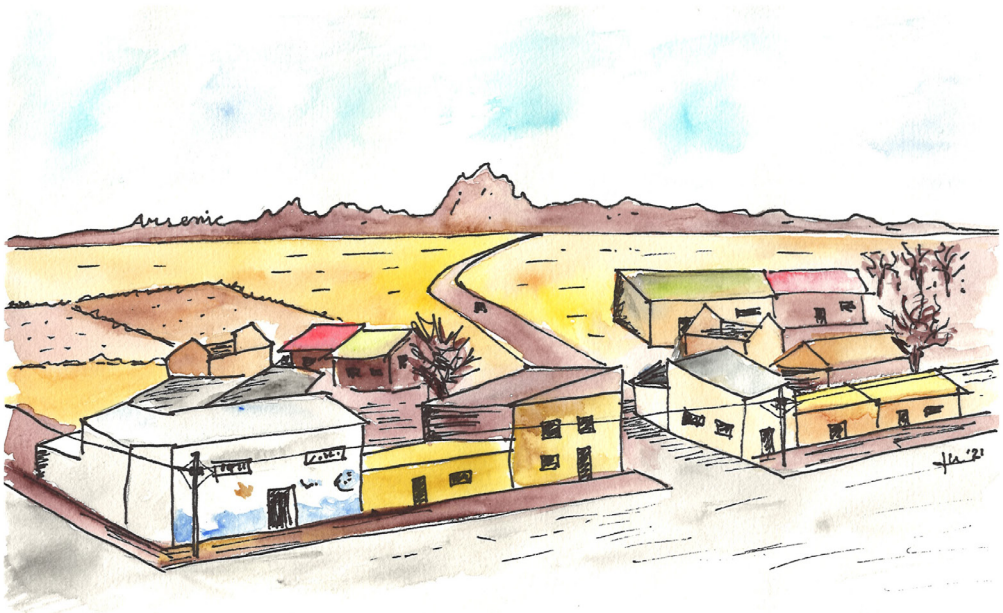
Thesis for doctoral degree (Ph.D.)
2021

Environmental arsenic exposure in humans: toxicity and adaptation in the Bolivian Andes



Jessica De Loma Olson

Karolinska Institutet

From the INSTITUTE OF ENVIRONMENTAL MEDICINE
Karolinska Institutet, Stockholm, Sweden

ENVIRONMENTAL ARSENIC EXPOSURE IN HUMANS: TOXICITY AND ADAPTATION IN THE BOLIVIAN ANDES

Jessica De Loma Olson

Karolinska Institutet

Stockholm 2021

All previously published papers were reproduced with permission from the publisher.
Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2021

ISBN 978-91-8016-228-9

Cover illustration: “Al final de la montaña” by Jessica De Loma. Watercolor, 2021.
Panoramic view of Pampa Aullagas, one of the villages in the Bolivian Andes included in this thesis work.

ENVIRONMENTAL ARSENIC EXPOSURE IN HUMANS: TOXICITY AND ADAPTATION IN THE BOLIVIAN ANDES

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Jessica De Loma Olson

The thesis will be defended in public at the Andreas Vesalius hall, Berzelius väg 3, Karolinska Institutet, Solna, Sweden

Friday, June 11th, 2021 at 13:00

Principal Supervisor:
Professor Karin Broberg
Karolinska Institutet
Institute of Environmental Medicine
Unit of Metals and Health

Co-supervisor(s):
Associate Professor Carina M. Schlebusch
Uppsala University
Department of Organismal Biology
Human Evolution Program

Anda R. Gliga
Karolinska Institutet
Institute of Environmental Medicine
Unit of Metals and Health

Opponent:
Professor Ana Navas-Acien
Columbia University
Department of Environmental Health Sciences

Examination Board:
Docent Åsa Sjöling
Karolinska Institutet
Department of Tumor and Cell Biology

Professor Erik Melén
Karolinska Institutet - Södersjukhuset
Department of Clinical Science and Education

Professor Markus Tamás
University of Gothenburg
Department of Chemistry and Molecular Biology

"The fuel for evolution is diversity, with natural selection leading to continuous adaptations and improvements in Nature's handiwork. Organisms (and organizations) without this diversity are doomed to extinction in a rapidly changing world."

Excerpt from Frances H. Arnold's speech at the Nobel Banquet, December 2018

To my parents, for teaching me
how to have a bright attitude
in the darkest moments.

ABSTRACT

Arsenic, a potent toxicant and carcinogen, is naturally present in soil and leaches into groundwater. More than 140 million people worldwide are exposed to arsenic through drinking water. How well humans metabolize arsenic is a susceptibility factor for arsenic toxicity: individuals with a less efficient arsenic metabolism are at higher risk of arsenic-related health effects, such as cancer or cardiovascular disease. In fact, positive selection of a more efficient arsenic metabolism phenotype has been described in the Argentinean Andes, where indigenous populations have been presumably exposed to arsenic for centuries. However, whether this genetic adaptation to arsenic has occurred elsewhere in the Andes was not clear.

The overall aim of this thesis was to assess the exposure, toxicity, and potential genetic adaptation to arsenic in the Bolivian Andes. For this, we recruited indigenous women from the Bolivian Andes living in 10 villages around Lake Poopó.

In **Paper I**, we described that these Bolivian women were exposed to arsenic with varying arsenic concentrations in urine (range 12–407 µg/L, median 65 µg/L). The women had on average an efficient arsenic metabolism compared to other populations across the world. Ethnicity, body weight, fish consumption, and tobacco smoking were identified as influencing their capacity to metabolize arsenic.

We then showed that these Bolivian women had molecular signs of arsenic toxicity by measuring four toxicity biomarkers in **Paper II**. Using multivariable-adjusted linear regression models, arsenic exposure was associated with longer telomeres and more copies of mitochondrial DNA in blood, two biomarkers for cancer risk, particularly in women with a less efficient arsenic metabolism. Urinary 4-hydroxy nonenal mercapturic acid, a metabolite of lipid peroxidation, was associated with increasing arsenic exposure. Urinary 8-oxo-2'-deoxyguanosine, a biomarker of oxidative DNA damage, showed discrepant results depending on the biomarker of exposure used: a positive association with urinary arsenic vs. a negative association with blood arsenic.

In **Paper III**, we found four putative cancer-related proteins in urine that were associated with arsenic exposure in blood: tumor necrosis factor ligand superfamily member 6, FASLG; seizure 6-like protein, SEZ6L; Ly6/PLAUR domain-containing protein 3, LYPD3; and tissue factor pathway inhibitor 2, TFPI2. Other factors influencing the variation of these proteins in urine were identified, including urinary osmolality, leukocytes in urine, and age.

By combining genotype-phenotype association analyses and genome-wide selection scans for positive selection in **Paper IV**, we identified genetic signatures of positive selection near *AS3MT*, the main arsenic methylating enzyme, in the indigenous Bolivian study group. These Bolivian communities have the highest frequency of protective *AS3MT* alleles associated with a more efficient arsenic metabolism described in the literature.

In conclusion, indigenous communities living in the Bolivian Andes were exposed to environmental arsenic and were genetically adapted to methylate arsenic more efficiently at a population level. Despite their efficient arsenic metabolism on average, arsenic exposure was still associated with toxicity biomarkers and changes in urinary cancer-related proteins, stressing the need to tackle the public health concern of arsenic in the Andes.

POPULAR SCIENCE SUMMARY

The Andes Mountains are one of the most extreme and harsh environments that humans inhabit. Humans have had to adapt over generations to cope with its high altitude, arid climate, and exposure to metals from the volcanic bedrock. Minerals rich in arsenic in the soil erode, and arsenic is released into drinking water sources. Exposure to arsenic during long periods can lead to cancer or cardiovascular disease in adults, higher child mortality and lower birth weight. Arsenic is eliminated from the human body with urine after it is metabolized. If the body is better at metabolizing and excreting arsenic, then there is less risk of it being toxic.

Certain regions of the Andes have had extreme levels of arsenic in water for centuries. In the 1990s, a group of indigenous women from the Argentinean Andes was described as uniquely efficient at metabolizing arsenic. As generations passed, those individuals who were better at eliminating arsenic from their bodies had a higher chance of surviving and reproducing, until today, where most people living in this region of Argentina are efficient at metabolizing arsenic. This capacity to metabolize arsenic better or worse is determined partly by genetics. The increase, after many generations, of genetic variants that are responsible for beneficial traits in a certain environment is known as “positive selection”.

Indigenous women from the Argentinean Andes have adapted to arsenic by metabolizing it more efficiently. But has this adaptation occurred as well in other regions of the Andes? To answer this question, we invited indigenous women from the Bolivian Andes living around Lake Poopó to participate in our study about the exposure, toxicity, and adaptation to arsenic.

In the first study, we analyzed how much these Bolivian women were exposed to arsenic and how well they could metabolize it. To do this, we measured different arsenic metabolites in their urine. We saw that the women living around Lake Poopó were exposed to arsenic via drinking water, and, on average, they were efficient at metabolizing arsenic. However, does arsenic still have toxic effects?

In the second and third study, we evaluated the toxicity of arsenic in these women. Since the study area includes remote villages with limited access to health care, we did not have information about how many women developed certain diseases. Instead, we measured disease and toxicity indicators at a molecular level known as biomarkers in urine and blood. In the second study, women who were more exposed to arsenic and were less efficient arsenic metabolizers had higher concentrations of biomarkers related to early signs of cancer, and oxidative damage to DNA and lipids. In the third study, we discovered that some proteins in urine related to processes that lead to cancer had different levels depending on the arsenic exposure of the women.

In the last study of this thesis, we found that genetic variants associated with a more efficient arsenic metabolism were very common in these Bolivian indigenous communities. We also detected signs of positive selection in their genomes related to arsenic metabolism.

In conclusion, this thesis shows that indigenous women living in the Bolivian Andes were exposed to arsenic and had, on average, an efficient arsenic metabolism thanks to genetic adaptation after multiple generations. However, there were indications that arsenic is toxic in these populations, so more efforts are needed to reduce arsenic exposure in this area.

RESUMEN DE DIVULGACIÓN CIENTÍFICA

La cordillera de los Andes es uno de los entornos más extremos y duros en los que vive el ser humano. A lo largo de generaciones, los humanos han tenido que adaptarse para hacer frente a su altura elevada, clima árido y exposición a metales del terreno volcánico. Los minerales ricos en arsénico en la tierra se erosionan y este elemento se libera en el agua que después es consumida por el ser humano. La exposición a arsénico durante largos periodos puede conllevar al desarrollo de cáncer o enfermedades cardiovasculares en adultos, a un incremento de la mortalidad infantil y a un menor peso en neonatos. El arsénico es eliminado del cuerpo humano a través de la orina después de ser metabolizado. Si el cuerpo puede metabolizar y excretar el arsénico de una manera más eficaz, entonces hay un menor riesgo de que sea tóxico.

Ciertas regiones de los Andes han tenido en sus aguas niveles de arsénico extremadamente elevados durante siglos. En la década de 1990 se describió a un grupo de mujeres indígenas de los Andes argentinos capaz de metabolizar el arsénico de manera excepcionalmente eficaz. A medida que pasaban las generaciones, aquellas personas que podían eliminar mejor el arsénico tenían una mayor probabilidad de sobrevivir y reproducirse, hasta la actualidad, en la que la mayoría de las personas que viven en esta región de Argentina metabolizan el arsénico de forma eficaz. Esta distinta capacidad de metabolizar el arsénico está determinada, en parte, por la genética. El incremento de variantes genéticas de características ventajosas tras generaciones se conoce como “selección positiva”.

Las mujeres indígenas de los Andes argentinos se han adaptado para metabolizar el arsénico de manera más eficaz. Sin embargo, ¿ha ocurrido también esta adaptación en otras regiones de los Andes? Para contestar a esta pregunta, invitamos a mujeres indígenas de los Andes bolivianos que vivían alrededor del lago Poopó a participar en nuestro estudio sobre la exposición, toxicidad y adaptación al arsénico.

En el primer estudio analizamos el grado de exposición y la capacidad de metabolización del arsénico de estas mujeres bolivianas. Para ello, medimos distintos metabolitos de arsénico en su orina. Se vio que las mujeres que vivían alrededor del lago Poopó estaban expuestas a arsénico a través del agua que consumían y, en promedio, podían metabolizar eficazmente el arsénico. A pesar de ello, ¿sigue teniendo el arsénico efectos tóxicos?

En el segundo y tercer estudio evaluamos la toxicidad del arsénico en estas mujeres. Al incluir esta zona pueblos remotos con acceso limitado a servicios de salud, no teníamos información acerca de cuántas mujeres desarrollaron ciertas enfermedades. En vez de eso, medimos indicadores de enfermedad y toxicidad a nivel molecular conocidos como biomarcadores en orina y sangre. En el segundo estudio, las mujeres que estaban más expuestas a arsénico y lo metabolizaban de manera menos eficaz tenían concentraciones más elevadas de marcadores relacionados con signos tempranos de cáncer y de daño oxidativo al ADN y lípidos. En el tercer estudio descubrimos que algunas proteínas en orina relacionadas con procesos que conllevan a cáncer tenían niveles distintos dependiendo de la exposición a arsénico de las mujeres.

En el último estudio de esta tesis encontramos que variantes genéticas asociadas con una mayor eficacia para metabolizar el arsénico eran muy comunes en estas comunidades indígenas de Bolivia. También hallamos en sus genomas signos de selección positiva relacionados con el metabolismo de arsénico.

En conclusión, esta tesis muestra que mujeres indígenas que viven en los Andes bolivianos estaban expuestas a arsénico y que, en promedio, metabolizaban el arsénico eficazmente gracias a la adaptación genética ocurrida durante múltiples generaciones. Sin embargo, sí que había signos de que el arsénico es tóxico en estas poblaciones, por lo que son necesarios más esfuerzos para reducir la exposición a arsénico en esta región.

POPULÄRVETENSKAPLIG SAMMANFATTNING

Bergskedjan Anderna är bland de mest extrema och ogästvänliga miljöer som människor bebor. Människor har behövt anpassa sig över generationer för att klara av dess höga höjder, torra klimat och metallrika, vulkaniska berggrund. Mineraler rika på arsenik i jorden eroderar och arsenik hamnar i dricksvattnet. Att exponeras för arsenik under långa perioder kan leda till cancer och hjärt- och kärlsjukdomar hos vuxna, samt ökad barnadödlighet och lägre födelsevikt. Arsenik elimineras från kroppen genom urinen efter att det har omvandlats i kroppen (metaboliserats). Om kroppen är bra på att metabolisera och utsöndra arsenik så minskar risken att det är toxiskt.

Vissa delar av Anderna har haft extremt höga arsenikhalter i vattnet under århundraden. På 1990-talet beskrevs en grupp kvinnor ur ursprungsbefolkningen i de argentinska Anderna som unikt effektiva på att metabolisera arsenik. Allteftersom generationerna fortskred har de individer som varit bättre på att eliminera arsenik från sina kroppar haft en större chans att överleva och fortplanta sig, tills idag när de flesta som bor i området metaboliserar arsenik effektivt. Förmågan att metabolisera arsenik mer eller mindre effektivt bestäms delvis av genetiken. Ökningen, efter många generationer, av genetiska variationer som bestämmer egenskaper som är fördelaktiga kallas “positiv selektion”.

Kvinnorna ur ursprungsbefolkningen i de argentinska Anderna har anpassat sig till arsenik genom att metabolisera det mer effektivt. Men har den här anpassningen skett även i andra delar av Anderna? För att besvara den frågan bjöd vi in kvinnor från ursprungsbefolkningen kring Poopósjön i de bolivianska Anderna att delta i vår studie om exponering för arsenik, toxicitet och anpassning.

Först studerade vi hur pass exponerade de bolivianska kvinnorna var för arsenik och hur väl de kunde metabolisera arsenik. För att göra detta mätte vi olika restprodukter (metaboliter) från arsenikomvandling i kvinnornas urin. Vi såg att kvinnorna runt Poopósjön exponerades för arsenik via dricksvattnet och överlag metaboliserade arsenik effektivt. Har arseniken trots detta toxiska effekter?

I den andra och tredje studien utvärderade vi arsenikens toxicitet hos dessa kvinnor. Eftersom studieområdet inkluderar avlägsna byar med begränsad tillgång till sjukvård, hade vi inte information om hur många kvinnor som utvecklade vissa sjukdomar. Istället mätte vi indikatorer på sjukdomar och toxicitet på molekylär nivå i urin och blod, så kallade biomarkörer. I den andra studien hade kvinnor som var mer exponerade för arsenik och metaboliserade arsenik mindre effektivt en högre koncentration av biomarkörer relaterade till tidiga tecken på cancer och oxidativa skador på DNA och lipider. I den tredje studien upptäckte vi att koncentrationen av vissa proteiner i urin relaterade till processer som leder till cancer var olika beroende på hur exponerade kvinnorna var för arsenik.

I avhandlingens fjärde och sista studie fann vi att genetiska variationer associerade med en mer effektiv arsenikmetabolism var mycket vanliga hos dessa bolivianska ursprungsgrupper. Vi fann även tecken på positiv selektion i deras arvsmassa relaterade till arsenikmetabolism.

Sammanfattningsvis visar denna avhandling att kvinnor ur ursprungsbefolkningen boende i de bolivianska Anderna var exponerade för arsenik och överlag hade en effektiv arsenikmetabolism tack vare genetisk adaption efter många generationer. Det finns dock indikationer på att arsenik är giftigt hos dessa populationer, så fler insatser behövs för att minska exponeringen för arsenik i området.

LIST OF SCIENTIFIC PAPERS

This thesis is based on the following papers, which are referred to in the text by their Roman numerals:

I. **De Loma, J.**, Tirado, N., Ascui, F., Levi, M., Vahter, M., Broberg, K., Gardon, J. (2019) Elevated arsenic exposure and efficient arsenic metabolism in indigenous women around Lake Poopó, Bolivia. *Science of the Total Environment*, 657:179-186.

II. **De Loma, J.**, Krais, A. M., Lindh, C. H., Mamani, J., Tirado, N., Gardon, J., Broberg, K. (-) Arsenic exposure and biomarkers for oxidative stress and telomere length in indigenous populations in Bolivia – modification by arsenic metabolism efficiency. *Submitted*.

III. **De Loma, J.**, Gliga, A.R., Levi, M., Ascui, F., Gardon, J., Tirado, N., Broberg, K. (2020) Arsenic Exposure and Cancer-Related Proteins in Urine of Indigenous Bolivian Women. *Frontiers in Public Health*, 8:605123.

IV. **De Loma, J.**, Vicente, M., Tirado, N., Ascui, F., Vahter, M., Gardon, J., Schlebusch, C. M., Broberg, K. (-) Human adaptation to arsenic exposure in Andean populations in Bolivia. *Manuscript*.

SCIENTIFIC PAPERS NOT INCLUDED IN THE THESIS

- **De Loma, J.**, Skröder, H., Raqib, R., Vahter, M., Broberg, K. (2018) Arsenite methyltransferase (AS3MT) polymorphisms and arsenic methylation in children in rural Bangladesh. *Toxicology and Applied Pharmacology*, 357:80-87.

- Skröder, H., Kippler, M., **De Loma, J.**, Raqib, R., Vahter, M. (2018) Predictors of selenium biomarker kinetics in 4-9-year-old Bangladeshi children. *Environment International*, 121(1):842-851.

- Torbøl Pedersen, J.*, **De Loma, J.***, Levi, M., Palmgren, M., Broberg, K. (2020) Predicted AS3MT Proteins Methylate Arsenic and Support Two Major Phylogenetic AS3MT Groups. *Chemical Research in Toxicology*, 33(12):3041-3047.

* These authors contributed equally.

CONTENTS

1 PREFACE....1
2 BACKGROUND....3
2.1 Arsenic....3
2.1.1 Forms, sources, and exposures....3
2.1.1.1 Exposure in Bolivia....4
2.1.2 Toxicokinetics in humans....5
2.1.2.1 Absorption and distribution....5
2.1.2.2 Metabolism....6
2.1.2.3 Excretion and biomarkers of exposure....6
2.1.3 Toxicity in humans....7
2.1.3.1 Acute toxicity....7
2.1.3.2 Chronic toxicity....8
2.1.3.3 Mechanism of toxicity....9
2.1.4 Metabolism efficiency as a susceptibility factor....9
2.1.5 Factors influencing arsenic metabolism....10
2.1.6 Genes involved in arsenic metabolism....11
2.1.6.1 Candidate-gene approach....11
2.1.6.2 Genome-wide approach....12
2.2 Population genetics....15
2.2.1 Population structure of Andean indigenous groups....15
2.2.2 Human adaptation....16
2.2.2.1 Processes driving adaptation....16
2.2.2.2 Examples of human adaptation in the Andean plateau....17
3 RESEARCH AIMS....19
4 MATERIALS AND METHODS....21
4.1 Study area....21
4.2 Study design and participants....22
4.3 Sampling and data collection....24
4.4 Exposure assessment....26
4.4.1 Arsenic in drinking water....26
4.4.2 Biomarkers of arsenic exposure....26
4.4.3 Urinary dilution correction....27
4.4.4 Other metals....28
4.5 Toxicity biomarkers....28
4.5.1 Telomere length and mitochondrial DNA copy number....29
4.5.2 8-oxo-dG and 4-HNE-MA....29
4.5.3 Identifying potential novel biomarkers....30
4.6 Genotyping....30
4.6.1 TaqMan allelic discrimination assay....31
4.6.2 SNP arrays....31
4.6.3 Targeted sequencing....31

4.7 Population structure....32
4.7.1 Principal component analysis....32
4.7.2 Ancestry estimation....32
4.8 Selection scans....33
4.8.1 iHS and XP-EHH....33
4.8.2 LSBL....34
4.9 Statistical analyses....35
5 RESULTS AND DISCUSSION....39
5.1 Characteristics of the study participants....39
5.2 Arsenic exposure assessment....41
5.3 Arsenic metabolism and influencing factors....43
5.4 Molecular effects of arsenic exposure....46
5.5 Genetic adaptation to arsenic exposure....49
5.6 Methodological considerations....52
5.7 Ethical considerations....54
5.8 Public health implications....55
6 CONCLUSIONS....57
7 FUTURE PERSPECTIVES....59
8 ACKNOWLEDGMENTS....61
9 REFERENCES....65

LIST OF ABBREVIATIONS

4-HNE-MA	4-hydroxy nonenal mercapturic acid
8-oxo-dG	8-oxo-2'-deoxyguanosine
As^{III}	Arsenite
AS3MT	Arsenite methyltransferase
As^{V}	Arsenate
ATP	Adenosine triphosphate
B-As	Arsenic concentration in blood
DMA or DMA^{V}	Dimethylarsinic acid
DMA^{III}	Dimethylarsinous acid
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
FASLG	Tumor necrosis factor ligand superfamily member 6
GSTO1	Glutathione S-transferase
GTEx	Genotype-Tissue Expression project database
GWAS	Genome-wide association studies
HBB	Hemoglobin beta
hg19	Human reference genome build version 37
HNO_3	Nitric acid
HPLC-HG-ICP-MS	High-performance liquid chromatography online with hydride generation and ICP-MS
IARC	International Agency for Research on Cancer
iAs	Inorganic arsenic as the sum of As^{III} and As^{V} in urine
ICP-MS	Inductively coupled plasma-mass spectrometry
iHS	Integrated haplotype score
Kb	Kilobase
LC-MS/MS	Liquid chromatography coupled with tandem mass spectrometry
LSBL	Locus-specific branch length
LYPD3	Ly6/PLAUR domain-containing protein 3
MMA or MMA^{V}	Monomethylarsonic acid

MMA^{III}	Monomethylarsonous acid
mtDNAcn	Mitochondrial DNA copy number
NGI	National Genomic Infrastructure (Uppsala, Sweden)
NPX	Normalized Protein eXpression
PCR	Polymerase chain reaction
PEA	Proximity Extension Assay
PNP	Purine nucleoside phosphorylase
SAC	San Antonio de los Cobres (Argentina)
SAM	S-adenosyl methionine
SEZ6L	Seizure 6-like protein
SMRT	Single-Molecule, Real-Time technology
SNP	Single nucleotide polymorphism
TFPI2	Tissue factor pathway inhibitor 2
TMAO	Trimethylarsine oxide
U-As	Urinary arsenic as the sum of iAs, MMA, and DMA concentrations
XP-EHH	Cross population extended haplotype homozygosity

1 PREFACE

Inorganic arsenic is one of the most potent carcinogens found in the environment. It is estimated that more than 140 million people around the world are exposed to inorganic arsenic via drinking water. Arsenic is naturally released into drinking water sources due to leaching processes. In the Andean plateau in South America, arsenic is easily released from the arsenic-rich volcanic bedrock. This creates a scenario in which indigenous populations have been, and are still today, exposed chronically over multiple generations to elevated arsenic concentrations via their drinking water.

In 1995, a group of indigenous Andean women living in the remote village of San Antonio de los Cobres in northwestern Argentina was described as having a distinctively efficient arsenic metabolism, a protective factor against arsenic toxicity (Vahter et al., 1995). This Andean region is characterized by very high arsenic concentrations in their drinking water (up to 200 µg/L) and in the nearby river (800 µg/L). Later studies focusing on this population revealed important genetic insights about the metabolism of arsenic. Variants in *AS3MT*, the gene coding for the arsenite methyltransferase, were associated with a more efficient arsenic methylation. These Andean women from San Antonio de los Cobres had a very high frequency of the protective *AS3MT* genetic variants compared to other populations worldwide (Engström et al., 2011; Schläwicke Engström et al., 2009; Schlebusch et al., 2015). This led to the first description of human adaptation to arsenic, supported by traits of genomic positive selection around *AS3MT* in these women (Schlebusch et al., 2015). This discovery directed us towards the current question approached in this thesis: has this adaptation to tolerate arsenic and metabolize it more efficiently occurred elsewhere?

Following the Andean mountain range, the Bolivian plateau has similar geological settings as in Argentina. Villages around Lake Poopó in Bolivia, situated in the central region of the highlands, are known to contain elevated arsenic concentrations in drinking water (Ormachea Muñoz et al., 2013; Ramos Ramos et al., 2012). Despite this, the extent to which these inhabitants around Lake Poopó were exposed to arsenic was not known. Therefore, one of the aims of this thesis was to characterize human exposure to environmental arsenic and the arsenic metabolism efficiency of indigenous Bolivians living around Lake Poopó.

Limited data exist on whether individuals with an efficient arsenic metabolism are less likely to develop arsenic-related health outcomes. The inherent difficulty to obtain accurate clinical data in these remote villages in the Andes stresses the need to identify and evaluate other biomarkers of toxicity in relation to arsenic. Therefore, this thesis work also focused on exploring the effects of arsenic in the Bolivian communities by using a combination of traditional and potential novel biomarkers of toxicity.

The work included in this thesis meets at the intersection of molecular epidemiology and population genetics. It hopes to help unravel how indigenous communities in the Andes have adapted to living in arsenic-rich environments and what consequences this long-lasting exposure has had for them in terms of toxicity and genetic adaptation.

2 BACKGROUND

2.1 ARSENIC

2.1.1 Forms, sources, and exposures

While arsenic is known in popular culture as a poison used to discreetly commit murders, humans are exposed unintentionally to this metalloid because of its ubiquitous presence in the environment. Arsenic forms are categorized as inorganic and organic (**Figure 1**). The inorganic forms include arsenite (As^{III}) and arsenate (As^{V}), which are the major forms in water. Organic forms are found primarily in food items as arsenosugars, arsenobetaine, arsenocholine, and arsenolipids; or are the products of the metabolism of inorganic arsenic, mostly as monomethylarsonic acid (MMA) and dimethylarsinic acid (DMA). Even inorganic arsenic and DMA are found in food.

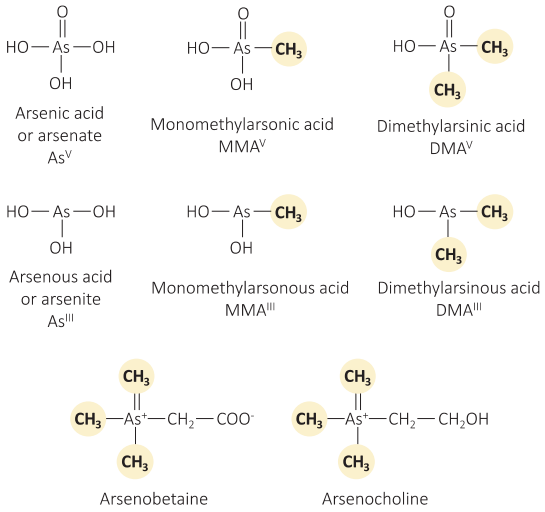


Figure 1. Summary of arsenic forms. The forms including a methyl group (-CH_3) are organic forms.

Arsenic is hardly ever found in nature in its elemental state. It is, however, a component of more than 200 minerals in its inorganic forms, often in combination with sulfur. The most common arsenic-containing minerals are realgar (or arsenic disulfide), arsenopyrite, and orpiment (or arsenic trisulfide). The erosion of these naturally occurring minerals is the main contributor to arsenic leaching into surface and groundwater. Anthropogenic activities, such as mining and the use of arsenic-containing fertilizers, also release arsenic into the environment.

For humans, drinking water is the main source of exposure for inorganic arsenic. In 1993, a guideline value of 10 µg/L of arsenic in drinking water was established (WHO, 2017). In the South American continent, this guideline value was incorporated in national regulations of several countries, including Bolivia, whilst other countries such as Uruguay, Peru, and Venezuela maintained the previous higher limit of 50 µg/L (as reviewed in Bundschuh et al., 2012).

Food is also a considerable source of arsenic. In fact, rice is an important exposure source of inorganic arsenic after drinking water, and it can also contain DMA (Kumarathilaka et al., 2019; Zhao et al., 2013). In certain types of seafood, such as crabs and mollusks, MMA and DMA have also been detected (Borak and Hosgood, 2007). Seaweed has also been reported to contain high concentrations of arsenosugars and inorganic arsenic (Molin et al., 2015; Taylor et al., 2017). Additionally, arsenic analogues of betaine and choline (arsenobetaine and arsenocholine), and various arsenosugars and arsenolipids are found mainly in marine animals, but no toxic effects of these organic forms have been clearly associated with human consumption. However, despite the generally accepted idea that organic arsenic forms are not toxic, the scarce supporting data highlight the need for further assessment (Bornhorst et al., 2020; Molin et al., 2015).

2.1.1.1 Exposure in Bolivia

Elevated arsenic concentrations in drinking water are widespread in Latin America (Bundschuh et al., 2012; Khan et al., 2020). In the Andes, the longest mountain range running north to south through the South American continent, arsenic is released mainly into the environment due to leaching processes of arsenic-rich volcanic rocks. The Andean plateau refers to the region with the widest surface of highland and covers southern Peru, western Bolivia, and northern areas of Chile and Argentina (Rupert and Hochachka, 2001). In some of these Andean countries, such as Argentina and Chile, human exposure to inorganic arsenic has been described in detail (Concha et al., 1998a; Francisca and Carro Pérez, 2009; Nicolli et al., 2010; Smith et al., 2018; Vahter et al., 1995). However, to date little is known about human exposure to arsenic in Bolivia.

