Arsenic Exposure and Cancer-Related Proteins in Urine of Indigenous Bolivian Women

Jessica De Loma[1]*, Anda R. Gliga[1], Michael Levi[1], Franz Ascui[2], Jacques Gardon[3], Noemi Tirado[4] and Karin Broberg[1]

[1] Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden, [2] Programa de Salud Familiar Comunitaria e Intercultural, Ministerio de Salud Bolivia, La Paz, Bolivia, [3] Hydrosciences Montpellier, Université de Montpellier, Institut de Recherche pour le Développement, Centre National de la Recherche Scientifique, Montpellier, France, [4] Genetics Institute, Universidad Mayor de San Andrés, La Paz, Bolivia

Indigenous people living in the Bolivian Andes are exposed through their drinking water to inorganic arsenic, a potent carcinogen. However, the health consequences of arsenic exposure in this region are unknown. The aim of this study was to evaluate associations between arsenic exposure and changes in cancer-related proteins in indigenous women ($n = 176$) from communities around the Andean Lake Poopó, Bolivia. Arsenic exposure was assessed in whole blood (B-As) and urine (as the sum of arsenic metabolites, U-As) by inductively coupled plasma-mass spectrometry (ICP-MS). Cancer-related proteins ($N = 92$) were measured in urine using the proximity extension assay. The median B-As concentration was 2.1 (range 0.60–9.1) ng/g, and U-As concentration was 67 (12–399) μg/L. Using linear regression models adjusted for age, urinary osmolality, and urinary leukocytes, we identified associations between B-As and four putative cancer-related proteins: FASLG, SEZ6L, LYPD3, and TFPI2. Increasing B-As concentrations were associated with lower protein expression of SEZ6L, LYPD3, and TFPI2, and with higher expression of FASLG in urine (no association was statistically significant after correcting for multiple comparisons). The associations were similar across groups with different arsenic metabolism efficiency, a susceptibility factor for arsenic toxicity. In conclusion, arsenic exposure in this region was associated with changes in the expression of some cancer-related proteins in urine. Future research is warranted to understand if these proteins could serve as valid biomarkers for arsenic-related toxicity.

Keywords: arsenic, Andes, Bolivia, FASLG, TFPI2, biomarker, carcinogenic

INTRODUCTION

Inorganic arsenic (iAs) is classified as a class I human carcinogen (1). Chronic exposure to iAs has been associated with multiple types of cancer, including cancers of the skin, lung, bladder, kidney, liver, and prostate (1–4). Exposure to iAs has also been associated with non-carcinogenic health effects such as cardiovascular disease, diabetes, and immunotoxicity (4–6).

Leakage of iAs from volcanic bedrocks into groundwater used as drinking water is one of the main contributors to human exposure, especially in Latin America where more than 14 countries present elevated iAs in drinking water (7). Arsenic-related cancer has been evaluated in some Latin American countries, mainly Chile, Argentina, and Mexico (6). In the Chilean city of Antofagasta

and in the Córdoba province of Argentina, studies have consistently found an increased lung, bladder, and kidney cancer mortality in relation to chronic exposure to iAs (8–10). Recently, iAs was linked to laryngeal cancer mortality in Chile (11), and chronic exposure to iAs and risk for breast cancer was evaluated in Latin America (12–14). In the Bolivian Andes, we have shown that women living around Lake Poopó are exposed to iAs (15). However, very little is known about cancer or any toxicity in relation to iAs exposure in Bolivia.

A growing body of knowledge supports the use of urinary proteins to detect early signs of cancer and other diseases (16–19). Although several epidemiological studies have investigated protein biomarkers associated with iAs in humans (20–22), the relation between urinary proteins and arsenic remains unexplored. To our knowledge, only one study has previously explored urinary proteins in relation to iAs exposure. In that study, limited to proteins within 2–10 kDa of size, highly exposed individuals (>100 µg/L U-As) from the United States and Chile had decreased expression of human beta-defensin 1 (HBD-1), an antimicrobial protein and potential tumor suppressor gene in urological cancers (23). The effect of iAs on urinary proteins has been overlooked in the scientific literature despite its potential for biomarker discovery. Therefore, the aim of this cross-sectional study was to evaluate the associations between putative cancer-related proteins in urine and arsenic exposure in indigenous individuals from the Bolivian Andes.

MATERIALS AND METHODS
Study Participants and Sample Collection
We recruited women living in villages around Lake Poopó, located in the southern region of the Bolivian highlands, at 3,686 m above sea level (15). This area is characterized by arsenic-rich soils, common in the Andean regions. Recruitment of 201 women took place during five field trips organized between September 2015 and November 2017 from a total of 10 villages around Lake Poopó. Women with no major health conditions were recruited on a voluntary basis (one woman with a brain tumor and tuberculosis was excluded). We recruited women from two ethnic groups, Aymara-Quechua and Uru, that inhabit this region. During recruitment, the women had similar eating habits and lifestyles in all villages, based on questionnaire data. Men were not recruited into the study because they usually worked away from their village leading to a different arsenic exposure. For this study, we started from a subset of women ($n = 182$) for which we had most complete biological data, including genetic data which is not presented in this study. Out of these women, we excluded five that presented extreme levels of nitrites in urine (based on urine strip determination described further on), and one woman without urine strip data. In total, 176 women were included in the current study, of which 149 identified themselves as Aymara-Quechua and 27 as Uru. This study was approved by the Comité Nacional de Bioética (Bolivia) and the Regional Ethic Committee of Karolinska Institutet (Sweden). Prior to recruitment and sampling, participants were informed orally and in written form about the project, and signed informed consent was obtained.

We carried out personal interviews including information such as age, ethnicity, weight, height, blood pressure (measured in sitting position), health status, and frequency of chewing coca leaves (15). Body mass index was not used since Andean populations are known to have higher values as a consequence of living at high altitude without them being associated with excess body fat (15, 24, 25). Spot urine samples were collected throughout the day during the field trips. The women were given instructions on how to collect urine samples, including wet wipe cleaning and mid-stream urine collection to minimize contamination. The urine samples were collected in 20 mL polyethylene bottles, previously confirmed as free of trace elements. Venous blood samples were obtained with BD Vacutainer Eclipse blood collection needles (Becton Dickinson, USA) in Trace Elements NH Sodium Heparin tubes (Vacuette, Greiner Bio, Austria), or in Lithium Heparin tubes (Vacuette) when the Sodium Heparin tubes were not available. Due to field trip limitations, it was not possible to separate plasma. Urine and blood samples were stored directly after sampling at −18°C in a portable freezer (ARB, Australia), and stored at Universidad Mayor de San Andrés (La Paz, Bolivia) at −20°C until further shipment. Samples were transported on dry ice to Karolinska Institutet (Stockholm, Sweden), where they were stored at −20°C for long-term storage. All urine and blood samples had similar freeze-thaw cycles.

Protein Measurements in Urine
Urine samples for protein measurements were thawed overnight at 4°C. After thoroughly mixing each sample, it was aliquoted, debris were separated by centrifugation (845 g for 10 min), and 50 µL of the supernatant was added to 96-well plates (Thermo Fisher Scientific, USA) covered with MicroAmp Clear Adhesive Film (Thermo Fisher Scientific). Samples were transported on ice from Karolinska Institutet (Stockholm, Sweden) to Olink Proteomics (Uppsala, Sweden).

