

Article

Adverse Effects of Arsenic Uptake in Rice Metabolome and Lipidome Revealed by Untargeted Liquid Chromatography Coupled to Mass Spectrometry (LC-MS) and Regions of Interest Multivariate Curve Resolution

Miriam Pérez-Cova ^{1,2}, Romà Tauler ¹  and Joaquim Jaumot ^{1,*} 

¹ Department of Environmental Chemistry, Instituto de Diagnóstico Ambiental y Estudios del Agua, Jordi Girona 18-26, 08034 Barcelona, Spain; mpcqam@idaea.csic.es (M.P.-C.); rtaqam@idaea.csic.es (R.T.)

² Department of Chemical Engineering and Analytical Chemistry, University of Barcelona, Diagonal 647, 08028 Barcelona, Spain

* Correspondence: joaquim.jaumot@idaea.csic.es

Abstract: Rice crops are especially vulnerable to arsenic exposure compared to other cereal crops because flooding growing conditions facilitates its uptake. Besides, there are still many unknown questions about arsenic's mode of action in rice. Here, we apply two untargeted approaches using liquid chromatography coupled to mass spectrometry (LC-MS) to unravel the effects on rice lipidome and metabolome in the early stages of growth. The exposure is evaluated through two different treatments, watering with arsenic-contaminated water and soil containing arsenic. The combination of regions of interest (ROI) and multivariate curve resolution (MCR) strategies in the ROIMCR data analyses workflow is proposed and complemented with other multivariate analyses such as partial least square discriminant analysis (PLS-DA) for the identification of potential markers of arsenic exposure and toxicity effects. The results of this study showed that rice metabolome (and lipidome) in root tissues seemed to be more affected by the watering and soil treatment. In contrast, aerial tissues alterations were accentuated by the arsenic dose, rather than with the watering and soil treatment itself. Up to a hundred lipids and 40 metabolites were significantly altered due to arsenic exposure. Major metabolic alterations were found in glycerophospholipids, glycerolipids, and amino acid-related pathways.

Keywords: rice; arsenic; metabolomics; lipidomics; LC-MS; ROIMCR



Citation: Pérez-Cova, M.; Tauler, R.; Jaumot, J. Adverse Effects of Arsenic Uptake in Rice Metabolome and Lipidome Revealed by Untargeted Liquid Chromatography Coupled to Mass Spectrometry (LC-MS) and Regions of Interest Multivariate Curve Resolution. *Separations* **2022**, *9*, 79. <https://doi.org/10.3390/separations9030079>

Academic Editor: Szymon Bocian

Received: 19 February 2022

Accepted: 16 March 2022

Published: 18 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Nowadays, there is a need for a better understanding of molecular processes that take place in cereal crops (e.g., rice (*Oryza sativa* L.), maize (*Zea mays* L.), or wheat (*Triticum aestivum* L.)), which are one of the major food staples worldwide [1]. This increasing knowledge leads to an enhancement in their growth, production, and quality [2]. Omics sciences (e.g., genomics, transcriptomics, proteomics, or metabolomics) have arisen as powerful tools for crop improvements from different biological perspectives. These technologies allow, for instance, studying developmental stages and breeding to increase yield, quality, and the bioavailability of nutrients [3–6]. Besides, these omics have been employed in the assessment of the effects of biotic stresses on crops, such as viruses and bacteria [7], but also environmental stressors such as high temperatures [8], salinity [9], or emerging contaminants (e.g., microplastics [10] and heavy metals [11,12]).

Among cereals, rice (*O. sativa* L.) is the most suitable candidate to sequence its DNA and perform genetic modifications, due to its small and well-mapped genome. Hence, rice is currently considered the second-best plant model, after *Arabidopsis thaliana*. Rice metabolome has been widely studied [13–15]. However, metabolite annotation is still a major challenge, due to the huge number of metabolites present in *O. sativa* L. and the

limitations of currently available databases [16]. From the different emerging contaminants, there is a special interest in evaluating the effects of heavy metals and metalloids on rice [17,18]. The reason is that metals tend to accumulate in different parts of the plant (e.g., in the leaves, roots, or stems), becoming a concern not only for the development of the organism itself, but also for human populations through the food chain [19]. Arsenic is a metalloid included in the top ten chemicals of major public health concern by the World Health Organization (WHO) [20]. Arsenic contamination has been related to natural sources like volcanism and geothermal activity, also in addition to anthropogenic sources such as industrial and agricultural activities [21–23]. The arsenic occurrence has been reported to be up to $183 \mu\text{g L}^{-1}$ in groundwaters [24] and 8 mg kg^{-1} in agricultural soils [25], but can also be highly absorbed by composts and biochars [26]. Besides, recent changes in land-use have led to an increase in arsenic levels, compromising human health [27]. The Food and Drug Administration (FDA) from the United States stated that rice presented the second highest arsenic levels, following seafood, and was the most consumed product with high arsenic content, due to its presence in many daily products [28]. Arsenic levels in polished rice have been reported up to 0.629 and 0.055 mg kg^{-1} , for total arsenic and inorganic arsenic content, respectively [29]. From the different arsenic forms, the inorganic is the most toxic. It is highly bioavailable because roots capture it and accumulate in the edible parts using the same transport system as silicon or phosphorus [30]. Therefore, the accumulation and translocation of arsenic is a major concern in rice crops [31–33].

Untargeted metabolomics (and lipidomics) based on liquid chromatography coupled to mass spectrometry (LC-MS) seem suitable approaches for the discovery of unknown metabolites and lipids, respectively [34]. On the one hand, LC-MS is a versatile analytical technique that allows the identification and quantification of a variety of metabolites, ranging from small and polar to big and non-polar, without the need for derivatization steps [35–37]. On the other hand, the goal of untargeted approaches is to have a qualitative profile of the major changes in metabolic pathways due to, for instance, exposure to certain stressors. Therefore, the focus is not on individual metabolites or lipids, but rather a global perspective without any a priori assumptions on the effects these exposures may have on specific biological pathways [38]. This work employed untargeted LC-MS metabolomic and lipidomic workflows to characterize rice metabolome and lipidome, respectively.

Here, the objective is to shed some light on the absorption and translocation (uptake) mechanisms of arsenic in rice together, with its potential impact on their metabolome and lipidome. Previous research in our group already addressed how arsenic affects rice lipidome when supplied by watering the crop [39]. This study aims to complement the previous work with a comparison with the scenario where arsenic content comes from contaminated soil. Therefore, arsenic was supplied through two main routes: watering with contaminated water or soil containing arsenic. In addition, this new study includes metabolomic as well as lipidomic analysis, in order to have a more global overview of arsenic exposure. Two analytical platforms based on LC-MS were employed for the analysis of polar and non-polar metabolites, including lipid species, that were affected by this contaminant. First, a compression strategy based on regions of interest (ROI) was performed to filter the data feature matrices. This step was followed by the application of multivariate curve resolution alternating least squares (MCR-ALS) analysis to resolve the elution and the spectra profiles of the compounds of interest and obtain the areas of the chromatographic peaks necessary to perform subsequent multivariate analyses. The untargeted metabolomic and lipidomic workflows proposed were able to discover potential markers of arsenic exposure and facilitate the identification of the metabolic pathways affected in the different treatments.

2. Materials and Methods

2.1. Chemicals and Reagents

Sodium arsenate dibasic heptahydrate ($\geq 98.0\%$), calcium carbonate (CaCO_3 , $\geq 99.0\%$), ammonium acetate ($\geq 99.0\%$), acetic acid ($\geq 95.0\%$), ammonium formate ($\geq 98.0\%$), and

formic acid ($\geq 95.0\%$) were purchased from Sigma-Aldrich (St. Louis, MO, USA). HPLC grade water, HPLC grade acetonitrile (AcN), HPLC grade methanol (MeOH), methyl tert-butyl ether (MTBE), and chloroform (CHCl_3) were supplied by Merck KGaA (Darmstadt, Germany). Water used for plant irrigation and preparing arsenic solutions was filtered through a $0.22 \mu\text{m}$ nylon filter and purified with an Elix 3 Milli-Q system (Millipore, Belford, MA, USA).

For the lipidomics study, thirteen lipid standards from several lipid families were used as extraction standards: 17:0 monoacylglycerol, 1,2,3-17:0 triglyceride, 17:1 lysophosphatidylethanolamine, 17:0 lysophosphatidylcholine, 17:0 lysophosphatidic acid, 17:0 lysophosphatidylglycerol, 17:0 lysophosphatidylserine, 17:0 cholesteryl ester, 16:0D31-18:1 phosphatidic acid, 16:0D31-18:1 phosphatidylethanolamine, 16:0D31-18:1 phosphatidyl glycerol, 16:0D31-18:1 phosphatidylcholine, and 16:0 D31-18:1 phosphatidylserine. Three sphingolipids were used as internal instrumental standards: N-lauroyl-D-erythro-sphingosylphosphorylcholine, N-(dodecanoyl)-1- β -glucosyl-sphing-4-ene, and N-(dodecanoyl)-sphing-4-enine. All these lipid standards were purchased from Avanti Polar Lipids (Alabaster, AL, US). For the metabolomics study, L-methionine sulfone and piperazine-1,4-bis (2-ethanesulfonic acid) (PIPES) were used as the extraction and internal instrumental standards, respectively, and were purchased from Sigma-Aldrich (St. Louis, MO, USA).

A stock solution of arsenic (V), from now on As (V), at $10,000 \mu\text{M}$, was prepared from the sodium arsenate salt. For the watering treatment, solutions containing 1 and $1000 \mu\text{M}$ of As (V) were prepared weakly by diluting the initial concentrated stock. For the soil treatment, solutions of 5 and 50 mg L^{-1} were prepared directly from the sodium arsenate salt. The solution containing $0.001 \mu\text{M}$ of As (V) used for watering the soil treatment harvest was prepared daily diluting from the initial concentrated stock.

The following abbreviations have been used to describe lipid families: lysophosphatidic acid (LPA), lysophosphatidylcholines (LPC), lysophosphatidylglycerol (LPG), lysophosphatidylethanolamine (LPE), lysophosphatidylserine (LPS), phosphatidic acid (PA), phosphatidylcholines (PC), phosphatidylglycerol (PG), phosphatidylinositols (PI), phosphatidylethanolamine (PE), lysophosphatidylserine (LPS), sphingomyelin (SM), ceramides (Cer), hexosylceramide (HexCer), monogalactosyldiacylglycerol (MGDG), digalactosyldiacylglycerol (DGDG), sulfolipid sulfoquinovosyldiacylglycerol (SQDG), diacylglycerol (DG), triacylglycerol (TG), fatty acid (FA), cholesteryl ester (CE), sterol lipid (ST), and eicosanoyl-EA (NAE).

