Journal Pre-proof

A new perspective on the toxicity of arsenic-contaminated soil: Tandem mass tag proteomics and metabolomics in earthworms

Ronggui Tang, Ping Lan, Changfeng Ding, Junsong Wang, Taolin Zhang, Xingxiang Wang

PII: S0304-3894(20)30814-1

DOI: https://doi.org/10.1016/j.jhazmat.2020.122825

Reference: HAZMAT 122825

To appear in: *Journal of Hazardous Materials*

Received Date: 10 February 2020

Revised Date: 23 April 2020

Accepted Date: 23 April 2020

Please cite this article as: { doi: https://doi.org/

A new perspective on the toxicity of arsenic-contaminated soil: tandem mass tag proteomics and metabolomics in earthworms

Ronggui Tang[a, b, f], Ping Lan[c], Changfeng Ding[a], Junsong Wang[e], Taolin Zhang[a], Xingxiang Wang[a*]

a. Key Laboratory of Soil Environment and Pollution Remediation, Institute of Soil Science, Chinese Academy of Sciences, Nanjing, 210008, China

b. State Key Laboratory of Subtropical Silviculture, Zhejiang A&F University, Hangzhou, 311300, China

c. State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing, 210008, China

e. Center for Molecular Metabolism, School of Environmental and Biological Engineering, Nanjing University of Science and Technology, Nanjing, 210014, China

f. University of the Chinese Academy of Sciences, Beijing, 100049, China

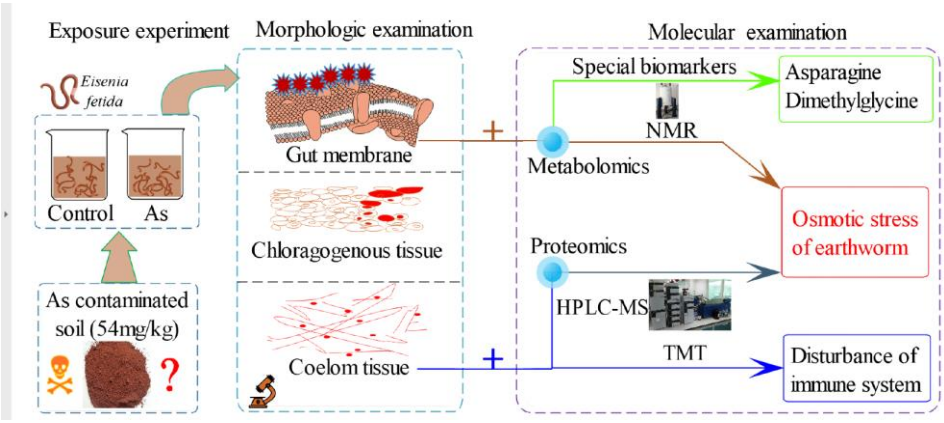
* Corresponding author.

Prof. Xingxiang Wang.

E-mail address: xxwang@issas.ac.cn.

Key Laboratory of Soil Environment and Pollution Remediation, Institute of Soil Science, Chinese Academy of Sciences, Nanjing, 210008, China

Graphical abstract



Highlights

1. Low-level As-contaminated soil induced the upregulation of earthworm profilin.
2. Glycine-tRNA ligase activity of earthworm was enhanced due to the As stress.
3. The osmotic stress of earthworms induced by As was confirmed using dual omics.
4. Asparagine and dimethylglycine were special biomarkers to As contamination.

Abstract

The toxicity of low-level arsenic (As)-contaminated soil is not well understood. An integrated proteomic and metabolomic approach combined with morphological examination was used to investigate the potential biological toxicity of As-contaminated soil based on an exposure experiment with the earthworm *Eisenia fetida*.

The results showed that the earthworm hindgut accumulated high As concentrations resulting in injury to the intestinal epithelia, chloragogenous tissues and coelom tissues. Furthermore, As-contaminated soil induced a significant increase in betaine levels and a decrease in dimethylglycine and myo-inositol levels in the earthworms, suggesting that the osmoregulatory metabolism of the earthworms may have been disturbed. The significantly altered levels of asparagine and dimethylglycine were proposed as potential biomarkers of As-contaminated soil. The upregulation of soluble calcium-binding proteins and profilin, the downregulation of sodium/potassium-transporting ATPase, and the proteins changes identified by gene ontology enrichment analysis confirmed that the earthworms suffered from osmotic stress. In addition, the significant changes in glycine-tRNA ligase activity and coelomic tissue injury revealed that As accumulation may disturb the earthworm immune system. This work provided new insight into the proteomic and metabolic toxicity of low-level As-contaminated soil ecosystems in earthworms, extended our knowledge of dual omics and highlighted the mechanisms underlying toxicity.

Keywords: Quantitative proteomic analysis, Nuclear magnetic resonance, *Eisenia fetida*, Multiomics, As contamination

1. Introduction

Inorganic arsenic (As) is widely distributed in soil and can be transported and bioaccumulated in the soil-plant-animal system via the food web [1]. Moreover, As is classified as a carcinogen by the U.S. Environmental Protection Agency and the International Agency for Research on Cancer [2] due to its cytotoxicity and

genotoxicity [3]. Some previous studies focused on the accumulation, biotransformation and general toxicity (mortality, sublethal effects, etc.) of As in earthworms as well as environmental factors (pH, soil organic matter, etc.) that influence As toxicity [4, 5]. Other studies have focused on the toxicity of As at the DNA level. Exposure to both high-level As-contaminated soil (up to 436 mg/kg) [6] and an environmentally equivalent As solution (2 μM) [7] induced oxidative DNA damage in living organisms such as earthworm (*Lumbricus terrestris*) coelomocytes, mouse cells and humans [8], based on the comet assay and epidemiological analysis [9]. However, nonspecific binding between As and DNA was observed [10], which suggested that As did not directly react with DNA and could indirectly induce gene toxicity [2]. Therefore, there was no explicit consensus on the mechanism of As toxicity at the DNA level.

DNA (gene) changes can lead to consequent protein alteration [11]. Exposure to As solution [12] or As in drinking water [13, 14] could decrease catalase and superoxide dismutase activities and increase malondialdehyde levels. In addition, As could bind to sulfhydryl groups in proteins, such as reduced cysteines, and enable alteration of the structures and functions of these proteins [15-18]. However, the responses of these typical antioxidant enzymes and lipid peroxide products represented only a small portion of the overall proteomic responses, and the increase or decrease in the levels of these proteins could not clearly explain the mechanism of action underlying As toxicity at the protein level. Proteomics has been defined as the large-scale study of the total proteins expressed by the genome of a given organism [19, 20]. Recently, proteomics based on the powerful tandem mass tag (TMT) technology has been widely used in medical [21, 22], agroforestry [23, 24] and environmental [25] research. Therefore, proteomics has potential applications in studies on the toxicity of pollutants at the whole-proteome level.

As exposure has adverse effects on DNA, and the binding mechanism between As and some proteins has also been elucidated. However, it is unclear whether alteration of these DNAs and proteins regulates downstream metabolites. Metabolomics could provide a quantitative understanding of the metabolite complements of integrated living systems [26]. Therefore, metabolomic analysis not only greatly facilitates the understanding of the interactions between different metabolic pathways but also augments and complements genetic and proteomic data (genomics and proteomics) if a combined omics approach is used. Several dual omics (proteomics/metabolomics or transcriptomics/metabolomics) studies have focused on the toxicity of pollutants in aquatic organisms and earthworms. For aquatic organisms, metabolomic analysis revealed that benzo[a]pyrene or dichlorodiphenyltrichloroethane disturbed energy metabolism in the gills of *Pinctada martensii* or the mussel *Perna viridis*, and this disturbance of energy metabolism was confirmed using proteomic analysis [27-29]. In addition, cadmium induces distinct oxidative stress, cellular injury, and disturbance of lipid metabolism in the larvae of the mussel *Mytilus galloprovincialis* based on proteomic and metabolomic analyses [30]. For earthworms, integrated transcriptomic and metabolomic analyses revealed that copper contamination induced overexpression of transcripts of enzymes involved in oxidative phosphorylation and disrupted the energy metabolism of the earthworm *Lumbricus rubellus* [31]. 2,2',4,4'-Tetrabromodiphenyl ether increased the ATP level in the earthworm *Eisenia fetida*, and the upregulation of ATP synthase was confirmed by metabolomic and proteomic analyses [32]. Therefore, combined use of different omics methods revealed disturbances in metabolites, proteins and genes involved in the same metabolic pathways, providing a powerful tool for elucidating the toxicity mechanisms of pollutants [33]. However, different omics methods also provided relatively independent data [30, 31], possibly because most protein analyses use two-dimensional gel electrophoresis (2-DE), which has poor experimental reproducibility and is not amenable to quantitative comparison of different gels or automation [34]. Nevertheless,

the TMT technique could overcome the above issues.