Urine samples were analyzed with the Proseek Multiplex Oncology II panel ($N = 92$ proteins) based on the Proximity Extension Assay (PEA) developed and performed by Olink Proteomics. The PEA is a dual-recognition immunoassay which uses a pair of antibodies labeled with DNA oligonucleotides specific for each pre-determined protein on the panel. When both labeled antibodies bind to the target protein, the oligonucleotide labels are in enough proximity to hybridize and create a PCR target sequence that will be quantified by real-time PCR. The Oncology II panel includes proteins belonging to biological process ontologies relevant to cancer development: angiogenesis ($n = 20$), apoptotic process ($n = 34$), cell adhesion ($n = 35$), cell differentiation ($n = 42$), cell motility ($n = 30$), cell proliferation ($n = 43$), cellular metabolic process ($n = 47$), cellular response to stress ($n = 23$), chemotaxis ($n = 14$), extracellular matrix organization ($n = 9$), immune response ($n = 27$), MAPK cascade ($n = 25$), proteolysis ($n = 19$), response to hypoxia ($n = 3$), and other gene ontology terms ($n = 9$). Methodological details, data processing, quality control and normalization are described by Assarsson et al. (26) and are available online at https://www.olink.com. Protein level data obtained by the Olink Proteomic facility are presented as Normalized Protein eXpression (NPX)

values, which have arbitrary units on a log2-scale. The intra-assay coefficient of variance (CV%) was below 20% for all proteins (only one protein above 15%), while the inter-assay CV% was below 30% for all proteins (only two proteins between 20 and 30%). The limit of detection (LOD) for each protein assay is defined as three times the standard deviation above background level based on internal controls. For downstream analyses, we included proteins that had more than 40% of the samples above the LOD, i.e., 45 proteins. Since this multiplex assay was initially developed to detect proteins in plasma, we expected that some of the selected proteins in the panel would not be present to the same extent in urine. Therefore, we were less stringent with the LOD cut-off compared to other studies that included proteins measured in plasma with more than 80–90% of the samples above LOD (27, 28).

Arsenic Exposure and Metabolism Efficiency Assessment

For this study, we used arsenic concentrations in whole blood (B-As), and not in urine, as a biomarker of exposure in order to avoid the potential co-excretion of arsenic and proteins in urine (29).

We previously described that the women in this region had elevated concentrations of arsenic and lithium in urine, and that arsenic and lithium concentrations in urine were correlated ($r_S = 0.47$, p-value < 0.001) (15). Therefore, in this study we also used lithium concentrations in whole blood (B-Li; correlation with B-As: $r_S = 0.44$, p-value < 0.001) to evaluate potential confounding.

The blood samples were prepared for inductively coupled plasma-mass spectrometry (ICP-MS; operating conditions in **Supplementary Table 1**) by a direct alkali dilution method (30). Briefly, blood samples were diluted 1:17–44 with an alkali solution consisting of 2% butanol (Honeywell Research Chemicals, Germany), 0.05% EDTA (Sigma-Aldrich, USA), 0.05% Triton X-100 (Sigma-Aldrich), 1% NH_4OH (Romil, UK), and 20 µg/g internal standards [45]Sc, [72]Ge, [103]Rh, [175]Lu, and [193]Ir. Before analysis, the diluted samples were sonicated for 5 min, and centrifuged at 694 g for 5 min. The Agilent 7900 ICP-MS (Agilent Technologies, Japan) equipped with an octopole reaction system (ORS) collision/reaction cell technology was used for measuring concentrations of arsenic and lithium. The LOD for each element was determined as three times the standard deviation of analyzed blanks (alkali solution) and as signal/noise = 3. The limit of quantification (LOQ) was determined as 10 times the standard deviation of analyzed blanks. The analysis precision was estimated by measuring two in-house blood samples in triplicate. As quality control, two commercially available whole blood reference materials were analyzed: Seronorm™ Trace Elements Whole Blood L-1 (LOT 1702821) and L-2 (LOT 1702825), and the obtained average arsenic values (1.9 ± 0.1 µg/kg and 10 ± 1 µg/kg) were in agreement with the reference values (2 ± 0.4 µg/kg and 11.6 ± 2.4 µg/kg). Certified values were converted from µg/L to µg/kg by dividing by the average density of blood (1.055 kg/L). In addition, the reference materials were spiked with 2–1385 µg/kg lithium, and on average 101% of the added lithium was recovered. Blanks and reference materials

were treated together with the collected whole blood samples and analyzed in the beginning, in the middle, and at the end of each analysis. For some individuals, blood samples for element analysis were collected in Lithium Heparin tubes instead of Trace Elements NH Sodium Heparin tubes (Vacuette, Greiner Bio), and therefore B-Li results could not be obtained for all individuals. Leach tests for both types of Vacuette tubes were performed, and no traces of arsenic were detected in neither of them (data not shown).

In the human body, iAs is metabolized via the one-carbon cycle by reducing As(V) to As(III), and methylating As(III) into methylarsonic acid (MMA) and further into dimethylarsinic acid (DMA). The metabolism is not complete, and all four arsenic species are to varying degrees excreted in urine (31). We previously measured arsenic in urine, both as total arsenic in urine (including organic forms such as arsenobetaine from seafood) and as the sum of iAs metabolite concentrations (iAs + MMA + DMA; U-As) in urine, and concluded that this study group is mainly exposed to inorganic forms (15).

Covariates

Urine test strips (Combur-7 Test strips, Roche, Switzerland) were used immediately after sample collection to determine urinary pH, glucose, ketones, leukocytes, nitrites, proteins, erythrocytes, and hemoglobin. These urine reagent strips are commonly used in clinical settings to detect diabetes and other kidney and urinary tract diseases in a semiquantitative manner. Results were obtained and graded on a discrete numerical scale (0, 1, 2, 3, or 4) by comparing the color of the patches on the strip with the colors on the label, according to the manufacturers indications.

To account for variations in urine dilution, we previously measured specific gravity and osmolality (15). In this study, we only included osmolality since it is strongly correlated with specific gravity in the study group ($r_S = 0.99$, p-value < 0.001), and because specific gravity can be influenced by the presence of proteins and glucose (32).

We evaluated if differences in storage time affected the relative protein concentrations by including the field trip as a covariate (~26 months between the first and the last sampling occasion).

We also assessed if the association between B-As and cancer-related proteins was influenced by the efficiency of arsenic metabolism, determined as the relative concentration (%) of the different iAs metabolites in urine. Higher fractions of iAs and MMA in urine are associated with higher risk of adverse health outcomes, and therefore a low efficiency of arsenic methylation is considered a susceptibility factor for arsenic toxicity (2, 33). We stratified the study group as below and equal, or above median %MMA.

Statistical Analyses

All analyses were performed with RStudio (version 1.1.423) using R (version 3.6.2). General characteristics of the study group that are continuous variables and protein levels as NPX values are presented as median and interquartile range (IQR), while categorical variables are presented as percentage or number of individuals per category. NPX values below LOD were excluded from the dataset.

Linear models of principal components were analyzed to assess the influence of different covariates on the overall protein variation. For this, we used the *prince* and *prince.plot* functions within the *swamp* package. These functions perform a principal component analysis for the overall protein expression variation and then perform a linear regression between each variable and principal component to evaluate their association. The heat map presents $-\log_{10}(p\text{-value})$ for the associations, and variables were hierarchically clustered using the *hclust* function. Since the *prince* function does not handle missing data, and imputation was not optimal due to limited input data, we used the NPX dataset substituting values below the LOD by the LOD value specific for each protein. Spearman correlation-tests were carried out between the covariates that were statistically significant in the linear model of principal component analyses. Covariates that were significantly associated with at least one of the 10 principal components and not correlated between each other, as assessed by Spearman correlation-test, were further included in linear regression models used to evaluate individual protein variation and to assess the relation between these proteins and B-As (see below).