A literature review on the problem of arsenic in Latin America highlighted two areas of concern in Bolivia (Bundschuh et al., 2012): the Pilcomayo River basin, in the western Andean region of Bolivia, due to its contamination from mining activities in the area; and the axis of the Titicaca-Desaguadero-Poopó-Salt basin, mostly being affected by naturally occurring arsenic and to some extent by mining activities. The Titicaca-Desaguadero-Poopó-Salt basin consists of the northern Lake Titicaca draining through the Desaguadero River into Lake Poopó. In the areas around Lake Poopó, arsenic concentrations in ground and surface water have been characterized in order to understand the geochemistry of the metalloid in this region (Garcia Moreno, 2006; Ormachea Muñoz et al., 2016, 2013; Ramos Ramos et al., 2012; Stassen et al., 2012; Van Den Bergh et al., 2010). In villages around Lake Poopó, up to 250 µg/L of arsenic have been described in water (Ormachea Muñoz et al., 2013; Ramos Ramos et al., 2012).

To properly estimate the health impacts of arsenic on humans, detailed exposure assessments must be carried out. In Bolivia, several studies have explored human exposure to arsenic. Archer et al. (2005) reported that total arsenic concentrations in urine ranged 11–891 µg/L in villages in the upstream area of Pilcomayo River with low levels of arsenic in drinking water. This study, however, only assessed total arsenic concentrations and therefore, the possibility of exposure to e.g., arsenobetaine or other less-toxic arsenic forms cannot be discarded. Moreover, Smolders et al. (2006) noted that people living downstream of Pilcomayo River (n = 83) were

exposed to inorganic arsenic based on arsenic concentrations in hair samples. New data on human exposure to inorganic arsenic in Bolivia are steadily emerging, including several studies exploring arsenic exposure in the mining city of Oruro (Barbieri et al., 2016; Goix et al., 2016, 2011; Pavilonis et al., 2017). However, little is still known about human exposure to inorganic arsenic in Bolivia. Therefore, a comprehensive analysis is necessary to shed light on arsenic exposure levels and to help estimate the magnitude of the problem. Such studies are also essential to carry out follow-up analyses on efficient arsenic detoxification, as in nearby study populations in Argentina (Vahter et al., 1995).

2.1.2 Toxicokinetics in humans

2.1.2.1 Absorption and distribution

In the general population, humans are exposed to arsenic mainly via ingestion; generally, only minute amounts of arsenic are inhaled. After an ingested dose of inorganic arsenic, 80–90% of the dose is absorbed in the gastrointestinal tract (Vahter and Norin, 1980). Both As^{III} and As^{V} are taken up in the gastrointestinal tract, and As^{V} is then reduced in the blood to its trivalent form (Vahter and Envall, 1983). Once reduced, As^{III} is transported to the liver, where it is metabolized (**Figure 2**).

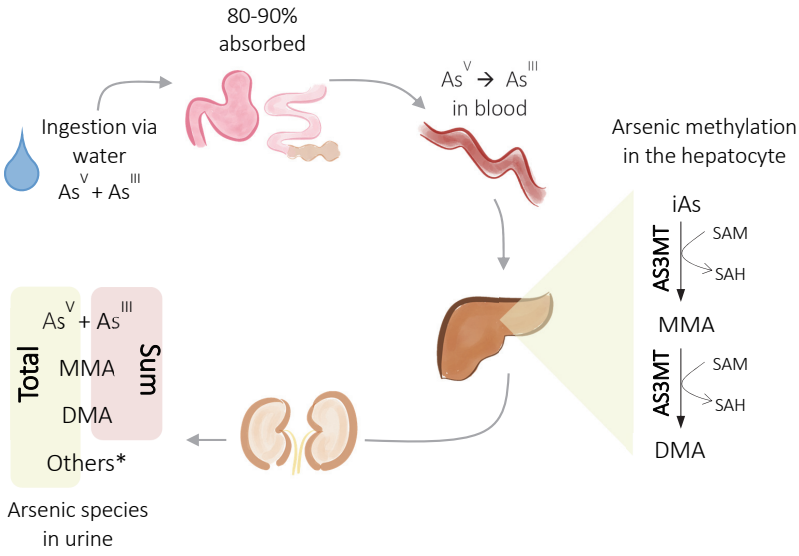


Figure 2. Kinetics of inorganic arsenic in humans after exposure via drinking water and simplified methylation pathway in the hepatocyte. *Note:* Others* refers to arsenic metabolites that are excreted in urine but are not the product of inorganic arsenic metabolism, e.g., arsenobetaine or arsenocholine. *Abbreviations:* As, arsenic; AS3MT, arsenite methyltransferase; DMA, dimethylarsinic acid; iAs, inorganic arsenic; MMA, monomethylarsonic acid; SAH, S-adenosyl homocysteine; SAM, S-adenosyl methionine.

Arsenic crosses biological membranes by different mechanisms. To achieve this, As^V is known to compete with inorganic phosphate for sodium-phosphate transporters during cellular uptake (Villa-Bellosta and Sorribas, 2008). In the case of trivalent forms, several membrane transport proteins are involved including the families of aquaporins and multidrug resistance proteins (as reviewed in Garbinski et al., 2019). In enterocytes, hepatocytes, and epithelial cells located in the intestine, liver, and kidneys respectively, several types of these proteins regulate the traffic of trivalent arsenic forms through membranes (as reviewed in Roggenbeck et al., 2016). In addition, arsenic is known to cross the placenta (Concha et al., 1998b), but more research is still required to understand this transport at a molecular level.

2.1.2.2 Metabolism

The metabolism of inorganic arsenic in humans takes place primarily in the liver. Arsenic is metabolized by a series of reduction and methylation reactions, resulting in the production of MMA and DMA (Vahter, 2000). AS3MT is the main arsenic methylating enzyme, which uses S-adenosyl methionine (SAM) from the one-carbon metabolism cycle as a methyl donor (Koller et al., 2020; Lin et al., 2002; Marafante and Vahter, 1984).

Not all absorbed inorganic arsenic is metabolically transformed in humans: non-methylated inorganic arsenic is excreted in urine along with the methylated products, MMA and DMA. Humans produce mainly MMA and DMA after the ingestion of inorganic arsenic. However, an oxidized trimethylated species (trimethylarsine oxide, TMAO) has also been detected in urine of humans exposed to DMA or arsenosugars in a controlled manner (Francesconi et al., 2002; Marafante et al., 1987). The reason why humans do not generally produce TMAO is probably due to the fast excretion of DMA resulting in low concentrations of DMA in hepatocytes (see Section 2.1.2.3 for more details). Interestingly, the methylation of MMA to DMA may also be inhibited by inorganic arsenic, limiting the amount of DMA in the liver (Lindberg et al., 2008a; Marafante et al., 1987). Other thioarsenic metabolites containing sulfur have been identified in human urine after the ingestion of arsenosugars (Francesconi et al., 2002; Raml et al., 2005). While thio-DMA was initially thought to be an exclusive metabolite of arsenosugars, it was recently identified in urine from Bangladeshi women exposed mainly to inorganic arsenic (Raml et al., 2007).

Although arsenic methylation seems to be present among most species except, for example, insects and angiosperms (Palmgren et al., 2017), there are major differences in arsenic metabolism and kinetics among mammalian species (Vahter, 1999). Thus, human studies are required to understand human arsenic metabolism.

2.1.2.3 Excretion and biomarkers of exposure

The half-life of inorganic arsenic in the human body is around 2-3 days (Buchet et al., 1981), and urine is the main route of excretion. Arsenic is also excreted in feces, sweat, nails, and hair but the extent of these routes is insignificant compared to that of urine (ATSDR, 2007). The accumulation of arsenic in hair and nails is due to the metalloid's affinity for sulfhydryl groups, which are abundant in the keratin of these tissues. Of all arsenic metabolites, DMA is most

rapidly cleared from the liver and excreted via urine (Gamble et al., 2007; Marafante et al., 1987). On average for humans, the arsenic metabolite fractions in urine are 10–30% unmethylated inorganic arsenic (iAs, defined as the sum of As^{III} and As^{V}), 10–20% MMA, and 60–70% DMA (Vahter, 1999). This indicates a marked variation among individuals regarding their capacity to methylate arsenic. In addition, rice consumption can also lead to an increased fraction of DMA in urine (Cascio et al., 2011; Meharg et al., 2014).

Arsenic concentration in urine is the most commonly used biomarker of exposure to assess recent intake (Concha et al., 2006; Marchiset-Ferlay et al., 2012). To evaluate exposure to inorganic arsenic over a longer period of time, arsenic can be measured in nails and hair, but caution must be taken in case of external contamination (Concha et al., 2006; Marchiset-Ferlay et al., 2012; Signes-Pastor et al., 2021). Arsenic in blood is not often used as a biomarker of exposure since it is quickly cleared from this matrix (ATSDR, 2007). However, in populations that are exposed chronically with a constant intake of inorganic arsenic, concentrations in urine and blood are known to correlate, and blood arsenic then serves as a valid biomarker of exposure (Concha et al., 1998a; Hall et al., 2006; Takayama et al., 2021). Distinguishing the different arsenic forms in blood has inherent methodological limitations as well, which discourages its use as a first-choice exposure biomarker. However, advances are being made in this field (Matoušek et al., 2017).

The prevailing method to determine total arsenic concentrations in urine is inductively coupled plasma-mass spectrometry (ICP-MS) (Chen et al., 2014; Francesconi and Kuehnelt, 2004). Regarding inorganic arsenic exposure assessment, it is critical to measure arsenic metabolites together (for instance by hydride generation) or individually (separating them with high-performance liquid chromatography), instead of assessing only total arsenic concentrations. The reason is that total arsenic also includes other organic forms, such as arsenobetaine, which do not generally contribute to arsenic toxicity. Further, by speciating the different arsenic metabolites, one may evaluate how efficient an individual is at metabolizing arsenic.

2.1.3 Toxicity in humans

2.1.3.1 Acute toxicity

Arsenic has been used historically as *the* poison of choice. In fact, during the Middle Ages it was the most common homicide method. Arsenic poisoning went easily undiscovered because inorganic arsenic is odorless and tasteless, and the acute poisoning symptoms − nausea, vomiting, diarrhea, and abdominal pain − are common to other diseases.

Despite its toxicity, or probably due to it, arsenic has also been used therapeutically for more than 2,400 years (Antman, 2001). For instance, during the 18th century, Fowler's solution, a potassium bicarbonate-based solution of arsenic trioxide, was used against malaria, syphilis, asthma, eczema, and psoriasis (Rohe, 1896; Scheindlin, 2005 from Hughes, 2011). In 1910, the magic bullet Salvarsan, discovered by Paul Ehrlich, was the main treatment against syphilis and trypanosomiasis until the discovery of penicillin (Antman, 2001).

Still today, arsenic has a place in modern medicine. Arsenic trioxide is employed as a treatment for specific types of acute promyelocytic leukemia, a cancer form characterized by an accumulation of promyelocytes. Already in 1878, Fowler's solution was reported to reduce white blood cell counts (Antman, 2001). In 2000, the U.S. Food and Drug Administration approved arsenic trioxide as a treatment against acute promyelocytic leukemia (Antman, 2001). Cases of this leukemia are characterized by a chromosomal translocation between chromosomes 15 and 17 that results in the fusion protein PML-RAR-α. Currently there is evidence that arsenic trioxide degrades this fusion protein, which could be related to the drug's selectivity in the treatment of acute promyelocytic leukemia cases (Chen et al., 1997, 2011; Shao et al., 1998).

2.1.3.2 Chronic toxicity

Although a guideline value of 10 µg/L of inorganic arsenic in drinking water is proposed, higher concentrations are estimated to affect at least 140 million people around the world (WHO, 2017). Given the large number of individuals exposed, inorganic arsenic poses a serious concern from a public health perspective.

Enough evidence exists for arsenic to be classified as a human carcinogen of group 1, according to the International Agency for Research on Cancer (IARC, 2012). Exposure to inorganic arsenic via ingestion is associated with multiple types of cancer, such as lung, bladder, and skin; and potentially with cancers of the kidney, liver, and prostate (IARC, 2012). Early signs of chronic arsenic toxicity include changes in skin pigmentation, particularly in palms and soles, followed by hyperkeratosis, a thickening of the skin in the form of patches (Guha Mazumder, 2003). These skin lesions can indicate a susceptibility to develop cancer in relation to arsenic exposure (Hsu et al., 2013).

Chronic exposure to inorganic arsenic via drinking water is also associated with several non-cancer outcomes. Multiple epidemiological studies show an association with cardiovascular and respiratory diseases, diabetes, immunotoxicity, and multiple types of child developmental adverse outcomes (Ahmed et al., 2014; Gardner et al., 2013; Hamadani et al., 2011; Huang et al., 2011; Raqib et al., 2017; Saha et al., 2012; Wang et al., 2007).

In utero exposure to arsenic has been associated with negative health outcomes at birth and later on in life. It has been associated with slightly lower birth weight, respiratory diseases, and overall increased risk of child mortality (Smith et al., 2013, 2012; Vahter, 2009). There is growing evidence that it is also related to an impaired immune function (Ahmed et al., 2014; Raqib et al., 2017). Furthermore, early-life exposure to arsenic has also been associated with a markedly increased risk of respiratory diseases, cancer and mortality in adulthood (Dauphiné et al., 2011; Roh et al., 2018).

2.1.3.3 Mechanism of toxicity

The potency of arsenic resides in its ability to interfere with a plethora of cellular processes through different molecular mechanisms of action. In addition, the oxidation state and the form of arsenic play a fundamental role in its toxicity.

The pentavalent form, As^{V}, is capable of mimicking phosphorus since both have similar physicochemical properties. In this way, As^{V} disrupts the mitochondrial respiratory chain by uncoupling ATP formation in a process termed "arsenolysis" (Doudoroff et al., 1947). This has been shown in cellular experiments where As^{V} depletes the formation of ATP in erythrocytes from rabbits (Delnomdedieu et al., 1994) and humans (Winski and Carter, 1998). However, the internal dose in the human body to As^{V} is relatively low since it is quickly reduced to As^{III}.

Trivalent forms including As^{III}, MMA^{III}, and DMA^{III} have a high affinity to react with thiol groups, for example those in sulfhydryl groups present in proteins (Hughes, 2002). Because of this, the trivalent forms are generally more toxic than their pentavalent counterparts are. Several *in vitro* studies suggest that MMA^{III} is the most toxic form (Hirano et al., 2004; Mass et al., 2001; Petrick et al., 2000; Wang et al., 2015). The increased toxicity of MMA compared to that of DMA is probably related to the fact that DMA is more readily excreted (Marafante et al., 1987).

The affinity of arsenic to react with sulfhydryl groups in proteins is associated with oxidative stress induction, abnormal cell proliferation, protein misfolding and aggregate formation, apoptosis induction, impairment of DNA repair mechanisms, epigenetic alterations, mitochondrial dysfunction, and modification of telomeres (Bhattacharjee et al., 2013; Chen et al., 2019; Hubaux et al., 2013; Hughes, 2002; Jacobson et al., 2012; Khairul et al., 2017). While arsenic is a known carcinogen, it has not been demonstrated to react directly with DNA (as reviewed in Cohen et al., 2016; Mass et al., 2001).

2.1.4 Metabolism efficiency as a susceptibility factor

There is a marked inter-individual variation regarding the efficiency to methylate arsenic (Vahter, 1999). More efficient arsenic metabolizers are individuals who present higher fractions of DMA in urine, and lower fractions of MMA. Not only are they better at methylating the element, but they also have higher excretion rates and less retention since DMA is more readily excreted (Gamble et al., 2007; Vahter, 1999). An indigenous population in the Andes living in the Argentinean village of San Antonio de los Cobres (SAC) had distinctively low MMA and high DMA fractions in urine indicating an efficient arsenic methylation (Vahter et al., 1995).

The efficiency to metabolize arsenic plays a central role in arsenic toxicity. This is mostly because the metabolism of arsenic acts both as a bioactivation and detoxification process: methylating iAs once converts it into a more toxic form (MMA), while methylating it twice (DMA) promotes its excretion. In fact, epidemiological studies have shown that a higher fraction of MMA in urine is associated with health effects related to arsenic such as arsenic-

induced skin lesions, cancer and cardiovascular effects (Ahsan et al., 2007; Antonelli et al., 2014; Chen et al., 2003; Lindberg et al., 2008b; Pierce et al., 2013; Steinmaus et al., 2006; Tseng et al., 2005). This supports the idea that a less efficient arsenic metabolism is a susceptibility factor for arsenic toxicity. Throughout this thesis, higher fractions of MMA^{V} in urine were considered an indicator of arsenic-related toxic outcomes, assuming it reflects a higher internal dose of MMA^{III}. However, lower MMA and higher DMA fractions in urine have also been linked to an increased incidence of diabetes, although further research is still needed to understand their causal role (Kuo et al., 2015; Mendez et al., 2016; Nizam et al., 2013).

2.1.5 Factors influencing arsenic metabolism

The variation in urinary arsenic metabolite patterns among individuals is known to be governed by several factors. Sex, age, smoking, nutritional status, the dose of inorganic arsenic itself, and genetics (see Section 2.1.6 for more details) modulate how well humans cope with arsenic in their bodies.

Women have a greater arsenic methylation capacity in comparison to men (Lindberg et al., 2008a; Tseng, 2009; Vahter et al., 2007). The underlying cause is thought to be estrogens. These primary female sex hormones enhance the endogenous synthesis of choline and betaine, the latter being involved in the re-methylation of homocysteine to methionine, in parallel to the folate-vitamin B_{12} cycle (Vahter, 2009). Increased methionine results in higher concentrations of SAM, the methyl donor within the one-carbon metabolism that is also used to methylate arsenic. The increased arsenic metabolism capacity of women is further enhanced during pregnancy (Gardner et al., 2011), supporting the contribution of estrogens to arsenic metabolism.

Children are also known to be more efficient at methylating arsenic than adults (Skröder Löveborn et al., 2016; Sun et al., 2007; Tseng, 2009). This is thought mainly to be due to the upregulation of the one-carbon metabolism for the production of SAM in children while they are growing (Hall et al., 2009), and/or due to children being exposed to fewer inhibiting factors, such as smoking.

Regarding modifiable factors, smoking is associated with an impaired one-carbon metabolism. Smokers have lower levels of the one-carbon cofactors folate, vitamin B_6 and B_{12}, and higher levels of homocysteine (O'Callaghan et al., 2002). This can lead to lower concentrations of SAM and a reduced arsenic methylation capacity. In fact, smokers present higher MMA fractions than non-smokers (Hopenhayn-Rich et al., 1996; Lindberg et al., 2010).

Multiple studies have also assessed how the nutritional status of humans affects their capacity to methylate arsenic. Primarily, diets deficient in protein, which means the methyl group intake is reduced, may result in a less efficient arsenic methylation (Kordas et al., 2016; Steinmaus et al., 2005; Vahter and Marafante, 1987), this effect being less pronounced in women (Li et al., 2008).

Elevated arsenic exposure levels can inhibit the methylation of inorganic arsenic to MMA and DMA. Epidemiological studies found that higher arsenic concentrations in urine were associated with lower DMA fractions, i.e., those individuals had an impaired arsenic methylation (Hopenhayn-Rich et al., 1996; Olmos et al., 2015; Skröder Löveborn et al., 2016; Sun et al., 2007). This idea is supported by *in vitro* studies showing that arsenic metabolism enzymes were inhibited by inorganic arsenic (Buchet and Lauwerys, 1988; Styblo et al., 1996).

2.1.6 Genes involved in arsenic metabolism

2.1.6.1 Candidate-gene approach

Little more than a decade ago, only few studies approached the idea that genetics could also contribute to the variability of arsenic metabolism. Since then, the knowledge pool has grown enormously, and today there is clear evidence that variations in genes related to arsenic metabolism affect the body's capacity to cope with inorganic arsenic. Not only are these genetic variants associated with arsenic metabolism efficiency, but some have also been shown to be associated with arsenic-related toxic outcomes.

The key enzyme in the methylation of inorganic arsenic is AS3MT (first annotated as cyt19), an arsenite methyltransferase capable of methylating inorganic arsenic, MMA, and DMA (Thomas et al., 2007). The genetic structure of *AS3MT* is composed of 11 exons over approximately 32,000 base pairs and it is located on chromosome 10 (10q24.32). The role of AS3MT in arsenic metabolism has been demonstrated in both *in vitro* and *in vivo* studies. Human urothelial cells that did not express *AS3MT* were not capable of methylating arsenic (Drobná et al., 2005), and *As3mt* knockout mice presented an impaired arsenic methylation capacity (Drobná et al., 2009). Regarding arsenic methylation variability within humans, variants in *AS3MT* have been shown to predict arsenic metabolite fractions in populations throughout the world (Antonelli et al., 2014; Engström et al., 2007; Ghosh et al., 2008; Hernández and Marcos, 2008). Despite this, it is still unclear which variants of *AS3MT* are causative. In a review by Antonelli et al. (2014), 360 articles published before 2013 found a total of 15 *AS3MT* variants associated with arsenic methylation. However, only three of these (rs3740390, rs11191439, and rs11191453) were influential across populations. Nevertheless, this heterogeneous number of variants associated with arsenic methylation throughout studies could be explained partly by the strong linkage disequilibrium that exists around *AS3MT* (Engström et al., 2011; Gomez-Rubio et al., 2010; Schläwicke Engström et al., 2009).

Most of the *AS3MT* polymorphisms associated with arsenic methylation are non-coding. This indicates that these genetic variants are likely to contribute to arsenic methylation differences by modulating the gene expression of *AS3MT*. The adrenal gland, spleen, brain, and liver are the organs with the highest expression of *AS3MT*, but its function in other organs apart from the liver is still unknown. Genetic variants associated with a more efficient arsenic methylation in the Andean population from SAC have been correlated with a reduced gene expression of *AS3MT* in peripheral blood (Engström et al., 2011, 2013). These seemingly counter-intuitive results may be explained by gene expression being measured in blood instead of in liver tissues,

the main methylating organ. Recently, a study based on tissue-specific gene expression data from the GTEx database (Genotype-Tissue Expression project) found that *AS3MT* variants associated with a less efficient arsenic methylation also had a reduced *AS3MT* expression in several tissues, but not in liver (Chernoff et al., 2020).

A longer list of genes involved has additionally been in the spotlight (as reviewed in Antonelli et al., 2014; Ghosh et al., 2008; Hernández and Marcos, 2008); for instance, enzymes in charge of the reduction of arsenic. One example is the MMA^V reductase, in humans encoded by the gene *GSTO1*, which is a glutathione S-transferase (Zakharyan et al., 2001). Glutathione S-transferases are a superfamily of enzymes that are crucial during metabolism because they catalyze the conjugation to reduced glutathione. Another candidate as a catalyzer for the reduction of pentavalent arsenic is the purine nucleoside phosphorylase enzyme, encoded by *PNP* (Németi et al., 2010). Variations in *PNP* have been associated with arsenic-related skin lesions (De Chaudhuri et al., 2008). Genetic variants in other methyltransferases, such as *N6AMT1*, have also been associated with arsenic methylation (Chen et al., 2017; de la Rosa et al., 2017; Harari et al., 2013). Arsenic methylation capacity is also affected by genetic variants related to the one-carbon metabolism enzymes (Hernández and Marcos, 2008; Schläwicke Engström et al., 2009). Lastly, a group of genes that could influence the toxicity of arsenic are those involved in DNA repair mechanisms that respond to the genetic insults induced by arsenic-related oxidative stress (Ghosh et al., 2008).

2.1.6.2 Genome-wide approach

With the development of technologies that enable researchers to produce and handle large amounts of data, the way we assess how genetics could influence arsenic metabolism has changed. It is now possible to evaluate thousands and millions of genetic polymorphisms at a time. This shift allows us to study the effect of genetics on arsenic metabolism and on arsenic-related toxic outcomes at a genome-wide level. The use of genome-wide association studies, also known by their abbreviation GWAS, introduces a new conceptual framework that is hypothesis-free. A summary of studies that have evaluated the genetic contribution of arsenic metabolism in humans using a genome-wide approach is included in **Table 1**. The main arsenic-related outcome assessed in these genome-wide studies has been skin lesions (Argos et al., 2018; Kibriya et al., 2017; Pierce et al., 2013, 2012).

Signals in *AS3MT* are the common denominator throughout these studies (Balakrishnan et al., 2017; Gao et al., 2015b; Pierce et al., 2012; Schlebusch et al., 2015, 2013). Other signals have also been detected in different regions of the genome apart from *AS3MT* (Gribble et al., 2015; Pierce et al., 2019; Schlebusch et al., 2015; Tellez-Plaza et al., 2013). Nevertheless, a functional characterization is still pending to validate the contribution and relation of these genes with arsenic kinetics and toxicity.

Table 1. Genome-wide association studies for arsenic metabolism efficiency or toxicity outcomes in humans.

Country	Study population	Genetic data	Participants included	n	Main results	Reference
Bangladesh	HEALS BEST	Illumina HumanCytoSNP-12 v2.1 ~300,000 SNPs	Adults	1,313	- GWAS signal for *AS3MT* for both MMA and DMA. - SNP rs9527 (from *AS3MT*) associated with skin lesions in a follow up with case-control subjects.	Pierce et al., 2012
Bangladesh	HEALS BEST	Illumina HumanCytoSNP-12 v2.1 ~300,000 SNPs	Adults	2,053	- Heritability estimation: 16% of DMA fraction was explained by the studied SNPs, when including distant relatives. - Strongest signal around *AS3MT*.	Gao et al., 2015
Bangladesh	HEALS	Illumina HumanCytoSNP-12 v2.1 ~300,000 SNPs and 4,829 CNVs	Adults (no skin lesions at enrollment)	2,171	- Higher risk of developing arsenic-induced skin lesions over time associated with genomic deletions in several genes (*OR5J2, GOLGA6L7P, APBA2, GALNTL5, VN1R31P, PHKG1P2, SGCZ,* and *ZNF658*).	Kibriya et al., 2017
Bangladesh	HEALS BEST	Illumina HumanCytoSNP-12 v2.1 ~300,000 SNPs	Adults	5,354	- Gene-environment interaction study using gene expression and DNA methylation data. Suggestive signal on chromosome 1 for SNP interaction with arsenic-induced skin lesions.	Argos et al., 2018
Bangladesh	HEALS BEST	Illumina exome array v1.1 ~240,000 SNPs	Adults	1,660 / 4,873	- Coding SNP in *FTCD* associated with all inorganic arsenic metabolite fractions, and with increased risk of arsenic-induced skin lesion.	Pierce et al., 2019
USA	Strong Heart Family Study	400 microsatellite markers	American Indian adults	2,907 / 487	- Heritability estimation: ~ 50% for all inorganic arsenic metabolite fractions. - Localized a QTL on chromosome 5 (for %iAs), 10 (for %DMA), and 11 (for %iAs and %MMA).	Tellez-Plaza et al., 2013
USA	Strong Heart Family Study	400 microsatellite markers	American Indian adults	2,189	- Localized a QTL on chromosome 10 explained partially by *AS3MT* variants. - Localized a QTL on chromosome 5 when assessing it with PCA.	Gribble et al., 2015
USA	Strong Heart Family Study	Illumina Cardio-MetaboChip ~200,000 SNPs	American Indian adults	2,428	- Variants in chr:10q24 associated with %iAs, %MMA, and %DMA.	Balakrishnan et al., 2017
Argentina	SAC and villages near Salta	806 microsatellite markers	Indigenous women in the Andes	346	- Higher frequency of protective *AS3MT* haplotype in Argentina compared to other populations, suggestive of positive selection.	Schlebusch et al., 2013
Argentina	SAC	Illumina Omni5M Chip ~4.3 million SNPs (GWAS with 1.3 million SNPs after filtering)	Indigenous women in the Andes	124	- Strong association between *AS3MT* SNPs and %MMA and %DMA. - Weaker association in chromosome 21 with arsenic metabolite fractions. Genes in the area: *LSS*, *MCM3AP*, and *YBEY*.	Schlebusch et al., 2015
Bolivia	Inhabitants near Lake Poopó	Illumina Omni5Exome Chip ~4.5 million SNPs (GWAS with 1.7 million SNPs after filtering)	Indigenous adults in the Andes	180 / 163	- No clear association for inorganic arsenic metabolites in the GWAS, but an association when focused on the protective *AS3MT* haplotype. - Signs of positive selection and highest frequency of protective *AS3MT* variants described.	**Paper IV**

Abbreviations: AS3MT, arsenite methyltransferase; BEST, Bangladesh vitamin E and selenium trial; CNV, copy number variation; DMA, dimethylarsinic acid; *FTCD,* formiminotransferase cyclodeaminase; GWAS, genome-wide association study; HEALS, health effects of arsenic longitudinal study; iAs, inorganic arsenic; MMA, monomethylarsonic acid; PCA, principal component analysis; QTL, quantitative trait locus; SAC, San Antonio de los Cobres (Argentina); SNP, single nucleotide polymorphism.

Heritability has also been evaluated to estimate how much of the arsenic metabolism variation is due to genetic variation. Using only close relatives, the heritability for %DMA estimated by Gao et al. (2015) was 63%, in line with the estimation of 59% from a family-based study (Tellez-Plaza et al., 2013). However, when Gao et al. (2015) included distant relatives, the heritability decreased to 16%. The authors suggested that this discrepancy could be due to rare variants (not included in the genotyping arrays) contributing to the variability in arsenic metabolism capacity. Overall, these studies published until today regarding human arsenic metabolism using genome-wide data have been carried out with study groups in Bangladesh, the United States, and the Argentinean Andes. This manifests the need to expand this type of studies in the future to new populations and regions around the world.

Discussions regarding arsenic and the human epigenome have also arisen during recent years. Multiple efforts have been directed towards grasping how epigenetics markers throughout the genome are affected by arsenic exposure and how these modulate arsenic metabolism capacity (Bailey et al., 2013; Bozack et al., 2020; Broberg et al., 2014; DiGiovanni et al., 2020; Engström et al., 2013; Kaushal et al., 2017; Somnath et al., 2017).

Interestingly, genome-wide data have also been used for arsenic studies in plant biology. The focus of these has been to understand how arsenic accumulates in grains, such as rice, to potentially give solutions on how to limit its accumulation in crops in order to avoid exposure (Chao et al., 2014; Hwang et al., 2017).

2.2 POPULATION GENETICS

Population genetics aims to understand past evolutionary processes (Ewens, 2010). By studying allele frequencies within and among populations, this field can potentially reconstruct demographic and selective processes that have shaped the gene pool of today's populations. Population genetics concerns the study of population structure and adaptation, among others.

2.2.1 Population structure of Andean indigenous groups

Population structure is defined as the departure from random mating in a population. For example, people often produce offspring with those nearby. This creates specific genetic patterns that correlate with the geographical distribution of those populations (Coop et al., 2009). In this sense, population structure is vital to assess the history of populations since it reflects demographic and migratory movements. It is important to note that population structure (or stratification) can affect association studies since it can act as a confounding factor (Bauchet, 2008).

The most widespread indigenous populations in the Andean plateau are the Aymara and the Quechua. These two ethnic groups are primarily linguistically defined, and they were the predominant languages of the Andean plateau during the arrival of the Spaniards (Sandoval et al., 2013). Yet in 1990, in the highlands of Ecuador, Peru, Bolivia, and Argentina, more than 6 million people spoke Quechua, while almost 2 million inhabitants near Lake Titicaca and La Paz spoke Aymara (Rupert and Hochachka, 2001).