Soil pollution is an urgent problem in China. Approximately 88.9% of As-contaminated soils are associated with low-level pollution [35]. However, the toxicity of low-level As-contaminated soil is poorly understood at the metabolic and proteomic levels. In addition, knowledge of this toxicity will be helpful for the development of government policy and will enhance public awareness. Therefore, the potential toxicity of the low-level As-contaminated soil needs to be elucidated. We hypothesized that the toxicity mechanisms of low-level As-contaminated soil could be effectively revealed by dual proteomic and metabolomic analyses in earthworms using TMT and nuclear magnetic resonance (NMR) technology. Therefore, the aims of the present study were to: (1) determine whether specific protein and metabolite expression signatures could be established for As-contaminated soil using the earthworm *E. fetida*; (2) explore whether specific metabolic disturbances in earthworms could be effectively identified using integrated proteomics and metabolomics; and (3) identify potential biomarkers at the metabolic level for the early detection of toxicity associated with As-contaminated soil.

2. Materials and methods

2.1. Soil and earthworm exposure

Surface soil of dry land (0-20 cm, agri-udic ferrosols, pH=4.84) was collected from Yingtan city (28°12′N, 116°56′E), Jiangxi Province, China. In May 2010, the exposed soils (exposed group) were spiked with 30 mg/kg As ($Na_3AsO_4.12H_2O$, analytical reagent). The control soil (control group) containing deionized water was prepared concurrently. For other details regarding sample preparation, including soil aging (5 years) and storage, earthworm exposure conditions (20 ℃, light intensity 600 lux, 80%

air humidity) and time (two weeks), please refer to our previous research [36]. Earthworm (*E. fetida*) exposure experiments were conducted according to the Organization for Economic Co-operation and Development guidelines in October, 2018. Finally, the measured concentrations of total As in the control and contaminated soils were 16.1 and 53.9 mg/kg, respectively.

2.2. Examination of the total As concentrations in earthworms

After 72 hours of excretion, the earthworm surface was thoroughly cleaned with deionized water (Fig. S1). To characterize the distribution of As in earthworms, the earthworms were divided into three parts, namely, the front, middle and back, based on the demarcation of the clitellum, for determination of the As concentrations. The total As concentrations in the earthworms was detected using X-ray fluorescence spectrometry (XRF, XLt 960, Thermo Scientific, Niton) [37, 38]. Certified reference materials (GBW07401, dark brown soil) were examined concurrently, and the recovery rates were 90.7-103.1%. The coefficient of variation (CV) of total As was less than 5%. The survival rates and fresh weights of the earthworms were also determined. The bioaccumulation factor (BAF) was calculated as the As concentrations in earthworms (dry weight) divided by that in soils.

2.3. Transcriptome examination

Transcriptome sequencing of the earthworm *E. fetida* was conducted to establish a protein database for mass spectrometry-based identification. Total RNA was extracted from earthworms *E. fetida* using TRIzol® reagent according the manufacturer's instructions (Invitrogen, Carlsbad, CA, USA). The integrity and purity of the total RNA was determined by a 2100 Bioanalyzer (Agilent Technologies, Inc., Santa Clara CA, USA), and the RNA was quantified using an ND-2000 instrument (NanoDrop, Thermo

Scientific, Wilmington, DE, USA). Only high-quality RNA samples (OD260/280=1.8~2.2, OD260/230≥2.0, RIN≥8.0, 28S:18S≥1.0, >2 μg) were used to construct a sequencing library. RNA purification, reverse transcription, library construction and sequencing were performed at Shanghai Majorbio Biopharm Biotechnology Co., Ltd. (Shanghai, China) according to the manufacturer's instructions (Illumina, San Diego, CA). Details of the analyses, including earthworm RNA-seq transcriptome library preparation, Illumina HiSeq X Ten sequencing, de novo assembly and annotation, differential expression analysis and functional enrichment, are shown in Text S1 (Supporting Information). Data analysis was performed using the free online Majorbio I-Sanger Cloud Platform (www.i-sanger.com). The protein files from the transcriptome coding sequence prediction results were applied to the following protein query libraries. The transcriptome sequences of *E. fetida* in the exposed and control groups were submitted to the NCBI database (BioProject ID: PRJNA625164).

2.4. Protein examination

Proteins were analyzed using liquid chromatography-tandem mass spectrometry (LC-MS/MS; Thermo Scientific) at Shanghai Applied Protein Technology Co., Ltd., as described below. For protein extraction, SDT butter (4% SDS, 1 mM DTT, 100 mM Tris-HCl, pH=7.6) was added to the earthworm samples, and the samples were then transferred to a 2 mL centrifuge tube with quartz sand. Then, the lysate was homogenized by an MP homogenizer (60 s, twice). The homogenate solution was sonicated and boiled for 15 min. After centrifugation at 14000×g for 40 min, the supernatant was filtered with 0.22 μm filters. The filtrate was quantified with the BCA Protein Assay Kit (Bio-Rad, USA). The sample was stored at -80°C [39]. The extracted proteins were separated on a 12.5% SDS-PAGE (polyacrylamide gel electrophoresis) gel (constant current of 14 mA for 90 min). Protein bands were visualized by Coomassie Blue R-250 staining. After filter-aided sample preparation (FASP digestion) [40], 100

μg of the peptide mixture of each sample was labeled using the TMT reagent according to the manufacturer's instructions (Thermo Fisher Scientific). The Pierce High pH Reversed-Phase Peptide Fractionation Kit (Thermo Scientific) was used to fractionate the TMT-labeled digest samples into 10 fractions by increasing acetonitrile step-gradient elution according to the manufacturer's instructions. Experiments were performed on a Q Exactive mass spectrometer coupled to an Easy nLC (Thermo Fisher Scientific). Each fraction was injected for nano-HPLC-MS/MS analysis. Detailed parameters of the HPLC-MS/MS analysis are shown in the Supporting Information (Text S2).

The differentially expressed proteins were screened based on a change in expression level of more than 1.2- or 0.83-fold (upregulated by more than 1.2-fold or downregulated by less than 0.83-fold) and a statistical significance level of $p \leq 0.05$. To obtain functional information regarding these altered proteins, gene ontology (GO) annotation of the target proteins was conducted using Blast2GO, which can be divided into the following four steps: blast, mapping, annotation notes and annotation augmentation. In addition, GO enrichment analyses were used to explore the impact of differentially expressed proteins on cell physiological processes based on Fisher's exact test. Only functional categories and pathways with p-values less than a threshold of 0.05 were considered significant. The Benjamini-Hochberg correlation was further applied to adjust the derived p-values. The relative expression data of the studied proteins were used to perform hierarchical clustering analysis using Cluster 3.0 (http://bonsai.hgc.jp/~mdehoon/software/cluster/software.htm) and Java Treeview software (http://jtreeview.sourceforge.net).

2.5. Morphological characterization and metabolite examination

Details of the morphological examination (hematoxylin-eosin (HE) analysis) and

NMR-based metabolic analysis were provided in our previous study [36]. Briefly, the metabolic analysis included the extraction, detection and identification of earthworm metabolites and a series of data analyses, such as univariate analysis (Student's *t*-test), multiple-component analysis (including principal component analysis 'PCA', orthogonal signal correction-partial least squares-discriminant analysis 'OSC-PLS-DA', cross-validation and permutation tests, the shared and unique structure plot 'SUS plot' analysis) and metabolic network analysis.

3. Results

3.1. Total As examination and absence of individual injury

No significant changes in survival rates or weights were observed between the control and exposed groups. The As concentrations gradually increased from the fronts to the backs of the earthworms in both the control and exposed groups. The As concentrations in the backs of earthworms in the exposed group was as high as 131.3 mg/kg. The BAF of different parts of earthworms in the control and exposed groups were greater than one (1.06-4.24), except for the front parts of earthworms in the exposed group (Fig. 1). In addition, the BAF of each part in the control group was higher than that in the exposed group.