To evaluate how much the selected covariates explained the variation of each protein, univariate linear regression analyses were performed. Each model included one protein (as dependent variable) and one covariate (as independent variable) at a time. The variance explained (R^2) was presented as stacked bar plots. To assess the association between B-As and protein expression, multivariable-adjusted linear analyses were performed, adjusting for the covariates that were significantly associated with the overall protein variation. Standardized beta coefficients were obtained for all covariates in the model and presented in stacked bar plots. One individual was excluded from these analyses since no B-As data was available. Similar analyses were performed also adjusting for B-Li as sensitivity analyses. The top associated proteins with B-As were further evaluated with Spearman pairwise correlations, with scatter plots against B-As including a stratification by arsenic metabolism efficiency, and with boxplots to compare between low and high exposed participants. The categorization between low and high exposed was done by splitting by the median B-As. Since the urinary proteins were not normally distributed, we performed Wilcoxon-tests to evaluate the differences between groups.

RESULTS

Characteristics of the Study Participants

Information on the women's anthropometric and lifestyle characteristics, urinary characteristics, and exposure biomarkers are included in **Table 1**. The exposure to iAs from drinking water of the current study group (U-As median 67 µg/L, range 12–399 µg/L; B-As median 2.1 ng/g, range 0.60–9.1 ng/g) is comparable to that of the whole study population ($N = 201$ women, U-As median 65 µg/L, range 12–407 µg/L) (15). Furthermore, U-As (as sum of metabolites, adjusted for average osmolality) and B-As were strongly correlated ($r_S = 0.85$, p-value < 0.001; **Supplementary Figure 1**) reflecting that both matrices are reliable biomarkers for iAs exposure in this study

TABLE 1 | Characteristics of the individuals included in the study ($n = 176$).

	n	**Median (IQR)[a]**
General characteristics		
Age (years)	176	36 (28–48)
Weight (kg)	176	60 (53–70)
Height (cm)	176	149 (146–153)
Coca chewing (yes)[b]	174	74%
Systolic blood pressure (mmHg)	175	100 (90–110)
Diastolic blood pressure (mmHg)	175	70 (60–70)
Urine characteristics		
Specific gravity	176	1.02 (1.01–1.02)
Osmolality (mOsm/kg)	176	764 (564–883)
Urine pH	176	5.5 (5.0–6.1)
Urine leukocytes (0/1/2/3)[c]	176	146/10/17/3
Exposures		
U-As (µg/L)[d]	176	67 (47–111)
iAs (%)	176	12 (8.6–17)
MMA (%)	176	7.7 (6.2–9.6)
DMA (%)	176	79 (74–84)
B-As (ng/g)[e]	175	2.1 (1.4–3.0)
B-Li (ng/g)[e]	159	18 (9.4–30)

[a] Data presented as median and interquartile range (IQR) unless otherwise stated.
[b] Percentage of individuals who reported chewing coca leaves regularly.
[c] Number of individuals in each category of urine leukocyte levels, classified according to the urine strips Combur 7 Test.
[d] Sum of urinary iAs metabolites adjusted for average urinary osmolality (727 mOsm/kg) of the total study population as described in De Loma et al. (15).
[e] Measured in whole blood.

group. Only two women reported consuming alcoholic beverages frequently (every other week), and seven reported smoking tobacco (yes/no). Based on the urine strips, no woman had ketones, two had elevated glucose, 10 nitrites, 16 proteins, 40 hemoglobin, and 12 erythrocytes in urine.
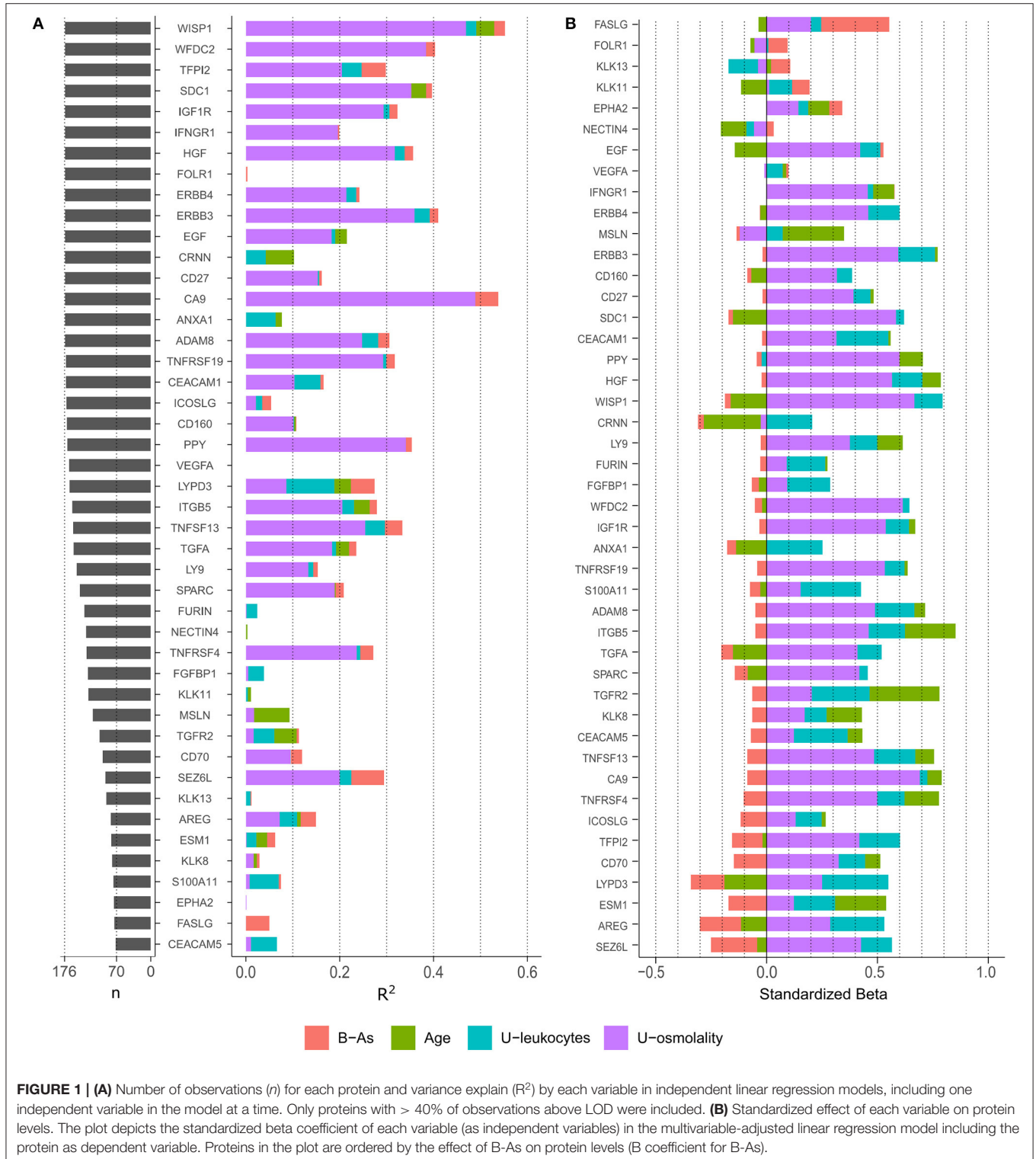
Covariates Explaining Protein Variation in Urine

Descriptive statistics of the 92 cancer-related proteins measured by the Multiplex Oncology II panel are shown in **Supplementary Table 2**. Out of the proteins initially measured, 45 proteins were further evaluated, i.e. those with > 40% of observations above LOD. To assess to which extent characteristics of the study participants were associated with variation of these proteins, we performed linear regression models of principal components (**Supplementary Figure 2**). Age, urinary osmolality, urinary leukocytes and urinary pH significantly explained the overall protein variation in urine. Urinary pH and osmolality were correlated ($r_S = -0.21$, p-value = 0.005), while none of the other covariates presented pairwise correlations, and therefore urinary pH was not included in further analyses. No significant associations were found for coca chewing, ethnicity, or sampling field trip. None of the other markers from the urine strips (glucose, nitrites, proteins, hemoglobin, nor erythrocytes) were associated with the overall protein variation of the cancer-related proteins. In addition, there

were no differences in protein expression between ethnicities (data not shown).