2.2. Plant Growth, Arsenic Treatments, and Extraction Protocols

2.2.1. General Growing Conditions and Harvesting

Plant growth and lipid extraction were performed using the procedure described elsewhere [39–41]. Briefly, rice seeds were obtained from the Centre for Research in Agricultural Genomics (CRAG, Bellaterra, Spain). Seeds were incubated for 48 h at $30 \text{ }^\circ\text{C}$ in an oven (J.P. Selecta) in a wet environment. After incubation, plants were grown on an Environmental Test Chamber MLE-352H (Panasonic) for 22 days, where cyclic environmental changes of temperature and light intensity were simulated, as shown in Supplementary Material A Figure S1. Soil employed for planting included a mixture of peat, vermiculite, fertilizer, and CaCO_3 . Plates containing different samples were placed in random order inside the chamber, and re-located each watering cycle, established at three times per week.

During the harvest, roots and aerial tissues (i.e., corresponding to the part of the plant above ground) were separated, quenched with liquid nitrogen, and kept at $-80 \text{ }^\circ\text{C}$ until extraction. Before extraction, rice samples were ground to a fine powder with a liquid nitrogen mortar and lyophilized to dryness for 24 h.

2.2.2. Watering and Soil Treatments

For the watering treatment, during the first 11 days, rice was irrigated with Milli-Q water. From that day until harvesting, plants were watered with 1 and $1000 \mu\text{M}$ of As (V) for the two concentration levels of exposure, and with Milli-Q water for control samples.

European legislation established the lowest concentration at 1 μM as it is the limit of the acceptable arsenic concentration in water [42]. The upper concentration was set at 1000 μM , a threshold established to ensure that the experiment was performed under sub-lethal arsenic concentration for the plant, based on previous studies [39].

For the soil treatment, two containers were prepared with 1 kg of soil two days before planting. Soil from the container was exposed to two arsenic concentration levels (5 and 50 mg L^{-1}). Once sowing, rice was irrigated for the whole growth period with a solution containing 0.001 μM of As (V). The lowest arsenic limit in this treatment was set at 5 mg L^{-1} as a maximum value of common arsenic leaches without toxic characteristics [43]. However, the background soil content of arsenic varies between 1 and 40 ppm, according to the US Food and Drug Administration (FDA) report [28]. The highest arsenic limit was established to 50 mg L^{-1} , as a considerably high arsenic content in the soil, slightly above the maximum frequently encountered levels. Table S1 summarizes the arsenic concentration levels selected in this work, expressed in μM for the sake of clarity. The two treatments are referred to with a W (watering) or an S (soil), followed by the concentration dose (L for low and H for high).

2.2.3. Lipid Extraction

A general lipid extraction for untargeted analysis was performed following a previous extraction protocol [39,44]. Briefly, 5 mg of the dried tissue were weighted in individual tubes for each replicate and dissolved in 1 mL of MTBE:MeOH (3:1). Extraction standards mix were added (10 μL at 20 μM , per sample), and then, the mixture was vortexed for 1 min and sonicated for 10 min. Afterwards, 0.5 mL of H_2O :MeOH (3:1) were added, vortexed for 1 min again, and centrifuged for 5 min at 14,500 rpm. The upper organic fraction was collected, whereas the lower aqueous phase was re-extracted with 0.65 mL of MTBE and 0.35 mL of MeOH: H_2O (1:0.85), vortexed for 1 min and centrifuged for 5 min at 2000 $\times g$. Next, combined organic phases were evaporated to dryness under nitrogen gas. Extracts were stored at $-20\text{ }^\circ\text{C}$ until analysis, and resuspended before injection with 250 μL of MeOH: H_2O (4:1). Finally, 10 μL of the internal standards mix at 20 μM were added to each sample.

2.2.4. Metabolite Extraction

For metabolite extraction, based on previous works from Ortiz-Villanueva et al. [45] and Navarro-Reig et al. [46], 40 mg of the dried tissue were weighted in individual tubes for each replicate and 1 mL of MeOH, and 50 μL of L-methionine sulfone (L-met) at 50 mg L^{-1} were added, acting as a surrogate. The mixture was vortexed for 1 min, and sonicated for 10 min, twice. Then, it was centrifuged at 14,500 rpm, and 750 μL of the supernatant were taken, and mixed with 500 μL of CHCl_3 and 400 μL H_2O . Next, it was vortex for 1 min, kept during 15 min at $4\text{ }^\circ\text{C}$, and centrifuged again for 20 min at 14,500 rpm. The aqueous fraction was collected and evaporated to dryness under nitrogen gas. Extracts were stored at $-20\text{ }^\circ\text{C}$ until analysis, and resuspended before injection with 450 μL of AcN: H_2O (1:1). Finally, 50 μL of 50 mg L^{-1} solution of the instrumental internal standard, PIPES, was added.

2.3. LC-MS Analysis

Five biological replicates were analyzed for each sample condition (control, low, and high exposure concentrations), each treatment (watering or soil treatments) and each extraction type (lipid or metabolite extractions). In total, 60 samples were analyzed in each analytical platform (lipidomics or metabolomics). In addition, quality control (QCs) pools composed of 70 (lipidomics) or 50 μL (metabolomics) of solution of each sample condition were prepared separately for different tissues (roots or aerial parts) and extraction types (lipid or metabolite extractions). QCs were repeatedly analyzed during the chromatographic batch every five samples.

2.3.1. Lipidomic Analysis

The lipidomic analysis was performed using a Waters Acquity UPLC system (Waters Corporation, MA, USA), connected to a Waters LCT Premier orthogonal accelerated time of flight mass spectrometer (Waters), operated in both positive and negative electrospray (ESI) ionization modes. Full scan spectra were acquired from 50 to 1500 Da at a scan cycle time of 0.3 s. The following parameters were set for positive ionization mode: capillary voltage, 3000.0 V; sample cone voltage, 50.0 V; desolvation temperature, 350.0 °C; source temperature, 100.0 °C; desolvation gas flow 600.0 L h⁻¹. The same parameters were also used for negative ionization mode, except capillary voltage, set to 2800.0 V instead.

The chromatographic column employed was a Kinetex C8 (100 × 2.1 mm, 1.7 μm) (Phenomenex) under the following conditions (already used in [47]): temperature at 30 °C, injection volume at 10 μL, and flow rate at 0.3 mL min⁻¹. Mobile phases selected were (A) MeOH 1 mM ammonium formate and (B) H₂O 2 mM ammonium formate, both at 0.2% formic acid. The gradient started at 80% A, increased to 90% A in 3 min, from 3 to 6 min remained at 90% A, changed to 99% A until minute 15, remained constant 1 min, and returned to initial conditions until minute 20.

2.3.2. Metabolomic Analysis

The metabolomic analysis was performed using a Waters Acquity UPLC system connected to a Q-Exactive (Thermo Fisher Scientific, Hemel Hempstead, UK) equipped with a quadrupole-Orbitrap mass analyzer. Electrospray (ESI) was used as an ionization source in both positive and negative ion modes. Full scan mass range was set from *m/z* 90 to 1000. The following parameters were set for positive ionization mode: electrospray voltage, 3000.0 V; sheath gas flow rate, 25 arbitrary units (a.u.); auxiliary gas flow rate, 10 a.u.; capillary temperature, 350 °C; and S-lens level, 60%. Negative ionization mode conditions were the same, except for the sheath gas flow rate, set to 40 a.u. All ion fragmentation (AIF) was performed with normalized collision energy (NCE) of 35 eV.

The column employed was a HILIC TSK gel amide-80 column (250 × 2.0 mm i.d., 5 μm) provided by Tosoh Bioscience (Tokyo, Japan), under the following experimental conditions (already employed in [45]): flow rate at 0.15 mL min⁻¹, at room temperature, and 5 μL injection volume. Mobile phases were (A) AcN and (B) 5 mM ammonium acetate, adjusted at pH 5.5 with acetic acid. The gradient employed was: starting conditions at 25% B, then increased until 30% B in 8 min; a 60% B was reached at 10 min, held for 2 min more and then back to 25% B until minute 14 min; lastly, a re-equilibration step was added and from 14 to 20 min at 25% B.

2.4. Data Analysis

2.4.1. Data Compression, Filtering, and Normalization

For data acquisition control and initial data preprocessing, MassLynx 4.1 (Waters Corporation, MA, USA) and Thermo Xcalibur 3.1.66.7 (Thermo Scientific, Hemel, UK) were used for lipidomics and metabolomics studies, respectively. In lipidomics analysis, LC-MS raw files (.raw) were converted to the 'common data format', cdf files using the DataBridge file converter tool available from MassLynx software suite. In the metabolomic analysis, raw files were converted into mzXML format using MS Convert GUI (Palo Alto, CA, USA) using the Proteowizard open-source software [48].

Raw data were then imported to MATLAB computer and visualization environment (Release 2020a, The Mathworks Inc, Natick, MA, US) and analyzed with the ROIMCR chemometrics strategy [49]. This approach was employed for data compression and filtering on one side, and for the resolution of the elution and mass spectra profiles of the different constituents (metabolites or lipids) present in the analyzed rice samples. More information about the ROIMCR approach can be found in Supplementary Material A Section S2. Briefly, spectral compression based on regions of interest (ROI) was performed through the MATLAB MSroi app [50]. ROI strategy allows significant data compression in the spectral dimension without losing their instrumental spectral accuracy. The approach

establishes an intensity threshold value, and MS signals below this threshold are discarded (considered noise). Two additional parameters should be defined, the mass error tolerance (related to the mass spectrometer maximum spectral resolution), and the minimum number of values (minimum number of MS signal occurrences) required to define a chromatographic peak across all the samples (which depends on the type of chromatographic column and conditions used). The main parameters used for the ROI procedure in this work are, briefly, mass error tolerance of 30 and 10 ppm (lipidomics and metabolomics analysis, respectively), a minimum signal factor of 2, a minimum occurrence value of 100, and ROIs were calculated with the median of the m/z values determined for each chromatographic peak. More details can be found in Supplementary Material A Table S2. The ROI approach provides two main outputs: a vector including the list of relevant m/z ROI values (according to the previously mentioned parameters selected for the analysis), and a data matrix with the MS intensities at the selected ROIs (for all considered retention times and samples).

Then, MCR-ALS was applied to the ROI feature data matrices obtained by the workflow described above. Details on the MCR-ALS procedure and parameters employed in the analysis of the data sets in this work are given in Supplementary Material A Section S2. Briefly, MCR-ALS is a bilinear model that decomposes the original data matrix into two-factor matrices related to the elution and spectral profile of the different components. Ideally, each component can be associated with lipid or metabolite constituents of the analyzed samples and possible contributions to the solvent and backgrounds instrument signals. The sample constituents can be identified using the information from the MCR-ALS resolved spectra profiles. MS signals from the same chemical compound, including multiple isotopic forms or adducts and possible mass and ion fragments, are merged in the same MCR-ALS component (i.e., componentization). On the other hand, quantitative information can be retrieved from the elution profiles of the resolved MCR-ALS components, by integrating the areas of their resolved chromatographic peaks. Hence, a data matrix containing the peak areas of each MCR-ALS component is one of the outputs of this method (i.e., component matrix). In this work, four peak area data matrices were obtained for each analytical platform (i.e., lipidomics and metabolomics). Each data matrix corresponded to a specific tissue of the rice plant (roots and aerial parts) and an electrospray ionization mode (positive and negative mode). Finally, these peak areas were normalized by the internal standards added before instrumental analysis, the surrogates employed to correct extraction losses and the dried weight of each replicate. QCs were used as an internal check of the data quality, obtaining similar values within each batch and between different batches. Therefore, no further normalization based on QCs was required.