The Uru live in the Andean plateau between present-day Bolivia and Peru, and they have been historically isolated from other surrounding communities (de la Barra Saavedra et al., 2011). In Peru, the majority of Urus live in the floating islands of Lake Titicaca, while in Bolivia they are separated in three independent communities. The Uru-Chipaya live near Lake Coipasa, the Uru-Irohito on the banks of Desaguadero River, and the Uru-Murato (or Uru-Poopó) live in several villages surrounding Lake Poopó (Sandoval et al., 2013).

Most research investigating the genetic history of populations in the Andes has included the main groups: Aymara and Quechua. The Aymara and Quechua ethnicities are fairly similar from a genetic perspective (Batai and Williams, 2014; Lindo et al., 2018). Furthermore, a study of mitochondrial DNA data from Bolivian individuals suggested that those populations living in the foothills were more closely related to Andean populations than to Amazonian ones (Corella et al., 2007). Regarding how the Andes were populated from a broader scope, a population structure study revealed that the Peruvian highlands might have been a crucial region during the early Paleoindian expansions (Gómez-Carballa et al., 2018). The suggested time for when populations diverted between low- and highland is estimated between 9,200 and 8,200 years ago (Lindo et al., 2018).

Whereas most Andean population structure research has delved into understanding the main ethnic groups, one study stands out focusing on the Uru ethnicity minority. In the work by Sandoval et al. (2013), the authors used Y-chromosome and mitochondrial DNA data of

388 Uru individuals from Peru and Bolivia compared to the uniparental markers of Aymara, Quechua, and Arawak populations to evaluate their genetic relationships. Their results showed that the Uru are more closely related to Aymara and Quechua from Lake Titicaca than to the Amazonian Arawaks. They also presented that the Uru populations from Peru and Bolivia are genetically differentiated. This study focused on uniparental markers and therefore, more studies including genome-wide data are needed.

2.2.2 Human adaptation

Humans live in countless and ever-changing habitats with a wide range of lifestyles. The fact that humans can inhabit such different, and sometimes harsh and seemingly inhospitable, geographic environments indicates that our species has a strong ability to adapt. At a genetic level, adaptation occurs when individuals from a specific population, over successive generations, inherit advantageous traits that increase their fitness and survival under specific conditions. Therefore, adaptation refers to traits sculpted by natural selection and is one of the driving forces of evolution (Mitton, 2002).

2.2.2.1 Processes driving adaptation

While adaptation regards the phenotype granting reproductive advantage, natural selection is the process driving and assuring that it is passed on to the coming generations. Therefore, adaptation is said to be a product of natural selection. However, selection is not the only driving force acting on and modulating the frequencies of alleles in populations. Selection along with mutation, gene flow, and genetic drift are four processes that do not work in isolation. These processes are mechanisms of evolution that generate changes in the frequencies of alleles (Andrews, 2010).

Mutation, which is a change of the nucleotide sequence, is the only force able to create new variation. The frequency of neutral mutations changes randomly in each generation. This is known as genetic drift, a process in which populations evolve due to alleles being fixated or eliminated by chance (Kimura, 1968). In addition, gene flow refers to the movement of genes in and out of a population, for instance due to migratory movements (Andrews, 2010).

When a phenotypic trait confers survival and reproductive advantages, the underlying genetic variant will be inherited with higher probability. Over time, if the advantage is maintained, the frequency of this variant will increase in the population. This variant is then under positive selection (Harris, 2013). On the contrary, when a variant is deleterious for the individual, the frequency of the allele decreases with time until it is eliminated in a process known as negative or purifying selection.

As mentioned above, if an allele is beneficial for the individual's survival and reproductive success, its frequency will increase in that population. At the same time, the variant subject to positive selection will *drag along* other neutral variants near it during recombination, forming a "haplotype" that is inherited together in a process termed *genetic hitchhiking*. As a result, these neutral variants nearby will also increase in frequency, and consequently, the variation in

the region surrounding the advantageous allele will decrease and haplotype homozygosity will increase. This process is referred to as a *selective sweep* (Harris, 2013).

Signatures of positive selection in the genome can be detected using different approaches, some of which are summarized in the text by Kimura and Ohashi (2013), and can be based on:

- site frequency within a population (e.g., the Tajima's *D* statistic),
- genetic differentiation between populations (e.g., 2-way and 3-way F_{ST} methods),
- haplotype variation within a population (e.g., Extended Haplotype Homozygosity [EHH] tests), or
- haplotype homozygosity between populations (e.g., the XP-EHH method).

2.2.2.2 Examples of human adaptation in the Andean plateau

The Andes Mountains, at an average altitude of 4,000 meters, offer a unique scenario to study human adaptation. The longest mountain range in the world is one of the harshest environments where humans have thrived. Much attention has been drawn to understanding how humans have adapted to the high-altitude hypoxia, especially in areas such as the Tibet and the Andes. Genetic signatures of positive selection have been identified in relation to hypoxia at high altitude in genes related to cellular oxygen sensing, defense against oxidative stress, modulation of pulmonary surfactant production, and cardiac muscle formation (Bigham et al., 2010; Jacovas et al., 2018; Lindo et al., 2018; Rupert and Hochachka, 2001; Simonson et al., 2015; Valverde et al., 2015). Interestingly, the role of DNA methylation as an epigenetic mechanism has also been studied in relation to how humans cope with high altitude. A genome-wide DNA methylation study comparing Peruvian populations with different histories of living at high or low altitudes identified differentially methylated regions in genes involved in the hypoxia-inducible factor pathway, production of red blood cells, and blood pressure (Childebayeva et al., 2021).

The Andean plateau is, however, not only characterized by extreme altitudes. The unique geochemistry of the Andes Mountains results in a setting in which indigenous populations have been exposed to toxic metals for thousands of years (Núñez et al. 1991). In fact, studies on Andean mummies from Chile revealed elevated arsenic concentrations in their hair and internal organs, confirming that the inhabitants of these areas have been exposed to arsenic over multiple generations (Echeverría et al., 2018; Kakoulli et al., 2014).

In the Argentinean highlands located at around 3,800 meters above the sea, women from San Antonio de los Cobres and surrounding villages were identified as having a unique and efficient arsenic metabolism, with low fractions of MMA in urine (Vahter et al., 1995). Nearly two decades after this discovery, with the further development of genetic analyses, women from the same region were found to have the highest frequency ever described of *AS3MT* polymorphisms associated with a more efficient arsenic metabolism. These women belonging mainly to the Kolla ethnic group presented high frequencies of *AS3MT* alleles linked to a more efficient arsenic metabolism (Schlebusch et al., 2013). It was later identified that these women

had genetic signatures of positive selection around the *AS3MT* region (Schlebusch et al., 2015). Shortly after, similar results were found in a different subset of men and women from the same village using complementary genome-wide selection scans (Eichstaedt et al., 2015). In a less comprehensive study, researchers evaluated the frequency of four variants of *AS3MT* in a population from the Atacama Desert in Chile (Apata et al., 2017). The authors concluded that this population had a high frequency of alleles associated with a more efficient arsenic metabolism; however, no genome-wide selection scans were used to assess traits of positive selection in their genome. A more exhaustive study of the Andean region and the inclusion of new areas will allow us to clarify if this adaptation event has occurred in more populations.

3 RESEARCH AIMS

The overall aim of this doctoral research project was to clarify the human exposure, toxicity, and possible adaptation to arsenic in the Andes Mountains. Specifically, the aims were to:

- Recruit a new study group of the main indigenous groups (Aymara-Quechua and Uru ethnicities) living in the Bolivian Andean highlands to assess their exposure to arsenic through drinking water and their arsenic metabolism efficiency (**Paper I**).

- Evaluate arsenic toxicity and how it may be modified by arsenic metabolism efficiency in the Bolivians (**Paper II** and **III**).

- Determine whether Bolivian populations from the Andes present signs of positive selection related to arsenic (**Paper IV**).

- Identify genetic variants that drive arsenic tolerance in Andean populations (**Paper IV**).

4 MATERIALS AND METHODS

This section summarizes and contextualizes the study area and design, as well as the methods used throughout this thesis. More technical and detailed descriptions of the methods are included in **Papers I–IV**.

4.1 STUDY AREA

The study area comprises numerous villages located around Lake Poopó, 300 km southeast of La Paz, Bolivia (**Figure 3**). Lake Poopó is located on the plateau of the Andes Mountains in Bolivia, also referred to as the Bolivian highlands or *altiplano*, approximately at 3,700 m above sea level (17S-20S; 66W-68W). Lake Poopó receives its water mainly from the Desaguadero River and through the smaller Lake Uru Uru. Lake Poopó and Lake Uru Uru are shallow saline lakes with a maximum depth of 3 and 1 m, respectively, depending on yearly precipitations. Due to climate change and misuse of water sources for agricultural activities, Lake Poopó was declared in a state of emergency in 2015 (Satgé et al., 2017). The area is characterized by a cold semi-arid climate, with pronounced temperature variations within the same day. The rainy season in the Bolivian highlands is between December and March. During these months, the heavy rain hinders the access to the study villages because of the precarious state of the roads.

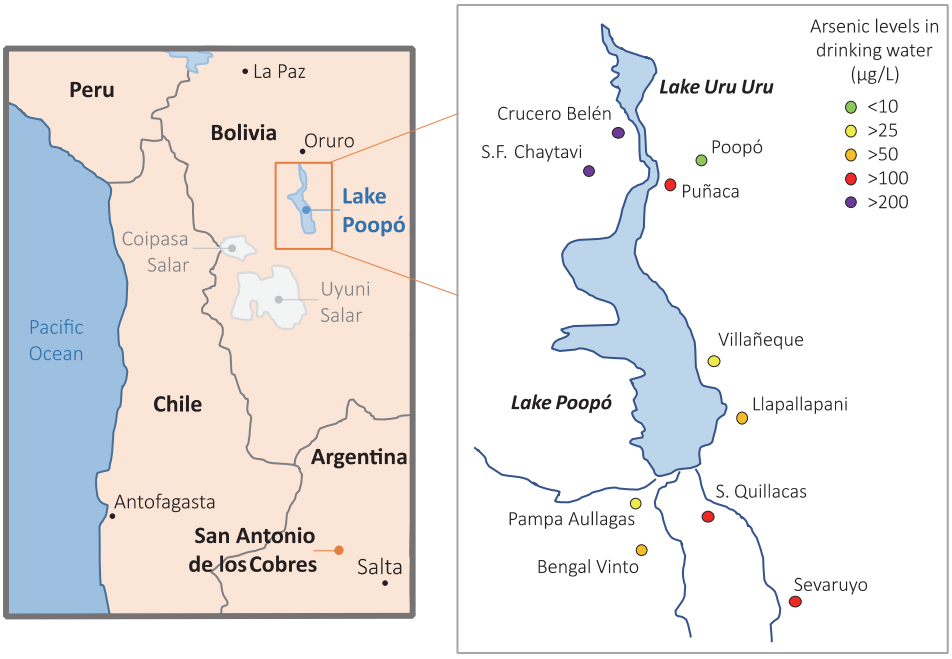


Figure 3. Map of the study area and villages included in the study. Arsenic concentrations in drinking water are the mean values of the sources sampled during the project, which are further described in **Paper I**.

Study participants were recruited from 10 villages situated around Lake Poopó, all located within the Oruro Department: Bengal Vinto, Crucero Belén, Llapallapani, Pampa Aullagas, Poopó, Puñaca, San Felipe de Chaytavi, Santuario de Quillacas, Sevaruyo, and Villañeque (**Figure 3** and **4**). The study area was preselected based on previous literature that identified elevated arsenic concentrations in water from villages of this region (Ormachea Muñoz et al., 2013; Ramos Ramos et al., 2012). Since then, further geochemical characterization of arsenic in groundwater has been described (Ormachea Muñoz et al., 2016).

4.2 STUDY DESIGN AND PARTICIPANTS

In total, 201 women living in 10 villages located around Lake Poopó were recruited into this cross-sectional study. The recruitment of study participants took place between September 2015 and November 2017 during five trips to conduct field work. However, additional visits and meetings were organized at the villages before, during, and after the recruitment period to inform about the arsenic problem and to explain the research project to the inhabitants of the region. Individuals were invited to participate in the study by the local doctors and nurses, during personal visits to the women's houses, or with the help of announcements on the local television. Since the access to these remote villages is limited, the study group resulted from a convenience sampling. Mainly women were invited to participate since men in this region often work outside the village where they live, impeding the proper characterization of their arsenic exposure. In addition, men are more prone to additional exposures such as tobacco and alcohol, which are known to affect the capacity to metabolize arsenic.

The Bolivian study group consists of individuals belonging to the Aymara-Quechua or Uru ethnicities. The Aymara and Quechua are the most extended ethnic groups in the Bolivian Andes. These two ethnic groups cohabit around Lake Poopó and are relatively similar from a genetic perspective (Batai and Williams, 2014; Lindo et al., 2018). Therefore, for the projects within this thesis, they were included as a combined Aymara-Quechua subgroup. The Uru around Lake Poopó, also known as Uru-Murato, live in Llapallapani, Villañeque, and Puñaca, as well as a minority in Poopó. Historically, the Uru have been isolated from their neighboring communities (de la Barra Saavedra et al., 2011). This segregation in the past between Uru and Aymara-Quechua villages allowed us to assess ethnicity based on where the participants, their parents, and grandparents lived. Both ethnic groups depend mainly on agriculture and the trade of handcrafts. In the past, the Uru relied on fishing. However, this was not the case during the recruitment period because of the lake's critical state.

In **Paper IV**, men belonging to the Uru communities (n = 5) were included in the study of genetic selection signals to increase the sample size (following the same sampling protocols as for the women) since these analyses are independent of the arsenic exposure and phenotypic endpoints. In addition, a subset of women (n = 53) from the Argentinean Andes, earlier described as having an efficient arsenic metabolism (Schlebusch et al., 2015), was included in **Paper IV** to evaluate the presence of novel genetic variants in relation to arsenic metabolism. These women were recruited in 2008 in San Antonio de los Cobres (SAC) located at 3,800 m

above sea level in the northern Argentinean Andes. A summary of the study design and participants of **Papers I–IV** is included in **Figure 5.**

Figure 4. Photographs of the study area included in this thesis taken during the field trips. View of (A) Pampa Aullagas (October 2016) and (B) Puñaca (October 2017). Health centers at (C) Sevaruyo (March 2017) and (D) Llapallapani (October 2017). Examples of systems used to obtain drinking water at (E) Santuario de Quillacas (October 2016) and (F) San Felipe de Chaytavi (October 2016). All photographs taken by Jessica De Loma.

4.3 SAMPLING AND DATA COLLECTION

Oral and written information about the study was presented to the participants, and after they gave their written informed consent, we interviewed them in Spanish. The questionnaire included information about their age, parity, time and place of their residency, residency location of their parents and grandparents, which water sources they used, general food and dietary habits, tobacco and alcohol consumption, coca chewing, and personal and family history of diseases. Ethnicity was based on the birthplace of the participants and complemented with information on where their parents and grandparents were from. During the interview, blood pressure was taken in sitting position, and weight and height were measured.

Spot urine samples were collected in plastic cups by the participants during the visit. Participants received instructions on how to proceed with wet-wipe cleaning and how to collect a mid-stream urine sample to avoid bacterial contamination of the urine sample. Directly after, we used urine sticks (Combur 7 Test strips, Roche) to obtain a semiquantitative measure of pH, and the presence of glucose, ketones, leukocytes, nitrites, proteins, erythrocytes, and hemoglobin in urine as indicators of diabetes or urinary tract infection. Urine samples were then transferred to 20 mL polyethylene bottles free of trace elements.

Venous blood samples were taken with BD Vacutainer Eclipse blood collection needles (Becton, Dickinson), and collected in EDTA tubes (Vacuette, Greiner Bio) for DNA extraction and in Trace Elements NH Sodium Heparin tubes (Vacuette) for the exposure assessment of arsenic and other metals. During one of the sampling visits, we did not have Trace Elements NH Sodium Heparin tubes, so we collected blood for metal exposure assessment in Lithium Heparin tubes (Vacuette) instead. Neither type of tube used for exposure assessment had traces of arsenic as confirmed by leach tests. Hemoglobin was measured in venous blood using HemoCue201+ (HemoCue) directly after the blood sampling. DNA was extracted from venous peripheral blood collected in EDTA tubes with the E.Z.N.A. Blood DNA Mini kit (OMEGA Bio-Tek).

All samples were stored at -18 °C in a portable freezer during the field work trips, generally 3 days long. Then, they were stored at -20 °C at the Institute of Genetics at Universidad Mayor de San Andrés, Bolivia until further shipment with dry ice to Karolinska Institutet, Sweden. All samples had similar freeze-thaw cycles.

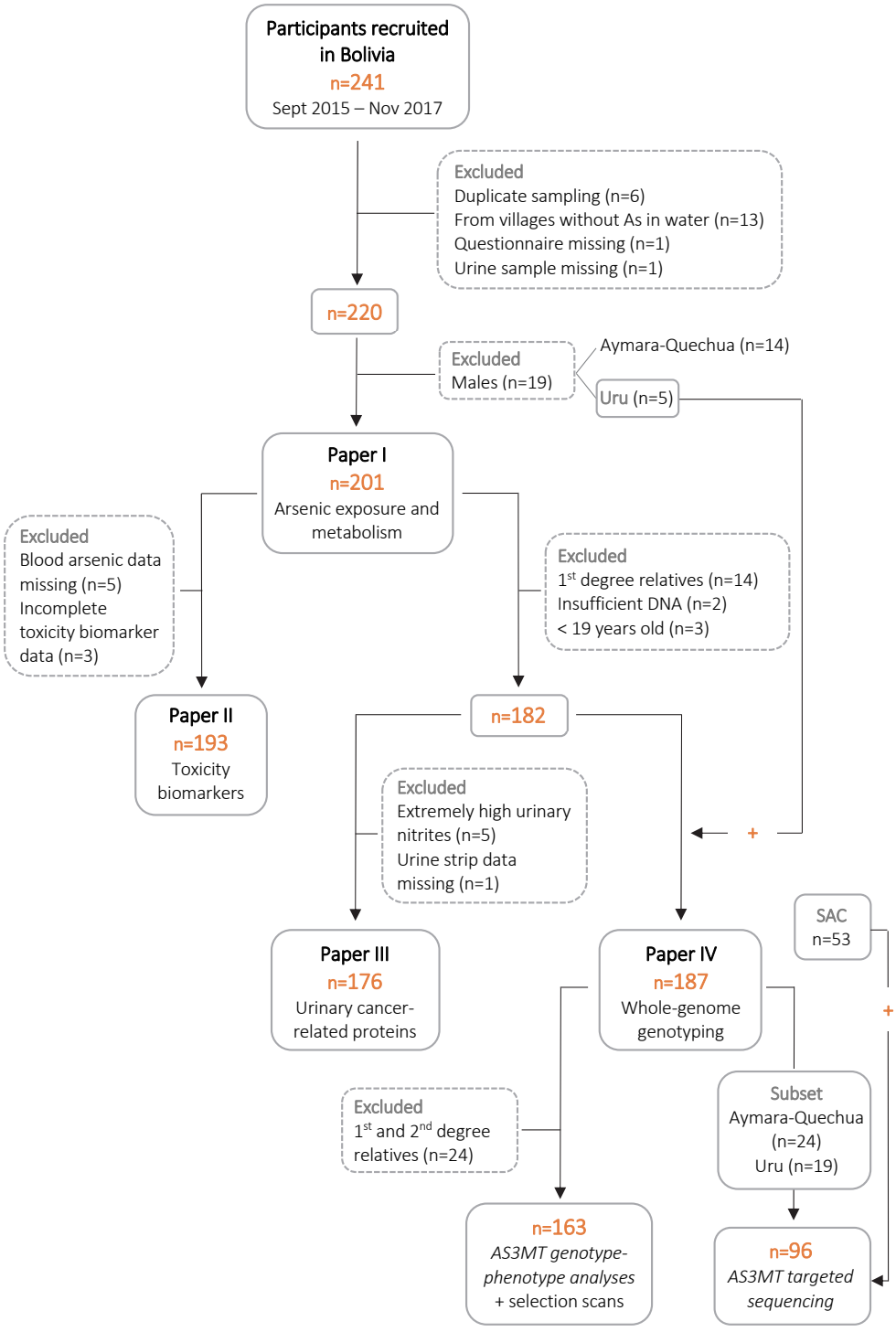


Figure 5. Study design and participants included in each scientific paper of this thesis. First-degree relatives excluded from the subset of **Paper I** were based on data obtained during the interviews. First- and second-degree relatives excluded in **Paper IV** were based on genetic relatedness.

4.4 EXPOSURE ASSESSMENT

All exposure assessment analyses were performed at the Unit of Metals and Health at the Institute of Environmental Medicine, Karolinska Institutet, in a laboratory specialized in trace element analysis.

4.4.1 Arsenic in drinking water

The initial inclusion of villages into the study was based on previous literature describing elevated arsenic concentrations in ground and surface water from villages surrounding Lake Poopó (Ormachea Muñoz et al., 2013; Ramos Ramos et al., 2012). To illustrate the problem of arsenic exposure to the inhabitants, an on-site quick test was used to determine arsenic based on a colorimetric assay (MQuant arsenic test, Merck).

To characterize the exposure to arsenic via drinking water of the study participants, arsenic concentrations were measured in several water sources from each village. For that, we collected water samples from the main drinking water sources as described by the participants during the interviews, including, for example, water from taps, wells, tanks, or rivers. Water coming from taps was left to run a couple of minutes before taking the sample to avoid the contamination by metals accumulated in the pipes. Water samples were collected in trace-element free 20 mL polyethylene bottles and stored and shipped with the other samples (see Section 4.3). At Karolinska Institutet, multielement analyses were performed using ICP-MS (Agilent 7700x, Agilent Technologies).

4.4.2 Biomarkers of arsenic exposure

Total arsenic concentrations in urine, including inorganic and organic arsenic forms, were measured by ICP-MS (Agilent 7700x). Urine samples were thawed overnight at 7 °C and diluted 1:10 with 1% HNO_3 (67% volume/volume).

The concentrations of different arsenic species that result from the metabolism of inorganic arsenic, i.e., iAs as the sum of As^{III} and As^{V}, MMA, and DMA, were measured in urine by high-performance liquid chromatography online with hydride generation and ICP-MS (HPLC-HG-ICP-MS; HPLC Agilent 1110 series; Hamilton column PRP-X100; ICP-MS Agilent 7500ce). The speciation of arsenic was used to (i) assess the exposure to inorganic arsenic, excluding other organic forms of arsenic which are associated with less toxicity such as arsenobetaine; and (ii) evaluate how efficiently arsenic is metabolized by calculating the relative fractions (percentage) of the metabolite concentrations in urine. The sum of inorganic arsenic metabolite concentrations in urine (As^{III} + As^{V} + MMA + DMA; abbreviated as U-As) was used as a biomarker of inorganic arsenic exposure.

Arsenic exposure was also assessed by concentrations of total arsenic in whole blood (B-As). These were measured by ICP-MS (Agilent 7900) after diluting the whole blood with an alkali solution, sonicating, and centrifuging the sample (Levi et al., 2018).

Plastic materials used during the analyses were acid-washed to avoid the contamination by trace elements. All quality control materials (commercial reference materials, in-house samples, and internal standards) used to assure the accuracy of the results are described in **Paper I** for the urine analyses and in **Paper III** for the blood analyses.

4.4.3 Urinary dilution correction

Collecting 24-h or first morning void urine samples is not feasible during the field trips due to practical limitations and low compliance rates. Instead, spot urine samples were collected, which serve as a proxy for arsenic concentrations since these measurements have been seen to correlate with those in 24-h urine samples, particularly when compensating for variations in the urine dilution (Yassine et al., 2011). By doing so, the differences in the concentrations that are not due to differences in exposure are corrected.

Urinary specific gravity and osmolality were measured in the Bolivian study group. Urinary osmolality was measured with a digital cryoscopic osmometer (OSMOMAT 030, Gonotec), and specific gravity with a digital refractometer (RD172, EUROMEX). In the Bolivian study group, U-As adjusted for specific gravity was highly correlated and in agreement with U-As adjusted for osmolality (**Figure 6**).

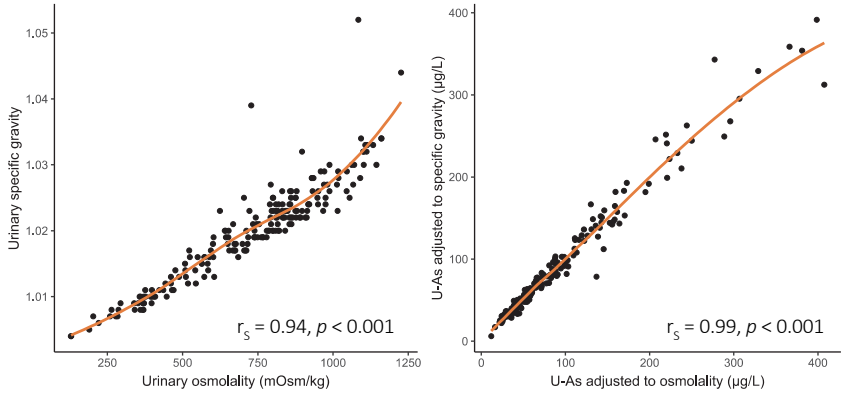


Figure 6. Scatter plots with loess line of two methods to correct for urinary dilution variation: specific gravity or osmolality. Among the three women with specific gravity values > 1.035, one had proteins in urine and the other two had high glucose (the only ones in the study group) based on the urine strips. *Abbreviation*: U-As, urinary arsenic as the sum of inorganic arsenic metabolite concentrations.

Urinary arsenic concentrations were adjusted for the mean urinary osmolality of the study group to correct for variations in urinary dilution throughout the studies. Although osmolality has not been widely used for research purposes, due mostly to the perception of it being costly, osmolality is less influenced by urinary glucose and proteins, and socio-demographic and medical factors than other correction methods (Parikh et al., 2002; Yeh et al., 2015). In fact, recent studies support the use of urinary osmolality as a correction method for the assessment of arsenic exposure (Middleton et al., 2019, 2016). On the contrary, creatinine concentrations are affected by variations in muscle mass, diet, age, gender, and ethnicity (Barr et al., 2005).

Particularly for arsenic, since its metabolism depends on the addition of methyl groups, correcting with urinary creatinine values is not recommended (Nermell et al., 2008).

To explore the toxicity of arsenic in this study group, two other toxicity biomarkers were quantified in urine, which make them susceptible to variations in urinary dilution. The effect of different urinary dilution correction methods on the associations between arsenic exposure and these toxicity biomarkers was explored in **Paper II** (see Section 4.5.2 for more details).

4.4.4 Other metals

The exposure in this study group to other metals and metalloids, including lithium, boron, strontium, and cesium, was assessed in drinking water and urine samples by measuring their concentrations by ICP-MS (Agilent 7700x, Agilent Technologies). The concentrations of lithium and boron in the drinking water samples collected around Lake Poopó were similar to those described in SAC, in the Argentinean Andes (Concha et al., 2010), but higher than previously reported around Lake Poopó (Ormachea Muñoz et al., 2013; Ramos Ramos et al., 2012). Additional information about these elements can be found in **Paper I**.

4.5 TOXICITY BIOMARKERS

A biological marker, or biomarker, is a characteristic that can be measured in a sample as an indicator of the state of an organism (Muller and Dieterle, 2009). Within environmental health research, these are classified as:

- Biomarker of exposure: how much is an organism exposed to a chemical?
- Biomarker of susceptibility: does the organism have an impaired capacity to respond to the chemical insult compared to others?
- Biomarker of effect: does the chemical alter the organism to any extent?

The biomarkers of effect included in **Paper II** reflect alterations of structures or functions at a biochemical level in relation to established mechanisms, such as via the induction of oxidative stress. In **Paper II**, a set of biomarkers was measured to evaluate the toxicity in relation to arsenic exposure. In this sense, the umbrella term “toxicity biomarkers” was used. Although broader in definition, and somewhat unspecific, it does not make assumptions regarding the type of toxicity exerted. In **Paper III**, an exploratory approach was taken to potentially identify novel biomarkers.

The use of biomarkers of effect is specifically valuable in the context of the Bolivian study area because there are no registries on disease outcomes nor extensive clinical records available in these rural communities. However, precisely because of this lack of clinical health outcomes, it is not possible to distinguish whether the changes observed are truly signs of toxicity or adaptive responses at a cellular level. Additional studies will be required to fill the gaps between the arsenic-related molecular effects and potential clinical outcomes.

4.5.1 Telomere length and mitochondrial DNA copy number

In **Paper II**, arsenic toxicity was evaluated by assessing if telomere length and mitochondrial DNA copy number (mtDNAcn), two toxicity biomarkers widely explored in relation to environmental toxicants, were associated with arsenic exposure.

Telomeres are nucleotide repetitions that protect the end of eukaryotic chromosomes. The machinery for DNA replication in eukaryotes cannot replicate the 3' end of the DNA strand. Therefore, each time a cell divides, telomeres are gradually shortened. To counteract this, telomeres are elongated by the telomerase (Blackburn, 2005). The gradual shortening of chromosomes may lead to genomic instability (Blackburn, 2005), while longer telomeres confer an extended replicative capacity to cells linked to carcinogenesis (McNally et al., 2019). In fact, the gene coding for the telomerase is overexpressed in the majority of cancers but it has very limited expression in most somatic cells after birth (Jafri et al., 2016).

Mitochondria, the main energy-producing organelle in eukaryotic cells, have multiple copies of extra-nuclear circular DNA. The mtDNAcn increases in response to oxidative stress (Lee et al., 2000), which can be generated naturally during cellular respiration by the mitochondria itself or due to external toxic insults. Therefore, it is used as a biomarker of oxidative stress.

Telomere length and mtDNAcn were determined using quantitative polymerase chain reaction (PCR) in peripheral blood leukocytes (Cawthon, 2002; Xing et al., 2008). Relative values of telomere length and mtDNAcn were calculated as the ratio with a single-copy gene (hemoglobin beta, *HBB*). Therefore, values have arbitrary units, which might limit the comparability of the results among different populations and studies. Furthermore, to account for inter-plate variation, the ratio for each sample and biomarker was normalized to a reference sample included in all run plates.

4.5.2 8-oxo-dG and 4-HNE-MA

Oxidative stress has multiple molecular targets at a cellular level, including DNA and lipids. As a biomarker for oxidative damage of DNA, 8-oxo-2'-deoxyguanosine (8-oxo-dG) concentrations were measured in urine (Loft and Møller, 2006). In addition, 4-hydroxy nonenal mercapturic acid (4-HNE-MA) concentrations were measured since it is a urinary metabolite product of lipid peroxidation (Dalleau et al., 2013). These two biomarkers of oxidative stress, 8-oxo-dG and 4-HNE-MA, were quantified in urine and studied in relation to arsenic exposure as part of **Paper II**.

Urinary concentrations of 8-oxo-dG and 4-HNE-MA were measured by liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) at the Division of Occupational and Environmental Medicine from the Department of Laboratory Medicine at Lund University, Sweden. There is no golden standard on how to adjust these two biomarkers in urine to correct for variations of urinary dilution. Therefore, in **Paper II**, two urinary dilution correction methods and approaches were evaluated in relation to these biomarkers: (i) urinary concentrations adjusted for specific gravity or osmolality, and (ii) specific gravity or osmolality

included as a covariate in the linear regression models. The associations between these two biomarkers of oxidative stress and arsenic exposure were not dependent on which correction method was used (described in **Paper II**). Considering that, and the fact that we adjusted U-As with urinary osmolality throughout this thesis work, these two biomarkers were also adjusted to the mean urinary osmolality of the study group.