To evaluate the protein variance explained by each covariate independently, we performed linear regression analyses including each covariate individually (age, urinary osmolality, urinary leukocytes, and B-As as the exposure of interest; **Figure 1A**). Urinary osmolality explained the variation of most proteins measured in urine, in some cases explaining up to 49%

FIGURE 1 | (A) Number of observations (*n*) for each protein and variance explain (R^2) by each variable in independent linear regression models, including one independent variable in the model at a time. Only proteins with > 40% of observations above LOD were included. **(B)** Standardized effect of each variable on protein levels. The plot depicts the standardized beta coefficient of each variable (as independent variables) in the multivariable-adjusted linear regression model including the protein as dependent variable. Proteins in the plot are ordered by the effect of B-As on protein levels (B coefficient for B-As).

of the variation. However, 13 proteins were not influenced by osmolality ($R^2 < 1\%$: FOLR1, CRNN, ANXA1, VEGFA, FURIN, NECTIN4, FGFBP1, KLK11, KLK13, ESM1, S100A11, EPHA2, and FASLG). Age explained the variation (R^2: 1.3–7.5%) of 11 proteins: WISP1, SCD1, EGF, CRNN, ANXA1, LYPD3, ITGB5, TGFA, MSLN, TGFR2, and ESM1. The only marker selected from the urine strip test was urinary leukocytes, as a proxy for inflammation of the urinary tract, which explained between 1 and 10% of the variation of 23 proteins of which LYPD3, ANXA1, and S100A11 showed the highest R^2 (10, 6.3, and 6.2% respectively). B-As explained 5–7% of the protein variation of TFPI2, LYPD3, SEZ6L, and FASLG.

Urinary Proteins Associated With Arsenic Exposure

We further examined how B-As was associated to cancer-related proteins in urine by multivariable linear regression models adjusted for age, urinary leukocytes, and urinary osmolality. Out of the 45 proteins, four were associated (p-value < 0.05) with B-As: 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 1B**, **Table 2**, and **Supplementary Table 3**). None of the associations were statistically significant after adjusting for multiple testing. The relative protein expression of SEZ6L, LYPD3, and TFPI2 decreased with higher B-As concentrations, while FASLG increased with increasing B-As concentrations (**Table 2** and **Figure 2A**). Differences in arsenic metabolism efficiency did not change the association between these top proteins and B-As (**Figure 2B**). In addition, we explored if these top proteins were significantly different between individuals with low or high arsenic exposure, by splitting by median B-As (**Figure 2C**). Individuals highly exposed to arsenic had significantly lower SEZ6L, LYPD3, and TFPI2 levels (p-value = 0.021, 0.0001, 0.004, respectively). Highly exposed individuals had higher levels of FASLG, although not significantly (p-value = 0.09).

We also explored the correlation between the top proteins that were associated with B-As. Spearman pairwise correlations between SEZ6L, LYPD3, and TFPI2 were observed, but not for FASLG (**Supplementary Figure 3**). We considered performing downstream analyses, e.g., pathway enrichment, but the number of proteins associated with B-As was not enough for such analyses to be suitable. Sensitivity analyses were performed including B-Li in the models. The effect estimates for the associations between B-As and FASLG, SEZ6L, LYPD3, and TFPI2 did not change more than 10%, although they were no longer significant (p-value < 0.05), probably due to the decrease in sample size (**Table 2** and **Supplementary Table 3**).

DISCUSSION

To our knowledge, this study is the first to evaluate the toxicity of environmental iAs exposure in individuals from Bolivia. We identified four proteins (FASLG, SEZ6L, LYPD3, and TFPI2) measured in urine that were associated with B-As in indigenous

TABLE 2 | Top 10 proteins in urine associated with B-As.

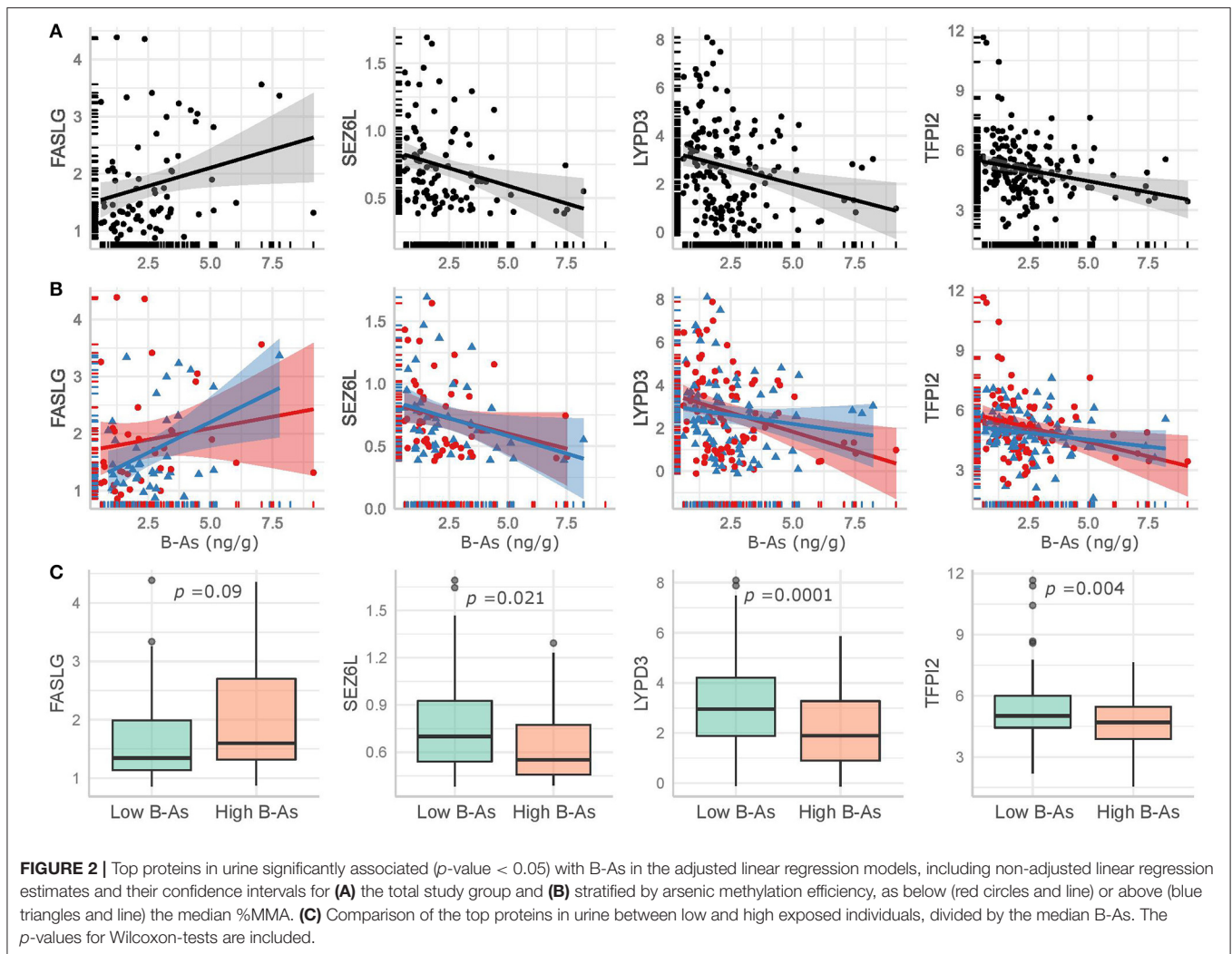| Protein	n	R² (%)	B (95% CI)	p-value
		Main analysis[a]		
FASLG	74	5	0.155 (0.034, 0.276)	0.013
SEZ6L	93	25	−0.038 (−0.072, −0.004)	0.028
LYPD3	166	23	−0.172 (−0.331, −0.014)	0.034
TFPI2	175	24	−0.125 (−0.249, −0.001)	0.049
AREG	82	15	−0.096 (−0.204, 0.012)	0.080
CA9	175	49	−0.05 (−0.112, 0.013)	0.119
ESM1	81	7	−0.067 (−0.152, 0.019)	0.124
ICOSLG	172	3	−0.097 (−0.225, 0.03)	0.134
CD70	98	11	−0.046 (−0.106, 0.015)	0.137
TNFRSF4	131	27	−0.066 (−0.164, 0.032)	0.183