2.4.2. Statistical Assessment, Exploratory Analysis, and Discovery of Markers of the Exposure

Chemometric analysis of the normalized peak areas of the different components resolved by MCR-ALS was performed with the PLS Toolbox 8.9.1 (Eigenvector Research Inc, Wenatchee, WA, US) under MATLAB (Release 2020b, The Mathworks Inc, Natick, MA, US) Different types of data analyses were applied for statistical assessment, exploratory analysis, and discrimination analysis of markers of the exposure.

The first step was the statistical assessment of the different rice sample treatments with ANOVA-simultaneous component analysis (ASCA) [51]. ASCA combines the multivariate analysis of variance, ANOVA, and simultaneous component analysis (SCA). The null hypothesis of ASCA is that the experimental factors from the experimental design have no effect on the observed results. ASCA was applied to the component matrices for both sample treatments (soil and watering), at the different concentration levels (high, low and control samples). Statistical assessment is performed by a permutation test considering 10,000 replicates.

Principal component analysis (PCA) [52] and hierarchical clustering analysis (HCA) [53] were used for the exploratory study of the effects produced by the different treatments

and conditions, on each of the MCR-ALS component peak area data matrices (related with the metabolites and lipids present in the analyzed samples). PCA describes the experimental data variation in a few components or contributions, explaining the most relevant information from the original variables. The scores plot visualizes the major trends in samples, clustering or discerning them according to their different levels of exposure compared to control samples. In this work, PCA was especially useful to analyze sample trends in an unsupervised manner (no prior information is provided about the different sample classes, i.e., type of treatment or arsenic concentration levels). Biological replicates are expected to cluster together, whereas control samples and samples at similar exposure concentration levels will hopefully cluster separately. On the other hand, HCA using a dendrogram (clustergram) representation allows visualizing trends in the different compounds (i.e., lipids/metabolites that cluster due to similar behavior) and in the samples (i.e., samples ordered by sample type). HCA was performed on data matrices with the fold changes (FC) in the logarithm scale. FC is a standard measurement in metabolomics to compare how much an original condition (control) has changed when related to another condition (exposed or treated). Thus, in untargeted type of data analysis, FCs are usually calculated as the ratio between areas of exposed samples divided by the areas of control samples. Therefore, FCs are expressed as relative abundances. In this case, the area of each component and each exposed replicate was divided by the mean value of the control samples. In this work, HCA was performed using only the more significant peak areas of the MCR-ALS resolved components from each dataset.

Finally, partial least squares discriminant analysis (PLS-DA) [54,55] was applied to the same data matrix of the peak areas of the MCR-ALS resolved components. PLS-DA is a useful and powerful approach for discriminating samples from a supervised perspective. Contrary to PCA, the model is built using information regarding the class membership of each sample (e.g., watering or soil treatment, and exposure levels). The analysis was performed by considering pairs of exposure concentrations (e.g., control samples versus lowest exposure level of the watering treatment, C vs. WL, etc.). A leave-one-out cross-validation method was applied. Variables important in projection (VIPs) of the PLS-DA models allow the identification of possible markers of arsenic exposure and unravel the uptake mechanisms by comparing the significant MCR-ALS components resulted from the different treatments, especially against control samples. Matthews correlation coefficient (MCC) was evaluated as an indicator of the quality of the binary classifications, ranging from -1 to 1 (1 represents a perfect model, -1 a wrong prediction, and 0 random predictions) [56]. In this work, the variables (peak areas of the different MCR-ALS components) associated with VIPs higher than 1.0 were considered relevant according to the various As (V) treatments. Each significant component was then associated with the most intense m/z value from their spectral profile for annotation purposes. Hence, only the compounds related to the relevant MCR-ALS components (with higher VIP values) were finally investigated and annotated.

For PCA and PLS-DA analysis, MCR-ALS resolved component peak area matrices were normalized with probabilistic quotient normalization (PQN) and autoscaled. For ASCA, the same peak areas data matrices were mean-centered. In HCA analysis, a logarithmic normalization was applied to the fold peak area changes before analysis.

2.4.3. Compound Identification

Relevant compounds (MCR-ALS components whose peak areas changed significantly between treatments) from the PLS-DA analysis were selected for identification, due to their implication in the metabolic changes caused by the different arsenic treatments. On one side, lipids were identified according to an in-house built database composed of a list of retention times (RT) associated with compounds frequently detected in plant matrices using the same LC-MS method employed in this work. Besides, LIPIDMAPS [57] and online spectral library human metabolite database [58] were also used for lipid annotation,

selecting the candidates that provided a lower error comparing m/z values of the mass spectra resolved by MCR-ALS and the theoretical one.

In addition, metabolites present in QC samples could be confirmed based on MS/MS spectral matches using public metabolite libraries from the MS-DIAL website [59]. Theoretical and experimental spectra provided by HMDB [58], Massbank [60], and Global Natural Product Social Molecular Networking (GNPS) [61] were compared with our experimental MS/MS data. Plantcyc online database [62] was employed to confirm whether the annotated compounds have been previously found in rice (*Oryza sativa* L.).

3. Results

3.1. Statistical Assessment and Exploratory Analysis of Arsenic Exposure

First, ASCA was employed to evaluate the statistical significance of the experimental treatment (watering/soil) and the As (V) concentration levels (high/low/control), as well as the interaction between these two factors. Both “treatment” factor, and the potential interaction between “treatment” and “As (V) concentration” resulted in being not significant, whereas the “As (V) concentration” factor (with all levels considered at a time) was statistically significant in all cases (i.e., in both tissues: roots and aerial parts, in both ionization modes: positive and negative, and in both platforms: lipidomics and metabolomics; for the eight datasets analyzed in total). Individual studies were also analyzed at two concentration levels (e.g., C vs. WL, etc.) or simultaneously at all concentration levels (e.g., C vs. WL vs. WH). In lipidomics datasets, all combinations (even at the lowest concentration level) were significant (with p -values between 0.003 and 0.0001), regardless of the two types of tissue (roots or aerial) and ionization (positive or negative) modes. Metabolomic datasets exhibited the same behavior (p -values ranging from 0.0346 to 0.0001), with some exceptions. Indeed, the lowest concentration level in soil treatment (C vs. SL) and, consequently, soil treatment in general (C vs. SL vs. SH) were not statistically significant, neither in aerial parts positive ionization nor in roots negative ionization for the metabolomic datasets. Besides, the lowest watering treatment (WL) was not significant in aerial parts positive ionization. In conclusion, aerial parts positive ionization metabolomic set was the least affected by the arsenic exposure in this study. The only clear significant factor for this data set was watering at the highest concentration level (with p -values C vs. WH: 0.0249, and C vs. WL vs. WH: 0.0045). All ASCA results are summarized in Table S3 in Supplementary Material A.

Second, PCA was applied to all datasets to visualize the effects of arsenic exposure. In all cases, more than 40% of all data variance was explained only with the first two components (PC1 and PC2) of the model. On one side, PCA scores plots of the lipidomic datasets showed a clear differentiation between control and exposed groups (samples were separated by PC1 in aerial tissues and by PC2 in roots), as expected and in agreement with previous ASCA results. In addition, a distinction between the two treatments (watering and soil) was found in the analysis of root samples. In contrast, aerial tissue samples were separated in PC2 accordingly to the concentration level, rather than with the treatment itself. Figure 1A,B summarize this trend for both tissues in negative ionization mode, although a similar tendency was obtained for positive mode as well (Figure S3A,B in Supplementary Material A). Hence, root lipids were affected differently according to how rice was exposed to arsenic (from soil or watering). However, this discernment was not present for lipids in aerial parts of the plant, which were more affected by the total arsenic content.

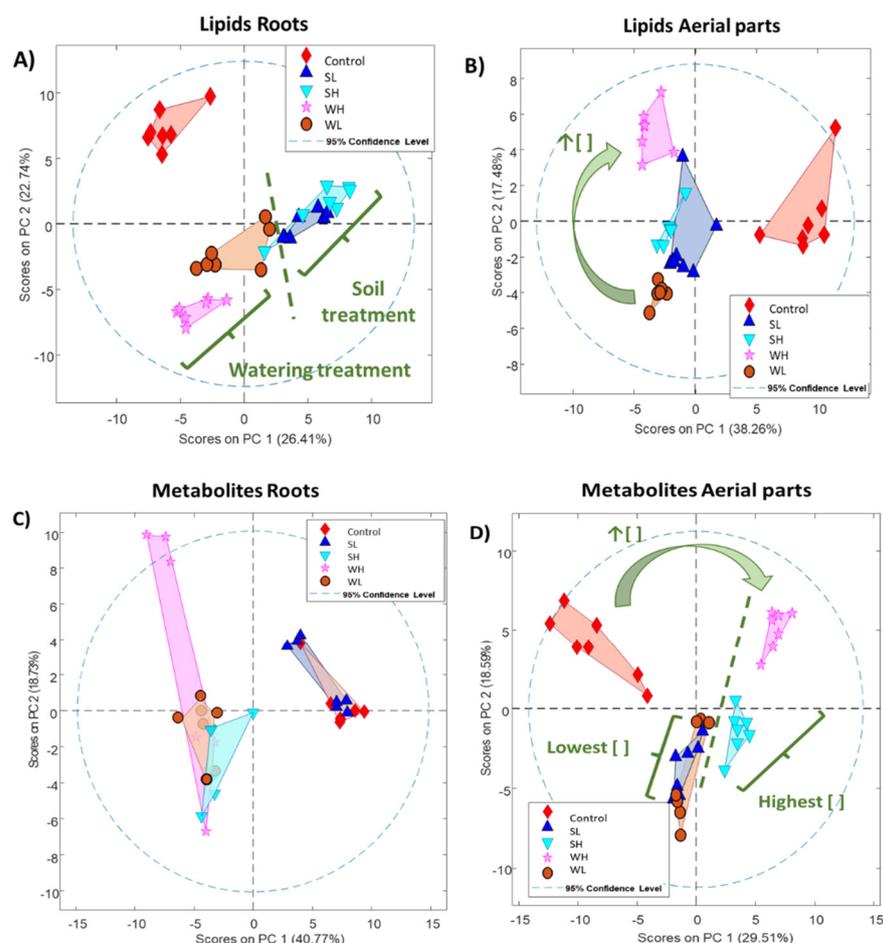


Figure 1. PCA score plots are shown for negative ionization mode obtained in the analysis of both tissues, roots, and aerial parts. (A,B) represent lipidomic analysis of roots and aerial tissues, respectively. Analogously, (C,D) refer to metabolomic analysis of roots and aerials. Both in lipidomic and metabolomic samples from root tissue are more affected by treatment rather than by concentration level, whereas aerial parts have the opposite scenario.