4.5.3 Identifying potential novel biomarkers

Sections 4.5.1 and 4.5.2 describe well-established biomarkers of toxicity that have been explored to different extents in the scientific literature in relation to arsenic. In **Paper III**, making use of a commercial multiplex proteomic technology, an exploratory study was conducted to evaluate how arsenic exposure was associated with a panel of proteins putatively related to cancer. Multiplexing technologies enable studies of a more exploratory nature expanding the capacity to identify novel biomarkers.

In **Paper III**, the Proximity Extension Assay (PEA) technology was performed by Olink Proteomics (Uppsala, Sweden). For this, 92 proteins related to cancer processes included in a predefined panel (Proseek Multiplex Oncology II panel) were measured in urine. For each targeted protein, two protein-specific antibodies are labeled with DNA oligonucleotides. As these antibodies bind to the protein of interest, the oligonucleotide labels are close enough to hybridize and form a double-strand DNA that is then amplified and quantified by real-time PCR. The results are relative protein expression levels in $\log_2$ scale (Normalized Protein eXpression values, NPX).

The advantages of this technology are the small sample volume required (1 µL), its increased sensitivity, and the dual recognition that minimizes cross-reactivity and non-specific signals. However, since protein levels are expressed as relative values, it is not possible to compare absolute concentrations between the different proteins and between other studies. This assay was initially developed to detect proteins in plasma samples, not in urine (Assarsson et al., 2014). Therefore, out of the 92 proteins initially measured, 45 proteins were explored in relation to arsenic, i.e., proteins with more than 40% of the observations above the limit of detection.

4.6 GENOTYPING

Genetic variations across the human genome are responsible, to some extent, for the phenotypic variations that are observed. Determining how a group of individuals differs genetically within or across populations helps identify which genetic variations explain differences in the phenotype and allows us to understand past demographic changes of a population. Single nucleotide polymorphisms (SNP) are changes in one base pair, and these are the most common type of genetic variation in the human genome. Different genotyping techniques, that is, determining which combination of alleles an individual has at a specific SNP locus, are described in the following sections based on their throughput capacity.

4.6.1 TaqMan allelic discrimination assay

The TaqMan allelic discrimination assay is based on the hybridization of DNA probes, followed by a quantitative PCR. Each SNP assay includes two allele-specific probes coupled with different fluorophores to genotype one specific SNP. This assay depends on the 5' exonuclease activity of the Taq DNA polymerase combined with fluorescence quenching: the probe has a fluorescent reporter molecule (fluorophore) and a quencher in proximity, preventing the fluorophore from emitting its signal. Once the probe is bound to the DNA template, the polymerase cleaves the probe as it replicates, physically separating the quencher and the fluorophore and allowing the latter to emit its fluorescence signal.

In **Paper II**, we genotyped SNP rs3740393, located in intron 6 of *AS3MT*, as it is known to predict arsenic metabolism efficiency (Agusa et al., 2011; Drobná et al., 2013; Engström et al., 2011). Genotyping was done by allelic discrimination using the LightCycler 480 II instrument (Roche).

4.6.2 SNP arrays

A SNP array is a DNA microarray where DNA probes, covering an extensive set of SNPs throughout the genome, are attached to a solid platform. These SNP arrays enable whole-genome genotyping following PCR-free protocols (Gunderson et al., 2005).

Genome-wide genotyping of autosomal chromosomes was performed on the Illumina Infinium Omni5Exome (4v1.3) at the SNP&SEQ Technology Platform (Uppsala, Sweden). More than 4 million SNPs were analyzed in 187 samples from the Bolivian study groups. The SNP data were aligned to the human reference genome build version 37 (hg19). Results based on genome-wide genotyping are described in **Paper IV**.

An advantage of SNP arrays is the small amount of DNA sample required to cover several million SNPs. However, it is worth noting that the SNPs contained in these arrays are identified mainly in human populations with ancestry from Europe, Asia, and West Africa (Clark et al., 2005). Therefore, the selection of SNPs in arrays may not reflect the genetic diversity of other less-represented populations, such as those from South America. This is referred to as ascertainment bias and must be considered when the aim is to identify SNPs that may be specific to certain populations, such as those potentially driving arsenic tolerance in Andean populations.

The genotype for SNP rs3740393 was determined with the TaqMan assay (**Paper II**) and within the SNP array (**Paper IV**) and had a 100% agreement. Map of allelic frequencies was made with the Geography of Genetic Variants browser (Marcus and Novembre, 2016).

4.6.3 Targeted sequencing

Although SNP arrays cover millions of SNP throughout the genome, the density of SNPs per genomic region is still low. In **Paper IV**, one of the aims was to investigate, by targeted

sequencing a candidate region near *AS3MT*, whether Andean populations had novel SNPs potentially missed due to the ascertainment bias of predefined SNP arrays.

The top SNP associated with urinary %MMA in Schlebusch et al. (2015) in SAC (rs486955 located at chr10:104546284 upstream of *AS3MT*, based on hg19) and the 2.5 kilobases (kb) on each side defined the candidate region to sequence. A total of 96 individuals were sequenced for this region: 24 Aymara-Quechua with high and low %MMA, and all 19 Uru and 53 individuals from SAC with enough DNA. An initial long-range PCR was performed to amplify the region selected and to incorporate primers with barcodes. By adding barcodes to the primers, the amplicons from each individual got a unique barcode, and all samples could be pooled together for the sequencing (data for each individual was then separated during the data processing). The Pacific Biosciences Sequel I system powered by Single-Molecule, Real-Time (SMRT) technology was used at the National Genomic Infrastructure (NGI)/Uppsala Genome Center (Uppsala, Sweden). Data were then analyzed following a bioinformatic pipeline proposed by PacBio, including mapping of the reads to a reference genome (hg19), quality control, and variant calling.

By barcoding the samples and pooling them together, a high throughput is achieved. Also, deep sequencing of the nucleotides within a defined region can detect known and novel variants, while requiring less data management resources than whole-genome sequencing. However, since this protocol depends on PCR amplification, it is susceptible to PCR-biased results. This means that the PCR amplification has a higher preference for one allele than the other in heterozygotic individuals, potentially leading to biased allelic frequencies.

4.7 POPULATION STRUCTURE

The comparative populations used to put the Bolivian study group into geographic and demographic contexts from a population structure perspective are described in **Paper IV**.

4.7.1 Principal component analysis

Principal component analysis is used to reduce the dimensions of big datasets into smaller subsets, or components, which can be more easily explored. It *reduces* the genetic variation into different components, which are then visualized in two- or three-dimensional plots. Individuals that are plotted closer to each other are genetically more similar than those that are further apart. Therefore, it is commonly used in population genetics to explore genetic distances between populations. Often, genetic affinities between populations are driven by geographic proximity (see Section 2.2.1) and therefore, the plots commonly reflect the geographic location of the populations.

4.7.2 Ancestry estimation

In order to quantify the population structure, a model-based estimation of ancestries was carried out with ADMIXTURE (Alexander et al., 2009). Based on a maximum likelihood, it infers the ancestry of populations through a clustering approach. It calculates the probability of the

genotypes observed when using different ancestry proportions. These results are dependent on the comparative populations included and should not be overinterpreted (Lawson et al., 2018).

4.8 SELECTION SCANS

Selection scans were performed in **Paper IV** to detect signals of positive selection across the genome. The aim was to determine whether the exposure to arsenic over generations in the Bolivian Andes has acted as a selective pressure; for example, by examining if alleles associated with an efficient arsenic metabolism were more frequent in the Bolivian populations than in comparative populations that have not endured the same environmental pressures. The process of how the haplotypes were phased, i.e., assigning each allele to the maternal or paternal chromosome, and the determination of ancestral alleles for the selection scans are described in **Paper IV**. First- and second-degree relatives were identified and excluded before performing the selection scans. The reason for this is that including close relatives will result in higher allelic frequencies due to inherited relatedness and not due to positive selection. The top 0.5% of the SNPs selected across the whole genome were further explored to find whether *AS3MT* was within 10 kb on each side of each SNP selected.

4.8.1 iHS and XP-EHH

If a new random mutation provides a beneficial trait to a certain environment, this mutation tends to be selected for and increases in frequency on a population level. These SNPs that are under selective pressures will drag along other neutral variants that are in linkage disequilibrium forming a haplotype. When the prevalence of this haplotype increases quickly in a population due to selection, recombination does not act fast enough to break down the haplotype and abnormally long haplotypes appear (Sabeti et al., 2002). Several methods explore the extent of haplotype homozygosity to detect long genomic regions of low genetic variability in relation to genome-wide levels.

The integrated haplotype score (iHS) detects *de novo* selection where the allele selected has substantially increased in frequency but has not reached fixation, i.e., the alternative allele is still present in the population (Voight et al., 2006). The length of the haplotype selected is then compared to the haplotype length of the alternative alleles. The cross population extended haplotype homozygosity test (XP-EHH) is based on the same principle but uses a comparative population to check the haplotypes of variants near fixation in one population compared to the other (Sabeti et al., 2007). In **Paper IV**, the Bolivian study groups were compared to Peruvians from the 1000 Genomes Project. These tests were performed using the R package *rehh* (Gautier and Vitalis, 2012).

These methods aim to detect hard sweeps of *de novo* mutations and recent occurrences of selection. Hard sweeps happen when a new mutation occurs, and the phenotypic trait linked to it is advantageous for individuals exposed to a certain environmental pressure (**Figure 7**). The sudden appearance of new advantageous mutations that get selected along with the neighboring neutral variants leads to one specific haplotype increasing in frequency in the population.

4.8.2 LSBL

While methods based on haplotype homozygosity such as iHS and XP-EHH detect easily hard sweeps, other methods based on fixation indexes are better at detecting soft sweeps and older selective events. Soft sweeps happen when a preexisting neutral genetic variation becomes advantageous after an environmental change (**Figure 7**). In these cases, the advantageous variant might be already present in different haplotypes backgrounds. The genetic pressures that increase the frequency of the advantageous SNP will also hitchhike the different haplotypes that the variant is present in. This means that more than one haplotype background will be favored. This could be the case of preexisting genetic variation that became beneficial as humans populated the Andes and encountered an arsenic-rich environment, which could have acted as a strong driver of selection.

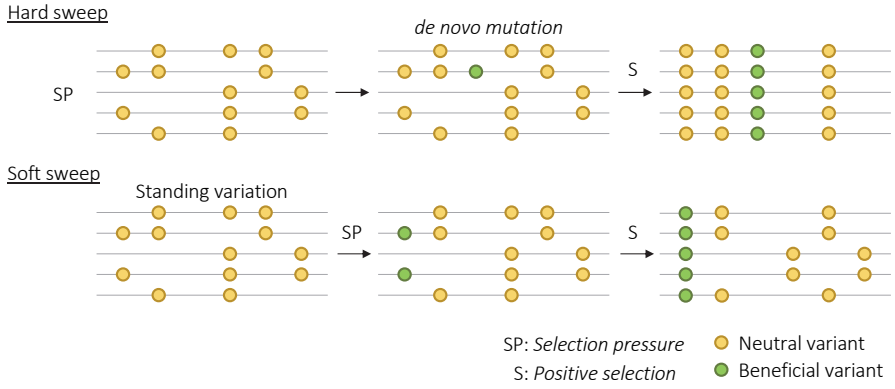


Figure 7. Graphical explanation of soft and hard selective sweeps. Each line represents a haploid genome, and each dot is a genetic variant. An example of selection pressure can be the exposure to arsenic.

Population differentiation is expected to be similar throughout the genome. However, if certain variants become beneficial due to a change in the environment, the genetic differentiation between populations will increase more in specific genomic regions when selected compared to the rest of the genome. Locus-specific branch length (LSBL) is a method based on fixation index (F_{ST}) values (Shriver et al., 2004).

Pairwise F_{ST} distances measure genetic differentiation between populations based on the heterozygosity at each locus. A higher F_{ST} value between two populations at a specific locus indicates a higher allele frequency difference at this locus. Three-way F_{ST} methods include a third "outgroup" population to help indicate in which population the frequency difference occurred, which is not possible with a pairwise F_{ST} measurement. F_{ST} values for the three-way comparison were calculated using the Weir and Cockerham's equation on PLINK between the Bolivian study groups (Aymara-Quechua and Uru separately), and Peruvians and Han Chinese (outgroup) from the 1000 Genomes Project. The divergence at each locus between the three populations was calculated as represented in **Figure 8**.

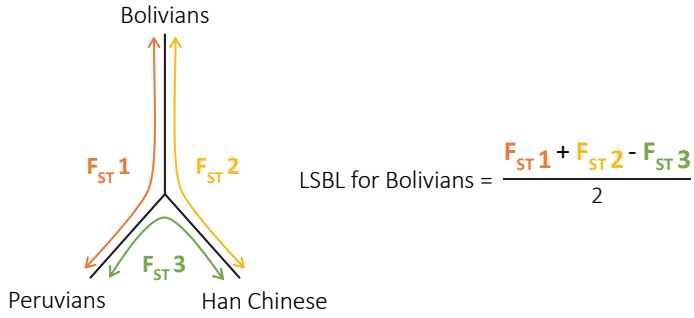


Figure 8. Locus-specific branch length (LSBL) comparisons performed in **Paper IV**. The F_{ST} estimate for each autosomal variant was calculated with the method by Weir and Cockerham. Data from Peruvians and Han Chinese were obtained from the 1000 Genomes Project. *Bolivians* refers to either the Aymara-Quechua or Uru, calculated separately.

4.9 STATISTICAL ANALYSES

The statistical analyses performed are described in detail in **Papers I–IV** and summarized in this section. All analyses were carried out in R using RStudio, with version 3.4.1 and 1.1.383, respectively, for **Paper I**; and version 3.6.2 and 1.1.423 for **Papers II–IV**.

Wilcoxon signed-rank test (for continuous variables) and the z-test of proportions (for categorical variables) were used to evaluate differences in general characteristics of the study group between ethnicities in **Paper I.** Wilcoxon tests were also performed to evaluate whether the toxicity biomarkers differed between ethnicities (**Paper II**), and whether relative urinary protein levels were different among individuals with high vs. low arsenic exposures (**Paper III**). Spearman's rank correlation tests were performed between (i) factors influencing U-As and its metabolites in urine (**Paper I**), (ii) urinary dilution correction methods (**Paper I**), (iii) arsenic concentrations in different matrices (**Paper II**), (iv) toxicity biomarkers and arsenic exposure (**Paper II**), (v) covariates that were associated with overall urinary protein variation (**Paper III**), and (vi) urinary proteins associated with B-As (**Paper III**).

Linear regression analyses were used through **Papers I–IV**, either with univariate or multivariable-adjusted models. Diagnostic plots of the residuals from the regression models were inspected visually, including residuals vs. fitted plots and Q-Q plots. Variables were $\log_2$-transformed when it increased the adjusted R^2 and improved the residual plots of the model, as was the case for 8-oxo-dG and 4-HNE-MA in **Paper II**. In **Paper I**, linear regression models were built to explore how U-As and iAs metabolites in urine were predicted by ethnicity, age, body weight, fish and meat consumption, and tobacco smoking. Then, the associations between arsenic exposure biomarkers and a set of toxicity biomarkers were studied in **Paper II**, further evaluating whether age, hemoglobin, ethnicity, chewing coca leaves, or smoking were also associated with the toxicity biomarkers. In **Paper III**, the extent to which arsenic exposure and other covariates explained the variation of urinary proteins was assessed with univariate linear regression models by evaluating R^2 (the percentage of the variance of the dependent variable

explained by the independent variable). Furthermore, the associations of B-As (and other covariates identified in the linear regression of principal components) with relative protein levels in urine were studied with multivariable-adjusted linear regression evaluating beta coefficients. Compared to B coefficients, beta coefficients account for differences in the units of the covariates by standardizing the variables. Then, the effect coefficients of the different covariates can be compared.

Then, the associations between genetic variants and the efficiency to metabolize arsenic in the Bolivian study group were explored. The relative arsenic metabolite fractions in urine were evaluated using linear regression analyses including SNP rs3740393 in *AS3MT* (**Paper II**) and a set of SNPs included within a protective haplotype around *AS3MT* (**Paper IV**). The genetic influence on arsenic metabolism efficiency was also assessed at a genome-wide level (**Paper IV**) using the R package *GenABEL* (v. 1.8) with functions *qtscore* and *egscore*, adjusting for potential underlying population structure.

Sensitivity analyses were performed to explore whether extreme values were driving the overall associations, whether different methods to correct for urinary dilution resulted in varying associations, and whether including other metals correlated previously with arsenic concentrations altered the associations.

Other methods carried out were the linear regression models of principal components to assess which factors explained the overall variation of urinary proteins in **Paper III**, and the population structure analyses (see Section 4.7) and selection scans (see Section 4.8) used in **Paper IV**.

The effect modification by arsenic metabolism efficiency on the associations between arsenic exposure and toxicity biomarkers was explored. First, by introducing a multiplicative interaction term between U-As or B-As and the rs3740393 genotype in relation to the toxicity biomarkers included in **Paper II** to evaluate their joint effect. Then, the linear regression models were stratified by arsenic metabolism efficiency, assessed by (i) rs3740393 genotype in **Paper II**, and (ii) the median of iAs metabolite fractions in urine in **Papers II** and **III**. An additive effect of the alleles was assumed when stratifying by genotype: individuals that were heterozygotic for the SNP were combined with those that were homozygotic for the less-frequent allele.

Generally, a *p*-value below 0.05 was considered as statistically significant. However, there is a higher probability of false positives given a fixed significance level as the number of test increases. A Bonferroni correction was used in **Paper III** since a high number of tests were performed. Nevertheless, this correction may be too stringent or conservative and assumes completely independent tests, which is not necessarily the case if the proteins measured in urine do not excrete independently or in the case of highly linked genetic variants. Considering this and the fact that the studies included in this thesis have a limited sample size (n = 201 at most), it is of foremost importance to evaluate the biological plausibility and relevance of the results in connection to their statistical significance by reporting all *p*-values and confidence intervals.

The commonly used statistical significant threshold ($p = 5 \times 10^{-8}$) for genome-wide association analyses (Fadista et al., 2016) and the Bonferroni correction were considered in **Paper IV**. However, the best approach for genome-wide data studies is still under discussion (Kaler and Purcell, 2019).

5 RESULTS AND DISCUSSION

5.1 CHARACTERISTICS OF THE STUDY PARTICIPANTS

General characteristics of the individuals included in **Papers I–IV** are found in **Table 2**, including arsenic concentrations in urine and blood, which are further described in Section 5.2.

Table 2. Characteristics of the study participants presented as median (minimum–maximum).

	Paper I			**Paper II**	**Paper III**	**Paper IV**		
	Combined ethnicities	Aymara-Quechua	Uru	Combined ethnicities	Combined ethnicities	Combined ethnicities	Uru men	SAC
n (women)	201	168	33	193	176	158 + 5 men	5 men	53
Age (years)	34 (14–85)	36 (15–85)	30 (14–65)	36 (14–85)	36 (16–85)	36 (18–85)	35 (27–59)	36 (14–71)
Height (cm)	149 (120–162)	140 (120–162)	149 (140–162)	149 (120–162)	149 (120–160)	150 (120–168)	161 (159–168)	152 (139–165)
Body weight (kg)	60 (36–97)	60 (36–89)	67 (47–97)	60 (36–97)	60 (37–97)	60 (37–89)	65 (58–83)	59 (41–89)
Hemoglobin (g/dL)	15 (8.2–20)	15 (8.2–19)	15 (11–20)	15 (8.2–20)	15 (8.2–20)	15 (8.2–19)	17 (15–17)	16 (9–18)
B-As (ng/g)	2.0 (0.60–9.2)	2.1 (0.60–9.2)	1.8 (0.70–3.5)	2.0 (0.60–9.2)	2.1 (0.60–9.2)	2.1 (0.60–9.2)	3.1 (0.66–3.3)	10.8 (3.0–21)
U-As total (µg/L)	75 (18–630)	76 (18–630)	67 (20–140)	74 (18–472)	78 (18–472)	77 (18–472)	66 (44–125)	365 (125–762)
U-As sum (µg/L)	65 (12–407)	66 (12–407)	55 (16–143)	64 (12–399)	67 (12–399)	67 (12–399)	68 (39–111)	268 (62–664)
%iAs	12 (3.2–34)	12 (3.2–34)	10 (4.9–23)	12 (3.2–34)	12 (3.2–34)	13 (3.2–34)	13 (8.4–20)	14 (3.3–29)
%MMA	7.7 (2.2–18)	8.1 (2.9–18)	5.9 (2.2–10)	7.7 (2.2–18)	7.7 (2.2–18)	7.7 (2.9–18)	7.0 (5.9–15)	8.7 (3.3–18)
%DMA	80 (54–91)	79 (54–91)	84 (68–91)	80 (54–91)	80 (54–91)	79 (54–91)	74 (67–85)	77 (62–93)

Notes: U-As total and sum were adjusted to mean urinary osmolality in the Bolivian study group and to specific gravity in SAC. Only 3 Uru men had B-As data. *Abbreviations:* B-As, arsenic concentration in blood; DMA, dimethylarsinic acid; iAs, inorganic arsenic; MMA, monomethylarsonic acid; SAC, San Antonio de los Cobres (Argentina); U-As sum, urinary arsenic as the sum of iAs, MMA, and DMA concentrations; U-As total, urinary arsenic as the total concentration of all forms.

The women living in villages around Lake Poopó rarely smoked tobacco or consumed alcohol. Only two women from the whole study group (n = 201 in **Paper I**) reported consuming alcoholic beverages every other week. Furthermore, only 3% of the women smoked tobacco (asked as yes/no in the questionnaire). On the contrary, 73% of the women chewed coca leaves, a longstanding practice among Andean communities as part of rituals and to cope with high altitude sickness, pain, and symptoms of exhaustion (Rivera et al., 2005).

On average, these Bolivian women had 3 children and 10% of them had more than 6. Around 75% of the study group identified themselves as farmers of crops or animals, artisans, merchants, or housekeepers. The rest worked either in health services or in schools.

The median body mass index of these women, described in **Paper I**, was 27 kg/m^2 (range 16–44 kg/m^2). Indigenous communities living in the Andes are shorter (half of the women in the study group were shorter than 149 cm) and have wider thoraxes than those in the lowlands, potentially as an adaptation to decreased atmospheric oxygen levels at high altitudes (Espinoza-Navarro et al., 2011; Harari et al., 2015; Macdonald et al., 2004). As a result, body mass indexes do not accurately estimate body fat in these populations, and therefore, were not used.

The population structure of the Bolivian study group, assessed separately for the individuals of Aymara-Quechua or Uru descent, was explored using genome-wide genotype data in **Paper IV**. The Aymara-Quechua and Uru study groups clustered in the principal component analysis with other Andean populations (**Figure 9**). Individuals belonging to the Uru ethnic group clustered furthest away possibly reflecting their historical isolation (de la Barra Saavedra et al., 2011). The level of European influence in the genomes of these Bolivian study groups was limited: 1.9% and 0.2% of the genetic diversity of Aymara-Quechua and Uru, respectively, was of European ancestry.

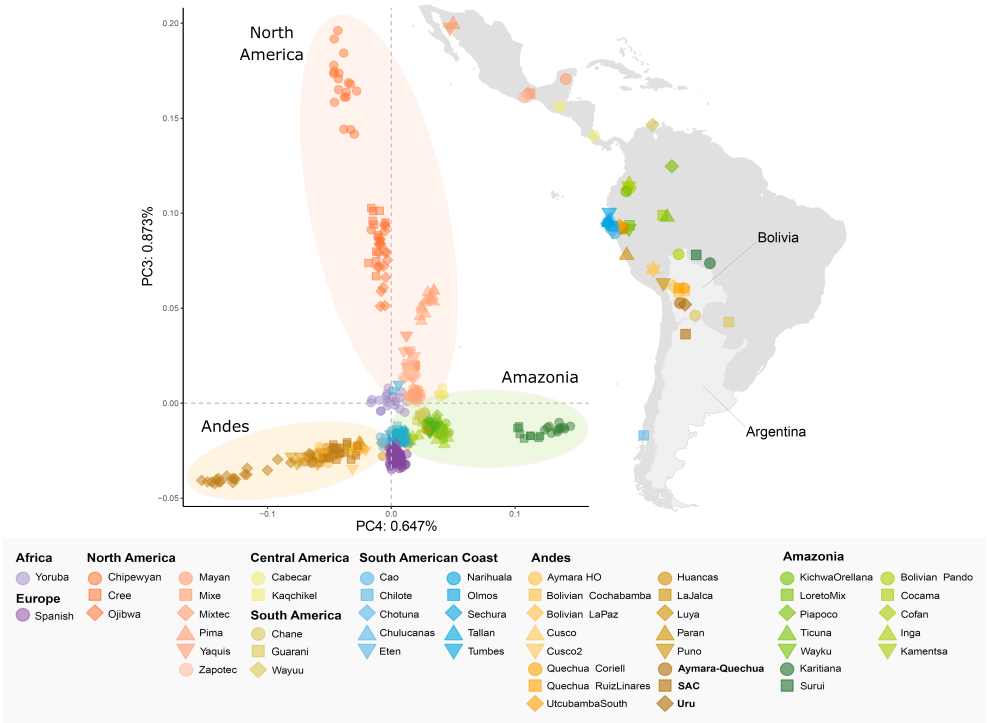


Figure 9. Principal component (PC) analyses for indigenous populations in Latin America. The Aymara-Quechua and Uru in Bolivia and the women from San Antonio de los Cobres (SAC) clustered together with other populations in the Andes (in brown). The Yoruba and Spanish populations (in purple) were included to remove African and Spanish admixture in PC1 and PC2 (not shown). Additional description of the method and populations is found in **Paper IV**.

5.2 ARSENIC EXPOSURE ASSESSMENT

The environmental exposure to arsenic in the inhabitants of villages situated around Lake Poopó, Bolivia was characterized by measuring concentrations of arsenic in urine, blood, and drinking water samples. Total concentrations of arsenic in urine can include both metabolites of inorganic arsenic (iAs, MMA, and DMA) and less-toxic organic forms (e.g., arsenobetaine) of this metalloid. By quantifying the arsenic metabolites separately, also known as arsenic speciation, and comparing the sum of them to total arsenic concentrations, it is possible to distinguish the major contributing forms to the overall arsenic exposure as done in **Paper I**.

In **Paper I**, the exposure to arsenic was assessed by concentrations in urine samples, the most used biomarker of arsenic exposure. The sum of inorganic arsenic metabolite concentrations in urine (referred to as U-As in the thesis) and total arsenic concentrations in urine were strongly correlated ($r_S = 0.99$, $p < 0.001$) and in agreement (the ratio between both measurements was close to 1). This suggests that the women living around Lake Poopó were exposed mainly to inorganic arsenic. Their exposure to inorganic arsenic was varied and elevated: U-As, adjusted to urinary osmolality, ranged between 12 and 407 µg/L with a median of 65 µg/L. In addition, there was a wide range of U-As across and within villages. Furthermore, the Aymara-Quechua women had significantly higher U-As compared to Uru (**Table 2**). Regarding the choice of urinary dilution correction method, a strong correlation ($r_S = 0.99$, $p < 0.001$) and agreement was found between adjusting U-As to osmolality or specific gravity as described in **Paper I** (**Figure 6**). However, U-As was adjusted to the mean urinary osmolality of the study group since it has been recently shown to be an optimal correction method for arsenic concentrations (Middleton et al., 2019, see Section 4.4.3).

The exposure to arsenic in these women was also evaluated in blood in **Papers II** and **III**. Arsenic concentrations in blood are generally considered as a biomarker of recent exposure because arsenic is quickly cleared from this matrix (Hughes, 2006). Nevertheless, in populations exposed chronically to arsenic, concentrations of arsenic in blood reach a steady state enabling the use of both urine and blood as valid matrices to evaluate arsenic exposure (Concha et al., 1998a; Hall et al., 2006; Takayama et al., 2021). The median B-As was 2.0 ng/g (range 0.6–9.2 ng/g) in these Bolivian women (n = 193, **Paper II**). The strong correlation between U-As (adjusted for osmolality) and B-As ($r_S = 0.85$, $p < 0.001$, **Figure 10**) supports a chronic exposure to arsenic in this study group, and it corroborates the validity of using B-As as a biomarker of exposure that does not require urinary dilution corrections.

Arsenic concentrations in water were also measured to characterize the study area and identify potential sources of arsenic exposure in **Paper I**. Arsenic in drinking water samples from mostly groundwater sources from the 10 villages included in the study ranged from 3.3 to 571 µg/L and did not exhibit temporal variations during the years the recruitment took place, further supporting a constant exposure to arsenic. Villages where Aymara-Quechua communities lived had higher water arsenic concentrations on average (median 130 µg/L, range 25–571 µg/L) than that of Uru villages (median 46 µg/L, range 3.3–222 µg/L), possibly explaining the difference in arsenic exposure between both ethnicities. In some villages, water

samples were collected from several sources and showed a wide variation in arsenic concentrations; for example, in Santuario de Quillacas, arsenic concentrations in water ranged between 25 and 176 µg/L. This variation in concentrations of arsenic in water within the same village, as well as potential differences in water consumption, might explain the variation of arsenic exposure of individuals living in the same village.

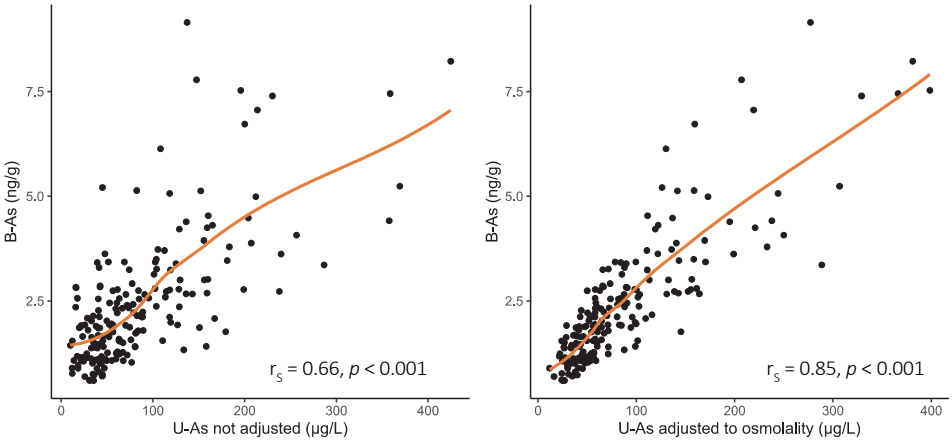


Figure 10. Scatter plots with loess line of B-As and U-As, not adjusted and adjusted to mean urinary osmolality to correct for urinary dilution variations. *Abbreviations:* B-As, arsenic concentration in blood; U-As, urinary arsenic as the sum of inorganic arsenic metabolite concentrations.