Data presented as B coefficients and 95% confidence intervals (CI) for the association with B-As.
No proteins passed the Bonferroni adjustment for multiple comparisons (p-value threshold = 0.05/45 = 0.001).
[a]Multivariate linear regression models adjusted for age, urinary osmolality, and urinary leukocytes.

women from the Bolivian Andes. Arsenic concentrations in blood were positively associated with relative protein expression of FASLG and inversely associated with SEZ6L, LYPD3, and TFPI2. These associations were not statistically significant when correcting for multiple comparisons likely due to limited sample size. When comparing individuals with high vs. low exposure to arsenic, SEZ6L, LYPD3, and TFPI2 were significantly decreased in the high exposure group, while FASLG was non-significantly increased. The individual proteins are discussed below including previous relations reported for arsenic, when known. So far, there is no knowledge about whether urinary levels of these proteins are predictive of cancer. Longitudinal research is needed to establish whether these proteins serve as early biomarkers of arsenic-related carcinogenicity, or if these associations are a consequence of an adaptive response to the exposure. This study also identified other influential factors for the variation in protein expression in urine, such as urinary dilution and age, which is important information for future use of these proteins as potential risk biomarkers.

FASLG binds to Fas receptors and regulates an apoptotic signaling pathway (34, 35). FASLG has a dual function; it both promotes and inhibits cell death depending on the cell type (36). Elevated serum concentrations of soluble FASLG have been detected in patients with leukemia, lymphoma and multiple solid tumors (37). Arsenic has been linked to FASLG protein expression, mainly in *in vitro* studies of the cancer drug arsenic trioxide, which is currently used for treatment of acute promyelocytic leukemia. Different leukemia cell lines exposed to arsenic trioxide showed increased gene or protein expression of FASLG (38–40). Human studies of environmental exposure to iAs and FASLG are few and inconclusive. Women highly exposed to iAs (median U-As 276 µg/L) from the Argentinean Andes presented higher *FASLG* gene expression in sorted T-cells compared to women with lower exposure (median U-As 65 µg/L)

FIGURE 2 | Top proteins in urine significantly associated (p-value < 0.05) with B-As in the adjusted linear regression models, including non-adjusted linear regression estimates and their confidence intervals for **(A)** the total study group and **(B)** stratified by arsenic methylation efficiency, as below (red circles and line) or above (blue triangles and line) the median %MMA. **(C)** Comparison of the top proteins in urine between low and high exposed individuals, divided by the median B-As. The p-values for Wilcoxon-tests are included.

(41). In contrast, five individuals from Mexico exposed to iAs (median 224.3 mg/g creatinine) showed lower gene expression of *FASLG* in peripheral blood mononuclear cells (including T cells, B cells, and NK cells) compared to five unexposed individuals (iAs < 2.2 mg/g creatinine) (42).

TFPI2 is an inhibitor of the tissue factor pathway involved in blood coagulation. Gene expression of *TFPI2* was decreased in samples of bladder tumors (43) and cervical tumors (44), while TFPI2 protein concentrations in serum were higher in ovarian clear cell adenocarcinoma patients (45, 46). Hypermethylation, generally associated with decreased gene expression (47), was detected for *TFPI2* in multiple types of cancer compared to normal adjacent tissues (48–52). A human prostate epithelial cell line transformed by chronic exposure to arsenite presented decreased *TFPI2* gene expression (53), in line with our findings. On the contrary, normal human lung cells exposed to sodium arsenite for 30–60 days showed an upregulation of *TFPI2* gene expression (54).

Two of the proteins identified, SEZ6L and LYPD3, have to our knowledge not been related to arsenic before. SEZ6L is a transmembrane protein in the endoplasmic reticulum and the cell membrane whose function is still unclear. In our study, lower urinary expression of SEZ6L was associated with increasing arsenic exposure. Interestingly, deletions in the *SEZ6L* gene have been found in primary lung tumor cells (55), and loss-of-function variants in *SEZ6L* have been associated with increased risk of lung cancer (56). *SEZ6L* is also reported as a fusion gene in some lung cancers, melanoma and skin cancer (57).

LYPD3 is a glycosylphosphatidylinositol-anchored urokinase receptor involved in cell-matrix interactions and metastasis (58, 59). In contrast to our findings of lower relative protein expression in relation arsenic exposure, enhanced protein expression of LYPD3 has been identified in tissues from urothelial cancers (59), breast cancers (60), melanoma (61), and lung cancers (62). Recently, high protein expression levels of LYPD3 in tumor tissues have been recognized as a biomarker of poor prognosis for lung cancer patients (63, 64).

We also identified factors that explained the overall variation of the 45 proteins measured in urine, such as age, urinary osmolality, and urinary leukocytes. This highlights the importance to take these factors into account when evaluating protein biomarkers in urine. Since urinary pH and osmolality

were correlated to some extent in this study group, we did not adjust the final models for urinary pH. A study with Sprague-Dawley rats did not show any effect of urinary pH on urinary proteins determined with urine test strips either (65). Although urinary osmolality is less affected by proteins in urine than other dilution markers such as specific gravity (66), it is not possible to discern the direction of the association between the proteins in this study and osmolality. Storage time of the samples may affect some proteins in plasma (67). In the current study, storage time did not explain the overall protein variation or influenced the top proteins associated with B-As. There were several women with menstrual bleeding, but this did not influence the proteins evaluated. Women with urinary tract infection, identified by the presence of leukocytes in urine, were expected considering the limited access to health care in the study area. However, this did not influence the variation of the proteins.

Regarding the choice of matrices for measuring exposure to arsenic, B-As is most suitable to assess recent exposure to iAs since it is rapidly cleared from this matrix (68). Still, the strong correlation between B-As and U-As concentrations in the current study group likely reflects a chronic exposure to iAs, and it allowed us to use B-As as a valid biomarker of exposure. Arsenic concentrations in drinking water from these villages did not present temporal variations during the 2 years of recruitment, supporting that these women were constantly exposed to iAs (15). By using two different biological matrices to measure exposure (blood) and effect biomarkers (urine), we avoided potential problems of co-excretion in urine, previously identified for other elements like cadmium (29). Furthermore, it would be valuable to analyze these cancer-related proteins in other matrices such as serum or plasma in order to correlate these with urine values, since no data on this is available. Unfortunately, the sampling of plasma from the study individuals was not optimal for protein analysis.

The detection of proteins in urine is commonly considered an indication of kidney disease. High levels of proteins in urine, also known as proteinuria, have been associated with exposure to arsenic (69). In the review by Zheng et al. (69), all studies evaluating proteinuria used U-As as a biomarker of exposure, therefore not being exempt from the potential problem of co-excretion between biomarkers of effect and exposure in the same matrix. The constant development of more advanced protein detection techniques in urine demonstrates that urine is more protein-rich than previously believed, even under normal conditions (70). This, and the fact that urine is an abundant and non-invasive sampling matrix, justifies the attempts to identify more disease biomarkers in urine. In fact, urinary protein biomarkers have been identified as early diagnostic markers for several cancer types (16–18, 71). Regarding the Proximity Extension Assay used in this study, only one other study has employed this technology to evaluate protein expression in urine (72). In Fellström et al. (72), they found an association between lipid markers in serum and inflammation- and cardiovascular-related urinary proteins in 75-year-old individuals, but urinary dilution was not adjusted for in the analyses.