3.2. Morphological examination and tissue injury

A complete transverse section of the earthworm hindgut is shown in Fig. 2. The control group showed a clear digestive cavity, intestinal epithelium, typhlosole, chloragogenous tissue, coelom cavity and coelom tissue. However, the exposed group showed partial shedding of the intestinal epithelium and chloragogenous tissue. Moreover, the structure of the coelom tissues in the exposed group was not distinct.

3.3. Protein examination and protein toxicity

In the present study, a total of 5457 proteins were identified in earthworms using the TMT-based quantitative proteomics approach (Fig. S2). The tree clustering method showed that 29 upregulated and 29 downregulated differentially expressed proteins were induced in response to As stress compared to the control treatment (Fig. S3). There were many unnamed proteins among the differentially expressed proteins due to incomplete genetic databases. The named upregulated proteins included soluble calcium-binding proteins, profilin, transcription initiation factor and 3-oxoacyl-[acyl-carrier-protein] synthase, and the named downregulated proteins included lacunin, intermediate filament protein, retinol dehydrogenase, transmembrane 9 superfamily member and sodium/potassium-transporting ATPase (Table S1). There were three identified ontologies, namely, biological process, molecular function and cellular component, which included the type and number of proteins (Fig. 3A). Metabolic process in biological process, catalytic activity and binding in molecular function, and membrane, cell and cell part in cellular component had large numbers of proteins (≥ 5). The molecular functions of the differentially expressed proteins were associated with catalytic activity, binding, structural molecule activity and transporter activity. In addition, GO enrichment analysis showed that response to salt stress, hyperosmotic salinity response, glycyl-tRNA aminoacylation and the response to osmotic stress in biological process, potassium-transporting ATPase activity, sodium:potassium exchanging ATPase activity, potassium (sodium) ion transmembrane transporter activity and glycine-tRNA ligase activity in molecular function, and intermediate filament (cytoskeleton) in cellular component had high enrichment factors with strong significance ($p<0.05$) (Fig. 3B).

3.4. Metabolite examination and metabolic toxicity

Univariate analysis of the metabolites was performed using the logarithm value of fold change between the control and exposed groups, as shown in Fig. S4. Among these metabolites, the levels of six metabolites, namely, alanine, asparagine, dimethylamine, betaine, lombricine and ATP significantly increased, whereas those of five metabolites, namely, dimethylglycine, myo-inositol, histidine, NAD^+/$NADP^+$ and NADPH/NADH significantly decreased. To reflect the weight characteristics of changing metabolites, multiple analyses were conducted. Unsupervised PCA showed separate trends for the control and exposed groups (Fig. S5). OSC-PLS-DA further illustrated the clear distinction between the control and exposed groups (Fig. 4A). The cross-validation and permutation test indicated that the OSC-PLS-DA model between the control and exposed groups was reliable, with R^2=0.84, Q^2=0.56 and p=0.006 (Fig. 4E, F). The s-plot showed that dimethylglycine, betaine, lombricine, alanine and asparagine, especially alanine and dimethylglycine, had a higher weight for (contribution to) the separation between the control and exposed groups (Fig. 4B). The loading plot showed that the metabolites with the most significant changes (peaks with red markers) were consistent with the results of the univariate analysis (Fig. 4C, D). In addition, the SUS plot showed that twenty groups of metabolites in the control group had a special correlation, while forty groups of metabolites in the exposed group had a special correlation. There were fourteen groups of metabolites with positive correlations, and a group of metabolites had negative correlations in the control and exposed groups (Fig. 5). These results indicated that earthworm metabolism in the exposed group was different from that in the control group. Metabolic network analysis showed that the osmoregulatory metabolism of earthworms could be disturbed by As-contaminated soil stress (Fig. S6).

4. Discussion

According to the national soil pollution investigation bulletin China, if the As concentrations is 1-3 times of the reference standard, it is generally called low-level pollution. We assumed that this reference standard was the second-grade standard value of the National Soil Environmental Quality Standard of China. The second-grade standard value of As is 40 mg/kg for the tested soil (pH<6.5) in the old version GB 15618-1995, and the risk screening values of As is also 40 mg/kg (pH<5.5) in the latest version GB15618-2018. Therefore, the As contaminated soil in this study was considered as low-level pollution (environmentally equivalent dose). Moreover, this study could represent actual soil pollution and have real environmental significance.

After 14 days of exposure, the survival rates and fresh weights of earthworms in the control and exposed groups showed no significant differences. However, the As concentrations of each part (front, middle and back) in earthworms in the exposed group was higher than that in the earthworm in the control group. Furthermore, the As concentrations in the back parts of the earthworms were approximately 4-fold and 1.7-fold higher than those in the front and middle parts of the earthworms, respectively. However, compared with the exposed group, the higher BAF in the control group may be related to toxic metal accumulation according to a plateau or saturation pattern [41]. Other studies also find that some toxic metals have a lower BAF with higher concentrations in soil [42-44]. Earthworms accumulate pollutants via passive epidermal uptake and active intestinal assimilation [45]. The latter pathway results in the desorption of pollutants adsorbed on soil particles, followed by absorption, transformation and redistribution in the earthworm's internal organs. The heterogeneous distribution of As concentrations in the different parts of the earthworm body suggested that As could accumulate via intestinal assimilation. Accumulation of mercury and cadmium in earthworms was also believed to occur via an intestinal assimilation pathway [36, 46, 47]. It was reported that some metals are not evenly distributed in different parts of earthworms [47, 48], such as Cu (front and middle), Se (front), Pb

(back) and Au (intestinal cavity) [49, 50]. Therefore, different metals may act in combination with specific target tissues [51]. The higher As concentrations in the hindgut of earthworms indicated that the hindgut could be the main target organ (tissue) of As, which may have potential toxicological significance despite the absence of individual level toxicity.

As-contaminated soil damaged the gut coelom tissues and might induce immune disturbance in earthworms. Morphological observations could provide straightforward qualitative evidence regarding tissue injuries in organisms [36, 52]. The morphological injury of chloragogenous tissues, coelom tissues and the intestinal epithelia in the earthworm backs indicated that As-contaminated soil induced toxic responses at the earthworm tissue or organ level. The As sequestration might be associated with the metallothionein-rich chloragogenous tissues [48] distributed in the hindgut of earthworms [53], which is consistent with the high As accumulation in the earthworm hindgut observed in our results. Homeostasis of the immune system is of great significance for the maintenance of organismal metabolic balance [54, 55]. Earthworm coelomocytes (immune cells) are part of the coelom tissue and circulate in fluid suspension in the coelom cavity [56-58]. It has been reported that coelomocytes are involved in the immune responses of the earthworm *L. rubellus* [59, 60]. Glycine-tRNA ligase (synthetase), which is known to interact with immune cells in organisms and boost the immune response [61], showed a high enrichment factor in the GO functional enrichment analysis (Fig. 3). In addition, As also causes the disturbances of the immune system in other organisms, such as juvenile mussels [62]. Therefore, the damaged coelom tissues (without distinct structures) indicated that the immune system of earthworms could be disturbed due to the stress of As accumulation. The lacunin protein modulated cell rearrangement, cell migration, and tissue remodeling [63], and the expression of this protein occurred mainly in the epithelial regions. The intermediate filament proteins were the main components of the cytoskeleton. The significant

downregulation of lacunin and intermediate filament protein indicated that As-contaminated soil could inhibit the expression of these proteins due to the injury of epithelial cells.