The arsenic concentrations in the water of each village (average if multiple sources were sampled) was weakly correlated with the U-As (adjusted to osmolality) of the inhabitants (r_S=0.30, p < 0.001). This was not surprising since the water samples collected did not represent the individual water consumption of each participant. Another common source of exposure to inorganic arsenic is rice, which also may contain some DMA (Meharg et al., 2009). However, rice collected from the main markets around Lake Poopó presented low arsenic concentrations (**Paper I**). Therefore, arsenic exposure via drinking water was still considered the main exposure source, particularly considering that we did not sample all water sources or thoroughly characterize individual water consumption. In addition, even harvested rainwater for human consumption, which is presumed to be free of arsenic, has been shown to have elevated arsenic concentrations (median 20 µg/L, range 0–50 µg/L) in collection systems implemented around Lake Poopó (Quaghebeur et al., 2019). The high arsenic concentrations in the rainwater collected were tracked to mineral dust particles with arsenic deposited on the roof, an important problem considering the arid climate of the study area. The presence of arsenic in dust indicates that other routes of exposure such as inhalation could be relevant in this population.

Papers I and **II** provide the most complete and thorough characterization of ongoing arsenic exposure in Bolivia up to date. Since **Paper I** was published in 2019, no new studies have evaluated arsenic exposure in humans in any other Bolivian region, highlighting the need for additional efforts.

5.3 ARSENIC METABOLISM AND INFLUENCING FACTORS

The efficiency to metabolize arsenic of the Bolivian women was evaluated in **Paper I**. Indigenous women living in villages around Lake Poopó had an efficient arsenic metabolism characterized by low urinary fractions of MMA, the form generally associated with most adverse health effects. The median relative fractions of inorganic arsenic metabolites in urine for this study group (when combining Aymara-Quechua and Uru women) were 12% iAs, 7.7% MMA, and 80% DMA (**Figure 11**). This distribution of metabolites in urine is different compared to the average across other populations with 10–30% iAs, 10–20% MMA, and 60−80% DMA (Vahter and Concha, 2001). The arsenic metabolism capacity in the Bolivian women was similar to that previously described in women from SAC in the Argentinean Andes with a median 7.5% MMA in urine (Schlebusch et al., 2015). No other population, apart from these in the Argentinean and Bolivian Andes, have been described as such efficient arsenic metabolizers. This opens the question to whether this efficient arsenic metabolism is specific to indigenous populations from the Andes or whether it is present elsewhere.

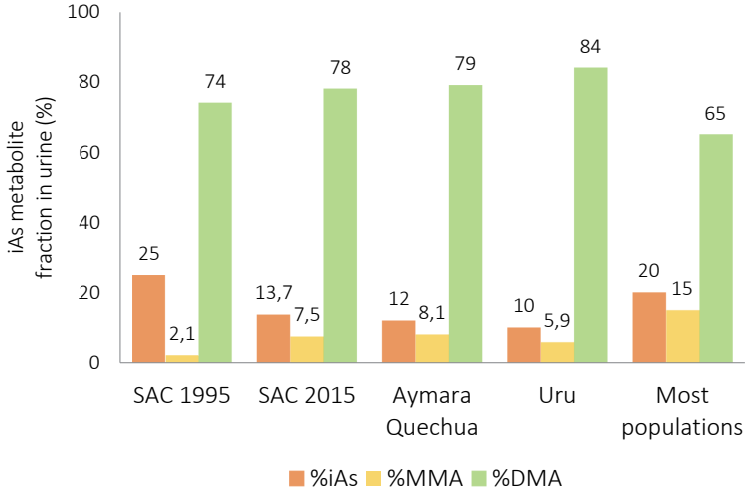


Figure 11. Distribution of inorganic arsenic metabolite fractions in urine in different populations. SAC 1995 and 2015 refer to women from San Antonio de los Cobres (SAC), Argentina described in Vahter et al. (1995) and Schlebusch et al. (2015), respectively. The Aymara-Quechua and Uru are described in **Paper I**, and *most populations* refers to an average across populations described in Vahter (1999). *Abbreviations:* DMA, dimethylarsinic acid; iAs, inorganic arsenic; MMA, monomethylarsonic acid.

What does it mean to have an efficient arsenic metabolism?

When humans are exposed to inorganic arsenic, for example through drinking water, it gets metabolized in the liver by adding one or two methyl groups to it. The three arsenic forms (monomethylated, dimethylated and remaining un-methylated arsenic) are excreted in urine. Generally, the dimethylated form (DMA) is better excreted, while the monomethylated one (MMA) is considered more toxic. Lower fractions of MMA and higher of DMA in urine are considered a sign of an efficient arsenic metabolism. Up until now, only indigenous groups in the Andes have been described as efficient arsenic metabolizers at a population level.

Factors known to influence the body's capacity to methylate arsenic were explored with multivariable-adjusted linear regression in **Paper I**. Body weight, the consumption of fish, smoking of tobacco, and ethnicity were associated with differences in arsenic metabolism efficiency in the Bolivian study group. Smoking tobacco, although highly infrequent in this study group, was associated with a less efficient arsenic methylation, i.e., higher %iAs and %MMA. On the contrary, fish consumption (reported as yes/no) and body weight (possibly reflecting more protein intake in this context) were associated with a more efficient arsenic metabolism. Higher arsenic exposure is known to impair the metabolism of arsenic (Li et al., 2008; Lindberg et al., 2008a); however, this was not the case in the Bolivian or Argentinean study groups. Interestingly, ethnicity was associated with all three inorganic arsenic metabolite fractions, even after adjusting for differences in arsenic exposure by including U-As in the models. On average, Uru women had 2.5 percentage unit (%) lower %iAs, 2.2 lower %MMA, and 4.7 higher %DMA, suggesting a more efficient metabolism than the Aymara-Quechua. All in all, these models including ethnicity, body weight, fish and tobacco consumption, and U-As explained between 11 and 26% of the variability in the arsenic metabolite fractions in urine, indicated by the R^2 estimates of the linear regression models. This indicates that other factors also determine the body's capacity to metabolize and excrete arsenic.

Ethnic differences in arsenic metabolism capacity have not been explored extensively. While differences in arsenic exposure have been identified across ethnic groups living in the United States (Jones et al., 2019), no significant differences in arsenic methylation efficiency were seen among non-Hispanic White, African American, Hispanic, and Chinese American groups exposed to low concentrations of arsenic in the United States (Balakrishnan et al., 2018). No differences in urinary %MMA were found between Hispanic and non-Hispanic individuals from the southwest United States and northwest Mexico either (Gomez-Rubio et al., 2011). In contrast, a study comparing Chinese populations found that Tibetans had a more efficient arsenic metabolism than Han and Hui Chinese individuals (Fu et al., 2014). However, the Tibetans in this study had high %MMA (median 25%) compared to other global populations.

Genetic polymorphisms in *AS3MT,* which codes for the main arsenic methylating enzyme, are known to predict arsenic methylation efficiency. In **Paper II**, SNP rs3740393 in *AS3MT*, previously associated with arsenic methylation (Agusa et al., 2011; Drobná et al., 2013; Engström et al., 2011), was studied in relation to arsenic metabolism efficiency in the Bolivian study group. Genotype CC of rs3740393 was associated with lower urinary %MMA, i.e., a more efficient arsenic metabolism. This finding is in line with other studies from, for example, Argentina and Bangladesh (Engström et al., 2011), Mexico (Drobná et al., 2013), and the United States (Balakrishnan et al., 2017).

The implication of *AS3MT* on arsenic metabolism was further explored throughout the whole gene. Polymorphisms along *AS3MT* are in high linkage disequilibrium (Engström et al., 2013; Gomez-Rubio et al., 2010). This means that SNPs within and near AS3MT are not inherited independently, but instead as a haplotype block. In the Argentinean women described previously as efficient arsenic metabolizers, 52 SNPs were associated at a genome-wide level with urinary %MMA forming what was defined as a protective haplotype (Schlebusch et al., 2015).

How is the *AS3MT* protective haplotype defined?

It is a combination of alleles that were associated with lower MMA fractions in urine in an indigenous population from San Antonio de los Cobres, in the Argentinean Andes. This combination of alleles located on the same chromosome, either maternal or paternal, is known as *haplotype*. It is considered *protective* because a lower MMA fraction in urine is linked to less adverse health effects related to the toxicity of arsenic. This protective haplotype includes alleles within *AS3MT*, the main methylating enzyme, as well as *WBP1L*, *CYP17A1*, and *CNNM2*, among others.

Using SNP array data from the Bolivian study group (n = 187), a protective haplotype based on the one from the Argentinean group was defined as a candidate region in **Paper IV**. From the 52 SNPs conforming the protective haplotype in Argentina, 35 SNPs were present in the Bolivian dataset after filtering and merging. This protective haplotype is located in chromosome 10 (chr10:104,315,215-10:104,852,121, hg19) and includes *AS3MT* among other genes (**Figure 12**). In **Paper IV**, SNPs within the protective haplotype were associated with arsenic metabolite fractions in the Bolivian study group, but to a lesser extent than in SAC. This could be explained by the Bolivian study group having a lower variation in urinary %MMA than the women in SAC, and due to higher-than-expected frequency of the protective variants. Furthermore, genome-wide association analyses did not find any clear signals or significant associations, beyond the significant threshold, with arsenic metabolite fractions.

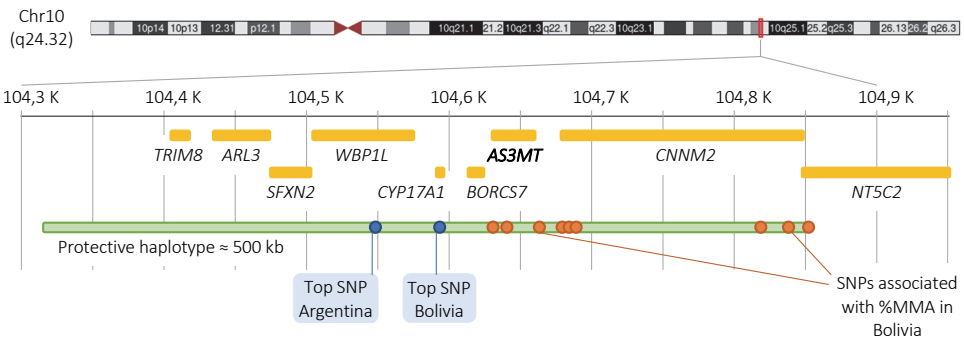


Figure 12. Genomic location of the protective haplotype associated with lower urinary %MMA in San Antonio de los Cobres, Argentina as described in Schlebusch et al. (2015). The protective haplotype is located at chr10:104,315,215-10:104,852,121 (based on the human reference genome version 37, hg19). Top SNP associated with urinary %MMA was rs486955 in Argentina (Schlebusch et al., 2015), and rs17115100 in Bolivia (**Paper IV**). Illustration of chromosome 10 obtained from UCSC Genome Browser (https://genome.ucsc.edu).

5.4 MOLECULAR EFFECTS OF ARSENIC EXPOSURE

Arsenic causes toxicity via a plethora of mechanisms of action at a molecular level. Driven mainly by the metalloid's capacity to bind to sulfhydryl groups, arsenic inhibits protein function and induces oxidative stress. The health consequences of arsenic exposure in the study area are widely unexplored and unknown. Considering the setting and how medical services are organized in these remote villages, thorough longitudinal studies are not currently feasible, and no clinical data are available. Instead, toxicity biomarkers were used to study the effects of arsenic. Despite these women at a population level being efficient at metabolizing arsenic, Bolivian women had molecular signs of arsenic-related toxicity, particularly in individuals with a less efficient metabolism.

In **Paper II**, the associations between arsenic exposure, assessed in two matrices (urine and blood), and traditional toxicity biomarkers were explored (**Figure 13**). The toxicity biomarkers included were telomere length and three related to oxidative stress, i.e., mtDNAcn, 8-oxo-dG, and 4-HNE-MA. Long telomeres and higher mtDNAcn were associated with arsenic exposure in the Bolivian study group. This association between arsenic exposure and longer telomeres has been previously described (Ameer et al., 2016; Gao et al., 2015a; Herlin et al., 2019; Jimenez Villarreal et al., 2019), as well as for higher mtDNAcn (Ameer et al., 2016; Sanyal et al., 2018).

Concentrations of 4-HNE-MA, a urinary metabolite of lipid peroxidation, were positively associated with U-As and B-As. Only two other studies have studied 4-HNE-MA in relation to arsenic exposure, and both found a positive association between U-As and 4-HNE metabolites in urine (He et al., 2020; Kozłowska et al., 2019).

Higher concentrations of 8-oxo-dG in urine, a biomarker of DNA damage induced by oxidative stress, were positively associated with U-As but inversely associated with B-As. A positive association between U-As and urinary 8-oxo-dG has been repeatedly described (Breton et al., 2007; Chou et al., 2014; Engström et al., 2010a, 2010b; Fujino et al., 2005; Kubota et al., 2006; Mar Wai et al., 2019). However, there are also studies reporting discrepant results depending on the biomarker of arsenic exposure used (Breton et al., 2007; Burgess et al., 2007). The positive association with U-As could be a consequence of co-excretion between the biomarker of exposure and toxicity in urine, previously described for other metals like cadmium (Wallin et al., 2014). It also highlights the need to properly understand whether 8-oxo-dG concentrations in urine reflect DNA damage. Based on these contradictory results, B-As was chosen as biomarker of exposure in **Paper III**.

The efficiency to metabolize arsenic was evaluated as an effect modifier for the association between arsenic exposure and these toxicity biomarkers in **Paper II**. The hypothesis was that individuals with a less efficient arsenic metabolism were more likely to show arsenic-related toxicity. Arsenic exposure was associated with longer telomeres and higher mtDNAcn particularly in women with a less efficient metabolism (**Figure 13**). This was previously described in the study group from SAC in the Argentinean Andes (Ameer et al., 2016). No clear effect of the arsenic metabolism capacity was seen for the urinary concentrations of 8-oxo-dG or 4-HNE-MA. Contrary to expected, efficient arsenic metabolizers, i.e., those with lower urinary %MMA, had higher urinary 4-HNE-MA concentrations with increasing B-As. It is speculated that there could be an interaction between the metabolic pathways of arsenic and 4-hydroxy nonenal since both depend on glutathione.

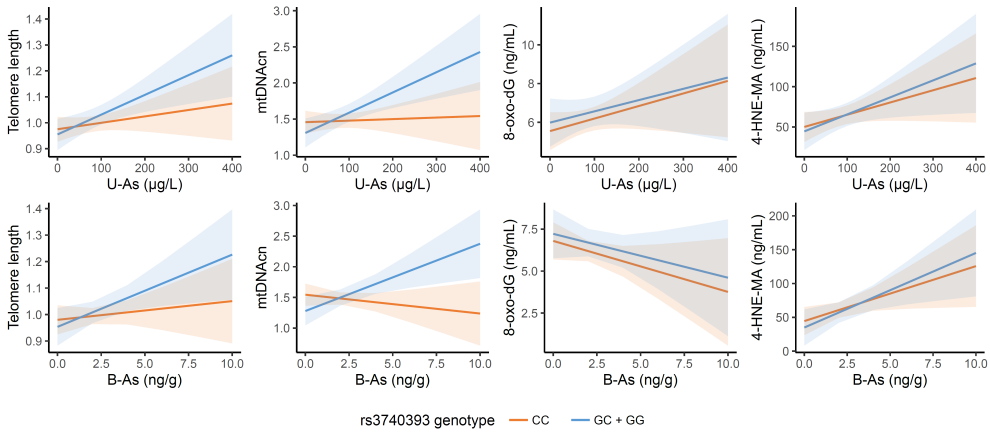


Figure 13. Predicted values (marginal effects of interaction terms) and 95% confidence intervals for adjusted linear regression models between toxicity biomarkers and arsenic exposure. Toxicity biomarkers (8-oxo-dG and 4-HNE-MA) and arsenic concentrations (U-As) measured in urine were corrected to mean urinary osmolality. Telomere length and mtDNAcn are relative values and have arbitrary units. Models were adjusted as described in **Paper II**. See Section 5.5 for a description of rs3740393. *Abbreviations:* 4-HNE-MA, 4-hydroxy nonenal mercapturic acid; 8-oxo-dG, 8-oxo-2'-deoxyguanosine; B-As, arsenic concentration in blood; mtDNAcn, mitochondrial DNA copy number; U-As, urinary arsenic as the sum of inorganic arsenic metabolite concentrations.

In **Paper III**, an exploratory approach was taken to evaluate whether a panel of 92 cancer-related proteins were associated with B-As. Out of the 92 proteins putatively related to cancer processes included in a predesigned panel, 45 proteins (those with more than 40% of observations above the limit of detection within each protein) were included in downstream analyses. Arsenic exposure, as B-As, was associated with the relative expression of four proteins measured in urine: tumor necrosis factor ligand superfamily member 6, FASLG; seizure 6-like protein, SEZ6L; Ly6/PLAUR domain-containing protein 3, LYPD3; and tissue factor pathway inhibitor 2, TFPI2 (**Figure 14**). Only FASLG has previously been linked to arsenic exposure in epidemiological studies. FASLG participates as a promotor and inhibitor of signaling pathways related to apoptosis, depending on the type of cell (Bossi et al., 2000). High concentrations of soluble FASLG in serum have been found in individuals with different types of cancer (Kadam and Abhang, 2016). *In vitro* studies have found an increased gene or protein expression of FASLG in leukemia cell lines treated with arsenic trioxide, a cancer drug currently used to treat acute promyelocytic leukemia (Toosi et al., 2018; Zeng et al., 2016; Zhu et al., 2003). Interestingly, the gene expression of *FASLG* was elevated in the indigenous women from SAC in the Argentinean Andes that were highly exposed to arsenic (Engström et al., 2017).

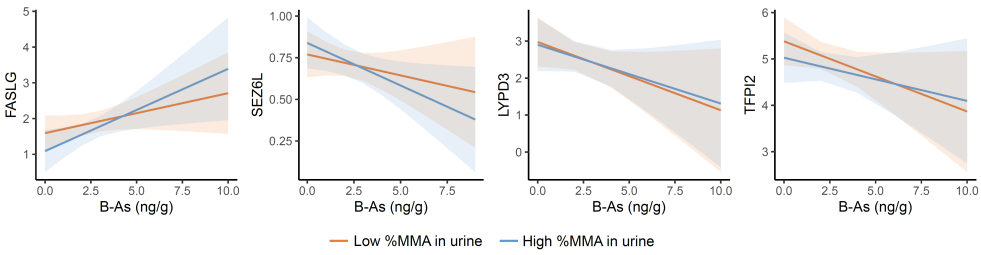


Figure 14. Predicted values (marginal effects of interaction terms) and 95% confidence intervals for adjusted linear regression models between the top four cancer-related proteins in urine associated with B-As. Models were adjusted as described in **Paper III**. Low urinary %MMA refers to individuals with %MMA below or equal the median, while high %MMA refers to above the %MMA median. *Abbreviations:* B-As, arsenic concentration in blood; MMA, monomethylarsonic acid.

Paper III is a cross-sectional study with no clinical outcomes included. Therefore, it is not possible to distinguish with certainty whether the findings are related to arsenic toxicity *per se* or are a consequence of an adaptive response that will not translate into adverse health outcomes. The associations between B-As and the four proteins were similar across arsenic metabolism efficiency strata. If stronger associations had been found among the less efficient arsenic metabolizers, which are more susceptible to arsenic toxicity, it would have supported that the changes in these proteins were because of arsenic toxicity. Nevertheless, the findings in **Paper III** serve as a starting point to determine novel biomarkers in future studies including clinical outcomes in relation to arsenic exposure. In addition, it promotes the potential that urine has as a non-invasive matrix in biomarker discovery.

Factors that influenced protein variation in urine were evaluated in **Paper III** using linear regression models of principal components. The overall protein variation was explained by differences in age, urinary osmolality, leukocytes in urine, and urinary pH. These factors were included as covariates when evaluating the associations between relative protein levels and B-As. Coca chewing, ethnicity, sampling date, and other data from the urine strip were not influencing factors. To evaluate how different factors explained the protein levels at an individual level, univariate models were used. The variation of most proteins measured in urine was explained by urinary osmolality, for some proteins explaining up to 49% but not being so influential for FASLG.

Studies on arsenic-related cancer in Latin America have focused mainly on Chile, Argentina, and Mexico (Khan et al., 2020). Very limited information is available regarding the health effects of arsenic in populations with an efficient arsenic metabolism such as the ones described in the Andes in **Paper I** and by Vahter et al. (1995). This, and the fact that cancer prevalence varies across populations (Lorenzo Bermejo et al., 2017), emphasizes the need for population-specific studies. Several outcomes have been studied in relation to arsenic exposure in the Argentinean women with a highly efficient arsenic metabolism living around SAC. However, these outcomes were all biomarkers of effects, and not clinical. For example, arsenic exposure was associated with an increased frequency of micronuclei in lymphocytes (Dulout et al., 1996); lower blood pressure, lower ratios of apolipoprotein B/A, and lower hemoglobin in blood (Ameer et al., 2015); and longer telomeres and higher mtDNAcn particularly in the subgroup of individuals with a less efficient arsenic metabolism (Ameer et al., 2016). These results, together with findings from **Papers II** and **III**, suggest that arsenic exposure can lead to toxic outcomes even in populations that are described, on average, as efficient arsenic metabolizers. However, no signs of hyperkeratosis or pigmentation changes of the palms were ever seen in these Argentinean women, which is interesting since these often are the first signs of chronic inorganic arsenic exposure (Concha et al., 1998a; Vahter et al., 1995).

5.5 GENETIC ADAPTATION TO ARSENIC EXPOSURE

The Bolivian women from the study group were shown to be efficient at methylating arsenic in **Paper I**. Then, in **Paper II**, the genetic variant rs3740393 in *AS3MT* was associated with arsenic metabolite fractions in urine: individuals with genotype CC had lower urinary %MMA, an indication of a more efficient arsenic metabolism. Furthermore, the C-allele of this SNP was at a higher frequency in the Bolivian study group (79%) than in other populations exposed chronically to iAs in Chile (68%; Apata et al., 2017) or Argentina (70%; Engström et al., 2011), and other populations worldwide (**Figure 15**). In **Paper IV**, the frequency of the *protective haplotype*, predefined based on SNPs associated with lower urinary %MMA in SAC (Schlebusch et al., 2015), was similar in SAC and in the Aymara-Quechua group (63% and 66%, respectively), but higher for the Uru group (75%).

The high frequency of alleles associated with a beneficial trait can be a consequence of strong selective pressures in the environment modulating the allelic frequencies in a population. Therefore, genome-wide selection scans were carried out to test for signs of positive selection across the genome in **Paper IV**. Signals near *AS3MT* were within the top selected regions throughout the genome in the Aymara-Quechua and Uru groups, supporting the hypothesis that these Bolivian study groups have genetically adapted to eliminate arsenic more efficiently.

Indications of a selection for increased detoxification of arsenic were first found in indigenous populations from the Andes living in Argentina (Schlebusch et al., 2015). Since then, smaller and less comprehensive studies with individuals across the Andes have reported changes in the frequency of alleles associated with a more efficient arsenic metabolism, as well as signs of selection near *AS3MT* (Apata et al., 2017; Eichstaedt et al., 2015; Jacovas et al., 2018). However, none of these studies included phenotypic data on arsenic methylation.

Genome-wide selection scans, as the name implies, serve to identify possible regions of the genome that have been subjected to positive selection. This alone does not necessarily translate to tolerance or adaptation, which require of a biological and environmental context to be able to interpret the results. The notion of adaptation to tolerate arsenic in indigenous populations in the Andes is supported by (i) the signals of positive selection near *AS3MT* in **Paper IV** that have led to (ii) the high frequencies of alleles associated with a more efficient arsenic metabolism, which (iii) is evident from phenotypic analyses in **Paper I**.

Adaptation takes place over the course of multiple generations. Therefore, putting arsenic exposure into an historical context helps understand how a potential adaptation to tolerate arsenic occurred. Evidence of environmental arsenic exposure for centuries in the Andes is supported by traces of arsenic in pre-Hispanic mummies from Chile (Arriaza et al., 2010; Echeverría et al., 2018; Kakoulli et al., 2014). Epidemiological studies indicating arsenic exposure is associated with higher child mortality and morbidity (Hopenhayn-Rich et al., 2000; Rahman et al., 2010; Shih et al., 2017; Smith et al., 2012) suggest arsenic as a strong selective pressure capable of shaping the survival chances and therefore, allowing beneficial traits to increase in frequency over generations. It is worth noting that the present exposure to arsenic in the Bolivian study groups may not reflect their exposure in the past, which could have been even higher and driving the selection.

The selection of a more efficient arsenic metabolizing phenotype in the Andean populations has likely risen from standing genetic variation. This means that the genetic variants responsible of the advantageous trait existed in the population before the individuals encountered the environmental stressor. This idea is supported by the fact that stronger selection signals are identified with allele frequency selection scans instead of haplotype-based methods (see Section 4.8 for more details). In addition, the protective haplotype is also observed in East Asian populations, although at a lower frequency (Schlebusch et al., 2015). The presence of the protective haplotype in other parts of the world supports the notion that the genetic variants under selection existed before the selective pressure modulated the allelic frequency.

Although signs of positive selection related to arsenic detoxification have been identified across populations in the Andes, the functional variant(s) and exact mechanism driving this selection are still unknown. A haplotype in high linkage disequilibrium spanning up to 600 kb around *AS3MT* has been detected in Argentinean (Schläwicke Engström et al., 2009) and Mexican populations (Gomez-Rubio et al., 2010). While this facilitates the detection of signals in whole-genome association studies, it hinders the identification of the functional variant(s).

In **Paper IV**, targeted sequencing around the top SNP associated with urinary %MMA in SAC was used to explore the presence of genetic variants not captured in the SNP array in Argentinean and Bolivian individuals. Twelve SNPs, out of which three were included in the SNP array, were found within the 4.8 kb region sequenced, and all had been previously identified in SNP databases. The associations of these SNPs with arsenic metabolism efficiency should be further evaluated.

Recently, a study including targeted sequencing data identified 13 rare variants (<1% allelic frequency) in the coding regions of *AS3MT* that were associated with a less efficient arsenic metabolism, indicated by lower %DMA, across three arsenic-exposed study groups (Delgado et al., 2021). It will be interesting to explore the consequences of these rare variants in coding regions on arsenic toxicity. However, human traits are usually complex and commonly explained by a collection of genetic variants, each of them contributing partially in an additive manner. Therefore, it could be that several non-coding genetic variants near *AS3MT* have a combined effect on, for example, the gene expression of *AS3MT*.

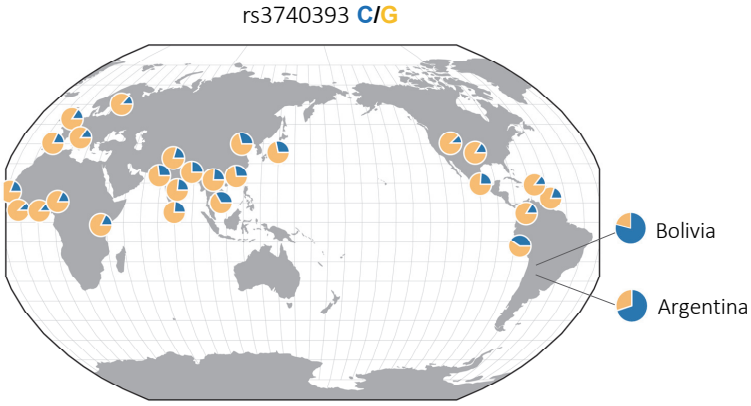


Figure 15. World map with allelic frequencies of SNP rs3740393 in *AS3MT*. Data from Argentina obtained from women from San Antonio de los Cobres as reported in Engström et al. (2011). Data from Bolivia is the combination of Aymara-Quechua and Uru women described in **Paper II**. Map made with the Geography of Genetic Variants browser (beta v0.4) developed by Marcus & Novembre (2016).

5.6 METHODOLOGICAL CONSIDERATIONS

The sample sizes of the studies included in this thesis are small, including between 163 and 201 individuals. Despite the apparent small sample size, recruitment took place in all Uru villages around Lake Poopó and in several Aymara-Quechua villages in the region. These villages are small in general; however, a precise estimate of the total population in each village was not available. When we asked at the local health centers for an estimation of the population size, in those villages that did have some type of register, we were warned that they were not accurate. This was because there was a large discrepancy between those officially registered in the village versus those that actually lived there and could have been reached. Recruiting in these remote rural villages of the Bolivian Andes is challenging. For example, the physical access via road to the villages was limited, all the materials and resources had to be transported to each village, and in some cases, we even had limited access to electricity. In this context, a convenience sampling was the only feasible solution. Therefore, an exact participation rate could not be calculated.

A potential source of bias, or systematic error, can occur during the selection of participants. However, it was not expected that women with higher levels of arsenic were more likely to participate since they did not have knowledge of their own arsenic exposure beforehand. In addition, since it was a cross-sectional study, causal relationships could not be concluded. Furthermore, increasing the sample size reduces random errors, i.e., those affecting the estimation of the effects. However, to overcome this, confidence intervals were included to indicate the precision of the estimates, and we were aware of this limitation when interpreting the results.

A novelty and strength of this thesis is the inclusion of several ethnic groups from the Bolivian Andes, the Aymara-Quechua and Uru, the latter highly underrepresented in the scientific literature. Ethnicity was assessed based on the location of residency of the participant and where their family members came from, which was feasible considering the limited contact between villages of different ethnicity. Ethnicity is a social construct, and its inclusion in epidemiological studies regarding environmental exposure has a set of considerations (Benmarhnia et al., 2021). In the projects included in this thesis, ethnicity served as a proxy for differences in lifestyle or exposures, as well as a reflection of distinct populations from a specific genetic perspective. Genetic differences between ethnicities might influence the capacity to metabolize arsenic, as well as other processes involved in how the metalloid exerts its toxicity, e.g., polymorphisms in genes related to oxidative stress management. In **Paper I**, ethnicity was included as a covariate to understand whether the capacity to methylate arsenic was different in each ethnic group. In **Paper II**, ethnicity was used as a confounder: both exposure (arsenic concentrations) and outcome (urinary concentrations of toxicity biomarkers) were different between ethnicities, and ethnicity is not a consequence of the arsenic exposure.

Arsenic exposure in this study group was characterized with accurate and reliable ICP-MS methods and the use of two matrices, blood and urine. Therefore, the risk of exposure misclassification, a type of information bias, is somewhat reduced. Furthermore, a wide range

of U-As permits the use of linear regression models with a quantitative exposure, especially since it was not possible to identify other villages in the region that were not exposed to arsenic.

Unidentified confounders can mask or falsely show associations, another type of systematic error. The fact that women in this study group did not smoke or consume alcohol to a great extent avoided additional potential confounders. However, it must be noted that other unidentified confounders can still exist. Another type of information bias can arise from how the variables related to food intake were defined. Our questionnaire was not intended to be a dietary questionnaire, and therefore, in **Paper I** the variables were used mostly to describe the general eating habits across the villages and ethnic groups.

In **Paper II**, and specifically **Paper III**, it is hard to disentangle whether the changes observed regarding the toxicity biomarkers and protein levels are in fact a toxic response or a consequence of adaptive mechanisms in the organism. Nevertheless, the findings still provide important insights by identifying potential signs of molecular toxicity and guide future research within biomarker discovery in urine. Furthermore, the exposure to other metals, pesticides or air pollutants can also contribute to oxidative stress and should be considered in future studies. As a pilot analysis, we evaluated 11 pesticide metabolites in urine of a subset of 59 Bolivian women and concluded that they were not highly exposed. However, it could be that the pesticides selected did not reflect the ones they used.