This study group in the Bolivian Andes has a markedly efficient arsenic metabolism capacity (15), which may influence the degree of arsenic toxicity. This, and the fact that cancer types differ between populations depending on their underlying genetic background and susceptibility (73), highlights the need for population-specific studies about arsenic toxicity. This work is a cross-sectional study of apparently healthy women, and therefore we cannot distinguish if the observed variations are adaptive or toxic responses to arsenic exposure. In addition, since iAs is also nephrotoxic (69), it is not possible to disentangle if the associations found are due to arsenic-induced renal toxicity altering protein excretion, or if these proteins serve as proxy for arsenic toxicity in other organs. More research is warranted about the relation between these proteins in different matrices and arsenic-related toxic outcomes for these proteins to be used as toxicity biomarkers or to elucidate the toxicity mechanisms of arsenic. The strengths of this study are the well-characterized arsenic exposure and the use of multiplex proteomic technologies to identify novel candidates to study arsenic-related health effects.

CONCLUSION

Using multiplex proteomic methods in urine samples, we identified four putative cancer-related proteins (FASLG, SEZ6L, LYPD3, and TFPI2) associated with arsenic exposure in women living around Lake Poopó, Bolivia. In order to clarify if these proteins represent early arsenic-related carcinogenic changes, follow-up studies are needed. By exploring cancer-related proteins in urine, we hope to contribute to the future development of disease and/or toxicity biomarkers with non-invasive sampling methods.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Comité Nacional de Bioética (Bolivia) and the Regional Ethic Committee of Karolinska Institutet (Sweden). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JG, NT, and KB planned and designed the research. JD, FA, JG, NT, and KB recruited the study participants and collected the data. ML performed the metal analysis. JD assisted by AG conducted the statistical analyses. JD and KB wrote the manuscript. All authors proofread and commented on previous versions of the manuscript and approved the final version of it.

FUNDING

Laboratory (France), the Eric Philip Sörensens Stiftelse (Sweden), the Swedish Research Council, Karolinska Institutet (Sweden), the Swedish International Development Cooperation Agency, and the Genetics Institute at Universidad Mayor de San Andrés (Bolivia).

ACKNOWLEDGMENTS

We warmly acknowledge the women who participated in the study. We also thank the Servicio Departamental de Salud de Oruro (SEDES) and the Salud Familiar Comunitaria e

Intercultural (SAFCI) program for their help during the field activities. We are grateful to everyone who assisted during the recruitment, including workers at the local health centers, and to Marina Cuti, Rolando Paz, and Josue Mamani at the Instituto de Genética, UMSA.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpubh. 2020.605123/full#supplementary-material

REFERENCES

1. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Arsenic, metals, fibres, and dusts. *IARC Monogr Eval Carcinog Risks Hum*. (2012) 100:11–465. Available online at: https://monographs.iarc.fr/wp-content/uploads/2018/06/mono100C.pdf

2. Antonelli R, Shao K, Thomas DJ, Sams R, II, Cowden J. AS3MT, GSTO, and PNP polymorphisms: impact on arsenic methylation and implications for disease susceptibility. *Environ Res*. (2014) 132:156–67. doi: 10.1016/j.envres.2014.03.012

3. Kuo CC, Moon KA, Wang SL, Silbergeld E, Navas-Acien A. The association of arsenic metabolism with cancer, cardiovascular disease, and diabetes: a systematic review of the epidemiological evidence. *Environ Health Perspect*. (2017) 125:087001. doi: 10.1289/EHP577

4. Minatel BC, Sage AP, Anderson C, Hubaux R, Marshall EA, Lam WL, et al. Environmental arsenic exposure: from genetic susceptibility to pathogenesis. *Environ Int*. (2018) 112:183–97. doi: 10.1016/j.envint.2017.12.017

5. Ferrario D, Gribaldo L, Hartung T. Arsenic exposure and immunotoxicity: a review including the possible influence of age and sex. *Curr Environ Heal Rep*. (2016) 3:1–12. doi: 10.1007/s40572-016-0082-3

6. Khan KM, Chakraborty R, Bundschuh J, Bhattacharya P, Parvez F. Health effects of arsenic exposure in Latin America: an overview of the past eight years of research. *Sci Total Environ*. (2020) 710:136071. doi: 10.1016/j.scitotenv.2019.136071

7. Bundschuh J, Nath B, Bhattacharya P, Liu CW, Armienta MA, Moreno López MV, et al. Arsenic in the human food chain: the Latin American perspective. *Sci Total Environ*. (2012) 429:92–106. doi: 10.1016/j.scitotenv.2011.09.069

8. Hopenhayn-Rich C, Biggs M, Lou Smith AH. Lung and kidney cancer mortality associated with arsenic in drinking water in Cordoba, Argentina. *Gt Br Int J Epidemiol*. (1998) 27:561–9. doi: 10.1093/ije/27.4.561

9. Smith AH, Goycolea M, Haque R, Biggs ML. Marked increase in bladder and lung cancer mortality in a region of northern chile due to arsenic in drinking water. *Am J Epidemiol*. (1998) 147:660–9. doi: 10.1093/oxfordjournals.aje.a009507

10. Hopenhayn-Rich C, Biggs M, Lou Fuchs A, Bergoglio R, Tello EE, Nicolli H, et al. Bladder cancer mortality associated with arsenic in drinking water in Argentina. *Epidemiology*. (1996) 7:117–24. doi: 10.1097/00001648-199603000-00003

11. Roh T, Steinmaus C, Marshall G, Ferreccio C, Liaw J, Smith AH. Age at exposure to arsenic in water and mortality 30-40 years after exposure cessation. *Am J Epidemiol*. (2018) 187:2297–305. doi: 10.1093/aje/kwy159

12. López-Carrillo L, Hernández-Ramírez RU, Gandolfi AJ, Ornelas-Aguirre JM, Torres-Sánchez L, Cebrian ME. Arsenic methylation capacity is associated with breast cancer in Northern Mexico. *Toxicol Appl Pharmacol*. (2014) 280:53–9. doi: 10.1016/j.taap.2014.07.013

13. Michel-Ramirez G, Recio-Vega R, Ocampo-Gomez G, Palacios-Sanchez E, Delgado-Macias M, Delgado-Gaona M, et al. Association between YAP expression in neoplastic and non-neoplastic breast tissue with arsenic urinary levels. *J Appl Toxicol*. (2017) 37:1195–202. doi: 10.1002/jat.3481

14. Smith AH, Marshall G, Yuan Y, Steinmaus C, Liaw J, Smith MT, et al. Rapid reduction in breast cancer mortality with inorganic arsenic in drinking water. *EBioMedicine*. (2014) 1:58–63. doi: 10.1016/j.ebiom.2014.10.005

15. De Loma J, Tirado N, Ascui F, Levi M, Vahter M, Broberg K, et al. Elevated arsenic exposure and efficient arsenic metabolism in indigenous women around Lake Poopó, Bolivia. *Sci Total Environ*. (2019) 657:179–86. doi: 10.1016/j.scitotenv.2018.11.473

16. Chakraborty A, Dasari S, Long W, Mohan C. Urine protein biomarkers for the detection, surveillance, and treatment response prediction of bladder cancer. *Am J Cancer Res*. (2019) 9:1104–17.