On the other side, similar exposure effects were observed for metabolomic datasets. Concentration levels differentiated metabolites from aerial parts, as shown in Figure 1D for negative mode (and in Figure S3D in Supplementary Material A for positive mode). Three clusters apart from controls were identified, one corresponding to the two lowest exposures (WL and SL), and the other two levels in an increasing order regarding its concentration (WH the most isolated). PCA score plots for metabolites in roots negative ionization (Figure 1C) were basically defined by the lack of differentiation between control and SL groups, separated in PC1 (also observed with ASCA). However, a closer look at PCAs by treatment (only considering C, WL, and WH) revealed that WH and WL clustered together separated from control samples in PC1, which explained 47% of the variance (data not shown). The same C-WL-WH trend was observed for roots positive ionization (Figure S3C in Supplementary Material A). Again, this fact confirmed that the treatment itself affected metabolites from roots more than by the different concentration levels.

3.2. MCR-ALS Component Selection and Annotation

PLS-DA models were built to identify the MCR-ALS components responsible for sample discrimination between different treatments and concentration levels (comparing control vs. treated samples by pairs). Thus, the chemical compounds associated with these components would be considered potential markers of arsenic exposure and helpful in unravelling metabolic changes caused in rice by this metalloid. Table S4 in Supplementary

Material A summarizes the total number of relevant VIPs > 1.0 and MCC values obtained for all datasets. Sample classification was excellent for all lipidomic datasets (MCC equal to 1.0 in all cases). Good discrimination was also achieved for all metabolomic datasets (MCC ranging from 0.7 to 1.0). The first 50 and 20 MCR-ALS components (for lipidomics and metabolomics analysis, respectively) with the higher VIP values for each pair of control vs. treated samples were considered significant, and therefore, contemplated for annotation.

Lipids were annotated using an in-house built retention time database in plant matrices and using the external aid of LIPIDMAPS [57] and HMDB [58]. Up to 100 significant lipids were annotated in total, considering both ionization modes. Metabolites MS/MS spectra from QCs samples were deconvoluted using MS-DIAL [63]. Up to 40 significant metabolites were annotated in total. Figure S4 in Supplementary Material A shows MS/MS spectrum match (experimental vs. theoretical) for L-tryptophan as an example. Tables S5 and S6 in Supplementary Material A list all parameters used in MS-DIAL analysis, employed exclusively for annotation purposes.

Annotation confidence corresponded to level 3 for lipids (no MS/MS information, only exact mass and retention time) and level 2 for metabolites (MS/MS, retention time, exact mass), according to the confidence level of compound annotation re-defined the Compound Identification workgroup of the Metabolomics Society in 2017 [64]. In these cases, when fragment ions were not detected under the mass range conditions of this study, the corresponding metabolites were only tentatively annotated (level 3). Supplementary Material B provides all significant annotated MCR-ALS components. Tables S7–S10 correspond to lipids grouped by tissue and ionization mode, whereas Tables S11–S14 are analogous for metabolites. In addition to compound information, each table furnishes details on which variables were significant for each of the treatments and concentration levels, fold change ratios (areas of all samples in one class divided by the mean area of control samples), and the global tendency of all replicates compared to controls (up/down).

3.3. Lipidomic Results

Annotated lipids, selected from those MCR-ALS resolved components whose peak areas were significant (higher VIPs from PLS-DA results, see above), mainly belonged to three lipid classes: glycerophospholipids (52%), glycerolipids (30%), and sphingolipids (12%). Figure 2A shows the proportion of significant lipids among each family in aerial parts, roots, and both tissues simultaneously. There is an increase of affected glycerolipids in roots (e.g., DGs, MGDGs, and DGDGs), in contrast with a slight increment in certain glycerophospholipids (e.g., PGs and PAs) in aerial tissues. Figure 2B depicts the number of significantly affected lipids is given for the different treatments. Most of the annotated lipids were found significant in all four treatments (WH, WL, SH, and SL) in comparison to control samples. Nevertheless, some specific compounds were related only to watering or soil exposure (e.g., four DGs suffered changes in roots when soil treatment was applied).

HCA was applied to the logarithm of the fold changes of the annotated MCR-ALS components to give a global perspective on lipid changes with the different treatments. There were primarily two clusters found for both tissues and ionization modes: upregulated (marked in red) and downregulated (marked in blue). HCA maps plus dendrograms (clustergrams) are included in Figure 3 for aerial positive ionization, and in Supplementary Figure S5 for aerial negative ionization, roots positive, and negative ionization. Besides, certain lipids in both aerial and root tissues exhibited differences regarding the treatments (watering and soil), although this effect was clearer in roots (in agreement with previous results from PCAs). Among the lipids that increased their concentration regarding controls in aerials, several LPCs, PAs, and DGs stood out, whereas PCs, PGs, and PSs generally decreased. Concerning roots, PCs have also reduced their concentrations, as well as DGDGs and MGDGs. Again, PAs abundances were also incremented due to arsenic exposure.

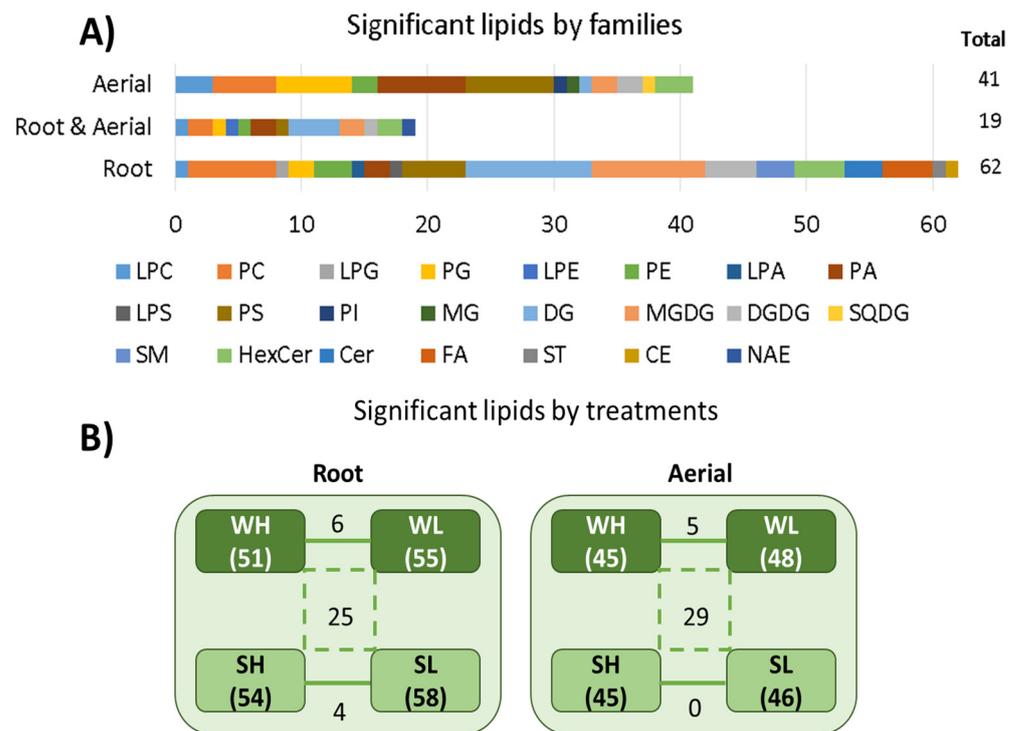


Figure 2. Representation of significant lipids annotated belonging to the different tissues, organized by (A) families and (B) specific treatments.

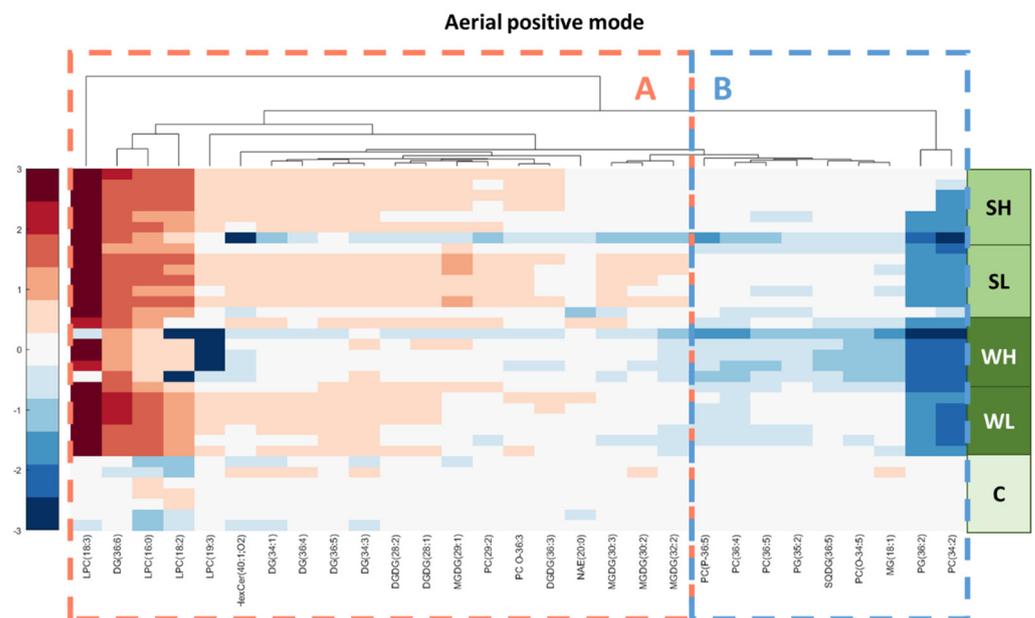


Figure 3. Hierarchical clustering heatmap applied to the logarithm of the fold changes of the annotated lipids for aerial positive ionization set. An intensity color bar is included on the left side of the figure, indicating the relative abundance of the lipid regarding control samples (higher abundance in red, lower abundance in blue). Two clusters are differentiated, with increasing abundance (A) and diminishing abundance (B).

