Kasaeian A, Gorabi AM, et al. Effect of Selenium Supplementation on Lipid Profile: A Systematic Review and Meta-Analysis. *Hormone and metabolic research = Hormon- und Stoffwechselforschung = Hormones et metabolisme*. 2018;50:715-727. doi: 10.1055/a-0749-6655.

Hei, Y.-j., Farahbakhshian, S., Chen, X., Battell, M.L., McNeill, J.H., 1998. Stimulation of MAP kinase and S6 kinase by vanadium and selenium in rat adipocytes. Mol. Cell. Biochem. 178, 367–375. https://doi.org/10.1023/A:1006819906820.

Houseman, E.A., Accomando, W.P., Koestler, D.C., Christensen, B.C., Marsit, C.J., Nelson, H.H., Wiencke, J.K., Kelsey, K.T., 2012. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinf. 13, 86. https://doi.org/10.1186/1471-2105-13-86.

Huang, Y.C., Combs Jr., G.F., Wu, T.L., Zeng, H., Cheng, W.H., 2022. Selenium status and type 2 diabetes risk. Arch. Biochem. Biophys. 730, 109400 https://doi.org/10.1016/j.abb.2022.109400.

Huang, J., Luo, R., Zheng, C., Cao, X., Zhu, Y., He, T., Liu, M., Yang, Z., Wu, X., Li, X., 2022. Integrative analyses identify potential key genes and calcium-signaling pathway in familial atrioventricular nodal reentrant tachycardia using whole-exome sequencing. Front. Cardiovasc. Med. 9, 910826 https://doi.org/10.3389/fcvm.2022.910826.

Jabłońska, E., Reszka, E., 2017. Selenium and epigenetics in cancer: Focus on DNA methylation. Adv. Cancer Res. 136, 193–234. https://doi.org/10.1016/bs.acr.2017.07.002.

Jiao, X., Wood, L.D., Lindman, M., Jones, S., Buckhaults, P., Polyak, K., Sukumar, S., Carter, H., Kim, D., Karchin, R., et al., 2012. Somatic mutations in the Notch, NF-KB, PIK3CA, and Hedgehog pathways in human breast cancers. Genes Chromosom. Cancer 51, 480–489. https://doi.org/10.1002/gcc.21935.

Kanehisa, M., Goto, S., 2000. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28, 27–30. https://doi.org/10.1093/nar/28.1.27.

Kaur, L., Pathak, M., Dan, S., Singhal, R., Kumar, V., Ojha, H., 2022. Chapter 15 - Protein kinase inhibitors from discovery to therapeutics. In: Hassan, M.I., Noor, S. (Eds.), Protein Kinase Inhibitors. Academic Press, pp. 435–456.

Khamkhash A, Srivastava V, Ghosh T, Akdogan G, Ganguli R, Aggarwal SJM. Mining-related selenium contamination in Alaska, and the state of current knowledge. 2017; 7:46.

Kryukov GV, Castellano S, Novoselov SV, Lobanov AV, Zehtab O, Guigó R, Gladyshev VN. Characterization of Mammalian Selenoproteomes. 2003;300:1439-1443. doi: doi:10.1126/science.1083516.

Larsson, C., Ali, M.A., Pandzic, T., Lindroth, A.M., He, L., Sjöblom, T., 2017. Loss of DIP2C in RKO cells stimulates changes in DNA methylation and epithelial-mesenchymal transition. BMC Cancer 17, 487. https://doi.org/10.1186/s12885-017-3472-5.

Lawrence, M.C., Jivan, A., Shao, C., Duan, L., Goad, D., Zaganjor, E., Osborne, J., McGlynn, K., Stippec, S., Earnest, S., et al., 2008. The roles of MAPKs in disease. Cell Res. 18, 436–442. https://doi.org/10.1038/cr.2008.37.

Lee, E.T., Welty, T.K., Fabsitz, R., Cowan, L.D., Le, N.-A., Oopik, A.J., Cucchiara, A.J., Savage, P.J., 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.