17. Grayson K, Gregory E, Khan G, Guinn BA. Urine biomarkers for the early detection of ovarian cancer—are we there yet? *Biomark Cancer*. (2019) 11:1179299X1983097. doi: 10.1177/1179299X19830977

18. Pedroza-Díaz J, Röthlisberger S. Advances in urinary protein biomarkers for urogenital and non-urogenital pathologies. *Biochem Medica*. (2015) 25:22–35. doi: 10.11613/BM.2015.003

19. Zhang H, Cao J, Li L, Liu Y, Zhao H, Li N, et al. Identification of urine protein biomarkers with the potential for early detection of lung cancer. *Sci Rep*. (2015) 5:11805. doi: 10.1038/srep11805

20. Guo Z, Hu Q, Tian J, Yan L, Jing C, Xie HQ, et al. Proteomic profiling reveals candidate markers for arsenic-induced skin keratosis. *Environ Pollut*. (2016) 218:34–8. doi: 10.1016/j.envpol.2016.08.035

21. Harezlak J, Wu MC, Wang M, Schwartzman A, Christiani DC, Lin X. Biomarker discovery for arsenic exposure using functional data. Analysis and Feature Learning of Mass Spectrometry Proteomic Data. *J Proteome Res*. (2008) 7:217–24. doi: 10.1021/pr070491n

22. Zhai R, Su S, Lu X, Liao R, Ge X, He M, et al. Proteomic profiling in the sera of workers occupationally exposed to arsenic and lead: identification of potential biomarkers. *BioMetals*. (2005) 18:603–13. doi: 10.1007/s10534-005-3001-x

23. Hegedus CM, Skibola CF, Warner M, Skibola DR, Alexander D, Lim S, et al. Decreased urinary beta-defensin-1 expression as a biomarker of response to arsenic. *Toxicol Sci*. (2008) 106:74–82. doi: 10.1093/toxsci/kfn104

24. Espinoza-Navarro A, Diaz J, Rodrígues H, Moreno A. 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) Efectos de la Altura Sobre Patrones Antropométricos y Fisiológicos en Población Aymará y no Aymará entre 18 y 65 Años de la Provincia de Parinacota, Chile (3.700 msnm). *Int J Morphol*. (2011) 29:34–40. doi: 10.4067/S0717-95022011000100005

25. Harari F, Langeén M, Casimiro E, Bottai M, Palm B, Nordqvist H, et al. Environmental exposure to lithium during pregnancy and fetal size: a longitudinal study in the Argentinean Andes. *Environ Int*. (2015) 77:48–54. doi: 10.1016/j.envint.2015.01.011

26. Assarsson E, Lundberg M, Holmquist G, Björkesten J, Bucht Thorsen S, Ekman D, et al. Homogenous 96-Plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. *PLoS ONE*. (2014) 9:e95192. doi: 10.1371/journal.pone.0095192

27. Alhamdow A, Lindh C, Albin M, Gustavsson P, Kan Tinnerberg H, Broberg K. Cardiovascular disease-related serum proteins in workers occupationally exposed to polycyclic aromatic hydrocarbons. *Toxicol Sci*. (2019) 171:235–46. doi: 10.1093/toxsci/kfz142

28. Harris SE, Cox SR, Bell S, Marioni RE, Prins BP, Pattie A, et al. Neurology-related protein biomarkers are associated with cognitive ability and brain volume in older age. *Nat Commun*. (2020) 11:1–12. doi: 10.1038/s41467-019-14161-7

29. Wallin M, Sallsten G, Lundh T, Barregard L. Low-level cadmium exposure and effects on kidney function. *Occup Environ Med.* (2014) 71:848–54. doi: 10.1136/oemed-2014-102279

30. Levi M, Hjelm C, Harari F, Vahter M. 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.* (2018) 53:81–7. doi: 10.1016/j.clinbiochem.2017.12.003

31. Vahter M. Mechanisms of arsenic biotransformation. *Toxicology.* (2002) 181–182:211–7. doi: 10.1016/S0300-483X(02)00285-8

32. Parikh CR, Gyamlani GG, Carvounis CP. Screening for microalbuminuria simplified by urine specific gravity. *Am J Nephrol.* (2002) 22:315–9. doi: 10.1159/000065220

33. Li H, Engström K, Vahter M, Broberg K. Arsenic exposure through drinking water is associated with longer telomeres in peripheral blood. *Chem Res Toxicol.* (2012) 25:2333–9. doi: 10.1021/tx300222t

34. Ruiz-Garcia R, Mora S, Lozano-Sanchez G, Martinez-Lostao L, Paz-Artal E, Ruiz-Contreras J, et al. Decreased activation-induced cell death by EBV-transformed B-cells from a patient with autoimmune lymphoproliferative syndrome caused by a novel FASLG mutation. *Pediatr Res.* (2015) 78:603–8. doi: 10.1038/pr.2015.170

35. Schneider P, Bodmer JL, Holler N, Mattmann C, Scuderi P, Terskikh A, et al. Characterization of Fas (Apo-1, CD95)-Fas ligand interaction. *J Biol Chem.* (1997) 272:18827–33. doi: 10.1074/jbc.272.30.18827

36. Bossi G, Stinchcombe JC, Page LJ, Griffiths GM. Sorting out the multiple roles of Fas ligand. *Eur J Cell Biol.* (2000) 79:539–43. doi: 10.1078/0171-9335-00077

37. Kadam CY, Abhang SA. Apoptosis markers in breast cancer therapy. In: *Advances in Clinical Chemistry.* Cambridge, MA: Academic Press Inc. (2016). p. 143–93. doi: 10.1016/bs.acc.2015.12.003

38. Toosi B, Zaker F, Alikarami F, Kazemi A, Teremmahi Ardestanii M. 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.* (2018) 102:428–37. doi: 10.1016/j.biopha.2018.03.009

39. Zeng Y, Weng G, Fan J, Li Z, Wu J, Li Y, et al. Curcumin reduces the expression of survivin, leading to enhancement of arsenic trioxide-induced apoptosis in myelodysplastic syndrome and leukemia stem-like cells. *Oncol Rep.* (2016) 36:1233–42. doi: 10.3892/or.2016.4944

40. Zhu J, Okumura H, Ohtake S, Nakamura S, Nakao S. The molecular mechanism of arsenic trioxide-induced apoptosis and oncosis in leukemia/lymphoma cell lines. *Acta Haematol.* (2003) 110:1–10. doi: 10.1159/000072407

41. Engström K, Wojdacz TK, Marabita F, Ewels P, Käller M, Vezzi F, et al. Transcriptomics and methylomics of CD4-positive T cells in arsenic-exposed women. *Arch Toxicol.* (2017) 91:2067–78. doi: 10.1007/s00204-016-1879-4

42. Salgado-Bustamante M, Ortiz-Pérez MD, Calderón-Aranda E, Estrada-Capetillo L, Niño-Moreno P, González-Amaro R, et al. Pattern of expression of apoptosis and inflammatory genes in humans exposed to arsenic and/or fluoride. *Sci Total Environ.* (2010) 408:760–7. doi: 10.1016/j.scitotenv.2009.11.016

43. Feng C, Ho Y, Sun C, Xia G, Ding Q, Gu B. TFPI-2 expression is decreased in bladder cancer and is related to apoptosis. *J Buon.* (2016) 21:1518–23.