Some potential markers of arsenic exposure had already been suggested in a previous lipidomics study [39]. In this previous work, the main aim was to develop a multidimensional chromatographic method and evaluate the effects only at the two concentration levels referred to as “watering treatment”. However, the coincident lipids from the prior

work and this study did not show an exclusive behavior for watering exposure, but they were instead associated with increasing concentration levels of arsenic (e.g., PA (36:2), PG (34:1), or PC (36:6), which were also markers of soil treatment). The present work also pointed out new lipids highly affected by arsenic exposure regardless of the treatment (e.g., LPC(18:3) in aerial tissues). Besides, the present study showed new insights regarding how arsenic can access rice, and allowed detecting potential markers of the different treatments. For instance, lipids that showed the same tendency in both tissues but accentuated in one of them (e.g., PS (39:3), especially decreased with watering treatments in aerials; or LPE (18:2), more affected by soil treatments in roots). Other lipids presented other remarkable changes. For instance, LPC (16:0) augmented in aerials tissues for the lower doses, but decreased in roots at the highest doses. In other cases, relevant lipids were only annotated for a single tissue (e.g., PG (34:2), PG (32:2), or DG (34:2) diminished due to watering treatment in roots). Nevertheless, further complementary targeted studies and MS/MS confirmation are necessary to assess the effects of arsenic exposure in rice lipidome completely.

3.4. Metabolomic Results

Significant annotated metabolites in Tables S11–S14 were previously detected in *Oryza sativa* L. according to the Plantcyc database [62]. Contrarily to lipids, no clear specific effect based on treatments (watering/soil) was detected in roots. In addition, control and SL groups for this tissue in negative mode cannot be distinguished, as already observed in PCA and ASCA. Most of the annotated metabolites were found in WH, SH, and WL groups, as shown in Figure 4A, (e.g., palmitic acid, allantoin, norvaline, succinic acid, tryptophan, and isoleucine). In addition, the arsenic effect in aerial tissues was dominated by its concentration level rather than by the treatment itself, as previously seen in PCA results. MetaboAnalyst pathway analysis [65] was performed to have a closer look into the metabolic pathways that could be affected by arsenic exposure in general (without tissue differentiation). Table 1 exhibits a detailed list of these pathways, ordered by decreasing significance, including the number of significant metabolites found for each pathway, their *p*-values and False Discovery Rate (FDR) results. Moreover, Figure 4B graphically displays the obtained results, pointing out the main pathways with letters. The five principal metabolic pathways altered by arsenic exposure in this study were amino acid related, i.e., aminoacyl-tRNA biosynthesis (A); alanine, aspartate, and glutamate metabolism (B); glycine, serine, and threonine metabolism (C); phenylalanine, tyrosine, and tryptophan biosynthesis (H); arginine and proline metabolism (I); arginine biosynthesis (G) (Figure 4B). A pathway comparison based on the analyzed tissue is provided in Figure 4C. Overlapping altered pathways were found for both roots and aerials, but the individual metabolites related to these pathways were not necessarily the same. For instance, some common metabolites in both tissues were tryptophan, phenylalanine, serine, proline, glutamine, shikimic acid, allantoin, and succinic acid. On the contrary, adenosine, palmitic acid, and betaine were exclusively detected as significant for roots, and dimethylglycine, pyroglutamic acid, and benzoic acid were only significant in aerials. A more in-depth characterization of the specific metabolic routes affected (e.g., via targeted analysis) could complement these findings and confirm metabolic changes in both tissues according to the treatments applied. Besides, larger spectral databases for secondary metabolites in plants are still lacking, which in the end, still limits their potential discovery to those with an already available MS/MS spectrum.

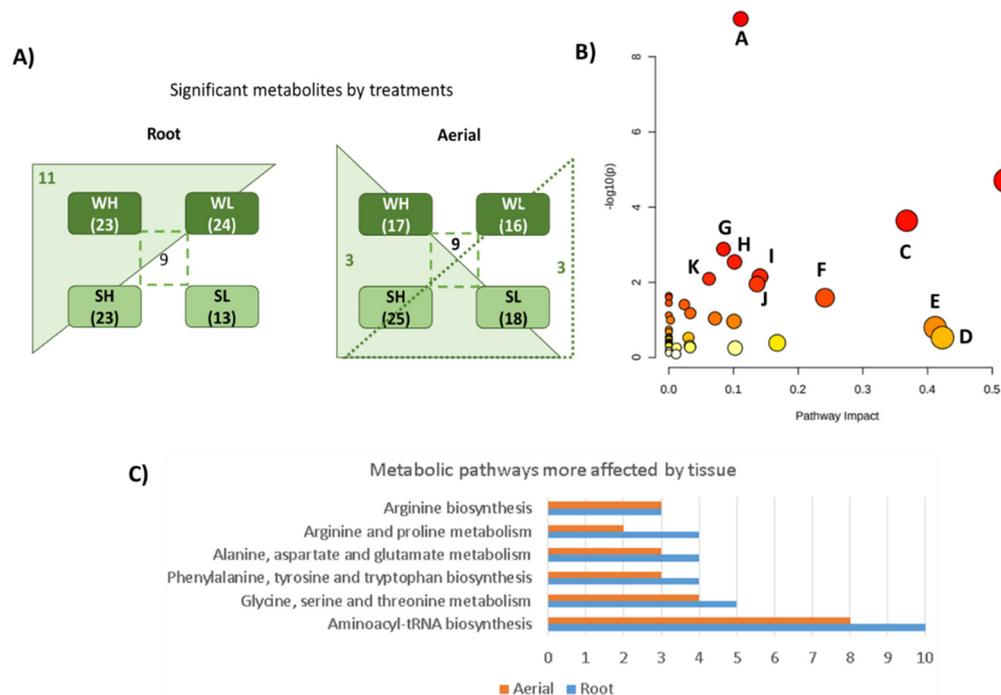


Figure 4. (A) Summary of the common metabolites expressed by the different treatments, with especial emphasis in the lower effect on SL group in roots. (B) Visualization of MetaboAnalyst results indicating the most affected metabolic pathways regarding arsenic exposure from Table 1, for the simultaneous analysis of all tissues and ionization modes. Letter code for the pathways corresponds to: (A) aminoacyl-tRNA biosynthesis; (B) alanine, aspartate, and glutamate metabolism; (C) glycine, serine, and threonine metabolism; (D) phenylalanine metabolism; (E) isoquinoline alkaloid biosynthesis; (F) tryptophan metabolism; (G) arginine biosynthesis; (H) phenylalanine, tyrosine, and tryptophan biosynthesis; (I) arginine and proline metabolism; (J) butanoate metabolism; (K) glyoxylate and dicarboxylate metabolism. (C) Comparison of the number of significant metabolites related to the main metabolic pathways affected by the exposure from (B), according to the tissue analyzed (roots vs. aerials).

Table 1. Metabolomic results from pathway analysis in Metaboanalyst online platform, for all tissues and ionization modes simultaneously. Metabolic pathways affected are ordered according to their significance.

Result from Pathway Analysis	Total	Expected	Hits	Raw p	−Log ₁₀ (p)	Holm Adjust	FDR	Impact
Aminoacyl-tRNA biosynthesis	46	1.31	12	9.86×10^{-10}	9.01	9.37×10^{-8}	9.37×10^{-8}	0.11
Alanine, aspartate, and glutamate metabolism	22	0.63	6	1.951×10^{-5}	4.71	1.83×10^{-3}	9.26×10^{-4}	0.52
Glycine, serine, and threonine metabolism	33	0.94	6	2.29×10^{-4}	3.64	2.13×10^{-2}	7.26×10^{-3}	0.37
Arginine biosynthesis	18	0.51	4	1.30×10^{-3}	2.89	1.20×10^{-1}	3.09×10^{-2}	0.08
Phenylalanine, tyrosine, and tryptophan biosynthesis	22	0.63	4	2.86×10^{-3}	2.54	2.61×10^{-1}	5.44×10^{-2}	0.10
Arginine and proline metabolism	28	0.80	4	7.08×10^{-3}	2.15	6.37×10^{-1}	1.09×10^{-1}	0.14
Glyoxylate and dicarboxylate metabolism	29	0.83	4	8.04×10^{-3}	2.09	7.16×10^{-1}	1.09×10^{-1}	0.06
Butanoate metabolism	17	0.48	3	1.11×10^{-2}	1.96	9.74×10^{-1}	1.31×10^{-1}	0.14
Valine, leucine, and isoleucine biosynthesis	22	0.63	3	2.27×10^{-2}	1.64	1.00	2.21×10^{-1}	0.00
Lysine biosynthesis	9	0.26	2	2.51×10^{-2}	1.60	1.00	2.21×10^{-1}	0.00
Tryptophan metabolism	23	0.66	3	2.56×10^{-2}	1.59	1.00	2.21×10^{-1}	0.24
Cyanoamino acid metabolism	26	0.74	3	3.54×10^{-2}	1.45	1.00	2.81×10^{-1}	0.00
Cysteine and methionine metabolism	46	1.31	4	3.90×10^{-2}	1.41	1.00	2.85×10^{-1}	0.02
Sulfur metabolism	15	0.43	2	6.59×10^{-2}	1.18	1.00	4.47×10^{-1}	0.03
Phenylpropanoid biosynthesis	35	1.00	3	7.48×10^{-2}	1.13	1.00	4.74×10^{-1}	0.00
beta-Alanine metabolism	18	0.51	2	9.10×10^{-2}	1.04	1.00	5.40×10^{-1}	0.07
Purine metabolism	63	1.80	4	1.01×10^{-1}	9.9710^{-1}	1.00	5.63×10^{-1}	0.00
Citrate cycle (TCA cycle)	20	0.57	2	1.09×10^{-1}	9.6210^{-1}	1.00	5.76×10^{-1}	0.10
Isoquinoline alkaloid biosynthesis	6	0.17	1	1.60×10^{-1}	7.9710^{-1}	1.00	7.98×10^{-1}	0.41
Galactose metabolism	27	0.77	2	1.78×10^{-1}	7.4910^{-1}	1.00	8.47×10^{-1}	0.00
Monobactam biosynthesis	8	0.23	1	2.07×10^{-1}	6.8410^{-1}	1.00	8.94×10^{-1}	0.00
Tropane, piperidine, and pyridine alkaloid biosynthesis	8	0.23	1	2.07×10^{-1}	6.8410^{-1}	1.00	8.94×10^{-1}	0.00
Valine, leucine, and isoleucine degradation	37	1.05	2	2.85×10^{-1}	5.4510^{-1}	1.00	1.00	0.00

Table 1. Cont.