Leek JT, Johnson WE, Parker HS, Fertig EJ, Jaffe AE, Storey JD, Zhang Y, Torres LCJRpv. sva: Surrogate variable analysis. 2019;3:882-883.

Levey, A.S., Stevens, L.A., Schmid, C.H., Zhang, Y.L., Castro 3rd, A.F., Feldman, H.I., Kusek, J.W., Eggers, P., Van Lente, F., Greene, T., et al., 2009. A new equation to estimate glomerular filtration rate. Ann. Intern. Med. 150, 604–612. https://doi.org/10.7326/0003-4819-150-9-200905050-00006.

Lewis, J., Hoover, J., MacKenzie, D., 2017. Mining and environmental health disparities in Native American Communities. Curr. Environ. Health Rep. 4, 130–141. https://doi.org/10.1007/s40572-017-0140-5.

Li, A., Zhou, Q., Mei, Y., Zhao, J., Zhao, M., Xu, J., Ge, X., Xu, Q., 2022. Novel strategies for assessing associations between selenium biomarkers and cardiometabolic risk factors: Concentration, visit-to-visit variability, or individual mean? Evidence from a repeated-measures study of older adults with high selenium. Front. Nutr. 9, 838613 https://doi.org/10.3389/fnut.2022.838613.

McCartney, D.L., Walker, R.M., Morris, S.W., McIntosh, A.M., Porteous, D.J., Evans, K.L., 2016. Identification of polymorphic and off-target probe binding sites on the Illumina Infinium MethylationEPIC BeadChip. Genomics Data 9, 22–24. https://doi.org/10.1016/j.gdata.2016.05.012.

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. https://doi.org/10.7326/0003-4819-159-10-201311190-00719.

Moon, K.A., Zhang, Y., Guallar, E., Francesconi, K.A., Goessler, W., Umans, J.G., Best, L. G., Howard, B.V., Devereux, R.B., Okin, P.M., et al., 2018. Association of low-moderate urine arsenic and QT interval: Cross-sectional and longitudinal evidence from the Strong Heart Study. Environ. Pollut. 240, 894–902. https://doi.org/10.1016/j.envpol.2018.04.129.

Muscatello J, Belknap A, Janz DJEp. Accumulation of selenium in aquatic systems downstream of a uranium mining operation in northern Saskatchewan, Canada. 2008;156:387-393.

Nassar, L.R., Barber, G.P., Benet-Pagès, A., Casper, J., Clawson, H., Diekhans, M., Fischer, C., Gonzalez, J.N., Hinrichs, A.S., Lee, B.T., et al., 2023. The UCSC Genome Browser database: 2023 update. Nucleic Acids Res. 51, D1188–D1195. https://doi.org/10.1093/nar/gkac1072.

Navas-Acien, A., Domingo-Relloso, A., Subedi, P., Riffo-Campos, A.L., Xia, R., Gomez, L., Haack, K., Goldsmith, J., Howard, B.V., Best, L.G., et al., 2021. Blood DNA methylation and incident coronary heart disease: Evidence from the Strong Heart Study. JAMA Cardiol. https://doi.org/10.1001/jamacardio.2021.2704 %J JAMA Cardiology.

Nilsson, E., Jansson, P.A., Perfilyev, A., Volkov, P., Pedersen, M., Svensson, M.K., Poulsen, P., Ribel-Madsen, R., Pedersen, N.L., Almgren, P., et al., 2014. Altered DNA methylation and differential expression of genes influencing metabolism and inflammation in adipose tissue from subjects with type 2 diabetes. Diabetes 63, 2962–2976. https://doi.org/10.2337/db13-1459.

O'Brien, K.M., Upson, K., Buckley, J.P., 2017. Lipid and creatinine adjustment to evaluate health effects of environmental exposures. Curr. Environ. Health Rep. 4, 44–50. https://doi.org/10.1007/s40572-017-0122-7.

Ogawa-Wong AN, Berry MJ, Seale LA. Selenium and Metabolic Disorders: An Emphasis on Type 2 Diabetes Risk. 2016;8:80.