44. Zhang Q, Zhang Y, Wang SZ, Wang N, Jiang WG, Ji YH, et al. Reduced expression of tissue factor pathway inhibitor-2 contributes to apoptosis and angiogenesis in cervical cancer. *J Exp Clin Cancer Res.* (2012) 31:1. doi: 10.1186/1756-9966-31-1

45. Arakawa N, Kobayashi H, Yonemoto N, Masuishi Y, Ino Y, Shigetomi H, et al. Clinical significance of tissue factor pathway inhibitor 2, a serum biomarker candidate for ovarian clear cell carcinoma. *PLoS ONE.* (2016) 11:e0165609. doi: 10.1371/journal.pone.0165609

46. Arakawa N, Miyagi E, Nomura A, Morita E, Ino Y, Ohtake N, et al. Secretome-based identification of TFPI2, A novel serum biomarker for detection of ovarian clear cell adenocarcinoma. *J Proteome Res.* (2013) 12:4340–50. doi: 10.1021/pr400282j

47. Clark SJ, Melki J. DNA methylation and gene silencing in cancer: which is the guilty party? *Oncogene.* (2002) 21:5380–7. doi: 10.1038/sj.onc.1205598

48. Dong Y, Tan Q, Tao L, Pan X, Pang L, Liang W, et al. Hypermethylation of TFPI2 correlates with cervical cancer incidence in the Uygur and Han populations of Xinjiang, China. *Int J Clin Exp Pathol.* (2015) 8:1844–54.

49. Hu H, Chen X, Wang C, Jiang Y, Li J, Ying X, et al. The role of TFPI2 hypermethylation in the detection of gastric and colorectal cancer. *Oncotarget.* (2017) 8:84054–65. doi: 10.18632/oncotarget.21097

50. Kim SY, Han YK, Song JM, Lee CH, Kang K, Yi JM, et al. Aberrantly hypermethylated tumor suppressor genes were identified in oral squamous cell carcinoma (OSCC). *Clin Epigenetics.* (2019) 11:116. doi: 10.1186/s13148-019-0715-0

51. Sun FK, Fan YC, Zhao J, Zhang F, Gao S, Zhao ZH, et al. Detection of TFPI2 methylation in the serum of hepatocellular carcinoma patients. *Dig Dis Sci.* (2013) 58:1010–5. doi: 10.1007/s10620-012-2462-3

52. Zhang JP, Yang S, Bin Xie YY, Chen XY, Zhao Y, He DZ, et al. Detection of methylated tissue factor pathway inhibitor 2 and human long DNA in fecal samples of patients with colorectal cancer in China. *Cancer Epidemiol.* (2012) 36:73–7. doi: 10.1016/j.canep.2011.04.006

53. Severson PL, Tokar EJ, Vrba L, Waalkes MP, Futscher BW. Coordinate H3K9 and DNA methylation silencing of ZNFs in toxicant-induced malignant transformation. *Epigenetics.* (2013) 8:1080–8. doi: 10.4161/epi.25926

54. Clancy HA, Sun H, Passantino L, Kluz T, Muñoz A, Zavadil J, et al. Gene expression changes in human lung cells exposed to arsenic, chromium, nickel or vanadium indicate the first steps in cancer. *Metallomics.* (2012) 4:784–93. doi: 10.1039/c2mt20074k

55. Nishioka M, Kohno T, Takahashi M, Niki T, Yamada T, Sone S, et al. Identification of a 428-kb homozygously deleted region disrupting the SEZ6L gene at 22q12.1 in a lung cancer cell line. *Oncogene.* (2000) 19:6251–60. doi: 10.1038/sj.onc.1204031

56. Gorlov IP, Meyer P, Liloglou T, Myles J, Boettger MB, Cassidy A, et al. Seizure 6-like (SEZ6L) gene and risk for lung cancer. *Cancer Res.* (2007) 67:8406–17. doi: 10.1158/0008-5472.CAN-06-4784

57. Mitelman F, Johansson B, Mertens F. *Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer.* (2020). Available online at: mitelmandatabase.isb-cgc.org (accessed July 24, 2020).

58. Hansen LV, Rdsvoll HG, Nielsen BS, Lund LR, Danø K, Jensen ON, et al. Structural analysis and tissue localization of human C4.4A: a protein homologue of the urokinase receptor. *Biochem J.* (2004) 380:845–57. doi: 10.1042/bj20031478

59. Smith BA, Kennedy WJ, Harnden P, Selby PJ, Trejdosiewicz LK, Southgate J. Identification of genes involved in human urothelial cell-matrix interactions: implications for the progression pathways of malignant urothelium. *Cancer Res.* (2001) 61:1678–85.

60. Fletcher GC, Patel S, Tyson K, Adam PJ, Schenker M, Loader JA, et al. hAG-2 and hAG-3, human homologues of genes involved in differentiation, are associated with oestrogen receptor-positive breast tumours and interact with metastasis gene C4.4a and dystroglycan. *Br J Cancer.* (2003) 88:579–85. doi: 10.1038/sj.bjc.6600740

61. Seiter S, Stassar M, Rappl G, Reinhold U, Tilgen W, Zöller M. Upregulation of C4.4 A expression during progression of melanoma. *J Invest Dermatol.* (2001) 116:344–7. doi: 10.1046/j.1523-1747.2001.01230.x

62. Würfel J, Seiter S, Stassar M, Claas A, Kläs R, Rösel M, et al. Cloning of the human homologue of the metastasis-associated rat C4.4A. *Gene.* (2001) 262:35–41. doi: 10.1016/S0378-1119(00)00515-1

63. Cohen AS, Khalil FK, Welsh EA, Schabath MB, Enkemann SA, Davis A, et al. Cell-surface marker discovery for lung cancer. *Oncotarget.* (2017) 8:113373–402. doi: 10.18632/oncotarget.23009

64. Hu P, Huang Y, Gao Y, Yan H, Li X, Zhang J, et al. Elevated expression of LYPD3 is associated with lung adenocarcinoma carcinogenesis and poor prognosis. *DNA Cell Biol.* (2020) 39:522–32. doi: 10.1089/dna.2019.5116

65. Reagan WJ, VanderLind B, Shearer A, Botts S. Influence of urine pH on accurate urinary protein determination in Sprague-Dawley rats. *Vet Clin Pathol.* (2007) 36:73–8. doi: 10.1111/j.1939-165X.2007.tb00185.x

66. Chadha V, Garg U, Alon US. Measurement of urinary concentration: a critical appraisal of methodologies. *Pediatr Nephrol.* (2001) 16:374–82. doi: 10.1007/s004670000551

67. Enroth S, Hallmans G, Grankvist K, Gyllensten U. Effects of long-term storage time and original sampling month on biobank plasma protein concentrations. *EBioMedicine.* (2016) 12:309–14. doi: 10.1016/j.ebiom.2016.08.038

68. Hughes MF. Biomarkers of exposure: a case study with inorganic arsenic. *Environ Health Perspect.* (2006) 114:1790–6. doi: 10.1289/ehp.9058

69. Zheng L, Kuo C-C, Fadrowski J, Agnew J, Weaver VM, Navas-Acien A. Arsenic and chronic kidney disease: a systematic review. *Curr Environ Heal Rep.* (2014) 1:192–207. doi: 10.1007/s40572-014-0024-x

70. Harpole M, Davis J, Espina V. Current state of the art for enhancing urine biomarker discovery. *Expert Rev Proteomics.* (2016) 13:609–26. doi: 10.1080/14789450.2016.1190651

71. Zhang C, Leng W, Sun C, Lu T, Chen Z, Men X, et al. Urine proteome profiling predicts lung cancer from control cases and other tumors. *EBioMedicine.* (2018) 30:120–8. doi: 10.1016/j.ebiom.2018.03.009

72. Fellström B, Helmersson-Karlqvist J, Lind L, Soveri I, Wu P-H, Ns Thulin M, et al. Associations between apolipoprotein A1, high-density lipoprotein cholesterol, and urinary cytokine levels in elderly males and females. *J Interf Cytokine Res.* (2020) 40:71–4. doi: 10.1089/jir.2019.0074

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

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.