As-contaminated soil induced osmotic and oxidative stress in earthworms. Profilin is a membrane protein that plays an important role in growth of the cytoskeleton and is known to be involved in oxidative stress [64]. Combined with the injury of earthworm gut epithelial cells, the significant upregulation of profilin could be associated with the oxidative stress of the earthworm gut membrane induced by As-contaminated soil stress. Upregulation of profilin was also observed in the mussel *M. galloprovincialis* under Cd stress [65]. Furthermore, profilin always bound some variants of membrane phospholipids, e.g., inositol triphosphate, that were involved in the regulation of calcium signals. The expression of soluble calcium-binding proteins that regulate the cellular calcium ion concentrations [66] significantly increased under As-contaminated soil stress. Therefore, the upregulation of profilin could be associated with the disturbance of the osmotic balance of the gut membrane. With regard to osmolytes, a significant decrease in the levels of myo-inositol and dimethylglycine and a significant increase in the levels of betaine and lombricine (a metabolite associated with membrane composition) directly indicated that the osmoregulatory metabolism of earthworms was disturbed by As-contaminated soil stress at the metabolic level. Moreover, the response to osmotic stress, the hyperosmotic response and the hyperosmotic salinity responses of earthworms exhibited high enrichment factors (0.33, 0.33 and 1, respectively) based on the GO functional enrichment analysis (Fig. 3), which further confirmed the osmotic stress of earthworms at the protein level. In addition, the significant downregulation of potassium/sodium-transporting ATPase activity ($p<0.001$) indirectly indicated the disturbance of gut membrane permeability. The osmotic stress induced by As was also observed in the clam *Ruditapes philippinarum* [67]. Therefore, the osmotic stress induced by As-contaminated soil in

earthworms was confirmed based on integrated proteomic and metabolomic data. Increased ATP levels could be an indicator of increased compensation for energy consumption during osmotic metabolism and oxidative stress (Fig. S6). Lombricine kinase plays a key role in the coupling of energy production and energy utilization in organisms [68] and is responsible for regulating cellular ATP levels [69]. The significant increase in lombricine and ATP levels indicated that the activity of lombricine kinase could be activated by the stress of As-contaminated soil. However, this needs to be confirmed in future studies. In addition, the predicted overall metabolic and protein response pattern of earthworms is depicted in Fig. 6. This pattern could improve our understanding of the mechanisms that underlie the molecular toxicity of low-level As-contaminated soil toward earthworms.

As induced the production of specific metabolites in different organisms. Notably, an alteration in the amount of protein involved in one metabolic pathway does not always result in a simultaneous alteration of metabolites because metabolites can be involved in several metabolic pathways [70]. The decrease in histidine levels and increase in betaine levels induced by As in the earthworm *E. fetida* were consistent with the trends observed in the clam *R. philippinarum*, while the changes in alanine and ATP levels were opposite between earthworms and clams [67]. The levels of some metabolites, such as asparagine, NADPH and dimethylglycine, specifically changed in earthworms. However, succinate and fumarate exhibited specific changes in clams under As stress [67]. In addition, As significantly disturbed the metabolomic profiles of mice, including the levels of amino acid derivatives, carnitine, fatty acids and glucuronide, via exposure to contaminated drinking water [71]. These results revealed that different organisms could generate either several specific metabolites or a few shared metabolites in response to As stress. Therefore, improved descriptions of the relationships between specific metals and the metabolic responses of specific organisms would improve our understanding of the mechanisms underlying the toxicity of metals. Furthermore,

metabolomics was used to identify target metabolites as potential biomarkers. Compared to our previous studies on low-level Pb-, Cd-, Hg- and Cr-contaminated soils [36, 52, 72, 73], specific changes in the asparagine and dimethylglycine levels of earthworms could be recommended as potential biomarkers of As-contaminated soil (Table S2).

Recent research has indicated that the major form of As in earthworm body tissues is As (Ⅲ) (> 77.9%), although the soil is spiked using Na_3AsO_4 (Ⅴ) [44]. Unfortunately, As species in soil and earthworms were not tested in the present study because the risk screening values for soil contamination of agricultural land in China only involved the total As concentrations and did not involve As species. However, different As species might affect on proteinic and motabolic toxicity of earthworms, thus we will strengthen related research in the future.

5. Conclusion

In summary, As mainly accumulated in the earthworm hindgut by the intestinal assimilation. As a result, high As accumulation in the hindgut might not only induce osmotic stress in earthworms, as indicated by a significant increase in alanine, asparagine, ATP and betaine levels and a decrease in dimethylglycine, myo-inositol, histidine and NAD^+/$NADP^+$ levels, but also might induce immune disturbance in earthworms, as indicated by damaged gut coelom tissues and significant enhancement of glycine-tRNA ligase activity. Furthermore, osmotic stress in the earthworms was confirmed using proteomics based on the upregulation of soluble calcium-binding proteins and profilin and downregulation of sodium/potassium-transporting ATPase. Asparagine and dimethylglycine were recommended as specific metabolic biomarkers of As contamination. For future research, the As species in earthworms and soil should

be characterized to further predict and explore which As species result in systematic changes among the genes, proteins and metabolites of earthworms using a novel three-dimensional spatial network-based visual analytical system [74] (http://www.omicsnet.ca./OmicsNet/faces/home.xhtml).

Credit author statement

RG Tang and XX Wang conceived the study and designed the experiments. RG Tang did the experiment, analyzed the data and made the figures with the helps of JS Wang, P Lan, CF Ding and TL Zhang. RG Tang wrote the manuscript. All authors discussed the results and commented on the manuscript.

Conflicts of intersect

There are no conflicts of interest to declare.

Acknowledgments

This research was supported by the National Key Research and Development Program of China (2016YFD0800400), the Natural Science Foundation of Zhejiang Province (LQ20D010003) and the National Natural Science Foundation of China (41501347). We are grateful to three anonymous reviewers for their insightful comments and suggestions for improving this manuscript.

References

[1] F. Yang, S. Xie, C. Wei, J. Liu, H. Zhang, T. Chen, J. Zhang, Arsenic characteristics in the terrestrial environment in the vicinity of the Shimen realgar mine, China, Science of the Total Environment, 626 (2018) 77-86.

[2] S. Shen, X.F. Li, W.R. Cullen, M. Weinfeld, X.C. Le, Arsenic binding to proteins, Chemical

Reviews, 113 (2013) 7769-7792.

[3] P. Bhattacharya, A.H. Welch, K.G. Stollenwerk, M.J. Mclaughlin, J. Bundschuh, G. Panaullah, Arsenic in the environment: biology and chemistry, Science of the Total Environment, 379 (2007) 109-120.

[4] A.A. Meharg, R.F. Shore, K. Broadgate, Edaphic factors affecting the toxicity and accumulation of arsenate in the earthworm *Lumbricus terrestris*, Environmental Toxicology and Chemistry, 17 (1998) 1124-1131.

[5] M. Button, G.R.T. Jenkin, C.F. Harrington, M.J. Watts, Arsenic biotransformation in earthworms from contaminated soils, Journal of Environmental Monitoring, 11 (2009) 1484-1491.

[6] M. Button, G.R.T. Jenkin, K.J. Bowman, C.F. Harrington, T.S. Brewer, G. Jones, M.J. Watts, DNA damage in earthworms from highly contaminated soils: assessing resistance to arsenic toxicity by use of the Comet assay, Mutation Research-genetic Toxicology and Environmental Mutagenesis, 696 (2010) 95-100.

[7] J. Bach, A. Sampayoreyes, R. Marcos, A. Hernandez, Ogg1 genetic background determines the genotoxic potential of environmentally relevant arsenic exposures, Archives of Toxicology, 88 (2013) 585-596.

[8] T.K. Hei, M. Filipic, Role of oxidative damage in the genotoxicity of arsenic, Free Radical Biology and Medicine, 37 (2004) 574-581.

[9] K. Engstrom, M. Vahter, G.M. Johansson, C.H. Lindh, F. Teichert, R. Singh, M. Kippler, B. Nermell, R. Raqib, U. Stromberg, Chronic exposure to cadmium and arsenic strongly influences concentrations of 8-oxo-7,8-dihydro-2'-deoxyguanosine in urine, Free Radical Biology and Medicine, 48 (2010) 1211-1217.

[10] K.T. Kitchin, K. Wallace, The role of protein binding of trivalent arsenicals in arsenic carcinogenesis and toxicity, Journal of Inorganic Biochemistry, 102 (2008) 532-539.

[11] J.C. Lindon, NMR-based metabolic phenotyping techniques and applications, Springer International Publishing, 2017.

[12] S. Altikat, K. Uysal, H.I. Kuru, M. Kavasoglu, G.N. Ozturk, A. Kucuk, The effect of arsenic on some antioxidant enzyme activities and lipid peroxidation in various tissues of mirror carp (*Cyprinus carpio carpio*), Environmental Science and Pollution Research, 22 (2015) 3212-3218.

[13] S. Sannadi, P.K. Kadeyala, R.R. Gottipolu, Reversal effect of monoisoamyl dimercaptosuccinic acid (MiADMSA) for arsenic and lead induced perturbations in apoptosis and antioxidant enzymes in developing rat brain, International Journal of Developmental Neuroscience, 31 (2013) 586-597.