Multiple testing also contributes to random error. The more tests performed, the higher the risk of chance findings, also known as false positives. However, for this to happen the tests should be independent, which is not entirely the case for the proteins measured in **Paper III** or the genetic variants explored in **Paper IV**. In any case, considerations for multiple testing were included in **Papers III** and **IV**. At the same time, there is also a risk for false negatives because of the low frequency of the alleles and the small sample size of these study groups. Consequently, special attention must be focused on interpreting the plausibility of the findings within a biological context (see Section 4.9).

5.7 ETHICAL CONSIDERATIONS

Ethical permits were granted from ethical committees in Sweden, Bolivia, and Argentina to perform the studies included in this thesis. The considerations in this section are described within the context of the Bolivian project since I personally participated in all processes of the study, from visiting the villages and helping recruit the participants, to data analyses and reporting of the results. However, they also apply to the part of the project carried out with data from Argentina included in **Paper IV** to identify novel *AS3MT* variants.

This cross-sectional study consists of people from three ethnic groups living around Lake Poopó: Aymara and Quechua (combined as Aymara-Quechua, see Section 4.2) and Uru. The Aymara and Quechua are the most predominant indigenous groups in the study area. On the contrary, the Uru communities around Lake Poopó are secluded to a few villages with a long history of social and economic exclusion. The interaction between researchers and study participants needs to rely on trust, clear understanding of the research intentions, and transparency. With this purpose, multiple socialization activities took place before, during, and after the recruitment, for example, information meetings and dynamic workshops with the inhabitants using visual aids such as the on-site arsenic detection kits. All this, while extending the length and difficulty of the recruitment, contributes to the connection between the participants and the researchers.

The project was explained orally and in written form to potential participants to guarantee that they understood the purpose of the study and what it entailed. This was especially important in cases where older or illiterate people wanted to participate. Before taking samples or interviewing the participants, oral and written consent was obtained. We explained that their participation was voluntary, and that they had the right to withdraw from the study without requiring an explanation and emphasizing that there were no negative consequences if doing so. Participants had the chance to ask questions to the researchers during the visits. The same villages were visited several times which allowed the participants to solve further doubts and discuss the results. In addition, contact information to the researchers in Bolivia was given, and we collaborated with a local doctor who was working in the area and could answer any questions while we were not there.

During the recruitment, blood and urine samples were taken, and we interviewed the participants. This could induce stress to the participants, especially if they were uncomfortable with needles, and implied that they invested part of their time to collaborate with us. In return, we did not offer gifts or economical compensation. Instead, we performed a quick on-site health assessment, and the results were directly given to the participants. This included determination of blood pressure, urinary tract infection analysis, glucose determination in urine as an indicator of diabetes, and hemoglobin data to assess iron status. Since health professionals and the medical services in the villages are normally very limited, this approach was considered more beneficial and relevant for the participants. Information about the arsenic concentrations in drinking water sources from the different villages was given to the health centers and local authorities. In addition, personalized letters with information about individual exposure to arsenic were prepared and delivered to the participants through the local health providers.

Finally, how to handle genetic information is of utmost importance and an ongoing debate. As mentioned to the participants, all data were presented only at a population level, and not individually. Genetic polymorphisms indicate to some extent the risk for arsenic toxicity at a population level. Recently, this topic has been delved into in Bangladesh, an area where arsenic has been a public health concern. Participants were asked whether they wanted to receive this genetic information, and interestingly, they did (Tamayo et al., 2020). However, SNP data alone should not be used to estimate personal risk to arsenic exposure, particularly considering that other factors also contribute. It should also be noted that in low resource villages, as the ones included in this thesis, receiving this genetic information might be detrimental since individual actions might not be able to be carried out to decrease their exposure to arsenic. Nonetheless, advantages and disadvantages of sharing this information must be weighed beforehand, and a thorough communication plan is required to avoid misinterpretation of the results.

5.8 PUBLIC HEALTH IMPLICATIONS

Arsenic exposure is a major public health concern that affects millions of people around the world. In the case of the indigenous communities in the Andes included in this thesis, our contact with them and our research findings hopefully contribute to:

- Addressing the lack of access to clean water, included within the Sustainable Development Goals of the United Nations. Data from **Paper I** can help plan the use of water sources by identifying which have less arsenic.
- Reporting data on arsenic concentrations in water and human exposure can attract other organizations to help reduce the metal exposure via drinking water.
- Supporting the introduction and development of new local policies concerning water treatments by contributing with exposure data.
- Increasing the awareness about arsenic as a public health concern in the communities.

It is worth noting that no clinical outcomes were included in this thesis to evaluate the toxicity of arsenic in these communities. This is due primarily to the lack of available clinical registries in these rural settings and the inherent limitations of the field work to perform longitudinal studies. Therefore, despite these populations showing a more efficient capacity to methylate arsenic, toxicity in these Andean populations cannot be excluded. In fact, as shown in **Papers II** and **III**, there are molecular signs related to arsenic toxicity that must not be overlooked. Furthermore, children, a vulnerable group to environmental pollutants, were not evaluated. As a result, even though these indigenous populations were described as efficient arsenic metabolizers, it is vital not to ignore the exposure to arsenic in this region further.

6 CONCLUSIONS

The conclusions of the work included in this thesis were the following:

- Indigenous women living around Lake Poopó, Bolivia are exposed to inorganic arsenic presumably through drinking water. These women had low fractions of MMA and high fractions of DMA in urine indicating an efficient arsenic metabolism. Factors such as body weight, fish consumption, tobacco smoking, and ethnicity predicted the capacity to metabolize arsenic in these women.

- Arsenic exposure, assessed in urine and blood, was associated with molecular signs of toxicity, and the association was modified by the arsenic metabolism efficiency – a known susceptibility factor. Arsenic exposure was associated with longer telomeres and higher mtDNAcn, particularly in women with a less efficient arsenic metabolism. A novel biomarker of lipid peroxidation, 4-HNE-MA, was positively associated with arsenic exposure. Conflicting results for 8-oxo-dG measured in urine accentuate the need to reconsider this biomarker of oxidative stress and the limitation of measuring both exposure and outcome in urine.

- Arsenic exposure assessed in blood was associated with higher relative urinary levels of FASLG, and lower of SEZ6L, LYPD3, and TFPI2, four proteins putatively related to cancer processes. However, future research is required to confirm whether these proteins can be used as biomarkers of early cancer in relation to arsenic exposure.

- Indigenous communities in the Bolivian Andes had genomic signs of positive selection near *AS3MT*, the main arsenic methylating enzyme in humans. The frequency of protective alleles near *AS3MT* associated with lower fractions of urinary MMA was the highest described so far. Taken together with their efficient arsenic metabolism phenotype, this supports the notion that exposure to arsenic during multiple generations has been a strong selective pressure in humans.

7 FUTURE PERSPECTIVES

In light of the current findings, additional efforts and attention should be focused towards:

- Exploring further environmental and genetic factors that explain the efficient arsenic metabolism of these Andean populations.

- Identifying the functional and potentially novel genetic variants near *AS3MT* that could have driven arsenic adaptation in these Andean populations.

- Elucidating the potential additive effect of multiple non-coding *AS3MT* variants on the expression of *AS3MT* and its consequence on arsenic metabolism efficiency.

- Monitoring and reducing the exposure to arsenic in the inhabitants of these villages.

- Characterizing additional arsenic exposure routes in this arid region, such as inhalation of dust.

- Investigating potential health effects of arsenic in Bolivian populations, such as the incidence of cancer, cardiovascular diseases, etc.

- Performing longitudinal studies to link potential protein biomarkers measured in urine with clinical data, such as cancer incidence.

8 ACKNOWLEDGMENTS

Working on this Ph.D. project has been one of the hardest adventures, personally and professionally, I have encountered so far. And yet, it has been an absolute pleasure to be part of this fascinating research, surrounded by the best teams I could have ever wished for. To everyone who I have met along this journey – *thank you*. When I came back to Sweden after my first trip to Bolivia in 2016, I remember telling my parents about it and describing it as a *life-changing* experience. And still today, I am certain this project has changed who I am as a researcher and as a person at levels I could have never imagined. It is incredible how small one can feel when we are exposed to unknown worlds to us, and how *adaptable* we realize we are once we are pushed out of our comfort zones.

This work was carried out with financial support from Karolinska Institutet, the Eric Philip Sörensens Stiftelse, the Swedish Research Council, the Swedish International Development Cooperation Agency, the Genetics Institute at Universidad Mayor de San Andrés, the Institut de Recherche pour le Développement, and the Hydrosciences Montpellier Laboratory.

To all the **strong women I met in Bolivia**. For all those times when we would shake hands before an interview, and you would be worried because my hands were so cold and wondered how I was handling the altitude; when I chocked on my saliva after chewing the coca leaves you invited me to in the middle of a meeting, and you just smiled back at me and we laughed; when you were scared of needles, and I tried to distract you with my colorful socks because I completely understand your fears. Thank you to everyone we met during our visits around Lake Poopó: for your warm welcome, and for believing and trusting us. It has been an honor for me to be able to work with you in your villages, get to know your culture, and hopefully contribute for you to get better drinking water quality. This is just the beginning.

To my main supervisor, **Karin Broberg**. What an adventure, right? Despite, and maybe also because of, all the tough times we have been through together, I am extremely grateful for this opportunity and having had you as my supervisor. Thank you for trusting and believing in me. We always had each other's backs – and that created the safest environment possible. You always had an open door for me – metaphorically and literally to all your homes – from the very first day. Sharing this journey with you has been a pleasure. It is impossible for me to imagine a more rewarding Ph.D. journey – thank you.

Carina Schlebusch, my co-supervisor: I wish the academic world were full of more people like you, and the non-academic world as well. Your curiosity and passion for your work, and your dedication to your team is inspiring. I still remember the days when I was haunting EBC with a sad face because I could not get any results, and you genuinely cared for me. Thank you for sharing your incredible knowledge with me, and for being so caring.

Anda Gliga, you are such an amazing scientist and friend. Thank you for stepping in as a co-supervisor, sharing all your knowledge, and helping me develop a critical-thinking mindset. Sharing a workspace with you is a blessing full of support and stimulating conversations.

To all my colleagues at the Unit of Metals and Health. Thank you to **Marie Vahter**, for sharing all your knowledge about arsenic and for teaching me so much more than you could imagine; **Maria Kippler**, for always keeping a positive attitude and giving me the chance to get involved in teaching – I think we made a great team; **Barbro Nermell**, for your contagious kindness and strength; **Michael Levi**, for your excellent lab skills and your eternal patience every time I had to order things; **Florencia Harari**, for all your hard-work in Argentina that was extremely inspiring to me; **Andréa Pons**, for being one of my first students with whom I had a great connection; **Andrea Cediel**, for caring so much for me since the first years in Sweden and sharing our passion for plants; **Daniela Andries** and **Puck Norell**, for your drive and passion; and to **Ulrike Dauter** and **Mathilde Delaval**, for contributing to such a nice environment.

To **The Metalligans**, who have been a cornerstone of my journey as a Ph.D. candidate. Having you by my side and having such a special bond has made it so much more fun and easier. I still remember when I went for the interview and on my way out, I stopped by your office to check who would be my potential colleagues. On the whiteboard you had a list of who wanted food "service" from the one and only Heleniz. I already knew then that I had to become part of that amazing group of coolness and kindness. To **Annachiara**, for always being ready to lend a helping hand when I needed it and for your fantastic language skills. To **Ayman**, for your inspiring strength and perseverance. Thank you for teaching me to be pragmatic when I needed it the most. To **Helena**, for including me as part of your family. Your eternal kindness, dedication, structure, and help have been, and will be, an important pilar for me. Working with you is fantastic, professionally and personally. To **Klara**, for being so true to yourself. Your creativity and care-free attitude are inspiring for me. I am not sure I could have done this without the four of you.

To everyone I had the pleasure to meet from the Institute of Environmental Medicine. Special thanks to **Jessika Barrón**; you are taking back amazing skills to Bolivia. To **Kristian Dreij**, **Emma Vincent**, and **Imran Ali** for giving me the opportunity to teach. And to all the teachers I have had during these years.

To the amazing team in Bolivia. I truly cherish all the memories we created together. **Noemi Tirado**, you have always treated me as family when I am there, and I am very thankful for that. Thank you for all the invaluable effort you put into making this project possible. To **Jacques Gardon**, for teaching me to be critical and taking care of me during the trips. To **Josue Mamani** and **Marion Gonzales**, for the special weeks we spent together in Stockholm and all your help in Bolivia; you are going to go so far with the passion that drives you. A **Marina Cuti** y **Rolando Paz**, por ser los mejores compañeros de trabajo de campo que una podría desear. A **Abdul Castillo**, por cuidarme tanto durante los viajes y por lo bien que lo pasamos juntos. Y un agradecimiento especial para **Franz Ascui**. Gracias por toda tu ayuda durante el proyecto en Bolivia, por cuidar de mí y por ser el compañero más divertido. Siempre me acordaré de cómo íbamos buscando sitios para comer mientras Jacques insistía en que teníamos que llegar a tiempo al siguiente pueblo.

To all members at the Human Evolution Program, at Uppsala University, specially to those in Carina's group. Thank you for welcoming me every time I sporadically visited you in Uppsala and sharing all your knowledge within population genetics with me. **Mário Vicente**, you ended up being one of the best things of my time in Uppsala. You have such a brilliant mind; thanks for letting me *borrow* it every time I needed it. I have had the best of times working with you, and I would do it all over again, even despite the coding. **Rickard Hammarén**, I appreciate how welcoming you have always been with me since the beginning.

To everyone I met in Copenhagen, thank you for being so tremendously fun – it is true that Danes (and non-Danes in Denmark) know how to party. To **Michael Palmgren**, for hosting me at the University of Copenhagen. Your vast knowledge, inside and out of the lab, and your passion are inspiring. I cannot wait to meet again on the next dance floor. To **Jesper Torbøl Pedersen**; working side by side with you is a pleasure. Thank you for being an amazing colleague and inviting me into your family every time I visited. I am happy in the end we did not go insane with the membranes.

To those at Lund University who taught me my first ever genotyping and were so friendly. Thank you, **Daniela Pineda**, for always being a source of help and good mood. To **Annette Krais**, thank you for all your help analyzing samples and your positive attitude from the distance. And to **Christian Lindh**, for sharing a Bolivian adventure and all your MS knowledge with me.

To my mentor, **Olle Bergman**. For always being there for me with a cheerful attitude and ready to listen, and for guiding me within the world of science communication.

To **Diana Ekman**, and the whole Bioinformatics Advisory Program and members, and **Vincent Van Hoef** for the chance to embrace the world of bioinformatics.

To the **Biomusicum** choir at Biomedicum, for being so welcoming and letting me explore another artistic side of mine. Singing with you helped me relax during challenging Ph.D. times.

To all my friends, near and far, that always managed to feel close at heart and made these years so fun. Special thanks to **Maria** and **Hayrie**, for your constant support and being like an extra family to me. Thank you for always taking so much care of me. To **Shuang**, for reconnecting thanks to a random course and always being there for me since then. To my beloved Spaniards and south Europeans that I met in Stockholm; thank you so much for your fresh energy and all the good times. To **Carolina**, your kindness and positive attitude always warm my heart. To **Paula**, **Gloria**, and **Andrea**, for making going back to Valencia feel like time has never passed. And to **Ana Perelló** and family, for including me as part of one of them and for *consistently* being there. Thank you also for your help editing the Spanish summary of this thesis.

To my family, friends of the family, and David's family and friends, for your encouragement and following my journey.

A special thank you to **David**. For your endless support, patience, care, and optimism. I really appreciate the pragmatic and kind approach you have towards life, and I learn from it every day. Sharing a home, a living room, a bedroom, and a full-time office, all in 20 m^2, was easy just because it was with you. Thank you for being invested in this thesis, looking at the same figure a million times at different resolutions and for helping me with the Swedish summary.

To my parents, **Sherry** and **Luis**. You both are a never-ending source of strength, courage, and motivation for me. The way you tackle life when it gets hard is a beautiful example of how humans can *adapt* with a determined, positive, and loving mindset. I am also thankful for all the times you have read and grammar-checked my documents – I learn from the best.

And to *past Jessica*, for thriving under – what at times felt like – the *harshest* conditions and not giving up. For knowing how to find support in the people around and when to stop and take one step back, in order to take a bigger leap forward after.

9 REFERENCES

Agency for Toxic Substances and Disease Registry (ATSDR). 2007. Toxicological Profile for Arsenic. Draft for Public Comment. Atlanta GA [updated 2007 August]. https://www.atsdr.cdc.gov/toxprofiles/tp2.pdf

Agusa, T., Fujihara, J., Takeshita, H., Iwata, H., 2011. Individual Variations in Inorganic Arsenic Metabolism Associated with AS3MT Genetic Polymorphisms. Int. J. Mol. Sci. 12, 2351–2382. https://doi.org/10.3390/ijms12042351

Ahmed, S., Moore, S.E., Kippler, M., Gardner, R., Hawlader, M.D.H., Wagatsuma, Y., Raqib, R., Vahter, M., 2014. Arsenic Exposure and Cell-Mediated Immunity in Pre-School Children in Rural Bangladesh. Toxicol. Sci. 141, 166–175. https://doi.org/10.1093/toxsci/kfu113

Ahsan, H., Chen, Y., Kibriya, M.G., Slavkovich, V., Parvez, F., Jasmine, F., Gamble, M. V., Graziano, J.H., 2007. Arsenic Metabolism, Genetic Susceptibility, and Risk of Premalignant Skin Lesions in Bangladesh. Cancer Epidemiol. Biomarkers Prev. 16, 1270–1278. https://doi.org/10.1158/1055-9965.EPI-06-0676

Alexander, D.H., Novembre, J., Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 19, 1655–1664. https://doi.org/10.1101/gr.094052.109

Ameer, S.S., Engström, K., Harari, F., Concha, G., Vahter, M., Broberg, K., 2015. The effects of arsenic exposure on blood pressure and early risk markers of cardiovascular disease: Evidence for population differences. Environ. Res. 140, 32–36. https://doi.org/10.1016/j.envres.2015.03.010

Ameer, S.S., Xu, Y., Engström, K., Li, H., Tallving, P., Nermell, B., Boemo, A., Parada, L.A., et al., 2016. Exposure to Inorganic Arsenic Is Associated with Increased Mitochondrial DNA Copy Number and Longer Telomere Length in Peripheral Blood. Front. Cell Dev. Biol. 4, 87. https://doi.org/10.3389/fcell.2016.00087

Andrews, C.A., 2010. Natural Selection, Genetic Drift, and Gene Flow Do Not Act in Isolation in Natural Populations. Nat. Educ. Knowl. 3(10):5.

Antman, K.H., 2001. Introduction: the history of arsenic trioxide in cancer therapy. Oncologist 6, 1–2. https://doi.org/10.1634/theoncologist.6-suppl_2-1

Antonelli, R., Shao, K., Thomas, D.J., Sams, R. 2nd, Cowden, J., 2014. AS3MT, GSTO, and PNP polymorphisms: Impact on arsenic methylation and implications for disease susceptibility. Environ. Res. 132, 156–167. https://doi.org/10.1016/j.envres.2014.03.012

Apata, M., Arriaza, B., Llop, E., Moraga, M., 2017. Human adaptation to arsenic in Andean populations of the Atacama Desert. Am. J. Phys. Anthropol. 163, 192–199. https://doi.org/10.1002/ajpa.23193

Archer, J., Hudson-Edwards, K.A., Preston, D.A., Howarth, R.J., Linge, K., 2005. Aqueous exposure and uptake of arsenic by riverside communities affected by mining contamination in the Río Pilcomayo basin, Bolivia. Mineral. Mag. 69, 719–736. https://doi.org/10.1180/0026461056950283

Argos, M., Tong, L., Roy, S., Sabarinathan, M., Ahmed, A., Islam, M.T., Islam, T., Rakibuz-Zaman, M., et al., 2018. Screening for gene–environment (G×E) interaction using omics data from exposed individuals: an application to gene-arsenic interaction. Mamm. Genome 29, 101–111. https://doi.org/10.1007/s00335-018-9737-8

Arriaza, B., Amarasiriwardena, D., Cornejo, L., Standen, V., Byrne, S., Bartkus, L., Bandak, B., 2010. Exploring chronic arsenic poisoning in pre-Columbian Chilean mummies. J. Archaeol. Sci. 37, 1274–1278. https://doi.org/10.1016/j.jas.2009.12.030

Assarsson, E., Lundberg, M., Holmquist, G., Björkesten, J., Bucht Thorsen, S., Ekman, D., Eriksson, A., Dickens, E.R., et al., 2014. Homogenous 96-Plex PEA Immunoassay Exhibiting High Sensitivity, Specificity, and Excellent Scalability. PLoS One 9, e95192. https://doi.org/10.1371/journal.pone.0095192

Bailey, K.A., Wu, M.C., Ward, W.O., Smeester, L., Rager, J.E., García-Vargas, G., Del Razo, L.M., Drobná, Z., et al., 2013. Arsenic and the Epigenome: Interindividual Differences in Arsenic Metabolism Related to Distinct Patterns of DNA Methylation. J. Biochem. Mol. Toxicol. 27, 106–115. https://doi.org/10.1002/jbt.21462

Balakrishnan, P., Jones, M.R., Vaidya, D., Tellez-Plaza, M., Post, W.S., Kaufman, J.D., Bielinski, S.J., Taylor, K., et al., 2018. Ethnic, geographic, and genetic differences in arsenic metabolism at low arsenic exposure: A preliminary analysis in the multi-ethnic study of atherosclerosis (MESA). Int. J. Environ. Res. Public Health 15. https://doi.org/10.3390/ijerph15061179

Balakrishnan, P., Vaidya, D., Franceschini, N., Voruganti, V.S., Gribble, M.O., Haack, K., Laston, S., Umans, J.G., et al., 2017. Association of Cardiometabolic Genes with Arsenic Metabolism Biomarkers in American Indian Communities: The Strong Heart Family Study (SHFS). Environ. Health Perspect. 125, 15–22. https://doi.org/10.1289/EHP251

Barbieri, F.L., Gardon, J., Ruiz-Castell, M., Paco V., P., Muckelbauer, R., Casiot, C., Freydier, R., Duprey, J.L., et al., 2016. Toxic trace elements in maternal and cord blood and social determinants in a Bolivian mining city. Int. J. Environ. Health Res. 26, 158–174. https://doi.org/10.1080/09603123.2015.1061114

Barr, D.B., Wilder, L.C., Caudill, S.P., Gonzalez, A.J., Needham, L.L., Pirkle, J.L., 2005. Urinary creatinine concentrations in the U.S. population: Implications for urinary biologic monitoring measurements. Environ. Health Perspect. 113, 192–200. https://doi.org/10.1289/ehp.7337

Batai, K., Williams, S.R., 2014. Mitochondrial variation among the Aymara and the signatures of population expansion in the central Andes. Am. J. Hum. Biol. 26, 321–30. https://doi.org/10.1002/ajhb.22507

Bauchet, M., 2008. Human Population Stratification and its Assessment by Microarray Genotyping. Encycl. Life Sci. (ELS). John Wiley Sons Ltd, Chichester. https://doi.org/10.1002/9780470015902.a0020797

Benmarhnia, T., Hajat, A., Kaufman, J.S., 2021. Inferential challenges when assessing racial/ethnic health disparities in environmental research. Environ. Heal. 20, 7. https://doi.org/10.1186/s12940-020-00689-5

Bhattacharjee, P., Banerjee, M., Giri, A.K., 2013. Role of genomic instability in arsenic-induced carcinogenicity. A review. Environ. Int. 53, 29–40. https://doi.org/10.1016/j.envint.2012.12.004

Bigham, A., Bauchet, M., Pinto, D., Mao, X., Akey, J.M., Mei, R., Scherer, S.W., Julian, C.G., et al., 2010. Identifying Signatures of Natural Selection in Tibetan and Andean Populations Using Dense Genome Scan Data. PLoS Genet. 6, e1001116. https://doi.org/10.1371/journal.pgen.1001116

Blackburn, E.H., 2005. Telomeres and telomerase: Their mechanisms of action and the effects of altering their functions. FEBS Lett. 579, 859–862. https://doi.org/10.1016/j.febslet.2004.11.036

Borak, J., Hosgood, H.D., 2007. Seafood arsenic: Implications for human risk assessment. Regul. Toxicol. Pharmacol. 47, 204–212. https://doi.org/10.1016/j.yrtph.2006.09.005

Bornhorst, J., Ebert, F., Meyer, S., Ziemann, V., Xiong, C., Guttenberger, N., Raab, A., Baesler, J., et al., 2020. Toxicity of three types of arsenolipids: species-specific effects in Caenorhabditis elegans. Metallomics 12, 794–798. https://doi.org/10.1039/d0mt00039f

Bossi, G., Stinchcombe, J.C., Page, L.J., Griffiths, G.M., 2000. Sorting out the multiple roles of Fas

ligand. Eur. J. Cell Biol. 79, 539–543. https://doi.org/10.1078/0171-9335-00077

Bozack, A.K., Domingo-Relloso, A., Haack, K., Gamble, M. V., Tellez-Plaza, M., Umans, J.G., Best, L.G., Yracheta, J., et al., 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. https://doi.org/10.1289/EHP6263

Breton, C. V., Kile, M.L., Catalano, P.J., Hoffman, E., Quamruzzaman, Q., Rahman, M., Mahiuddin, G., Christiani, D.C., 2007. GSTM1 and APE1 genotypes affect arsenic-induced oxidative stress: A repeated measures study. Environ. Heal. 6, 39. https://doi.org/10.1186/1476-069X-6-39

Broberg, K., Ahmed, S., Engström, K., Hossain, M.B., Jurkovic Mlakar, S., Bottai, M., Grandér, M., Raqib, R., et al., 2014. Arsenic exposure in early pregnancy alters genome-wide DNA methylation in cord blood, particularly in boys. J. Dev. Orig. Health Dis. 5, 288–298. https://doi.org/10.1017/S2040174414000221

Buchet, J.P., Lauwerys, R., 1988. Role of thiols in the in-vitro methylation of inorganic arsenic by rat liver cytosol. Biochem. Pharmacol. 37, 3149–3153. https://doi.org/10.1016/0006-2952(88)90313-9

Buchet, J.P., Lauwerys, R., Roels, H., 1981. Urinary Excretion of Inorganic Arsenic and Its Metabolites After Repeated Ingestion of Sodium Metaarsenite by Volunteers. Int. Arch. Int Arch Occup Env. Heal. 48, 111–118.

Bundschuh, J., Litter, M.I., Parvez, F., Román-Ross, G., Nicolli, H.B., Jean, J.S., Liu, C.W., López, D., et al., 2012. One century of arsenic exposure in Latin America: A review of history and occurrence from 14 countries. Sci. Total Environ. 429, 2–35. https://doi.org/10.1016/j.scitotenv.2011.06.024

Burgess, J.L., Meza, M.M., Josyula, A.B., Poplin, G.S., Kopplin, M.J., McClellen, H.E., Stürup, S., Clark Lantz, R., 2007. Environmental Arsenic Exposure and Urinary 8-OHdG in Arizona and Sonora. Clin. Toxicol. 45, 490–498. https://doi.org/10.1080/15563650701354119

Cascio, C., Raab, A., Jenkins, R.O., Feldmann, J., Meharg, A.A., Haris, P.I., 2011. The impact of a rice based diet on urinary arsenic. J. Environ. Monit. 13, 257–265. https://doi.org/10.1039/c0em00482k

Cawthon, R.M., 2002. Telomere measurement by quantitative PCR. Nucleic Acids Res. 30, e47. https://doi.org/10.1093/nar/30.10.e47

Chao, D.Y., Chen, Y., Chen, J., Shi, S., Chen, Z., Wang, C., Danku, J.M., Zhao, F.J., et al., 2014. Genome-wide Association Mapping Identifies a New Arsenate Reductase Enzyme Critical for Limiting Arsenic Accumulation in Plants. PLoS Biol. 12, e1002009. https://doi.org/10.1371/journal.pbio.1002009

Chen, G.Q., Shi, X.G., Tang, W., Xiong, S.M., Zhu, J., Cai, X., Han, Z.G., Ni, J.H., et al., 1997. Use of Arsenic Trioxide (As2O3) in the Treatment of Acute Promyelocytic Leukemia (APL): I. As2O3 Exerts Dose-Dependent Dual Effects on APL Cells. Blood 89, 3345–3353.

Chen, M., Ma, L.Y., Chen, X.W., 2014. New procedures for arsenic speciation: A review. Talanta 125, 78–86. https://doi.org/10.1016/j.talanta.2014.02.037

Chen, Q.Y., Desmarais, T., Costa, M., 2019. Metals and Mechanisms of Carcinogenesis. Annu. Rev. Pharmacol. Toxicol. 6, 537–554. https://doi.org/10.1146/annurev-pharmtox-010818-021031

Chen, S.J., Zhou, G.B., Zhang, X.W., Mao, J.H., De Thé, H., Chen, Z., 2011. From an old remedy to a magic bullet: molecular mechanisms underlying the therapeutic effects of arsenic in fighting leukemia. Blood 117, 6425–6437. https://doi.org/10.1182/blood-2010-11-283598

Chen, X., Guo, X., He, P., Nie, J., Yan, X., Zhu, J., Zhang, L., Mao, G., et al., 2017. Interactive

Influence of N6AMT1 and As3MT Genetic Variations on Arsenic Metabolism in the Population of Inner Mongolia, China. Toxicol. Sci. 155, 124–134. https://doi.org/10.1093/toxsci/kfw181

Chen, Y.C., Su, H.J., Guo, Y.L., Hsueh, Y.M., Smith, T.J., Ryan, L.M., Lee, M.S., Christiani, D.C., 2003. Arsenic methylation and bladder cancer risk in Taiwan. Cancer Causes Control 14, 303–310. https://doi.org/10.1023/A:1023905900171

Chernoff, M., Tong, L., Demanelis, K., Ahsan, H., Vander Griend, D., Pierce, B.L., 2020. Genetic determinants of reduced arsenic metabolism efficiency in the 10q24.32 region are associated with reduced AS3MT expression in multiple human tissue types. Toxicol. Sci. 176, 382–395. https://doi.org/10.1093/toxsci/kfaa075

Childebayeva, A., Goodrich, J.M., Leon-Velarde, F., Rivera-Chira, M., Kiyamu, M., Brutsaert, T.D., Dolinoy, D.C., Bigham, A.W., 2021. Genome-Wide Epigenetic Signatures of Adaptive Developmental Plasticity in the Andes. Genome Biol. Evol. 13, evaa239. https://doi.org/10.1093/gbe/evaa239

Chou, W.C., Chung, Y.T., Chen, H.Y., Wang, C.J., Ying, T.H., Chuang, C.Y., Tseng, Y.C., Wang, S.L., 2014. Maternal arsenic exposure and DNA damage biomarkers, and the associations with birth outcomes in a general population from Taiwan. PLoS One 9, e86398. https://doi.org/10.1371/journal.pone.0086398

Clark, A.G., Hubisz, M.J., Bustamante, C.D., Williamson, S.H., Nielsen, R., 2005. Ascertainment bias in studies of human genome-wide polymorphism. Genome Res. 15, 1496–1502. https://doi.org/10.1101/gr.4107905

Cohen, S.M., Chowdury, A., Arnold, L.L., 2016. Inorganic arsenic: A non-genotoxic carcinogen. J. Environ. Sci. 49, 28–37. https://doi.org/10.1016/j.jes.2016.04.015

Concha, G., Broberg, K., Grandér, M., Cardozo, A., Palm, B., Vahter, M., 2010. High-Level Exposure to Lithium, Boron, Cesium, and Arsenic via Drinking Water in the Andes of Northern Argentina. Environ. Sci. Technol. 44, 6875–6880. https://doi.org/10.1021/es1010384

Concha, G., Nermell, B., Vahter, M., 2006. Spatial and temporal variations in arsenic exposure via drinking-water in Northern Argentina. J. Heal. Popul. Nutr. 24, 317–326.