Ohira, S.-I., Kirk, A.B., Dyke, J.V., Dasgupta, P.K., 2008. Creatinine adjustment of spot urine samples and 24 h excretion of iodine, selenium, perchlorate, and thiocyanate. Environ. Sci. Tech. 42, 9419–9423. https://doi.org/10.1021/es8020952.

Oliver-Williams, C., Howard, A.G., Navas-Acien, A., Howard, B.V., Tellez-Plaza, M., Franceschini, N., 2018. Cadmium body burden, hypertension, and changes in blood pressure over time: Results from a prospective cohort study in American Indians. J. Am. Soc. Hypertens. 12 (426–437), e429.

Oo SM, Misu H, Saito Y, Tanaka M, Kato S, Kita Y, Takayama H, Takeshita Y, Kanamori T, Nagano TJSr. Serum selenoprotein P, but not selenium, predicts future hyperglycemia in a general Japanese population. 2018;8:1-10.

Pang, Y., Peng, R.D., Jones, M.R., Francesconi, K.A., Goessler, W., Howard, B.V., Umans, J.G., Best, L.G., Guallar, E., Post, W.S., et al., 2016. Metal mixtures in urban and rural populations in the US: The multi-ethnic study of atherosclerosis and the Strong Heart Study. Environ. Res. 147, 356–364. https://doi.org/10.1016/j.envres.2016.02.032.

Peters, U., Takata, Y., 2008. Selenium and the prevention of prostate and colorectal cancer. Mol. Nutr. Food Res. 52, 1261–1272. https://doi.org/10.1002/mnfr.200800103.

Phipson B, Maksimovic J, Oshlack AJB. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. 2016;32:286-288.

Phiri, F.P., Ander, E.L., Lark, R.M., Bailey, E.H., Chilima, B., Gondwe, J., Joy, E.J.M., Kalimbira, A.A., Phuka, J.C., Suchdev, P.S., et al., 2020. Urine selenium concentration is a useful biomarker for assessing population level selenium status. Environ. Int. 134, 105218 https://doi.org/10.1016/j.envint.2019.105218.

Pichler G, Grau-Perez M, Tellez-Plaza M, Umans J, Best L, Cole S, Goessler W, Francesconi K, Newman J, Redon JJCCI. Association of arsenic exposure with cardiac geometry and left ventricular function in young adults: evidence from the Strong Heart Family Study. 2019;12:e009018.

Rashtchizadeh, N., Karimi, P., Dehgan, P., Salimi, M.M., 2015. Effects of selenium in the MAPK signaling cascade. J. Cardiovasc. Thoracic Res. 7, 107–112. https://doi.org/10.15171/jcvtr.2015.23.

Rayman, M.P., 2012. Selenium and human health. Lancet 379, 1256–1268. https://doi.org/10.1016/S0140-6736(11)61452-9.

Rayman, M.P., Stranges, S., Griffin, B.A., Pastor-Barriuso, R., Guallar, E., 2011. Effect of supplementation with high-selenium yeast on plasma lipids: a randomized trial. Ann. Intern. Med. 154, 656–665. https://doi.org/10.7326/0003-4819-154-10-201105170-00005.

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GKJNar. limma powers differential expression analyses for RNA-sequencing and microarray studies. 2015; 43:e47-e47.

Rojas, D., Rager, J.E., Smeester, L., Bailey, K.A., Drobná, Z., Rubio-Andrade, M., Stýblo, M., García-Vargas, G., Fry, R.C., 2015. Prenatal arsenic exposure and the epigenome: Identifying sites of 5-methylcytosine alterations that predict functional changes in gene expression in newborn cord blood and subsequent birth outcomes. Toxicol. Sci. : Off. J. Soc. Toxicol. 143, 97–106. https://doi.org/10.1093/toxsci/kfu210.

Sanz Alaejos, M., Díaz, R.C., 1993. Urinary selenium concentrations. Clin. Chem. 39, 2040–2052. https://doi.org/10.1093/clinchem/39.10.2040.