[14] P.K. Kadeyala, S. Sannadi, R.R. Gottipolu, Alterations in apoptotic caspases and antioxidant enzymes in arsenic exposed rat brain regions: Reversal effect of essential metals and a chelating agent, Environmental Toxicology & Pharmacology, 36 (2013) 1150-1166.

[15] K. Marapakala, C. Packianathan, A.A. Ajees, D.S. Dheeman, B. Sankaran, P. Kandavelu,

B.P. Rosen, A disulfide-bond cascade mechanism for arsenic(III) S-adenosylmethionine methyltransferase, Journal of Electronic Materials, 67 (2016) 571-586.

[16] A.A. Ajees, K. Marapakala, C. Packianathan, B. Sankaran, B.P. Rosen, Structure of an As(III) S-adenosylmethionine methyltransferase: Insights into the mechanism of arsenic biotransformation, Biochemistry, 51 (2012) 5476-5485.

[17] M. Schuliga, S. Chouchane, E.T. Snow, Upregulation of glutathione-related genes and enzyme activities in cultured human cells by sublethal concentrations of inorganic arsenic, Toxicological Sciences An Official Journal of the Society of Toxicology, 70 (2002) 183-192.

[18] H. Castillomichel, J. Hernandezviezcas, K.M. Dokken, M.A. Marcus, J.R. Peraltavidea, J.L. Gardeatorresdey, Localization and speciation of arsenic in soil and desert plant *Parkinsonia florida* using μXRF and μXANES, Environmental Science & Technology, 45 (2011) 7848-7854.

[19] S. Naabyhansen, M.D. Waterfield, R. Cramer, Proteomics - post-genomic cartography to understand gene function, Trends in Pharmacological Sciences, 22 (2001) 376-384.

[20] N.L. Anderson, N.G. Anderson, Proteome and proteomics: new technologies, new concepts, and new words, Electrophoresis, 19 (2010) 1853-1861.

[21] R. Lin, Y. Mo, H. Zha, Z. Qu, P. Xie, Z.J. Zhu, Y. Xu, Y. Xiong, K.L. Guan, CLOCK acetylates ASS1 to drive circadian rhythm of ureagenesis, Molecular Cell, 68 (2017) 198-209.

[22] X. Zhou, L. Liu, X. Lan, D. Cohen, Y. Zhang, A.V. Ravindran, S. Yuan, P. Zheng, D.

Coghill, L. Yang, Polyunsaturated fatty acids metabolism, purine metabolism and inosine as potential independent diagnostic biomarkers for major depressive disorder in children and adolescents, Molecular Psychiatry, (2018).

[23] X. Cao, D. Song, M. Yang, N. Yang, Q. Ye, D. Tao, B. Liu, R. Wu, X. Yue, Comparative analysis of whey N-glycoproteins in human colostrum and mature milk using quantitative glycoproteomics, Journal of Agricultural & Food Chemistry, 65 (2017) 10360-10367.

[24] Z. Ma, L. Zhu, T. Song, Y. Wang, Q. Zhang, Y. Xia, M. Qiu, Y. Lin, H. Li, L. Kong, A paralogous decoy protects *Phytophthora sojae* apoplastic effector PsXEG1 from a host inhibitor, Science, 355 (2017) 710-714.

[25] P. Zhang, X.Y. Xu, Y.P. Chen, M.Q. Xiao, B. Feng, K.X. Tian, Y.H. Chen, Y.Z. Dai, Protein corona between nanoparticles and bacterial proteins in activated sludge: Characterization and effect on nanoparticle aggregation, Bioresource Technology, 250 (2017) 10-16.

[26] J.K. Nicholson, J.C. Lindon, Systems biology: Metabonomics, Nature, 455 (2008) 1054-1056.

[27] H. Chen, Q. Song, X. Diao, H. Zhou, Proteomic and metabolomic analysis on the toxicological effects of Benzo[a]pyrene in pearl oyster *Pinctada martensii*, Aquatic Toxicology, 175 (2016) 81-89.

[28] H. Chen, X. Diao, H. Wang, H. Zhou, An integrated metabolomic and proteomic study of toxic effects of Benzo[a]pyrene on gills of the pearl oyster *Pinctada martensii*, Ecotoxicology &

Environmental Safety, 156 (2018) 330-336.

[29] Q. Song, H. Chen, Y. Li, H. Zhou, Q. Han, X. Diao, Toxicological effects of benzo(a)pyrene, DDT and their mixture on the green mussel Perna viridis revealed by proteomic and metabolomic approaches, Chemosphere, 144 (2016) 214-224.

[30] H. Wu, L. Xu, C. Ji, D. Yu, Proteomic and metabolomic responses in D-shape larval mussels *Mytilus galloprovincialis* exposed to cadmium and arsenic, Fish & Shellfish Immunology, 58 (2016) 514-520.

[31] J.G. Bundy, J.K. Sidhu, F. Rana, D.J. Spurgeon, C. Svendsen, J.F. Wren, S.R. Sturzenbaum, A.J. Morgan, P. Kille, 'Systems toxicology' approach identifies coordinated metabolic responses to copper in a terrestrial non-model invertebrate, the earthworm *Lumbricus rubellus*, BMC Biology, 6 (2008) 1-21.

[32] C. Ji, H. Wu, L. Wei, J. Zhao, H. Lu, J. Yu, Proteomic and metabolomic analysis of earthworm *Eisenia fetida* exposed to different concentrations of 2,2 ',4,4 '-tetrabromodiphenyl ether, Journal of Proteomics, 91 (2013) 405-416.

[33] A. D'Alessandro, C. Marrocco, V. Zolla, M. D'Andrea, L. Zolla, Meat quality of the longissimus lumborum muscle of Casertana and Large White pigs: metabolomics and proteomics intertwined, Journal of Proteomics, 75 (2012) 610-627.

[34] A.J. Thompson, J. Schafer, K. Kuhn, S. Kienle, J. Schwarz, G. Schmidt, Neumann, Thomas, C. Hamon, Tandem mass tags: A novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS, Analytical Chemistry, 75 (2003) 1895-1904.

[35] Ministry of Environmental Protection and Land and Resources People's Republic of China, National soil pollution investigation bulletin China (In chinese) (2014). (http://www.gov.cn/xinwen/2014-04/17/content_2661765.htm)

[36] R. Tang, C. Ding, F. Dang, Y. Ma, J. Wang, T. Zhang, X. Wang, NMR-based metabolic toxicity of low-level Hg exposure to earthworms, Environmental Pollution, 239 (2018) 428-437.

[37] H.J. Sánchez, C. Pérez, R.D. Pérez, XRF analysis of arsenic uptaking in mice organs and tissues, Nuclear Instruments & Methods in Physics Research, 124 (1997) 140-142.

[38] R.C. Studinski, F.E. Mcneill, D.R. Chettle, J.M. O'Meara, Estimation of a method detection limit for an in vivo XRF arsenic detection system, Physics in Medicine & Biology, 50 (2005) 521-530.

[39] Y. Zhu, H. Xu, H. Chen, J. Xie, M. Shi, B. Shen, X. Deng, C. Liu, X. Zhan, C. Peng, Proteomic analysis of solid pseudopapillary tumor of the pancreas reveals dysfunction of the endoplasmic reticulum protein processing pathway, Molecular & Cellular Proteomics, 13 (2014) 2593-2603.

[40] J.R. Wisniewski, A. Zougman, N. Nagaraj, M. Mann, Universal sample preparation method for proteome analysis, Nature Methods, 6 (2009) 359-362.

[41] A.L.M. Augustsson, T.E. Uddh-Soederberg, K.J. Hogmalm, M.E.M. Filipsson, Metal uptake by homegrown vegetables-the relative importance in human health risk assessments at contaminated sites, Environmental Research, 138 (2015) 181-190.

[42] C. Ding, Y. Ma, X. Li, T. Zhang, X. Wang, Determination and validation of soil thresholds for cadmium based on food quality standard and health risk assessment, Science of the Total Environment, 619-620 (2018) 700-706.

[43] S.K. Gaw, N.D. Kim, G.L. Northcott, A.L. Wilkins, G. Robinson, Uptake of ΣDDT, arsenic, cadmium, copper, and lead by lettuce and radish grown in contaminated horticultural soils, Journal of Agricultural & Food Chemistry, 56 (2008) 6584-6593.