Result from Pathway Analysis	Total	Expected	Hits	Raw p	$-\text{Log}_{10}(\text{p})$	Holm Adjust	FDR	Impact
Nitrogen metabolism	12	0.34	1	2.94×10^{-1}	5.3110^{-1}	1.00	1.00	0.00
Phenylalanine metabolism	12	0.34	1	2.94×10^{-1}	5.3110^{-1}	1.00	1.00	0.42
Pyrimidine metabolism	38	1.08	2	2.96×10^{-1}	5.2910^{-1}	1.00	1.00	0.03
Nicotinate and nicotinamide metabolism	13	0.37	1	3.15×10^{-1}	5.0210^{-1}	1.00	1.00	0.00
Cutin, suberine, and wax biosynthesis	14	0.40	1	3.34×10^{-1}	4.76×10^{-1}	1.00	1.00	0.00
Sphingolipid metabolism	17	0.48	1	3.90×10^{-1}	4.09×10^{-1}	1.00	1.00	0.00
Ascorbate and aldarate metabolism	18	0.51	1	4.08×10^{-1}	3.90×10^{-1}	1.00	1.00	0.00
Tyrosine metabolism	18	0.51	1	4.08×10^{-1}	3.90×10^{-1}	1.00	1.00	0.17
Fructose and mannose metabolism	20	0.57	1	4.42×10^{-1}	3.55×10^{-1}	1.00	1.00	0.00
Propanoate metabolism	20	0.57	1	4.42×10^{-1}	3.55×10^{-1}	1.00	1.00	0.00
Carbon fixation in photosynthetic organisms	21	0.60	1	4.58×10^{-1}	3.39×10^{-1}	1.00	1.00	0.00
Zeatin biosynthesis	21	0.60	1	4.58×10^{-1}	3.39×10^{-1}	1.00	1.00	0.00
Biosynthesis of unsaturated fatty acids	22	0.63	1	4.73×10^{-1}	3.25×10^{-1}	1.00	1.00	0.00
Fatty acid elongation	23	0.66	1	4.89×10^{-1}	3.11×10^{-1}	1.00	1.00	0.00
Pantothenate and CoA biosynthesis	23	0.66	1	4.89×10^{-1}	3.11×10^{-1}	1.00	1.00	0.03
Phosphatidylinositol signaling system	26	0.74	1	5.32×10^{-1}	2.74×10^{-1}	1.00	1.00	0.03
Glutathione metabolism	27	0.77	1	5.45×10^{-1}	2.63×10^{-1}	1.00	1.00	0.01
Inositol phosphate metabolism	28	0.80	1	5.59×10^{-1}	2.53×10^{-1}	1.00	1.00	0.10
Ubiquinone and other terpenoid-quinone biosynthesis	35	1.00	1	6.41×10^{-1}	1.93×10^{-1}	1.00	1.00	0.00
Fatty acid degradation	37	1.05	1	6.62×10^{-1}	1.79×10^{-1}	1.00	1.00	0.00
Flavonoid biosynthesis	47	1.34	1	7.49×10^{-1}	1.25×10^{-1}	1.00	1.00	0.00
Fatty acid biosynthesis	56	1.60	1	8.08×10^{-1}	9.23×10^{-2}	1.00	1.00	0.01

4. Discussion

Previous studies in the literature have evaluated arsenic species accumulation and translocation from roots to shoots or grains [31,33,66]. For instance, specific transfer factors have been measured to understand how several arsenic species were transported from roots to other plant parts [66]. As a general conclusion, the higher the uptake, the more arsenic content translocates to the grains. Moreover, elevated concentrations of arsenic had a negative effect on plant development. These facts were in agreement with our study, where higher concentration levels of exposure caused severer changes in the phenotype (lighter colors in the aerial parts and darker in the roots, especially for the WH group), and also in the lipidome and metabolome, as already discussed in the previous sections. The work presented here also demonstrates that regardless of the treatment tested (watering and soil), arsenic alters metabolic pathways in both tissues (roots and aerals), leading to severe damage in the whole plant, including the grain. Furthermore, the arsenic tendency to translocate from roots to shoots, and from shoots to grain, poses a threat on human populations, which take in this element and biomagnifies through the food chain [32].

Since rice is cultivated in flooded conditions, where arsenic mobility is higher [33], the plant is susceptible to uptake arsenic from two main sources, i.e., contaminated groundwaters [24] and contaminated agricultural soils [25]. This study demonstrates that both scenarios of contamination threaten the development and growth of rice. The effects in the roots lipidome differed with the two tested treatments (watering and soil), whereas aerals seem more affected by the total arsenic dosage supplied. Hence, the findings in the study draw attention to the importance of arsenic sources when proposing detoxification strategies.

Untargeted metabolomics is a useful tool to assess metal and metalloid toxicity in model organisms, such as plants [18]. This omic approach provides a snapshot of what happens at the cellular level at the moment of the harvest, which means real-time information on arsenic exposure. Main lipidic changes detected in this study were related to key alterations in glycerophospholipids and glycerolipids, which are the dominant lipid families in rice [67]. The first group is a principal component of biological membranes in animals and plants, but also a major class of lipid in rice grain, often related to its quality and nutritional significance [68]. More specifically, lysoglycerophospholipids seem to be particularly vulnerable to environmental changes, which is in agreement with the accentuated alterations found in the current study for some LPCs and LPEs. There was also a link between some of our significant annotated metabolites and glycerophospholipid pathways (e.g., serine). The second group is also found in plant cell membranes [69], and has also been linked to photosynthesis, especially glyceroglycolipids (e.g., MGDG, DGDG) [70,71]. Among glycerolipids, DGs may play a crucial role in rice. Some studies suggested that DGs derived from phosphatidic acid, and that glycerophospholipid synthesis in rice may be linked to 1,2-diacylglycerol pathways [68]. Therefore, according to our study, arsenic exposure is damaging key lipidomic pathways related to important functions in plant cells and also linked to the grain's quality.

Our metabolomic study found major effects in amino acid-related pathways, such as aminoacyl-tRNA biosynthesis, alanine, aspartate, and glutamate metabolism or glycine, serine, and threonine metabolism. These molecules are known for their essential role in the development, growth, and stress responses of plants, as they have been related to their immune system [72]. A recent review from Guo et al. summarizes current knowledge of what is the function of amino acids in rice as signal molecules, how are they transported from the roots, and how these molecules regulate plant architecture and defense against abiotic stresses, specifically mentioning the role of proline, glycine, glutamate, and glutamine in stress responses [73]. Thus, our results suggest that rice defense mechanism against arsenic involves alterations in the amino acid-related pathways, as a response of their immune system against this contaminant.

Although our study was only performed at one harvesting time, further information could be obtained with time-course experiments at several sampling times. Besides,

untangling the mode of action of this metalloid leads us to potentially identify novel detoxification mechanisms. Lastly, not all arsenic species exhibit the same toxicity and translocation [31], and the cultivar tolerance is also a key aspect to consider. For instance, a previous study evaluated the role of amino acids and thiolic ligands on arsenite tolerance in rice [74]. Results were dependent on the cultivar's tolerance. Although in our work we supplied arsenic (V) solution to rice instead of arsenic (III), both studies have in common that amino acids were altered due to this contaminant regardless of the arsenic species provided. Besides, amino acids are not only involved in rice response against arsenic stress, but could also be employed for detoxification purposes, as well as other molecules such as thiol ligands [74]. Further metabolomic studies in combination with speciation analysis could shed more light on arsenic uptake mechanism and way of action.

5. Conclusions

Untargeted lipidomic and metabolomic approaches have allowed increasing our current knowledge on arsenic exposure in rice at early stages of growth (up to 22 days), through two different treatments (watering with arsenic contaminated water or growth in contaminated soil). Metabolic and lipidomic alterations caused by As (V) treatment were present in both root and aerial tissues. The application of a proper chemometrics workflow has allowed the proposal of potential markers of these metabolic disturbances.

Specifically, arsenic impact in the roots lipidome differed according to the treatment, watering, and soil contaminations, which revealed that the nature of the arsenic source produced a different type of effects on this root tissue. In addition, severe damage in the metabolome (and lipidome) was also found in aerial tissues, confirming the presence of adverse effects due to arsenic exposure throughout the whole plant (and eventually, to the rice grains). In contrast to roots, adverse arsenic effects in aerials were more related to the arsenic dose rather than the treatment itself. Some of the lipids most affected by arsenic exposure belonged to the following lipid families: LPCs, PAs, DGs, PGs, and PSs in aerial tissues, and PCs, PAs, DGDGs, MGDGs in roots. Regarding the metabolomic alterations, the comparison of significant changes in roots and aerial metabolomes showed a considerable overlapping of biochemical pathway alterations between both tissues, although the affected metabolites did not necessarily coincide in both cases. Most of the metabolic pathways disturbed were amino acid related. In rice, amino acids have been previously associated with defense mechanisms against abiotic stresses and they also play a key role in plant immune system. These changes in amino acids may be a consequence of stress response of rice to defend itself against arsenic exposure, and they could be used as an indication of arsenic detoxification.

In conclusion, untargeted analysis has proven to be a powerful tool to generate hypothesis regarding the modes of action of toxics such as arsenic, because it furnishes a wider overview of metabolic changes; in this case, of adverse effects caused by As (V) solutions in rice lipidome and metabolome. Targeted analysis of the potential markers found in previous untargeted analysis can confirm and validate the proposed discoveries. Other complementary omic analysis, such as transcriptomics, will complementarily improve the characterization and role of specific metabolites and lipids involved in arsenic exposure.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/separations9030079/s1>, Figure S1: Experimental conditions employed for rice growing in the chamber MLE-352H: light intensity and temperature; Figure S2: Scheme of the ROIMCR workflow; Figure S3: PCA score plots positive ionization mode obtained for lipidomic results for roots (A) and aerial parts (B), and metabolomic results for roots (C) and aerial parts (D); Figure S4: Experimental MS/MS spectrum of L-tryptophan compared to a theoretical MS/MS spectrum for the same compound in aerial tissues with ESI (+); Figure S5: Hierarchical clustering heatmaps applied to the logarithm of the fold changes of the annotated lipids for aerial in positive mode, roots in positive and negative modes, respectively; Table S1: Summary of the As (V) concentration levels employed in this study; Table S2: ROI parameters employed and MCR components obtained for each dataset; Table S3: Statistical results from ASCA; Table S4: PLS-DA

results: no. of Variables Important in Projection > 1.0 and Matthew Correlation Coefficients; Table S5: MS-DIAL parameters used in metabolomic annotation; Table S6: Experiment file used in MS method type section from start a project window in MS-DIAL; Table S7: Lipid identification of the most relevant features responsible for the changes induced by arsenic exposure on rice in root tissues, in positive ionization mode; Table S8: Lipid identification of the most relevant features responsible for the changes induced by arsenic exposure on rice in root tissues, in negative ionization mode; Table S9: Lipid identification of the most relevant features responsible for the changes induced by arsenic exposure on rice in aerial tissues, in positive ionization mode; Table S10: Lipid identification of the most relevant features responsible for the changes induced by arsenic exposure on rice in aerial tissues, in negative ionization mode; Table S11: Metabolite identification of the most relevant features responsible for the changes induced by arsenic exposure on rice in root tissues, in positive ionization mode; Table S12: Metabolite identification of the most relevant features responsible for the changes induced by arsenic exposure on rice in root tissues, in negative ionization mode; Table S13: Metabolite identification of the most relevant features responsible for the changes induced by arsenic exposure on rice in aerial tissues, in positive ionization mode; Table S14: Metabolite identification of the most relevant features responsible for the changes induced by arsenic exposure on rice in aerial tissues, in negative ionization mode [75,76].