Scheer, J., Findenig, S., Goessler, W., Francesconi, K.A., Howard, B., Umans, J.G., Pollak, J., Tellez-Plaza, M., Silbergeld, E.K., 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. Analyt. Methods : Adv. Methods Applic. 4, 406–413. https://doi.org/10.1039/c2ay05638k.

Sharmasarkar S, Vance GFJJoeq. Soil and plant selenium at a reclaimed uranium mine. 2002;31:1516-1521.

Slowly, M., Domingo-Relloso, A., Santella, R.M., Haack, K., Fallin, D.M., Terry, M.B., Rhoades, D.A., Herreros-Martinez, M., Garcia-Esquinas, E., Cole, S.A., et al., 2024. Blood DNA methylation and liver cancer in American Indians: Evidence from the Strong Heart Study. Cancer Causes Control : CCC 35, 661–669. https://doi.org/10.1007/s10552-023-01822-8.

Smith LD, Garg U. Chapter 17 - Disorders of trace metals. In: Garg U, Smith LD, eds. *Biomarkers in Inborn Errors of Metabolism*. San Diego: Elsevier; 2017:399-426.

Song, M., Kumaran, M.N., Gounder, M., Gibbon, D.G., Nieves-Neira, W., Vaidya, A., Hellmann, M., Kane, M.P., Buckley, B., Shih, W., et al., 2018. Phase I trial of selenium plus chemotherapy in gynecologic cancers. Gynecol. Oncol. 150, 478–486. https://doi.org/10.1016/j.ygyno.2018.07.001.

Spallholz, J.E., Palace, V.P., Reid, T.W., 2004. Methioninase and selenomethionine but not Se-methylselenocysteine generate methylselenol and superoxide in an in vitro chemiluminescent assay: implications for the nutritional carcinostatic activity of selenoamino acids. Biochem. Pharmacol. 67, 547–554. https://doi.org/10.1016/j.bcp.2003.09.004.

Speckmann, B., Grune, T., 2015. Epigenetic effects of selenium and their implications for health. Epigenetics 10, 179–190. https://doi.org/10.1080/15592294.2015.1013792.

Stapleton, S.R., Garlock, G.L., Foellmi-Adams, L., Kletzien, R.F., 1997. Selenium: potent stimulator of tyrosyl phosphorylation and activator of MAP kinase. BBA 1355, 259–269. https://doi.org/10.1016/s0167-4889(96)00140-1.

Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, Stein TI, Nudel R, Lieder I, Mazor Y, et al. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Current protocols in bioinformatics*. 2016;54:1.30.31-31.30.33. doi: 10.1002/cpbi.5.

Stoddart ML, Jarvis B, Blake B, Fabsitz RR, Howard BV, Lee ET, Welty TK. Recruitment of American Indians in epidemiologic research: the Strong Heart Study. *American Indian and Alaska native mental health research (Online)*. 2000;9:20-37. doi: 10.5820/aian.0903.2000.20.

Strang KH, Golde TE, Giasson BI. MAPT mutations, tauopathy, and mechanisms of neurodegeneration. *Laboratory investigation; a journal of technical methods and pathology*. 2019;99:912-928. doi: 10.1038/s41374-019-0197-x.

Stranges S, Laclaustra M, Ji C, Cappuccio FP, Navas-Acien A, Ordovas JM, Rayman M, Guallar EJTJon. Higher selenium status is associated with adverse blood lipid profile in British adults. 2010;140:81-87.

Team RC. *R: A language and environment for statistical computing. R Foundation for Statistical Computing;* Vienna, Austria: 2020.

Tellez-Plaza, M., Guallar, E., Fabsitz, R.R., Howard, B.V., Umans, J.G., Francesconi, K.A., Goessler, W., Devereux, R.B., Navas-Acien, A., 2013. Cadmium exposure and incident peripheral arterial disease. J. Circ.: Cardiovasc. Quality Outcomes 6, 626–633.