[44] H. Wang, D. Zhu, G. Li, F. Zheng, J. Ding, P.J. O'Connor, Y. Zhu, X. Xue, Effects of arsenic on gut microbiota and its biotransformation genes in earthworm *Metaphire sieboldi*, Environmental Science & Technology, 53 (2019) 3841-3849.

[45] T. Jager, R. Baerselman, E. Dijkman, A.C. de Groot, E.A. Hogendoorn, J.A. De, J.A. Kruitbosch, W.J. Peijnenburg, Availability of polycyclic aromatic hydrocarbons to earthworms (*Eisenia andrei*, Oligochaeta) in field-polluted soils and soil-sediment mixtures, Environmental Toxicology & Chemistry, 22 (2010) 767-775.

[46] F. Dang, J. Zhao, D. Zhou, Uptake dynamics of inorganic mercury and methylmercury by the earthworm *Pheretima guillemi*, Chemosphere, 144 (2016) 2121-2126.

[47] A.J. Morgan, J. Mosselmans, J. Charnock, A. Bennett, C. Winters, M. O'Reilly, P. Fisher, J. Andre, M. Turner, P. Gunning, In situ metal imaging and Zn ligand-speciation in a soil-dwelling sentinel: complementary electron microprobe and synchrotron microbeam X-ray analyses, Environmental Science & Technology, 47 (2012) 1073-1081.

[48] A.J. Morgan, P. Kille, A. Bennett, M. O'Reilly, P. Fisher, J.M. Charnock, Pb and Zn

imaging and in situ speciation at the geogenic/biogenic interface in sentinel earthworms using electron microprobe and synchrotron micro-focus X-ray spectroscopy, Environmental Pollution, 173 (2013) 68-74.

[49] J.M. Unrine, S.E. Hunyadi, O.V. Tsyusko, W. Rao, W.A. Shoults-Wilson, P.M. Bertsch, Evidence for bioavailability of Au nanoparticles from soil and biodistribution within earthworms (*Eisenia fetida*), Environmental Science & Technology, 44 (2010) 8308-8313.

[50] J. Yuan, L. Liqiang, Distribution analysis of K, Ca, Cu, Zn and Pb in earthworm by in situ synchrotron radiation micro X-ray fluorescence (In Chinese), Chinese Journal of Analytical Chemistry, 44 (2016) 792-798.

[51] C.F. Calhôa, M.S. Monteiro, A.M. Soares, R.M. Mann, The influence of metal speciation on the bioavailability and sub-cellular distribution of cadmium to the terrestrial isopod, *Porcellio dilatatus*, Chemosphere, 83 (2011) 531-537.

[52] R. Tang, C. Ding, Y. Ma, J. Wang, T. Zhang, X. Wang, Time-dependent responses of earthworms to soil contaminated with low levels of lead as detected using ^{1}H NMR metabolomics, RSC Advances, 7 (2017) 34170-34181.

[53] A. Franchini, M. Marchetti, The effects of okadaic acid on *Enchytraeus crypticus* (Annelida: Oligochaeta), Invertebrate Survival Journal, 3 (2006) 111-117.

[54] S.H. Gokhan, Inflammation and metabolic disorders, Nature, 444 (2006) 860-867.

[55] R.L. Greer, X. Dong, A.C. Moraes, R.A. Zielke, G.R. Fernandes, E. Peremyslova, S.

Vasquezperez, A.A. Schoenborn, E.P. Gomes, A.C. Pereira, Akkermansia muciniphilamediates negative effects of IFNγ on glucose metabolism, Nature Communications, 7 (2016) 13329.

[56] J. Homa, E. Olchawa, S.R. Stürzenbaum, A.J. Morgan, B. Plytycz, Early-phase immunodetection of metallothionein and heat shock proteins in extruded earthworm coelomocytes after dermal exposure to metal ions, Environmental Pollution, 135 (2005) 275-280.

[57] E. Olchawa, M. Bzowska, S.R. Stürzenbaum, A.J. Morgan, B. Plytycz, Heavy metals affect the coelomocyte-bacteria balance in earthworms: environmental interactions between abiotic and biotic stressors, Environmental Pollution, 142 (2006) 373-381.

[58] E. Kauschkea, E.L. Cooper, Coelomic fluid proteins as basic components of innate immunity in earthworms, European Journal of Soil Biology, 43 (2007) S110-S115.

[59] A.J. Morgan, S.R. Stürzenbaum, C. Winters, G.W. Grime, N.A. Aziz, P. Kille, Differential metallothionein expression in earthworm (*Lumbricus rubellus*) tissues, Ecotoxicology & Environmental Safety, 57 (2004) 11-19.

[60] S.R. Stürzenbaum, O. Georgiev, A.J. Morgan, P. Kille, Cadmium detoxification in earthworms: from genes to cells, Environmental Science & Technology, 38 (2004) 6283-6289.

[61] M.C. Park, T. Kang, D. Jin, J.M. Han, S.B. Kim, Y.J. Park, K. Cho, Y.W. Park, M. Guo, W. He, Secreted human glycyl-tRNA synthetase implicated in defense against ERK-activated tumorigenesis, Proceedings of the National Academy of Sciences of the United States of

America, 109 (2012) 4035-4036.

[62] D. Yu, C. Ji, J. Zhao, H. Wu, Proteomic and metabolomic analysis on the toxicological effects of As (III) and As (V) in juvenile mussel *Mytilus galloprovincialis*, Chemosphere, 150 (2016) 194-201.

[63] J.B. Nardi, R. Martos, K.K. Walden, D.J. Lampe, H.M. Robertson, Expression of lacunin, a large multidomain extracellular matrix protein, accompanies morphogenesis of epithelial monolayers in *Manduca sexta*, Insect Biochemistry & Molecular Biology, 29 (1999) 883-897.

[64] C. Ji, H. Wu, L. Wei, J. Zhao, iTRAQ-based quantitative proteomic analyses on the gender-specific responses in mussel *Mytilus galloprovincialis* to tetrabromobisphenol A, Aquatic Toxicology, 157 (2014) 30-40.

[65] L. Xu, X. Peng, D. Yu, C. Ji, J. Zhao, H. Wu, Proteomic responses reveal the differential effects induced by cadmium in mussels *Mytilus galloprovincialis* at early life stages, Fish & Shellfish Immunology, 55 (2016) 510-515.

[66] P. Thiruketheeswaran, E. Kiehl, J. D'Haese, Soluble calcium-binding proteins (SCBPs) of the earthworm *Lumbricus terrestris*: molecular characterization and localization by FISH in muscle and neuronal tissue, Histochemistry & Cell Biology, 146 (2016) 1-10.

[67] H. Wu, X. Zhang, Q. Wang, L. Li, C. Ji, X. Liu, J. Zhao, X. Yin, A metabolomic investigation on arsenic-induced toxicological effects in the clam *Ruditapes philippinarum* under different salinities, Ecotoxicology & Environmental Safety, 90 (2013) 1-6.

[68] W.R. Ellington, Evolution and physiological roles of phosphagen systems, Annual Review of Physiology, 63 (2001) 289-325.

[69] D.J. Bush, O. Kirillova, S.A. Clark, O. Davulcu, F. Fabiola, Q. Xie, T. Somasundaram, W.R. Ellington, M.S. Chapman, The structure of lombricine kinase: implications for phosphagen conformational changes, Journal of Biological Chemistry, 286 (2011) 9338-9350.

[70] C. Ji, H. Wu, L. Wei, J. Zhao, J. Yu, Proteomic and metabolomic analysis reveal gender-specific responses of mussel *Mytilus galloprovincialis* to 2,2',4,4'-tetrabromodiphenyl ether (BDE 47), Aquatic Toxicology, 140 (2013) 449-457.

[71] K. Lu, R. Abo, K.A. Schlieper, M.E. Graffam, S.S. Levine, J.S. Wishnok, J.A. Swenberg, S.R. Tannenbaum, J.G. Fox, Arsenic exposure perturbs the gut microbiome and its metabolic profile in mice: an integrated metagenomics and metabolomics analysis, Environmental Health Perspectives, 122 (2014) 284-291.