Concha, G., Nermell, B., Vahter, M., 1998a. Metabolism of inorganic arsenic in children with chronic high arsenic exposure in northern Argentina. Environ. Health Perspect. 106, 355–359. https://doi.org/10.1289/ehp.98106355

Concha, G., Vogler, G., Lezcano, D., Nermell, B., Vahter, M., 1998b. Exposure to inorganic arsenic metabolites during early human development. Toxicol. Sci. 44, 185–190. https://doi.org/10.1006/toxs.1998.2486

Coop, G., Pickrell, J.K., Novembre, J., Kudaravalli, S., Li, J., Absher, D., Myers, R.M., Cavalli-Sforza, L.L., et al., 2009. The Role of Geography in Human Adaptation. PLoS Genet. 5, e1000500. https://doi.org/10.1371/journal.pgen.1000500

Corella, A., Bert, F., Pérez-Pérez, A., Gené, M., Turbón, D., 2007. Mitochondrial DNA diversity of the Amerindian populations living in the Andean Piedmont of Bolivia: Chimane, Moseten, Aymara and Quechua. Ann. Hum. Biol. 34, 34–55. https://doi.org/10.1080/03014460601075819

Dalleau, S., Baradat, M., Guéraud, F., Huc, L., 2013. Cell death and diseases related to oxidative stress:4-hydroxynonenal (HNE) in the balance. Cell Death Differ. 20, 1615–1630. https://doi.org/10.1038/cdd.2013.138

Dauphiné, D.C., Ferreccio, C., Guntur, S., Yuan, Y., Hammond, S.K., Balmes, J., Smith, A.H., Steinmaus, C., 2011. Lung function in adults following in utero and childhood exposure to arsenic in drinking water: Preliminary findings. Int. Arch. Occup. Environ. Health 84, 591–600. https://doi.org/10.1007/s00420-010-0591-6

De Chaudhuri, S., Ghosh, P., Sarma, N., Majumdar, P., Sau, T.J., Basu, S., Roychoudhury, S., Ray, K., et al., 2008. Genetic variants associated with arsenic susceptibility: Study of purine nucleoside phosphorylase, arsenic (+3) methyltransferase, and glutathione S-transferase omega genes. Environ. Health Perspect. 116, 501–505. https://doi.org/10.1289/ehp.10581

de la Barra Saavedra, S.Z., Lara Barrientos, G.M., Coca Cru, R.O., 2011. Exclusión y subalternidad de los urus del lago Poopó Discriminación en la relación mayorías y minorías étnicas. Serie Informes de Investigación PIEB, La Paz.

de la Rosa, R., Steinmaus, C., Akers, N.K., Conde, L., Ferreccio, C., Kalman, D., Zhang, K.R., Skibola, C.F., et al., 2017. Associations between arsenic (+3 oxidation state) methyltransferase (AS3MT) and N-6 adenine-specific DNA methyltransferase 1 (N6AMT1) polymorphisms, arsenic metabolism, and cancer risk in a chilean population. Environ. Mol. Mutagen. 58, 411–422. https://doi.org/10.1002/em.22104

Delgado, D.A., Chernoff, M., Huang, L., Tong, L., Chen, L., Jasmine, F., Shinkle, J., Cole, S.A., et al., 2021. Rare, Protein-Altering Variants in AS3MT and Arsenic Metabolism Efficiency: A Multi-Population Association Study. Environ. Health Perspect. 129, 047007. https://doi.org/10.1289/EHP8152

Delnomdedieu, M., Basti, M.M., Styblo, M., Otvos, J.D., Thomas, D.J., 1994. Complexation of arsenic species in rabbit erythrocytes. Chem. Res. Toxicol. 7, 621–627. https://doi.org/10.1021/tx00041a006

DiGiovanni, A., Demanelis, K., Tong, L., Argos, M., Shinkle, J., Jasmine, F., Sabarinathan, M., Rakibuz-Zaman, M., et al., 2020. Assessing the impact of arsenic metabolism efficiency on DNA methylation using Mendelian randomization. Environ. Epidemiol. 4, e083. https://doi.org/10.1097/EE9.0000000000000083

Doudoroff, M., Barker, H.A., Hassid, W.Z., 1947. Studies with bacterial sucrose phosphorylase. III. Arsenolytic decomposition of sucrose and of glucose-1-phosphate. J. Biol. Chem. 170, 147–150.

Drobná, Z., Del Razo, L.M., García-Vargas, G.G., Sánchez-Peña, L.C., Barrera-Hernández, A., Stýblo, M., Loomis, D., 2013. Environmental exposure to arsenic, AS3MT polymorphism and prevalence of diabetes in Mexico. J. Expo. Sci. Environ. Epidemiol. 23, 151–155. https://doi.org/10.1038/jes.2012.103

Drobná, Z., Naranmandura, H., Kubachka, K.M., Edwards, B.C., Herbin-Davis, K., Styblo, M., Le, X.C., Creed, J.T., et al., 2009. Disruption of the arsenic (+3 oxidation state) methyltransferase gene in the mouse alters the phenotype for methylation of arsenic and affects distribution and retention of orally administered arsenate. Chem. Res. Toxicol. 22, 1713–20. https://doi.org/10.1021/tx900179r

Drobná, Z., Waters, S.B., Devesa, V., Harmon, A.W., Thomas, D.J., Stýblo, M., 2005. Metabolism and toxicity of arsenic in human urothelial cells expressing rat arsenic (+3 oxidation state)-methyltransferase. Toxicol. Appl. Pharmacol. 207, 147–159. https://doi.org/10.1016/j.taap.2004.12.007

Dulout, F.N., Grillo, C.A., Seoane, A.I., Maderna, C.R., Nilsson, R., Vahter, M., Darroudi, F., Natarajan, A.T., 1996. Chromosomal aberrations in peripheral blood lymphocytes from native Andean women and children from northwestern Argentina exposed to arsenic in drinking water. Mutat. Res. 370, 151–158. https://doi.org/10.1016/S0165-1218(96)00060-2

Echeverría, J., Niemeyer, H.M., Muñoz, L., Uribe, M., 2018. Arsenic in the hair of mummies from agro-ceramic times of Northern Chile (500 BCE–1200 CE). J. Archaeol. Sci. Reports 21, 175–182. https://doi.org/10.1016/J.JASREP.2018.07.008

Eichstaedt, C., Antao, T., Cardona, A., Pagani, L., Kivisild, T., Mormina, M., 2015. Positive selection of AS3MT to arsenic water in Andean populations. Mutat Res 780, 97–102. https://doi.org/10.1016/j.mrfmmm.2015.07.007

Engström, K., Vahter, M., Mlakar, S.J., Concha, G., Nermell, B., Raqib, R., Cardozo, A., Broberg, K., 2011. Polymorphisms in Arsenic (+III Oxidation State) Methyltransferase (AS3MT) Predict Gene Expression of AS3MT as Well as Arsenic Metabolism. Environ. Health Perspect. 119, 182–188. https://doi.org/10.1289/ehp.1002471

Engström, K., Wojdacz, T.K., Marabita, F., Ewels, P., Käller, M., Vezzi, F., Prezza, N., Gruselius, J., et al., 2017. Transcriptomics and methylomics of CD4-positive T cells in arsenic-exposed women. Arch. Toxicol. 91, 2067–2078. https://doi.org/10.1007/s00204-016-1879-4

Engström, K.S., Broberg, K., Concha, G., Nermell, B., Warholm, M., Vahter, M., 2007. Genetic polymorphisms influencing arsenic metabolism: Evidence from Argentina. Environ. Health Perspect. 115, 599–605. https://doi.org/10.1289/ehp.9734

Engström, K.S., Hossain, M.B., Lauss, M., Ahmed, S., Raqib, R., Vahter, M., Broberg, K., 2013. Efficient Arsenic Metabolism — The AS3MT Haplotype Is Associated with DNA Methylation and Expression of Multiple Genes Around AS3MT. PLoS One 8, e53732. https://doi.org/10.1371/journal.pone.0053732

Engström, K.S., Vahter, M., Johansson, G., Lindh, C.H., Teichert, F., Singh, R., Kippler, M., Nermell, B., et al., 2010a. Chronic exposure to cadmium and arsenic strongly influences concentrations of 8-oxo-7,8-dihydro-2′-deoxyguanosine in urine. Free Radic. Biol. Med. 48, 1211–1217. https://doi.org/10.1016/j.freeradbiomed.2010.02.004

Engström, K.S., Vahter, M., Lindh, C., Teichert, F., Singh, R., Concha, G., Nermell, B., Farmer, P.B., et al., 2010b. Low 8-oxo-7,8-dihydro-2′-deoxyguanosine levels and influence of genetic background in an Andean population exposed to high levels of arsenic. Mutat. Res. 683, 98–105. https://doi.org/10.1016/j.mrfmmm.2009.10.014

Espinoza-Navarro, A., Diaz, J., Rodrígues, H., Moreno, A., 2011. Effects of Altitude on Anthropometric and Physiological Patterns in Aymara and Non-Aymara Population Between 18 and 65 Years in the Province of Parinacota Chile (3.700 masl). Int. J. Morphol 29, 34–40. https://doi.org/10.4067/S0717-95022011000100005

Ewens, W.J., 2010. Population Genetics: Overview. eLS. John Wiley Sons Ltd, Chichester. https://doi.org/10.1002/9780470015902.a0001737.pub2

Fadista, J., Manning, A.K., Florez, J.C., Groop, L., 2016. The (in)famous GWAS P-value threshold revisited and updated for low-frequency variants. Eur. J. Hum. Genet. 24, 1202–1205. https://doi.org/10.1038/ejhg.2015.269

Francesconi, K.A., Kuehnelt, D., 2004. Determination of arsenic species: A critical review of methods and applications, 2000-2003. Analyst 129, 373–395. https://doi.org/10.1039/b401321m

Francesconi, K.A., Tanggaar, R., McKenzie, C.J., Goessler, W., 2002. Arsenic metabolites in human urine after ingestion of an arsenosugar. Clin. Chem. 48, 92–101.

Francisca, F.M., Carro Pérez, M.E., 2009. Assessment of natural arsenic in groundwater in Cordoba Province, Argentina. Environ. Geochem. Health 31, 673–682. https://doi.org/10.1007/s10653-008-9245-y

Fu, S., Wu, J., Li, Y., Liu, Y., Gao, Y., Yao, F., Qiu, C., Song, L., et al., 2014. Urinary arsenic metabolism in a Western Chinese population exposed to high-dose inorganic arsenic in drinking water: Influence of ethnicity and genetic polymorphisms. Toxicol. Appl. Pharmacol. 274, 117–123. https://doi.org/10.1016/J.TAAP.2013.11.004

Fujino, Y., Guo, X., Liu, J., Matthews, I.P., Shirane, K., Wu, K., Kasai, H., Miyatake, M., et al., 2005. Chronic arsenic exposure and urinary 8-Hydroxy-2 0-deoxyguanosine in an arsenic-affected area in Inner Mongolia, China. J. Expo. Anal. Environ. Epidemiol. 15, 147–152. https://doi.org/10.1038/sj.jea.7500381

Gamble, M. V, Liu, X., Slavkovich, V., Pilsner, J.R., Ilievski, V., Factor-Litvak, P., Levy, D., Alam, S., et al., 2007. Folic acid supplementation lowers blood arsenic. Am. J. Clin. Nutr. 86, 1202–1209. https://doi.org/10.1093/ajcn/86.4.1202

Gao, J., Roy, S., Tong, L., Argos, M., Jasmine, F., Rahaman, R., Rakibuz-Zaman, M., Parvez, F., et al., 2015a. Arsenic exposure, telomere length, and expression of telomere-related genes among Bangladeshi individuals. Environ. Res. 136, 462–469. https://doi.org/10.1016/j.envres.2014.09.040

Gao, J., Tong, L., Argos, M., Bryan, M.S., Ahmed, A., Rakibuz-Zaman, M., Kibriya, M.G., Jasmine, F., et al., 2015b. The Genetic Architecture of Arsenic Metabolism Efficiency:A SNP-Based Heritability Study of Bangladeshi Adults. Environ. Health Perspect. 123, 985–992. https://doi.org/10.1289/ehp.1408909

Garbinski, L.D., Rosen, B.P., Chen, J., 2019. Pathways of arsenic uptake and efflux. Environ. Int. 126, 585–597. https://doi.org/10.1016/j.envint.2019.02.058

Garcia Moreno, M.E., 2006. Transport of arsenic and heavy metals to Lake Poopo, Bolivia: Natural leakage and anthropogenic effects. Water Resour. Eng. Lund Univ.

Gardner, R., Nermell, B., Kippler, M., Grandér, M., Li, L., Ekström, E.C., Rahman, A., Lönnerdal, B., et al., 2011. Arsenic methylation efficiency increases during the first trimester of pregnancy independent of folate status. Reprod. Toxicol. 31, 210–218. https://doi.org/10.1016/j.reprotox.2010.11.002

Gardner, R.M., Kippler, M., Tofail, F., Bottai, M., Hamadani, J., Grandér, M., Nermell, B., Palm, B., et al., 2013. Environmental Exposure to Metals and Children's Growth to Age 5 Years: A Prospective Cohort Study. Am. J. Epidemiol. 177, 1356–1367. https://doi.org/10.1093/aje/kws437

Gautier, M., Vitalis, R., 2012. Rehh An R package to detect footprints of selection in genome-wide SNP data from haplotype structure. Bioinformatics 28, 1176–1177. https://doi.org/10.1093/bioinformatics/bts115

Ghosh, P., Banerjee, M., Giri, A.K., Ray, K., 2008. Toxicogenomics of arsenic: Classical ideas and recent advances. Mutat. Res. 659, 293–301. https://doi.org/10.1016/J.MRREV.2008.06.003

Goix, S., Point, D., Oliva, P., Polve, M., Duprey, J.L., Mazurek, H., Guislain, L., Huayta, C., et al., 2011. Influence of source distribution and geochemical composition of aerosols on children exposure in the large polymetallic mining region of the Bolivian Altiplano. Sci. Total Environ. 412–413, 170–184. https://doi.org/10.1016/j.scitotenv.2011.09.065

Goix, S., Uzu, G., Oliva, P., Barraza, F., Calas, A., Castet, S., Point, D., Masbou, J., et al., 2016. Metal concentration and bioaccessibility in different particle sizes of dust and aerosols to refine metal exposure assessment. J. Hazard. Mater. 317, 552–562. https://doi.org/10.1016/j.jhazmat.2016.05.083

Gómez-Carballa, A., Pardo-Seco, J., Brandini, S., Achilli, A., Perego, U.A., Coble, M.D., Diegoli, T.M., Álvarez-Iglesias, V., et al., 2018. The peopling of South America and the trans-Andean gene flow of the first settlers. Genome Res. 28, 767–779. https://doi.org/10.1101/gr.234674.118

Gomez-Rubio, P., Meza-Montenegro, M.M., Cantu-Soto, E., Klimecki, W.T., 2010. Genetic association between intronic variants in AS3MT and arsenic methylation efficiency is focused on a large linkage disequilibrium cluster in chromosome 10. J. Appl. Toxicol. 30, 260–270. https://doi.org/10.1002/jat.1492

Gomez-Rubio, P., Roberge, J., Arendell, L., Harris, R.B., O'Rourke, M.K., Chen, Z., Cantu-Soto, E., Meza-Montenegro, M.M., et al., 2011. Association between body mass index and arsenic methylation efficiency in adult women from southwest U.S. and northwest Mexico. Toxicol. Appl. Pharmacol. 252, 176–182. https://doi.org/10.1016/j.taap.2011.02.007

Gribble, M.O., Voruganti, V.S., Cole, S.A., Haack, K., Balakrishnan, P., Laston, S.L., Tellez-Plaza, M., Francesconi, K.A., et al., 2015. Linkage Analysis of Urine Arsenic Species Patterns in the Strong Heart Family Study. Toxicol. Sci. 148, 89–100. https://doi.org/10.1093/toxsci/kfv164

Guha Mazumder, D.N., 2003. Chronic Arsenic Toxicity: Clinical Features, Epidemiology, and Treatment: Experience in West Bengal. J. Environ. Sci. Heal. 38, 141–163. https://doi.org/10.1081/ESE-120016886

Gunderson, K.L., Steemers, F.J., Lee, G., Mendoza, L.G., Chee, M.S., 2005. A genome-wide scalable SNP genotyping assay using microarray technology. Nat. Genet. 37, 549–554. https://doi.org/10.1038/ng1547

Hall, M., Chen, Y., Ahsan, H., Slavkovich, V., van Geen, A., Parvez, F., Graziano, J., 2006. Blood arsenic as a biomarker of arsenic exposure: Results from a prospective study. Toxicology 225, 225–233. https://doi.org/10.1016/j.tox.2006.06.010

Hall, M.N., Liu, X., Slavkovich, V., Ilievski, V., Pilsner, J.R., Alam, S., Factor-Litvak, P., Graziano, J.H., et al., 2009. Folate, Cobalamin, Cysteine, Homocysteine, and Arsenic Metabolism among Children in Bangladesh. Environ. Health Perspect. 117, 825–831. https://doi.org/10.1289/ehp.0800164

Hamadani, J., Tofail, F., Nermell, B., Gardner, R., Shiraji, S., Bottai, M., Arifeen, S., Huda, S., et al., 2011. Critical windows of exposure for arsenic-associated impairment of cognitive function in pre-school girls and boys: a population-based cohort study. Int. J. Epidemiol. 40, 1593–1604. https://doi.org/10.1093/ije/dyr176

Harari, F., Engström, K., Concha, G., Colque, G., Vahter, M., Broberg, K., 2013. N-6-adenine-specific DNA methyltransferase 1 (N6AMT1) polymorphisms and arsenic methylation in Andean women. Environ. Health Perspect. 121, 797–803. https://doi.org/10.1289/ehp.1206003

Harari, F., Langeén, M., Casimiro, E., Bottai, M., Palm, B., Nordqvist, H., Vahter, M., 2015. Environmental exposure to lithium during pregnancy and fetal size: A longitudinal study in the Argentinean Andes. Environ. Int. 77, 48–54. https://doi.org/10.1016/j.envint.2015.01.011

Harris, E.E., 2013. Gene Evolution and Human Adaptation. eLS. John Wiley Sons, Ltd Chichester. https://doi.org/10.1002/9780470015902.a0020773.pub2

He, Y., Zou, L., Luo, W., Yi, Z., Yang, P., Yu, S., Liu, N., Ji, J., et al., 2020. Heavy metal exposure, oxidative stress and semen quality: Exploring associations and mediation effects in reproductive-aged men. Chemosphere 244, 125498. https://doi.org/10.1016/j.chemosphere.2019.125498

Herlin, M., Broberg, K., Malin Igra, A., Li, H., Harari, F., Vahter, M., 2019. Exploring telomere length in mother-newborn pairs in relation to exposure to multiple toxic metals and potential modifying effects by nutritional factors. BMC Med. 17, 77. https://doi.org/10.1186/s12916-019-1309-6

Hernández, A., Marcos, R., 2008. Genetic variations associated with interindividual sensitivity in the response to arsenic exposure. Pharmacogenomics 9, 1113–1132. https://doi.org/10.2217/14622416.9.8.1113

Hirano, S., Kobayashi, Y., Cui, X., Kanno, S., Hayakawa, T., Shraim, A., 2004. The accumulation and toxicity of methylated arsenicals in endothelial cells: important roles of thiol compounds. Toxicol. Appl. Pharmacol. 198, 458–467. https://doi.org/10.1016/j.taap.2003.10.023

Hopenhayn-Rich, C., Biggs, M. Lou, Smith, A.H., Kalman, D.A., Moore, L.E., 1996. Methylation study of a population environmentally exposed to arsenic in drinking water. Environ. Health Perspect. 104, 620–628. https://doi.org/10.1289/ehp.96104620

Hopenhayn-Rich, C., Browning, S.R., Hertz-Picciotto, I., Ferreccio, C., Peralta, C., Gibb, H., 2000. Chronic arsenic exposure and risk of infant mortality in two areas of Chile. Environ. Health

Perspect. 108, 667–673. https://doi.org/10.1289/ehp.00108667

Hsu, L.I., Chen, G.S., Lee, C.H., Yang, T.Y., Chen, Y.H., Wang, Y.H., Hsueh, Y.M., Chiou, H.Y., et al., 2013. Use of Arsenic-Induced Palmoplantar Hyperkeratosis and Skin Cancers to Predict Risk of Subsequent Internal Malignancy. Am. J. Epidemiol. 177, 202–212. https://doi.org/10.1093/aje/kws369

Huang, C.F., Chen, Y.W., Yang, C.Y., Tsai, K.S., Yang, R. Sen, Liu, S.H., 2011. Arsenic and diabetes: Current perspectives. Kaohsiung J. Med. Sci. 27, 402–410. https://doi.org/10.1016/j.kjms.2011.05.008

Hubaux, R., Becker-Santos, D.D., Enfield, K.S., Rowbotham, D., Lam, S., Lam, W.L., Martinez, V.D., 2013. Molecular features in arsenic-induced lung tumors. Mol. Cancer 12. https://doi.org/10.1186/1476-4598-12-20

Hughes, M.F., 2006. Biomarkers of exposure: a case study with inorganic arsenic. Environ. Health Perspect. 114, 1790–1796. https://doi.org/10.1289/ehp.9058

Hughes, M.F., 2002. Arsenic toxicity and potential mechanisms of action. Toxicol. Lett. 133, 1–16. https://doi.org/10.1016/S0378-4274(02)00084-X

Hwang, S.G., Chapagain, S., Lee, J.W., Han, A.R., Jang, C.S., 2017. Genome-wide transcriptome profiling of genes associated with arsenate toxicity in an arsenic-tolerant rice mutant. Plant Physiol. Biochem. 120, 40–51. https://doi.org/10.1016/j.plaphy.2017.09.019

Jacobson, T., Navarrete, C., Sharma, S.K., Sideri, T.C., Ibstedt, S., Priya, S., Grant, C.M., Christen, P., et al., 2012. Arsenite interferes with protein folding and triggers formation of protein aggregates in yeast. J. Cell Sci. 125, 5073–5083. https://doi.org/10.1242/jcs.107029

Jacovas, V.C., Couto-Silva, C.M., Nunes, K., Lemes, R.B., de Oliveira, M.Z., Salzano, F.M., Bortolini, M.C., Hünemeier, T., 2018. Selection scan reveals three new loci related to high altitude adaptation in Native Andeans. Sci. Rep. 8, 12733. https://doi.org/10.1038/s41598-018-31100-6

Jafri, M.A., Ansari, S.A., Alqahtani, M.H., Shay, J.W., 2016. Roles of telomeres and telomerase in cancer, and advances in telomerase-targeted therapies. Genome Med. 8. https://doi.org/10.1186/s13073-016-0324-x

Jimenez Villarreal, J., Murillo Ortiz, B., Martinez Garza, S., Rivas Armendáriz, D.I., Boone Villa, V.D., Carranza Rosales, P., Betancourt Martínez, N.D., Delgado Aguirre, H., et al., 2019. Telomere length analysis in residents of a community exposed to arsenic. J. Biochem. Mol. Toxicol. 33, e22230. https://doi.org/10.1002/jbt.22230

Jones, M.R., Tellez-Plaza, M., Vaidya, D., Grau-Perez, M., Post, W.S., Kaufman, J.D., Guallar, E., Francesconi, K.A., et al., 2019. Ethnic, geographic and dietary differences in arsenic exposure in the multi-ethnic study of atherosclerosis (MESA). J. Expo. Sci. Environ. Epidemiol. 29, 310–322. https://doi.org/10.1038/s41370-018-0042-0

Kadam, C.Y., Abhang, S.A., 2016. Apoptosis Markers in Breast Cancer Therapy. Adv. Clin. Chem. 74, 143–193. https://doi.org/10.1016/bs.acc.2015.12.003

Kakoulli, I., Prikhodko, S. V., Fischer, C., Cilluffo, M., Uribe, M., Bechtel, H.A., Fakra, S.C., Marcus, M.A., 2014. Distribution and Chemical Speciation of Arsenic in Ancient Human Hair Using Synchrotron Radiation. Anal. Chem. 86, 521–526. https://doi.org/10.1021/ac4024439

Kaler, A.S., Purcell, L.C., 2019. Estimation of a significance threshold for genome-wide association studies. BMC Genomics 20, 618. https://doi.org/10.1186/s12864-019-5992-7

Kaushal, A., Zhang, H., Karmaus, W.J.J., Everson, T.M., Marsit, C.J., Karagas, M.R., Tsai, S.F., Wen, H.J., et al., 2017. Genome-wide DNA methylation at birth in relation to in utero arsenic

exposure and the associated health in later life. Environ. Heal. 16, 50. https://doi.org/10.1186/s12940-017-0262-0

Khairul, I., Wang, Q.Q., Jiang, Y.H., Wang, C., Naranmandura, H., 2017. Metabolism, toxicity and anticancer activities of arsenic compounds. Oncotarget 8, 23905–23926. https://doi.org/10.18632/oncotarget.14733

Khan, K.M., Chakraborty, R., Bundschuh, J., Bhattacharya, P., Parvez, F., 2020. Health effects of arsenic exposure in Latin America: An overview of the past eight years of research. Sci. Total Environ. 710, 136071. https://doi.org/10.1016/j.scitotenv.2019.136071

Kibriya, M.G., Jasmine, F., Parvez, F., Argos, M., Roy, S., Paul-Brutus, R., Islam, T., Ahmed, A., et al., 2017. Association between genome-wide copy number variation and arsenic-induced skin lesions: a prospective study. Environ. Health 16, 75. https://doi.org/10.1186/s12940-017-0283-8

Kimura, M., 1968. Evolutionary rate at the molecular level. Nature 217, 624–626. https://doi.org/10.1038/217624a0

Kimura, R., Ohashi, J., 2013. Identifying Regions of the Human Genome that Exhibit Evidence of Positive Selection. eLS. John Wiley Sons, Ltd Chichester. https://doi.org/10.1002/9780470015902.a0020850.pub2

Koller, B.H., Snouwaert, J.N., Douillet, C., Jania, L.A., El-Masri, H., Thomas, D.J., Stýblo, M., 2020. Arsenic Metabolism in Mice Carrying a BORCS7/AS3MT Locus Humanized by Syntenic Replacement. Environ. Health Perspect. 128, 087003. https://doi.org/10.1289/EHP6943

Kordas, K., Queirolo, E.I., Mañay, N., Peregalli, F., Hsiao, P.Y., Lu, Y., Vahter, M., 2016. Low-level arsenic exposure: nutritional and dietary predictors in first-grade Uruguayan children. Environ. Res. 147, 16–23. https://doi.org/10.1016/j.envres.2016.01.022

Kozłowska, L., Janasik, B., Nowicka, K., Wąsowicz, W., 2019. A urinary metabolomics study of a Polish subpopulation environmentally exposed to arsenic. J. Trace Elem. Med. Biol. 54, 44–54. https://doi.org/10.1016/j.jtemb.2019.03.009

Kubota, R., Kunito, T., Agusa, T., Fujihara, J., Monirith, I., Iwata, H., Subramanian, A., Tana, T.S., et al., 2006. Urinary 8-hydroxy-2′-deoxyguanosine in inhabitants chronically exposed to arsenic in groundwater in Cambodia. J. Environ. Monit. 8, 293–299. https://doi.org/10.1039/b513652k

Kumarathilaka, P., Seneweera, S., Ok, Y.S., Meharg, A., Bundschuh, J., 2019. Arsenic in cooked rice foods: Assessing health risks and mitigation options. Environ. Int. 127, 584–591. https://doi.org/10.1016/j.envint.2019.04.004

Kuo, C.C., Howard, B. V, Umans, J.G., Francesconi, K.A., Goessler, W., Lee, E., Guallar, E., Navas-Acien, A., 2015. Arsenic Exposure, Arsenic Metabolism, and Incident Diabetes in the Strong Heart Study. Diabetes Care 38, 620–627. https://doi.org/10.2337/dc14-1641

Lawson, D.J., van Dorp, L., Falush, D., 2018. A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. Nat. Commun. 9, 1–11. https://doi.org/10.1038/s41467-018-05257-7

Lee, H.C., Yin, P.H., Lu, C.Y., Chi, C.W., Wei, Y.H., 2000. Increase of mitochondria and mitochondrial DNA in response to oxidative stress in human cells. Biochem. J 348, 425–432.