Tellez-Plaza M, Tang W-y, Shang Y, Umans JG, Francesconi KA, Goessler W, Ledesma M, Leon M, Laclaustra M, Pollak JJEhp. Association of global DNA methylation and global DNA hydroxymethylation with metals and other exposures in human blood DNA samples. 2014;122:946-954.

The Gene Ontology resource: enriching a GOld mine. *Nucleic acids research*. 2021;49: D325-d334. doi: 10.1093/nar/gkaa1113.

Tian F-Y, Everson TM, Lester B, Punshon T, Jackson BP, Hao K, Lesseur C, Chen J, Karagas MR, Marsit CJJEi. Selenium-associated DNA methylation modifications in placenta and neurobehavioral development of newborns: an epigenome-wide study of two US birth cohorts. 2020;137:105508.

Tinggi, U., 2008. Selenium: its role as antioxidant in human health. Environ. Health Prev. Med. 13, 102–108. https://doi.org/10.1007/s12199-007-0019-4.

Triche Jr., T.J., Weisenberger, D.J., Van Den Berg, D., Laird, P.W., Siegmund, K.D., 2013. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. Nucleic Acids Res. 41, e90.

Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, et al. Proteomics. Tissue-based map of the human proteome. *Science (New York, NY)*. 2015;347:1260419. doi: 10.1126/science.1260419.

Urbano, T., Filippini, T., Lasagni, D., De Luca, T., Grill, P., Sucato, S., Polledri, E., Djeukeu Noumbi, G., Malavolti, M., Santachiara, A., et al., 2021. Association of urinary and dietary selenium and of serum selenium species with serum alanine aminotransferase in a healthy Italian population. Antioxidants (Basel Switzerland) 10. https://doi.org/10.3390/antiox10101516.

US Department of Health Human Services. Agency for Toxic Substances and Disease Registry-ATSDR. TOXICOLOGICAL PROFILE FOR SELENIUM. https://www.atsdr.cdc.gov/ToxProfiles/tp92.pdf. 2003. Accessed November 2022.

Vinceti M, Filippini T, Del Giovane C, Dennert G, Zwahlen M, Brinkman M, Zeegers MP, Horneber M, D'Amico R, Crespi CM. Selenium for preventing cancer. *The Cochrane database of systematic reviews*. 2018;1:Cd005195. doi: 10.1002/14651858.CD005195.pub4.

Walsh, R., Bezzina, C.R., 2021. ALPK3: A full spectrum cardiomyopathy gene? Eur. Heart J. 42, 3074–3077. https://doi.org/10.1093/eurheartj/ehab415.

Welty, T.K., Lee, E.T., Yeh, J., Cowan, L.D., Go, O., Fabsitz, R.R., Le, N.A., Oopik, A.J., Robbins, D.C., Howard, B.V., 1995. Cardiovascular disease risk factors among American Indians. The Strong Heart Study. Am. J. Epidemiol. 142, 269–287. https://doi.org/10.1093/oxfordjournals.aje.a117633.

Wu, W., Jiang, S., Zhao, Q., Zhang, K., Wei, X., Zhou, T., Liu, D., Zhou, H., Zeng, Q., Cheng, L., et al., 2018. Environmental exposure to metals and the risk of hypertension: A cross-sectional study in China. Environ. Pollut. 233, 670–678. https://doi.org/10.1016/j.envpol.2017.10.111.

Zhao, D., Domingo-Relloso, A., Tellez-Plaza, M., Nigra, A.E., Valeri, L., Moon, K.A., Goessler, W., Best, L.G., Ali, T., Umans, J.G., et al., 2022. High level of selenium exposure in the Strong Heart Study: A cause for incident cardiovascular disease? Antioxid. Redox Signal. https://doi.org/10.1089/ars.2022.0029.

Zhao, Y., Zhang, H., Hao, D., Wang, J., Zhu, R., Liu, W., Liu, C., 2022. Selenium regulates the mitogen-activated protein kinase pathway to protect broilers from hexavalent chromium-induced kidney dysfunction and apoptosis. Ecotoxicol. Environ. Saf. 239, 113629 https://doi.org/10.1016/j.ecoenv.2022.113629.