[72] R. Tang, C. Ding, Y. Ma, J. Wang, T. Zhang, X. Wang, Metabolic responses of *Eisenia Fetida* to individual Pb and Cd contamination in two types of soils, Scientific Reports, 7 (2017) 13110.

[73] R. Tang, X. Li, Y. Mo, Y. Ma, C. Ding, J. Wang, T. Zhang, X. Wang, Toxic responses of metabolites, organelles and gut microorganisms of *Eisenia fetida* in a soil with chromium contamination, Environmental Pollution, 251 (2019) 910-920.

[74] G. Zhou, J. Xia, OmicsNet: a web-based tool for creation and visual analysis of biological networks in 3D space, Nucleic Acids Research, 46 (2018) W514-W522.

Figure captions

Fig. 1. Arsenic concentrations and bioaccumulation factor (BAF) of the front, middle and back parts of earthworms between the control and exposed groups.

Fig. 2. Hematoxylin-eosin (HE) staining analysis of the control and exposed groups. Panel (A) shows a clear transverse section of the earthworm gut, including the digestive cavity (Dc), intestinal epithelium (Ie), typhlosole (Ty), chloragogenous tissue (Ct), coelom cavity (Cc), and coelom tissue (Cot). Panel (B) indicates the damaged Ct (red arrows), Ie (green arrow) and Cot (azury arrows) in the exposed group.

Fig. 3. Panel A shows the number of differentially expressed proteins in the biological process, molecular function and cellular component categories. Panel B shows the GO functional enrichment analysis of the three ontologies, namely, biological process, molecular function and cellular component. The ordinate shows the number of differentially expressed proteins in each GO functional classification. The colors of the bar graph based on the adjusted Benjamini-Hochberg method indicate the significance of the GO functional classification. The color gradient represents the *p* value. The color gradient changes from green to red. The closer the color is to red, the higher the significance level. The enrichment factor ($\leq$1) is shown above the bar chart. The enrichment factor indicates the proportion of differentially expressed proteins that account for all identified proteins annotated to a GO functional classification.

Fig. 4. Orthogonal signal correction-partial least squares-discriminant analysis (OSC-PLA-DA) of the NMR data of earthworm metabolites after exposure to low-level arsenic-contaminated soil. (A) Score plot where each point represents one sample and each ellipse corresponds to a confidence interval of 95%. (B) S-plot where points with different colors and shapes represent different variables (metabolites); the farther away from the center a variable is, the higher contribution of the variable to the separation of the groups. (C and D) Corresponding loading plots (0.70-4.45 and 5.38-9.40 ppm) color coded according to the correlation coefficients from blue (low coefficients) to red (high coefficients). Positive and negative peaks indicate decreased and increased metabolite levels in the treated groups. (E and F) Scatter plots of statistical cross-validation and histograms for permutation test scores of OSC-PLA-DA models. The red arrow indicates the performance based on the original labels, significant for a p value < 0.05. The R^2 (total explained variation) and Q^2 (predictability of the model) values were used to confirm the validity of these models.

Fig. 5. SUS plot of earthworm metabolites in the control and exposed (As) groups. All the shared metabolites: the metabolites scattered across the red dotted line represent those that were positively correlated (++/−−), while the metabolites across the green dotted line represent those that were negatively correlated (+−/−+).The metabolites in the red/green boxes across the plot axes are either positive or negative for that particular model and represent the “unique structures”.

Fig. 6. Sketch of the toxicity responses of earthworms in low-level arsenic-

contaminated soil. The blue and red text represents the downregulation and upregulation of metabolites and proteins under arsenic stress.

Fig. 1.

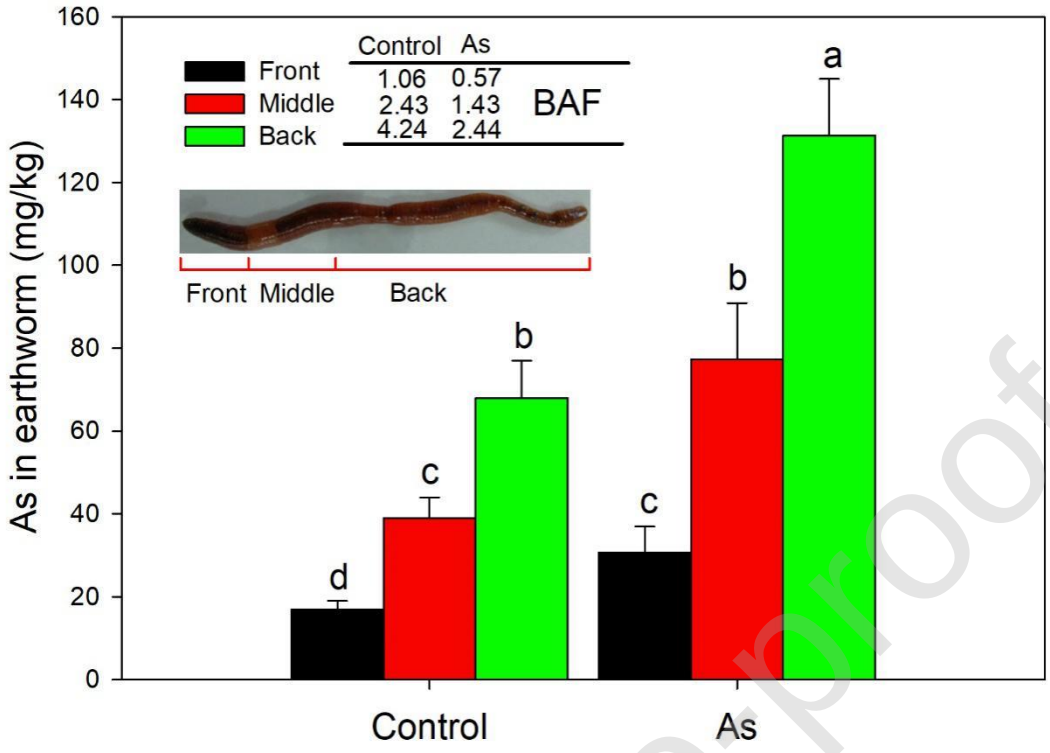


Fig. 2.

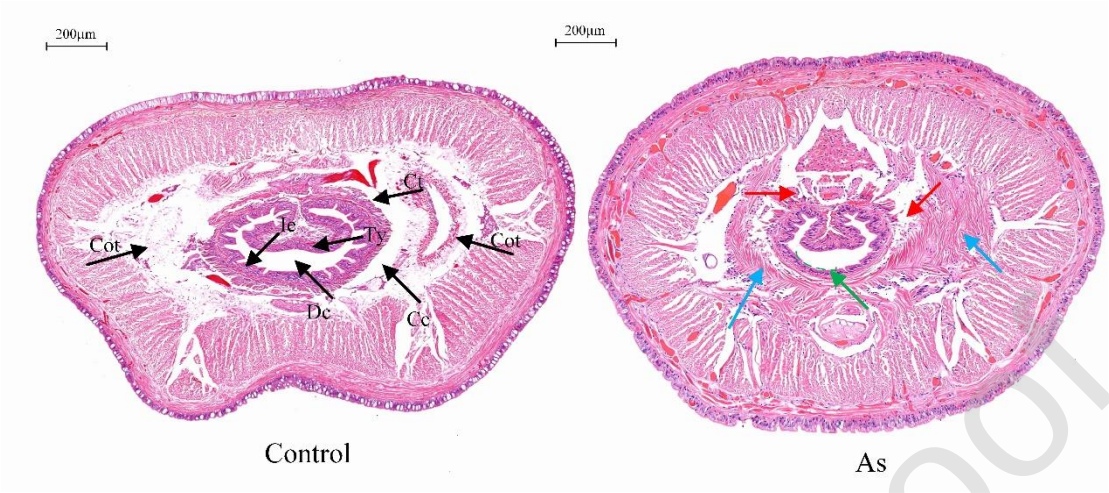


Fig. 3.