Levi, M., Hjelm, C., Harari, F., Vahter, M., 2018. ICP-MS measurement of toxic and essential elements in human breast milk. A comparison of alkali dilution and acid digestion sample preparation methods. Clin. Biochem. 53, 81–87. https://doi.org/10.1016/j.clinbiochem.2017.12.003

Li, L., Ekström, E.C., Goessler, W., Lönnerdal, B., Nermell, B., Yunus, M., Rahman, A., El Arifeen, S., et al., 2008. Nutritional status has marginal influence on the metabolism of inorganic arsenic

in pregnant Bangladeshi women. Environ. Health Perspect. 116, 315–321. https://doi.org/10.1289/ehp.10639

Lin, S., Shi, Q., Nix, F.B., Styblo, M., Beck, M.A., Herbin-Davis, K.M., Hall, L.L., Simeonsson, J.B., et al., 2002. A Novel S-Adenosyl-L-methionine:Arsenic(III) Methyltransferase from Rat Liver Cytosol. J. Biol. Chem. 277, 10795–10803. https://doi.org/10.1074/jbc.M110246200

Lindberg, A.L., Ekström, E.C., Nermell, B., Rahman, M., Lönnerdal, B., Persson, L.Å., Vahter, M., 2008a. Gender and age differences in the metabolism of inorganic arsenic in a highly exposed population in Bangladesh. Environ. Res. 106, 110–120. https://doi.org/10.1016/j.envres.2007.08.011

Lindberg, A.L., Rahman, M., Persson, L.Å., Vahter, M., 2008b. The risk of arsenic induced skin lesions in Bangladeshi men and women is affected by arsenic metabolism and the age at first exposure. Toxicol. Appl. Pharmacol. 230, 9–16. https://doi.org/10.1016/j.taap.2008.02.001

Lindberg, A.L., Sohel, N., Rahman, M., Persson, L.Å., Vahter, M., 2010. Impact of Smoking and Chewing Tobacco on Arsenic-Induced Skin Lesions. Environ. Health Perspect. 118, 533–538. https://doi.org/10.1289/ehp.0900728

Lindo, J., Haas, R., Hofman, C., Apata, M., Moraga, M., Verdugo, R.A., Watson, J.T., Llave, C.V., et al., 2018. The genetic prehistory of the Andean highlands 7000 years BP though European contact. Sci. Adv. 4, eaau4921. https://doi.org/10.1126/sciadv.aau4921

Loft, S., Møller, P., 2006. Oxidative DNA damage and human cancer: Need for cohort studies. Antioxidants Redox Signal. 8, 1021–1031. https://doi.org/10.1089/ars.2006.8.1021

Lorenzo Bermejo, J., Boekstegers, F., González Silos, R., Marcelain, K., Baez Benavides, P., Barahona Ponce, C., Müller, B., Ferreccio, C., et al., 2017. Subtypes of Native American ancestry and leading causes of death: Mapuche ancestry-specific associations with gallbladder cancer risk in Chile. PLoS Genet. 13, e1006756. https://doi.org/10.1371/journal.pgen.1006756

Macdonald, B., Johns, T., Gray-Donald, K., Receveur, O., 2004. Ecuadorian Andean women's nutrition varies with age and socioeconomic status. Food Nutr. Bull. 25, 239–247. https://doi.org/10.1177/156482650402500303

Mar Wai, K., Umezaki, M., Mar, O., Umemura, M., Watanabe, C., 2019. Arsenic exposure through drinking Water and oxidative stress Status: A cross-sectional study in the Ayeyarwady region, Myanmar. J. Trace Elem. Med. Biol. 54, 103–109. https://doi.org/10.1016/j.jtemb.2019.04.009

Marafante, E., Vahter, M., 1984. The effect of methyltransferase inhibition on the metabolism of [74As]arsenite in mice and rabbits. Chem. Biol. Interact. 50, 49–57. https://doi.org/10.1016/0009-2797(84)90131-5

Marafante, E., Vahter, M., Norin, H., Envall, J., Sandström, M., Christakopoulos, A., Ryhage, R., 1987. Biotransformation of dimethylarsinic acid in mouse, hamster and man. J. Appl. Toxicol. 7, 111–117. https://doi.org/10.1002/jat.2550070207

Marchiset-Ferlay, N., Savanovitch, C., Sauvant-Rochat, M.P., 2012. What is the best biomarker to assess arsenic exposure via drinking water? Environ. Int. 39, 150–171. https://doi.org/10.1016/j.envint.2011.07.015

Marcus, J.H., Novembre, J., 2016. Visualizing the geography of genetic variants. Bioinformatics 33, btw643. https://doi.org/10.1093/bioinformatics/btw643

Mass, M.J., Tennant, A., Roop, B.C., Cullen, W.R., Styblo, M., Thomas, D.J., Kligerman, A.D., 2001. Methylated Trivalent Arsenic Species Are Genotoxic. Chem. Res. Toxicol. 14, 355–361. https://doi.org/10.1021/tx0002511

Matoušek, T., Wang, Z., Douillet, C., Musil, S., Stýblo, M., 2017. Direct Speciation Analysis of

Arsenic in Whole Blood and Blood Plasma at Low Exposure Levels by Hydride Generation-Cryotrapping-Inductively Coupled Plasma Mass Spectrometry. Anal. Chem. 89, 9633–9637. https://doi.org/10.1021/acs.analchem.7b01868

McNally, E.J., Luncsford, P.J., Armanios, M., 2019. Long telomeres and cancer risk: The price of cellular immortality. J. Clin. Invest. 129, 3474–3481. https://doi.org/10.1172/JCI120851

Meharg, A.A., Williams, P.N., Adomako, E., Lawgali, Y.Y., Deacon, C., Villada, A., Cambell, R.C.J., Sun, G., et al., 2009. Geographical Variation in Total and Inorganic Arsenic Content of Polished (White) Rice. Environ. Sci. Technol. 43, 1612–1617. https://doi.org/10.1021/es802612a

Meharg, A.A., Williams, P.N., Deacon, C.M., Norton, G.J., Hossain, M., Louhing, D., Marwa, E., Lawgalwi, Y., et al., 2014. Urinary excretion of arsenic following rice consumption. Environ. Pollut. 194, 181–187. https://doi.org/DOI: 10.1016/j.envpol.2014.07.031

Mendez, M.A., González-Horta, C., Sánchez-Ramírez, B., Ballinas-Casarrubias, L., Cerón, R.H., Morales, D.V., Baeza Terrazas, F.A., Ishida, M.C., et al., 2016. Chronic exposure to arsenic and markers of cardiometabolic risk: A cross-sectional study in Chihuahua, Mexico. Environ. Health Perspect. 124, 104–111. https://doi.org/10.1289/ehp.1408742

Middleton, D.R.S., Watts, M.J., Lark, R.M., Milne, C.J., Polya, D.A., 2016. Assessing urinary flow rate, creatinine, osmolality and other hydration adjustment methods for urinary biomonitoring using NHANES arsenic, iodine, lead and cadmium data. Environ. Heal. 15, 68. https://doi.org/10.1186/s12940-016-0152-x

Middleton, D.R.S., Watts, M.J., Polya, D.A., 2019. A comparative assessment of dilution correction methods for spot urinary analyte concentrations in a UK population exposed to arsenic in drinking water. Environ. Int. 130, 104721. https://doi.org/10.1016/j.envint.2019.03.069

Mitton, J.B., 2002. Adaptation and Natural Selection: Overview. eLS. John Wiley Sons Ltd, Chichester. https://doi.org/10.1038/npg.els.0001706

Molin, M., Ulven, S.M., Meltzer, H.M., Alexander, J., 2015. Arsenic in the human food chain, biotransformation and toxicology – Review focusing on seafood arsenic. J. Trace Elem. Med. Biol. 31, 249–259. https://doi.org/10.1016/j.jtemb.2015.01.010

Muller, P.Y., Dieterle, F., 2009. Tissue-specific, non-invasive toxicity biomakers: Translation from preclinical safety assessment to clinical safety monitoring. Expert Opin. Drug Metab. Toxicol. 5, 1023–1038. https://doi.org/10.1517/17425250903114174

Németi, B., Regonesi, M.E., Tortora, P., Gregus, Z., 2010. Polynucleotide Phosphorylase and Mitochondrial ATP Synthase Mediate Reduction of Arsenate to the More Toxic Arsenite by Forming Arsenylated Analogues of ADP and ATP. Toxicol. Sci. 117, 270–281. https://doi.org/10.1093/toxsci/kfq141

Nermell, B., Lindberg, A.L., Rahman, M., Berglund, M., Persson, L.Å., El Arifeen, S., Vahter, M., 2008. Urinary arsenic concentration adjustment factors and malnutrition. Environ. Res. 106, 212–218. https://doi.org/10.1016/j.envres.2007.08.005

Nicolli, H.B., Bundschuh, J., García, J.W., Falcón, C.M., Jean, J.S., 2010. Sources and controls for the mobility of arsenic in oxidizing groundwaters from loess-type sediments in arid/semi-arid dry climates - Evidence from the Chaco-Pampean plain (Argentina). Water Res. 44, 5589–5604. https://doi.org/10.1016/j.watres.2010.09.029

Nizam, S., Kato, M., Yatsuya, H., Khalequzzaman, M., Ohnuma, S., Naito, H., Nakajima, T., 2013. Differences in urinary arsenic metabolites between diabetic and non-diabetic subjects in Bangladesh. Int. J. Environ. Res. Public Health 10, 1006–1019. https://doi.org/10.3390/ijerph10031006

O'Callaghan, P., Meleady, R., Fitzgerald, T., Graham, I., European COMAC group, 2002. Smoking

and plasma homocysteine. Eur. Heart J. 23, 1580–1586. https://doi.org/10.1053/euhj.2002.3172

Olmos, V., Navoni, J., Calcagno, M., Sassone, A., Villaamil Lepori, E., 2015. Influence of the level of arsenic (As) exposure and the presence of T860C polymorphism in human As urinary metabolic profile. Hum. Exp. Toxicol. 34, 170–178. https://doi.org/10.1177/0960327114533574

Ormachea Muñoz, M., Garcia Aróstegui, J.L., Bhattacharya, P., Sracek, O., Garcia Moreno, M.E., Kohfahl, C., Quintanilla Aguirre, J., Hornero Diaz, J., et al., 2016. Geochemistry of naturally occurring arsenic in groundwater and surface-water in the southern part of the Poopó Lake basin, Bolivian Altiplano. Groundw. Sustain. Dev. 2–3, 104–116. https://doi.org/10.1016/j.gsd.2016.04.002

Ormachea Muñoz, M., Wern, H., Johnsson, F., Bhattacharya, P., Sracek, O., Thunvik, R., Quintanilla, J., Bundschuh, J., 2013. Geogenic arsenic and other trace elements in the shallow hydrogeologic system of Southern Poopó Basin, Bolivian Altiplano. J. Hazard. Mater. 262, 924–940. https://doi.org/10.1016/J.JHAZMAT.2013.06.078

Palmgren, M., Engström, K., Hallström, B.M., Wahlberg, K., Søndergaard, D.A., Sall, T., Vahter, M., Broberg, K., 2017. AS3MT-mediated tolerance to arsenic evolved by multiple independent horizontal gene transfers from bacteria to eukaryotes. PLoS One 12, e0175422. https://doi.org/10.1371/journal.pone.0175422

Parikh, C.R., Gyamlani, G.G., Carvounis, C.P., 2002. Screening for microalbuminuria simplified by urine specific gravity. Am. J. Nephrol. 22, 315–319. https://doi.org/10.1159/000065220

Pavilonis, B., Grassman, J., Johnson, G., Diaz, Y., Caravanos, J., 2017. Characterization and risk of exposure to elements from artisanal gold mining operations in the Bolivian Andes. Environ. Res. 154, 1–9. https://doi.org/10.1016/j.envres.2016.12.010

Petrick, J.S., Ayala-Fierro, F., Cullen, W.R., Carter, D.E., Vasken Aposhian, H., 2000. Monomethylarsonous Acid (MMA(III)) Is More Toxic Than Arsenite in Chang Human Hepatocytes. Toxicol. Appl. Pharmacol. 163, 203–207. https://doi.org/10.1006/taap.1999.8872

Pierce, B.L., Kibriya, M.G., Tong, L., Jasmine, F., Argos, M., Roy, S., Paul-Brutus, R., Rahaman, R., et al., 2012. Genome-wide association study identifies chromosome 10q24.32 variants associated with arsenic metabolism and toxicity phenotypes in Bangladesh. PLoS Genet. 8, e1002522. https://doi.org/10.1371/journal.pgen.1002522

Pierce, B.L., Tong, L., Argos, M., Gao, J., Farzana, J., Roy, S., Paul-Brutus, R., Rahaman, R., et al., 2013. Arsenic metabolism efficiency has a causal role in arsenic toxicity: Mendelian randomization and gene-environment interaction. Int. J. Epidemiol. 42, 1862–1871. https://doi.org/10.1093/IJE/DYT182

Pierce, B.L., Tong, L., Dean, S., Argos, M., Jasmine, F., Rakibuz-Zaman, M., Sarwar, G., Shahriarid, H., et al., 2019. A missense variant in FTCD is associated with arsenic metabolism and toxicity phenotypes in Bangladesh. PLoS Genet. 15, e1007984. https://doi.org/10.1371/journal.pgen.1007984

Quaghebeur, W., Mulhern, R.E., Ronsse, S., Heylen, S., Blommaert, H., Potemans, S., Valdivia Mendizábal, C., Terrazas García, J., 2019. Arsenic contamination in rainwater harvesting tanks around Lake Poopó in Oruro, Bolivia: An unrecognized health risk. Sci. Total Environ. 688, 224–230. https://doi.org/10.1016/j.scitotenv.2019.06.126

Rahman, A., Persson, L.Å., Nermell, B., El Arifeen, S., Ekström, E.C., Smith, A.H., Vahter, M., 2010. Arsenic exposure and risk of spontaneous abortion, stillbirth, and infant mortality. Epidemiology 21, 797–804. https://doi.org/10.1097/EDE.0b013e3181f56a0d

Raml, R., Goessler, W., Traar, P., Ochi, T., Francesconi, K.A., 2005. Novel thioarsenic metabolites in human urine after ingestion of an arsenosugar, 2′,3′-dihydroxypropyl 5-deoxy-5-dimethylarsinoyl- β-D-riboside. Chem. Res. Toxicol. 18, 1444–1450.

https://doi.org/10.1021/tx050111h

Raml, R., Rumpler, A., Goessler, W., Vahter, M., Li, L., Ochi, T., Francesconi, K.A., 2007. Thio-dimethylarsinate is a common metabolite in urine samples from arsenic-exposed women in Bangladesh. Toxicol. Appl. Pharmacol. 222, 374–380. https://doi.org/10.1016/j.taap.2006.12.014

Ramos Ramos, O.E., Cáceres, L.F., Ormachea Muñoz, M.R., Bhattacharya, P., Quino, I., Quintanilla, J., Sracek, O., Thunvik, R., et al., 2012. Sources and behavior of arsenic and trace elements in groundwater and surface water in the Poopó Lake Basin, Bolivian Altiplano. Env. Earth Sci 66, 793–807. https://doi.org/10.1007/s12665-011-1288-1

Raqib, R., Ahmed, S., Ahsan, K.B., Kippler, M., Akhtar, E., Roy, A.K., Lu, Y., El Arifeen, S., et al., 2017. Humoral Immunity in Arsenic-Exposed Children in Rural Bangladesh: Total Immunoglobulins and Vaccine-Specific Antibodies. Environ. Health Perspect. 125, 067006. https://doi.org/10.1289/EHP318

Rivera, M.A., Aufderheide, A.C., Cartmell, L.W., Torres, C.M., Langsjoen, O., 2005. Antiquity of Coca-Leaf Chewing in the South Central Andes: A 3,000 Year Archaeological Record of Coca-Leaf Chewing from Northern Chile. J. Psychoactive Drugs 37, 455–458. https://doi.org/10.1080/02791072.2005.10399820

Roggenbeck, B.A., Banerjee, M., Leslie, E.M., 2016. Cellular arsenic transport pathways in mammals. J. Environ. Sci. 49, 38–58. https://doi.org/10.1016/j.jes.2016.10.001

Roh, T., Steinmaus, C., Marshall, G., Ferreccio, C., Liaw, J., Smith, A.H., 2018. Age at Exposure to Arsenic in Water and Mortality 30–40 Years After Exposure Cessation. Am. J. Epidemiol. 187, 2297–2305. https://doi.org/10.1093/aje/kwy159

Rupert, J.L., Hochachka, P.W., 2001. Genetic approaches to understanding human adaptation to altitude in the Andes. J. Exp. Biol. 204, 3151–60. https://doi.org/10.1242/jeb.204.18.3151

Sabeti, P.C., Reich, D.E., Higgins, J.M., Levine, H.Z.P., Richter, D.J., Schaffner, S.F., Gabriel, S.B., Platko, J. V., et al., 2002. Detecting recent positive selection in the human genome from haplotype structure. Nature 419, 832–837. https://doi.org/10.1038/nature01140

Sabeti, P.C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X., Byrne, E.H., et al., 2007. Genome-wide detection and characterization of positive selection in human populations. Nature 449, 913–918. https://doi.org/10.1038/nature06250

Saha, K.K., Engström, A., Hamadani, J.D., Tofail, F., Rasmussen, K.M., Vahter, M., 2012. Pre- and Postnatal Arsenic Exposure and Body Size to 2 Years of Age: A Cohort Study in Rural Bangladesh. Environ. Health Perspect. 120, 1208–1214. https://doi.org/10.1289/ehp.1003378

Sandoval, J.R., Lacerda, D.R., Jota, M.S.A., Salazar-Granara, A., Paulo, P., Vieira, R., Acosta, O., Cuellar, C., et al., 2013. The Genetic History of Indigenous Populations of the Peruvian and Bolivian Altiplano: The Legacy of the Uros. PLoS One 8, e73006. https://doi.org/10.1371/journal.pone.0073006

Sanyal, T., Bhattacharjee, P., Bhattacharjee, S., Bhattacharjee, P., 2018. Hypomethylation of mitochondrial D-loop and ND6 with increased mitochondrial DNA copy number in the arsenic-exposed population. Toxicology 408, 54–61. https://doi.org/10.1016/j.tox.2018.06.012

Satgé, F., Espinoza, R., Zolá, R.P., Roig, H., Timouk, F., Molina, J., Garnier, J., Calmant, S., et al., 2017. Role of climate variability and human activity on Poopó Lake droughts between 1990 and 2015 assessed using remote sensing data. Remote Sens. 9, 218. https://doi.org/10.3390/rs9030218

Schläwicke Engström, K., Nermell, B., Concha, G., Strömberg, U., Vahter, M., Broberg, K., 2009. Arsenic metabolism is influenced by polymorphisms in genes involved in one-carbon

metabolism and reduction reactions. Mutat. Res. 667, 4–14. https://doi.org/10.1016/j.mrfmmm.2008.07.003

Schlebusch, C.M., Gattepaille, L.M., Engström, K., Vahter, M., Jakobsson, M., Broberg, K., 2015. Human adaptation to arsenic-rich environments. Mol. Biol. Evol. 32, 1544–1555. https://doi.org/10.1093/molbev/msv046

Schlebusch, C.M., Lewis, C.M., Vahter, M., Engström, K., Tito, R.Y., Obregón-Tito, A.J., Huerta, D., Polo, S.I., et al., 2013. Possible positive selection for an arsenic-protective haplotype in humans. Environ. Health Perspect. 121, 53–58. https://doi.org/10.1289/ehp.1205504

Shao, W., Fanelli, M., Ferrara, F.F., Riccioni, R., Rosenauer, A., Davison, K., Lamph, W.W., Waxman, S., et al., 1998. Arsenic Trioxide as an Inducer of Apoptosis and Loss of PML/RARα Protein in Acute Promyelocytic Leukemia Cells. J. Natl. Cancer Inst. 90, 124–133. https://doi.org/10.1093/jnci/90.2.124

Shih, Y.H., Islam, T., Hore, S.K., Sarwar, G., Shahriar, M.H., Yunus, M., Graziano, J.H., Harjes, J., et al., 2017. Associations between prenatal arsenic exposure with adverse pregnancy outcome and child mortality. Environ. Res. 158, 456–461. https://doi.org/10.1016/j.envres.2017.07.004

Shriver, M.D., Kennedy, G.C., Parra, E.J., Lawson, H.A., Sonpar, V., Huang, J., Akey, J.M., Jones, K.W., 2004. The genomic distribution of population substructure in four populations using 8,525 autosomal SNPs. Hum. Genomics 1, 274–286. https://doi.org/10.1186/1479-7364-1-4-274

Signes-Pastor, A.J., Gutiérrez-González, E., García-Villarino, M., Rodríguez-Cabrera, F.D., López-Moreno, J.J., Varea-Jiménez, E., Pastor-Barriuso, R., Pollán, M., et al., 2021. Toenails as a biomarker of exposure to arsenic: A review. Environ. Res. 195, 110286. https://doi.org/10.1016/j.envres.2020.110286

Simonson, T.S., Huff, C.D., Witherspoon, D.J., Prchal, J.T., Jorde, L.B., 2015. Adaptive genetic changes related to haemoglobin concentration in native high-altitude Tibetans. Exp. Physiol. 100, 1263–1268. https://doi.org/10.1113/EP085035

Skröder Löveborn, H., Kippler, M., Lu, Y., Ahmed, S., Kuehnelt, D., Raqib, R., Vahter, M., 2016. Arsenic Metabolism in Children Differs From That in Adults. Toxicol. Sci. 152, 29–39. https://doi.org/10.1093/toxsci/kfw060

Smith, A.H., Marshall, G., Liaw, J., Yuan, Y., Ferreccio, C., Steinmaus, C., 2012. Mortality in young adults following in utero and childhood exposure to arsenic in drinking water. Environ. Health Perspect. 120, 1527–1531. https://doi.org/10.1289/ehp.1104867

Smith, A.H., Marshall, G., Roh, T., Ferreccio, C., Liaw, J., Steinmaus, C., 2018. Lung, Bladder, and Kidney Cancer Mortality 40 Years After Arsenic Exposure Reduction. J. Natl. Cancer Inst. 110, 241–249. https://doi.org/doi: 10.1093/jnci/djx201

Smith, A.H., Yunus, M., Khan, A.F., Ercumen, A., Yuan, Y., Smith, M.H., Liaw, J., Balmes, J., et al., 2013. Chronic respiratory symptoms in children following in utero and early life exposure to arsenic in drinking water in Bangladesh. Int. J. Epidemiol. 42, 1077–1086. https://doi.org/10.1093/ije/dyt120

Somnath, P., Pritha, B., Ashok K., G., Pritha, B., 2017. Arsenic toxicity and epimutagenecity: the new LINEage. Biometals 30, 505–515. https://doi.org/10.1007/s10534-017-0021-2

Stassen, M.J.M., Preeker, N.L., Ragas, A.M.J., Van De Ven, M.W.P.M., Smolders, A.J.P., Roeleveld, N., 2012. Metal exposure and reproductive disorders in indigenous communities living along the Pilcomayo River, Bolivia. Sci. Total Environ. 427–428, 26–34. https://doi.org/10.1016/j.scitotenv.2012.03.072

Steinmaus, C., Bates, M.N., Yuan, Y., Kalman, D., Atallah, R., Rey, O.A., Biggs, M.L., Hopenhayn, C., et al., 2006. Arsenic Methylation and Bladder Cancer Risk in Case-Control Studies in

Argentina and the United States. J. Occup. Environ. Med. 48, 478–488. https://doi.org/10.1097/01.jom.0000200982.28276.70

Steinmaus, C., Carrigan, K., Kalman, D., Atallah, R., Yuan, Y., Smith, A.H., 2005. Dietary intake and arsenic methylation in a U.S. population. Environ. Health Perspect. 113, 1153–1159. https://doi.org/10.1289/EHP.7907

Styblo, M., Delnomdedieu, M., Thomas, D.J., 1996. Mono- and dimethylation of arsenic in rat liver cytosol in vitro. Chem. Biol. Interact. 99, 147–164. https://doi.org/10.1016/0009-2797(95)03666-0

Sun, G., Xu, Y., Li, X., Jin, Y., Li, B., Sun, X., 2007. Urinary arsenic metabolites in children and adults exposed to arsenic in drinking water in Inner Mongolia, China. Environ. Health Perspect. 115, 648–652. https://doi.org/10.1289/ehp.9271

Takayama, Y., Masuzaki, Y., Mizutani, F., Iwata, T., Maeda, E., Tsukada, M., Nomura, K., Ito, Y., et al., 2021. Associations between blood arsenic and urinary arsenic species concentrations as an exposure characterization tool. Sci. Total Environ. 750, 141517. https://doi.org/10.1016/j.scitotenv.2020.141517

Tamayo, L.I., Lin, H., Ahmed, A., Shahriar, H., Hasan, R., Sarwar, G., Eunus, H.M., Ahsan, H., et al., 2020. Research Participants' Attitudes towards Receiving Information on Genetic Susceptibility to Arsenic Toxicity in Rural Bangladesh. Public Health Genomics 23, 69–75. https://doi.org/10.1159/000505632

Taylor, V., Goodale, B., Raab, A., Schwerdtle, T., Reimer, K., Conklin, S., Karagas, M.R., Francesconi, K.A., 2017. Human exposure to organic arsenic species from seafood. Sci. Total Environ. 580, 266–282. https://doi.org/10.1016/j.scitotenv.2016.12.113

Tellez-Plaza, M., Gribble, M.O., Voruganti, V.S., Francesconi, K.A., Goessler, W., Umans, J.G., Silbergeld, E.K., Guallar, E., et al., 2013. Heritability and preliminary genome-wide linkage analysis of arsenic metabolites in urine. Environ. Health Perspect. 121, 345–51. https://doi.org/10.1289/ehp.1205305

Thomas, D.J., Li, J., Waters, S.B., Xing, W., Adair, B.M., Drobna, Z., Devesa, V., Styblo, M., 2007. Arsenic (+3 Oxidation State) Methyltransferase and the Methylation of Arsenicals. Exp Biol Med 232, 3–13.

Toosi, B., Zaker, F., Alikarami, F., Kazemi, A., Teremmahi Ardestanii, M., 2018. VS-5584 as a PI3K/mTOR inhibitor enhances apoptotic effects of subtoxic dose arsenic trioxide via inhibition of NF-κB activity in B cell precursor-acute lymphoblastic leukemia. Biomed. Pharmacother. 102, 428–437. https://doi.org/10.1016/j.biopha.2018.03.009

Tseng, C.H., 2009. A review on environmental factors regulating arsenic methylation in humans. Toxicol. Appl. Pharmacol. 235, 338–350. https://doi.org/10.1016/j.taap.2008.12.016

Tseng, C.H., Huang, Y.K., Huang, Y.L., Chung, C.J., Yang, M.H., Chen, C.J., Hsueh, Y.M., 2005. Arsenic exposure, urinary arsenic speciation, and peripheral vascular disease in blackfoot disease-hyperendemic villages in Taiwan. Toxicol. Appl. Pharmacol. 206, 299–308. https://doi.org/10.1016/j.taap.2004.11.022

Vahter, M., 2009. Effects of Arsenic on Maternal and Fetal Health. Annu. Rev. Nutr. 29, 381–399. https://doi.org/10.1146/annurev-nutr-080508-141102

Vahter, M., 2000. Genetic polymorphism in the biotransformation of inorganic arsenic and its role in toxicity. Toxicol. Lett. 112–113, 209–217. https://doi.org/10.1016/S0378-4274(99)00271-4

Vahter, M., 1999. Methylation of inorganic arsenic in different mammalian species and population groups. Sci. Prog. 82 (Pt 1), 69–88. https://doi.org/10.1177/003685049908200104

Vahter, M., Åkesson, A., Lidén, C., Ceccatelli, S., Berglund, M., 2007. Gender differences in the disposition and toxicity of metals. Environ. Res. 104, 85–95. https://doi.org/10.1016/j.envres.2006.08.003

Vahter, M., Concha, G., 2001. Role of metabolism in arsenic toxicity. Pharmacol. Toxicol. 89, 1–5. https://doi.org/10.1111/j.1600-0773.2001.890101.x

Vahter, M., Concha, G., Nermell, B., Nilsson, R., Dulout, F., Natarajan, A.T., 1995. A unique metabolism of inorganic arsenic in native Andean women. Eur. J. Pharmacol. 293, 455–462. https://doi.org/https://doi.org/10.1016/0926-6917(95)90066-7

Vahter, M., Envall, J., 1983. In vivo reduction of arsenate in mice and rabbits. Environ. Res. 32, 14–24. https://doi.org/10.1016/0013-9351(83)90187-1

Vahter, M., Marafante, E., 1987. Effects of low dietary intake of methionine, choline or proteins on the biotransformation of arsenite in the rabbit. Toxicol. Lett. 37, 41–46. https://doi.org/10.1016/0378-4274(87)90165-2

Vahter, M., Norin, H., 1980. Metabolism of 74 As-labeled trivalent and pentavalent inorganic arsenic in mice. Environ. Res. 21, 446–457. https://doi.org/10.1016/0013-9351(80)90049-3

Valverde, G., Zhou, H., Lippold, S., de Filippo, C., Tang, K., López Herráez, D., Li, J., Stoneking, M., 2015. A novel candidate region for genetic adaptation to high altitude in Andean populations. PLoS One 10, e0125444. https://doi.org/10.1371/journal.pone.0125444

Van Den Bergh, K., Du Laing, G., Montoya, J.C., De Deckere, E., Tack, F.M.G., 2010. Arsenic in drinking water wells on the Bolivian high plain: Field monitoring and effect of salinity on removal efficiency of iron-oxides- containing filters. J. Environ. Sci. Heal. - Part A Toxic/Hazardous Subst. Environ. Eng. 45, 1741–1749. https://doi.org/10.1080/10934529.2010.513262

Villa-Bellosta, R., Sorribas, V., 2008. Role of rat sodium/phosphate cotransporters in the cell membrane transport of arsenate. Toxicol. Appl. Pharmacol. 232, 125–134. https://doi.org/10.1016/j.taap.2008.05.026

Voight, B.F., Kudaravalli, S., Wen, X., Pritchard, J.K., 2006. A Map of Recent Positive Selection in the Human Genome. PLoS Biol. 4, e72. https://doi.org/10.1371/journal.pbio.0040072

Wallin, M., Sallsten, G., Lundh, T., Barregard, L., 2014. Low-level cadmium exposure and effects on kidney function. Occup. Environ. Med. 71, 848–854. https://doi.org/10.1136/oemed-2014-102279

Wang, C.H., Hsiao, C.K., Chen, C.L., Hsu, L.I., Chiou, H.Y., Chen, S.Y., Hsueh, Y.M., Wu, M.M., et al., 2007. A review of the epidemiologic literature on the role of environmental arsenic exposure and cardiovascular diseases. Toxicol. Appl. Pharmacol. 222, 315–326. https://doi.org/10.1016/j.taap.2006.12.022

Wang, Q.Q., Lan, Y.F., Rehman, K., Jiang, Y.H., Maimaitiyiming, Y., Zhu, D.Y., Naranmandura, H., 2015. Effect of Arsenic Compounds on the in Vitro Differentiation of Mouse Embryonic Stem Cells into Cardiomyocytes. Chem. Res. Toxicol. 28, 351–353. https://doi.org/10.1021/tx500286t

Winski, S.L., Carter, D.E., 1998. Arsenate toxicity in human erythrocytes: characterization of morphologic changes and determination of the mechanism of damage. J. Toxicol. Environ. Health 53, 345–55. https://doi.org/10.1080/009841098159213

World Health Organization (WHO). 2017. Guidelines for drinking-water quality: fourth edition incorporating first addendum, 4th ed + 1st add. https://apps.who.int.iris/handle/10665/254637

Xing, J., Chen, M., Wood, C.G., Lin, J., Spitz, M.R., Ma, J., Amos, C.I., Shields, P.G., et al., 2008. Mitochondrial DNA content: Its genetic heritability and association with renal cell carcinoma. J.

Natl. Cancer Inst. 100, 1104–1112. https://doi.org/10.1093/jnci/djn213

Yassine, H., Kimzey, M.J., Galligan, M.A., Gandolfi, A.J., Stump, C.S., Lau, S.S., 2011. Adjusting for Urinary Creatinine Overestimates Arsenic Concentrations in Diabetics. Cardiorenal Med. 2, 26–32. https://doi.org/10.1159/000334225

Yeh, H.C., Lin, Y.S., Kuo, C.C., Weidemann, D., Weaver, V., Fadrowski, J., Neu, A., Navas-Acien, A., 2015. Urine osmolality in the US population: Implications for environmental biomonitoring. Environ. Res. 136, 482–490. https://doi.org/10.1016/j.envres.2014.09.009

Zakharyan, R.A., Sampayo-Reyes, A., Healy, S.M., Tsaprailis, G., Board, P.G., Liebler, D.C., Aposhian, H. V., 2001. Human monomethylarsonic acid (MMA(V)) reductase is a member of the glutathione-S-transferase superfamily. Chem. Res. Toxicol. 14, 1051–1057. https://doi.org/10.1021/tx010052h

Zeng, Y., Weng, G., Fan, J., Li, Z., Wu, J., Li, Y., Zheng, R., Xia, P., et al., 2016. Curcumin reduces the expression of survivin, leading to enhancement of arsenic trioxide-induced apoptosis in myelodysplastic syndrome and leukemia stem-like cells. Oncol. Rep. 36, 1233–1242. https://doi.org/10.3892/or.2016.4944

Zhao, F.J., Zhu, Y.G., Meharg, A.A., 2013. Methylated arsenic species in rice: Geographical variation, origin, and uptake mechanisms. Environ. Sci. Technol. 47, 3957–3966. https://doi.org/10.1021/es304295n

Zhu, J., Okumura, H., Ohtake, S., Nakamura, S., Nakao, S., 2003. The Molecular Mechanism of Arsenic Trioxide-Induced Apoptosis and Oncosis in Leukemia/Lymphoma Cell Lines. Acta Haematol 110, 1–10. https://doi.org/10.1159/000072407