Author Contributions: Conceptualization, M.P.-C., R.T. and J.J.; Methodology, M.P.-C., R.T. and J.J.; Formal analysis, M.P.-C. and J.J.; Investigation, M.P.-C. and J.J.; Writing—original draft, M.P.-C.; Writing—review and editing, M.P.-C., R.T. and J.J.; Visualization, M.P.-C. and J.J.; Supervision, R.T. and J.J.; Project administration, J.J.; Funding acquisition, R.T. and J.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Grant CTQ2017-82598-P and CEX2018-000794-S funded by MCIN/AEI/ 10.13039/501100011033, and supported from the Catalan Agency for Management of University and Research Grants (AGAUR, Grant 2017SGR753). MPC was funded by a predoctoral FPU 16/02640 scholarship from the Spanish Ministry of Education and Vocational Training (MEFF).

Data Availability Statement: The data presented in this study are openly available in Digital.CSIC (<https://digital.csic.es/handle/10261/261849>, accessed on 17 March 2022) and Zenodo (<https://zenodo.org/record/6222067>, accessed on 17 March 2022).

Conflicts of Interest: The authors declare no conflict of interest. In addition, the authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Awika, J.M. Major Cereal Grains Production and Use around the World. *ACS Symp. Ser.* **2011**, *1089*, 1–13. [[CrossRef](#)]
2. Kaur, B.; Sandhu, K.S.; Kamal, R.; Kaur, K.; Singh, J.; Röder, M.S.; Muqaddasi, Q.H. Omics for the Improvement of Abiotic, Biotic, and Agronomic Traits in Major Cereal Crops: Applications, Challenges, and Prospects. *Plants* **2021**, *10*, 1989. [[CrossRef](#)]
3. Tarpley, L.; Duran, A.L.; Kebrom, T.H.; Sumner, L.W. Biomarker Metabolites Capturing the Metabolite Variance Present in a Rice Plant Developmental Period. *BMC Plant Biol.* **2005**, *5*, 8. [[CrossRef](#)] [[PubMed](#)]
4. Xue, J.; Lu, D.; Wang, S.; Lu, Z.; Liu, W.; Wang, X.; Fang, Z.; He, X. Integrated Transcriptomic and Metabolomic Analysis Provides Insight into the Regulation of Leaf Senescence in Rice. *Sci. Rep.* **2021**, *11*, 14083. [[CrossRef](#)] [[PubMed](#)]
5. Hall, R.D.; Brouwer, I.D.; Fitzgerald, M.A. Plant Metabolomics and Its Potential Application for Human Nutrition. *Physiol. Plant.* **2008**, *132*, 162–175. [[CrossRef](#)]
6. Yang, Y.; Saand, M.A.; Huang, L.; Abdelaal, W.B.; Zhang, J.; Wu, Y.; Li, J.; Sirohi, M.H.; Wang, F. Applications of Multi-Omics Technologies for Crop Improvement. *Front. Plant Sci.* **2021**, *12*, 1846. [[CrossRef](#)]
7. Thi, K.; Vo, X.; Rahman, M.; Rahman, M.; Trinh, T.; Kim, S.T.; Jeon, J.-S. Proteomics and Metabolomics Studies on the Biotic Stress Responses of Rice: An Update. *Rice* **2021**, *14*, 30. [[CrossRef](#)]
8. Glaubitz, U.; Li, X.; Schaedel, S.; Erban, A.; Sulpice, R.; Kopka, J.; Hinch, D.K.; Zuther, E. Integrated Analysis of Rice Transcriptomic and Metabolomic Responses to Elevated Night Temperatures Identifies Sensitivity- and Tolerance-Related Profiles. *Plant Cell Environ.* **2017**, *40*, 121–137. [[CrossRef](#)] [[PubMed](#)]
9. Gupta, P.; De, B. Metabolomics Analysis of Rice Responses to Salinity Stress Revealed Elevation of Serotonin, and Gentic Acid Levels in Leaves of Tolerant Varieties. *Plant Signal. Behav.* **2017**, *12*, e1335845. [[CrossRef](#)] [[PubMed](#)]
10. Wu, X.; Hou, H.; Liu, Y.; Yin, S.; Bian, S.; Liang, S.; Wan, C.; Yuan, S.; Xiao, K.; Liu, B.; et al. Microplastics Affect Rice (*Oryza sativa* L.) Quality by Interfering Metabolite Accumulation and Energy Expenditure Pathways: A Field Study. *J. Hazard. Mater.* **2022**, *422*, 126834. [[CrossRef](#)] [[PubMed](#)]

11. Saeed, M.; Quraishi, U.M.; Malik, R.N. Arsenic Uptake and Toxicity in Wheat (*Triticum aestivum* L.): A Review of Multi-Omics Approaches to Identify Tolerance Mechanisms. *Food Chem.* **2021**, *355*, 129607. [[CrossRef](#)] [[PubMed](#)]
12. Jamla, M.; Khare, T.; Joshi, S.; Patil, S.; Penna, S.; Kumar, V. Omics Approaches for Understanding Heavy Metal Responses and Tolerance in Plants. *Curr. Plant Biol.* **2021**, *27*, 100213. [[CrossRef](#)]
13. Oikawa, A.; Matsuda, F.; Kusano, M.; Okazaki, Y.; Saito, K. Rice Metabolomics. *Rice* **2008**, *1*, 63–71. [[CrossRef](#)]
14. Kim, T.J.; Kim, S.Y.; Park, Y.J.; Lim, S.H.; Ha, S.H.; Park, S.U.; Lee, B.; Kim, J.K. Metabolite Profiling Reveals Distinct Modulation of Complex Metabolic Networks in Non-Pigmented, Black, and Red Rice (*Oryza sativa* L.) Cultivars. *Metabolites* **2021**, *11*, 367. [[CrossRef](#)] [[PubMed](#)]
15. Kusano, M.; Yang, Z.; Okazaki, Y.; Nakabayashi, R.; Fukushima, A.; Saito, K. Using Metabolomic Approaches to Explore Chemical Diversity in Rice. *Mol. Plant* **2015**, *8*, 58–67. [[CrossRef](#)] [[PubMed](#)]
16. Yang, Z.; Nakabayashi, R.; Okazaki, Y.; Mori, T.; Takamatsu, S.; Kitanaka, S.; Kikuchi, J.; Saito, K. Toward Better Annotation in Plant Metabolomics: Isolation and Structure Elucidation of 36 Specialized Metabolites from *Oryza sativa* (Rice) by Using MS/MS and NMR Analyses. *Metabolomics* **2014**, *10*, 543–555. [[CrossRef](#)]
17. Liu, C.; Lan, M.M.; He, E.K.; Yao, A.J.; Wang, G.B.; Tang, Y.T.; Qiu, R.L. Phenomic and Metabolomic Responses of Roots to Cadmium Reveal Contrasting Resistance Strategies in Two Rice Cultivars (*Oryza sativa* L.). *Soil Ecol. Lett.* **2021**, *3*, 220–229. [[CrossRef](#)]
18. Booth, S.C.; Workentine, M.L.; Weljie, A.M.; Turner, R.J. Metabolomics and Its Application to Studying Metal Toxicity. *Metallomics* **2011**, *3*, 1142–1152. [[CrossRef](#)] [[PubMed](#)]
19. Feng, Z.; Ji, S.; Ping, J.; Cui, D. Recent Advances in Metabolomics for Studying Heavy Metal Stress in Plants. *TrAC—Trends Anal. Chem.* **2021**, *143*, 116402. [[CrossRef](#)]
20. World Health Organization Web Page. Available online: <https://www.who.int/news-room/fact-sheets/detail/arsenic> (accessed on 14 February 2022).
21. Shankar, S.; Shanker, U. Shikha Arsenic Contamination of Groundwater: A Review of Sources, Prevalence, Health Risks, and Strategies for Mitigation. *Sci. World J.* **2014**, *2014*, 304524. [[CrossRef](#)] [[PubMed](#)]
22. Bundschuh, J.; Schneider, J.; Alam, M.A.; Niazi, N.K.; Herath, I.; Parvez, F.; Tomaszewska, B.; Guilherme, L.R.G.; Maity, J.P.; López, D.L.; et al. Seven Potential Sources of Arsenic Pollution in Latin America and Their Environmental and Health Impacts. *Sci. Total Environ.* **2021**, *780*, 146274. [[CrossRef](#)] [[PubMed](#)]
23. Ren, S.; Song, C.; Ye, S.; Cheng, C.; Gao, P. The Spatiotemporal Variation in Heavy Metals in China’s Farmland Soil over the Past 20 years: A Meta-Analysis. *Sci. Total Environ.* **2022**, *806*, 150322. [[CrossRef](#)] [[PubMed](#)]
24. Nilkarnjanakul, W.; Watchalayann, P.; Chotpantararat, S. Spatial Distribution and Health Risk Assessment of As and Pb Contamination in the Groundwater of Rayong Province, Thailand. *Environ. Res.* **2022**, *204*, 111838. [[CrossRef](#)]
25. Varol, M.; Gündüz, K.; Sünbül, M.R. Pollution Status, Potential Sources and Health Risk Assessment of Arsenic and Trace Metals in Agricultural Soils: A Case Study in Malatya Province, Turkey. *Environ. Res.* **2021**, *202*, 111806. [[CrossRef](#)]
26. Lima, J.Z.; Ferreira da Silva, E.; Patinha, C.; Durães, N.; Vieira, E.M.; Rodrigues, V.G.S. Sorption of Arsenic by Composts and Biochars Derived from the Organic Fraction of Municipal Solid Wastes: Kinetic, Isotherm and Oral Bioaccessibility Study. *Environ. Res.* **2022**, *204*, 111988. [[CrossRef](#)]
27. Li, Y.; Bi, Y.; Mi, W.; Xie, S.; Ji, L. Land-Use Change Caused by Anthropogenic Activities Increase Fluoride and Arsenic Pollution in Groundwater and Human Health Risk. *J. Hazard. Mater.* **2021**, *406*, 124337. [[CrossRef](#)] [[PubMed](#)]
28. US Food and Drug Administration. *Arsenic in Rice and Rice Products Risk Assessment Report*; Center for Food Safety and Applied Nutrition of the Food and Drug Administration, 2016; Volume 1, pp. 1–284. Available online: <http://www.fda.gov/Food/FoodScienceResearch/RiskSafetyAssessment/default.htm> (accessed on 15 March 2022).
29. Roel, A.; Campos, F.; Verger, M.; Huertas, R.; Carracelas, G. Regional Variability of Arsenic Content in Uruguayan Polished Rice. *Chemosphere* **2022**, *288*, 132426. [[CrossRef](#)] [[PubMed](#)]
30. Zhao, F.J.; McGrath, S.P.; Meharg, A.A. Arsenic as a Food Chain Contaminant: Mechanisms of Plant Uptake and Metabolism and Mitigation Strategies. *Annu. Rev. Plant Biol.* **2010**, *61*, 535–559. [[CrossRef](#)] [[PubMed](#)]
31. Mishra, S.; Mattusch, J.; Wennrich, R. Accumulation and Transformation of Inorganic and Organic Arsenic in Rice and Role of Thiol-Complexation to Restrict Their Translocation to Shoot. *Sci. Rep.* **2017**, *7*, 40522. [[CrossRef](#)] [[PubMed](#)]
32. Kumar, S.; Dubey, R.S.; Tripathi, R.D.; Chakrabarty, D.; Trivedi, P.K. Omics and Biotechnology of Arsenic Stress and Detoxification in Plants: Current Updates and Prospective. *Environ. Int.* **2015**, *74*, 221–230. [[CrossRef](#)] [[PubMed](#)]
33. Tuli, R.; Chakrabarty, D.; Trivedi, P.K.; Tripathi, R.D. Recent Advances in Arsenic Accumulation and Metabolism in Rice. *Mol. Breed.* **2010**, *26*, 307–323. [[CrossRef](#)]
34. Perez de Souza, L.; Alosekh, S.; Naake, T.; Fernie, A. Mass Spectrometry-Based Untargeted Plant Metabolomics. *Curr. Protoc. Plant Biol.* **2019**, *4*, e20100. [[CrossRef](#)]
35. Akram, M.I.; Vincent, I.M.; Siddiqui, A.J.; Musharraf, S.G. Polymeric Hydrophilic Interaction Liquid Chromatography Coupled with Orbitrap Mass Spectrometry and Chemometric Analysis for Untargeted Metabolite Profiling of Natural Rice Variants. *J. Cereal Sci.* **2017**, *73*, 165–173. [[CrossRef](#)]
36. Xiao, R.; Ma, Y.; Zhang, D.; Qian, L. Discrimination of Conventional and Organic Rice Using Untargeted LC-MS-Based Metabolomics. *J. Cereal Sci.* **2018**, *82*, 73–81. [[CrossRef](#)]