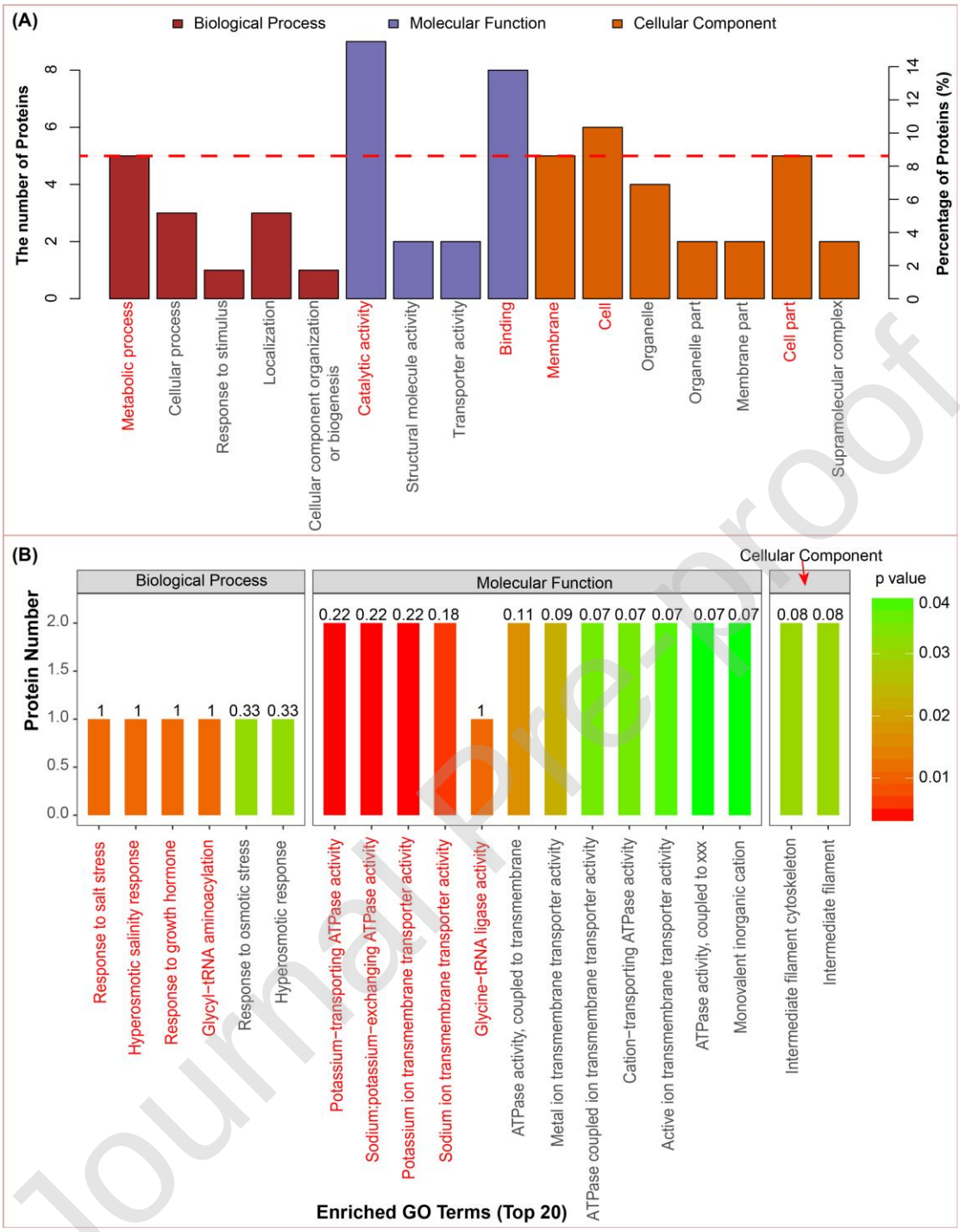


Fig. 4.

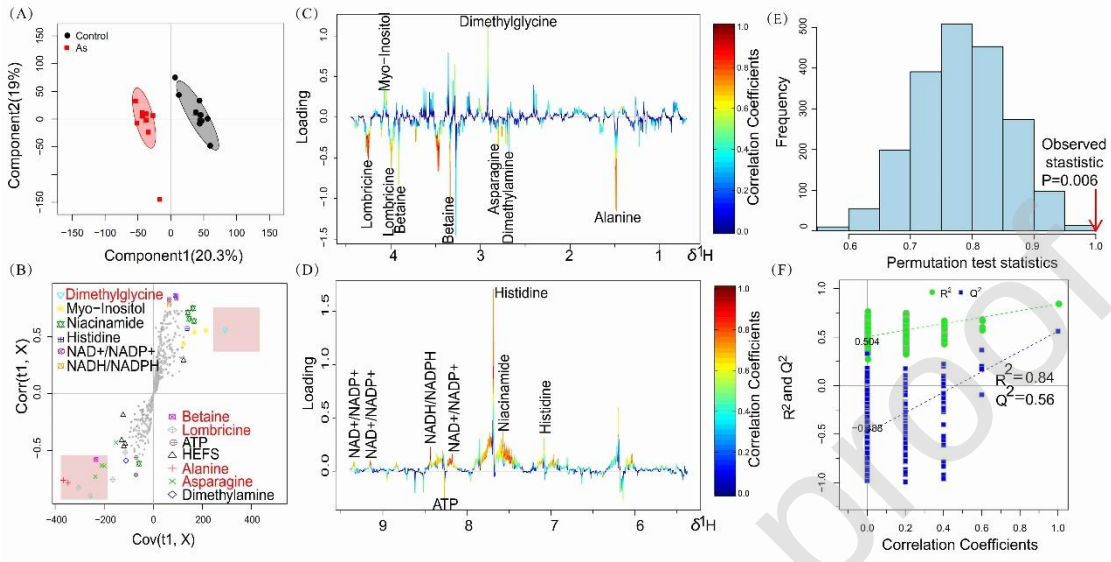


Fig. 5.

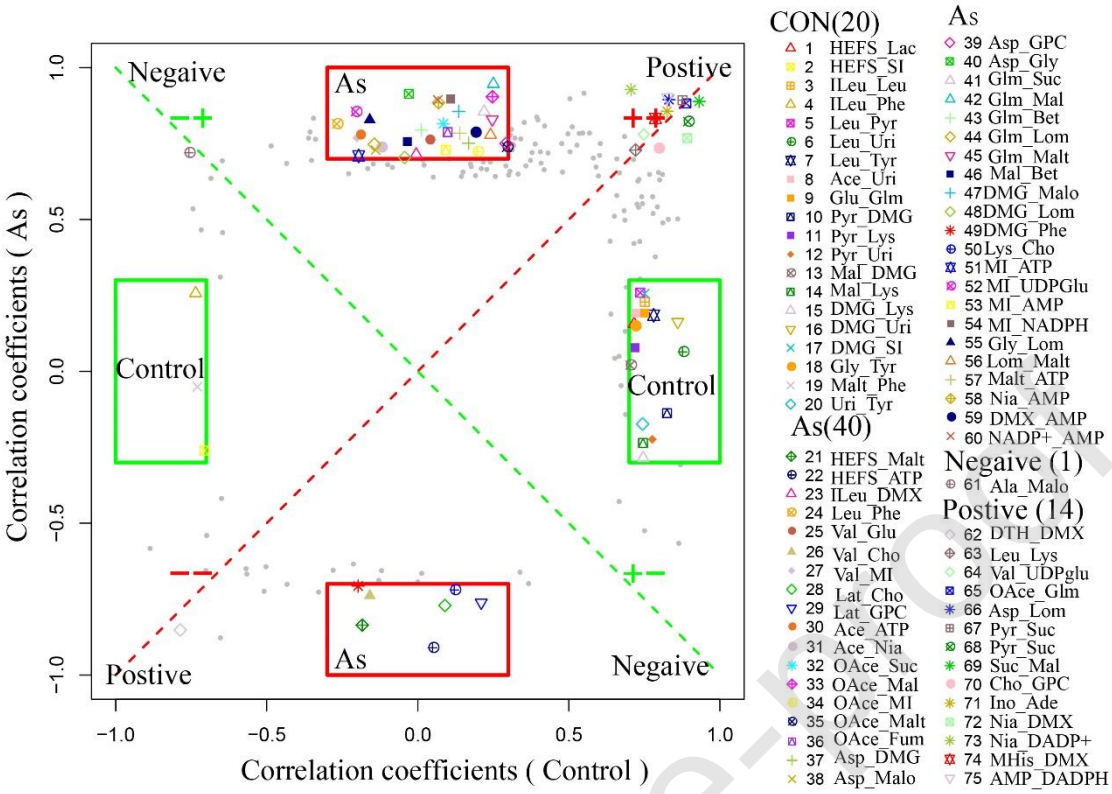


Fig. 6.