37. Concepcion, J.C.T.; Calingacion, M.; Garson, M.J.; Fitzgerald, M.A. Lipidomics Reveals Associations between Rice Quality Traits. *Metabolomics* **2020**, *16*, 54. [CrossRef]
38. Navas-Iglesias, N.; Carrasco-Pancorbo, A.; Cuadros-Rodríguez, L. From Lipids Analysis towards Lipidomics, a New Challenge for the Analytical Chemistry of the 21st Century. Part II: Analytical Lipidomics. *TrAC—Trends Anal. Chem.* **2009**, *28*, 393–403. [CrossRef]
39. Navarro-Reig, M.; Jaumot, J.; Tauler, R. An Untargeted Lipidomic Strategy Combining Comprehensive Two-Dimensional Liquid Chromatography and Chemometric Analysis. *J. Chromatogr. A* **2018**, *1568*, 80–90. [CrossRef] [PubMed]
40. Matyash, V.; Liebisch, G.; Kurzchalia, T.V.; Shevchenko, A.; Schwudke, D. Lipid Extraction by Methyl-Tert-Butyl Ether for High-Throughput Lipidomics. *J. Lipid Res.* **2008**, *49*, 1137–1146. [CrossRef]
41. Navarro-Reig, M.; Jaumot, J.; Piña, B.; Moyano, E.; Galceran, M.T.; Tauler, R. Metabolomic Analysis of the Effects of Cadmium and Copper Treatment in: *Oryza sativa* L. Using Untargeted Liquid Chromatography Coupled to High Resolution Mass Spectrometry and All-Ion Fragmentation. *Metabolomics* **2017**, *9*, 660–675. [CrossRef]
42. 2006/118/EC. Directive 2006/118/EC of the European Parliament and of the Council of 12 December 2006 on the Protection of Groundwater against Pollution and Deterioration. Available online: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32006L0118> (accessed on 15 March 2022).
43. Akhter, H.; Cartledge, F.K.; Miller, J.; McLearn, M. Treatment of Arsenic-Contaminated Soils. I: Soil Characterization. *J. Environ. Eng.* **2000**, *126*, 999–1003. [CrossRef]
44. Menéndez-Pedriza, A.; Jaumot, J.; Bedia, C. Lipidomic analysis of single and combined effects of polyethylene microplastics and polychlorinated biphenyls on human hepatoma cells. *J. Hazard. Mater.* **2022**, *421*, 126777. [CrossRef]
45. Ortiz-Villanueva, E.; Jaumot, J.; Martínez, R.; Navarro-Martín, L.; Piña, B.; Tauler, R. Assessment of Endocrine Disruptors Effects on Zebrafish (*Danio Rerio*) Embryos by Untargeted LC-HRMS Metabolomic Analysis. *Sci. Total Environ.* **2018**, *635*, 156–166. [CrossRef] [PubMed]
46. Navarro-Reig, M.; Jaumot, J.; García-Reiriz, A.; Tauler, R. Evaluation of Changes Induced in Rice Metabolome by Cd and Cu Exposure Using LC-MS with XCMS and MCR-ALS Data Analysis Strategies. *Anal. Bioanal. Chem.* **2015**, *407*, 8835–8847. [CrossRef] [PubMed]
47. Puig-Castellví, F.; Bedia, C.; Alfonso, I.; Piña, B.; Tauler, R. Deciphering the Underlying Metabolomic and Lipidomic Patterns Linked to Thermal Acclimation in *Saccharomyces Cerevisiae*. *J. Proteome Res.* **2018**, *17*, 2034–2044. [CrossRef] [PubMed]
48. Chambers, M.C.; Maclean, B.; Burke, R.; Amode, D.; Ruderman, D.L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egerton, J.; et al. A Cross-Platform Toolkit for Mass Spectrometry and Proteomics. *Nat. Biotechnol.* **2012**, *30*, 918–920. [CrossRef]
49. Gorrochategui, E.; Jaumot, J.; Tauler, R. ROIMCR: A Powerful Analysis Strategy for LC-MS Metabolomic Datasets. *BMC Bioinform.* **2019**, *20*, 256. [CrossRef] [PubMed]
50. Pérez-Cova, M.; Bedia, C.; Stoll, D.R.; Tauler, R.; Jaumot, J. MSroi: A Pre-Processing Tool for Mass Spectrometry-Based Studies. *Chemom. Intell. Lab. Syst.* **2021**, *215*, 104333. [CrossRef]
51. Smilde, A.K.; Jansen, J.J.; Hoefsloot, H.C.J.; Lamers, R.J.A.N.; van der Greef, J.; Timmerman, M.E. ANOVA-Simultaneous Component Analysis (ASCA): A New Tool for Analyzing Designed Metabolomics Data. *Bioinformatics* **2005**, *21*, 3043–3048. [CrossRef]
52. Jolliffe, I.T.; Morgan, B. Principal Component Analysis and Exploratory Factor Analysis. *Stat. Methods Med. Res.* **1992**, *1*, 69–95. [CrossRef] [PubMed]
53. Mangiameli, P.; Chen, S.K.; West, D. A Comparison of SOM Neural Network and Hierarchical Clustering Methods. *Eur. J. Oper. Res.* **1996**, *93*, 402–417. [CrossRef]
54. Gromski, P.S.; Muhamadali, H.; Ellis, D.I.; Xu, Y.; Correa, E.; Turner, M.L.; Goodacre, R. A Tutorial Review: Metabolomics and Partial Least Squares-Discriminant Analysis—A Marriage of Convenience or a Shotgun Wedding. *Anal. Chim. Acta* **2015**, *879*, 10–23. [CrossRef]
55. Lee, L.C.; Liong, C.Y.; Jemain, A.A. Partial Least Squares-Discriminant Analysis (PLS-DA) for Classification of High-Dimensional (HD) Data: A Review of Contemporary Practice Strategies and Knowledge Gaps. *Analyst* **2018**, *143*, 3526–3539. [CrossRef] [PubMed]
56. Chicco, D.; Jurman, G. The Advantages of the Matthews Correlation Coefficient (MCC) over F1 Score and Accuracy in Binary Classification Evaluation. *BMC Genom.* **2020**, *21*, 6. [CrossRef] [PubMed]
57. Fahy, E.; Sud, M.; Cotter, D.; Subramaniam, S. LIPID MAPS Online Tools for Lipid Research. *Nucleic Acids Res.* **2007**, *35*, W606–W612. [CrossRef] [PubMed]
58. Wishart, D.S.; Guo, A.C.; Oler, E.; Wang, F.; Anjum, A.; Peters, H.; Dizon, R.; Sayeeda, Z.; Tian, S.; Lee, B.L.; et al. HMDB 5.0: The Human Metabolome Database for 2022. *Nucleic Acids Res.* **2022**, *50*, D622–D631. [CrossRef]
59. MS-DIAL Web Page. Available online: <http://Prime.Psc.Riken.Jp/Compms/MS-DIAL/Main.Html> (accessed on 14 February 2022).
60. Schulze, T.; Meier, R.; Alygizakis, N.; Schymanski, E.; Bach, E.; LI, D.H.; lauperbe; raalizadeh; Tanaka, S.; Witting, M. *MassBank/MassBank-Data: Release Version 2021.12*. 2021. Available online: <https://doi.org/10.5281/ZENODO.5775684> (accessed on 17 March 2022).

61. Wang, M.; Carver, J.J.; Phelan, V.V.; Sanchez, L.M.; Garg, N.; Peng, Y.; Nguyen, D.D.; Watrous, J.; Kaponov, C.A.; Luzzatto-Knaan, T.; et al. Sharing and Community Curation of Mass Spectrometry Data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34*, 828–837. [[CrossRef](#)] [[PubMed](#)]
62. Hawkins, C.; Ginzburg, D.; Zhao, K.; Dwyer, W.; Xue, B.; Xu, A.; Rice, S.; Cole, B.; Paley, S.; Karp, P.; et al. Plant Metabolic Network 15: A Resource of Genome-Wide Metabolism Databases for 126 Plants and Algae. *J. Integr. Plant Biol.* **2021**, *63*, 1888–1905. [[CrossRef](#)] [[PubMed](#)]
63. Tsugawa, H.; Ikeda, K.; Takahashi, M.; Satoh, A.; Mori, Y.; Uchino, H.; Okahashi, N.; Yamada, Y.; Tada, I.; Bonini, P.; et al. A Lipidome Atlas in MS-DIAL 4. *Nat. Biotechnol.* **2020**, *38*, 1159–1163. [[CrossRef](#)] [[PubMed](#)]
64. Blaženović, I.; Kind, T.; Ji, J.; Fiehn, O. Software Tools and Approaches for Compound Identification of LC-MS/MS Data in Metabolomics. *Metabolites* **2018**, *8*, 31. [[CrossRef](#)] [[PubMed](#)]
65. Pang, Z.; Chong, J.; Zhou, G.; de Lima Morais, D.A.; Chang, L.; Barrette, M.; Gauthier, C.; Jacques, P.É.; Li, S.; Xia, J. MetaboAnalyst 5.0: Narrowing the Gap between Raw Spectra and Functional Insights. *Nucleic Acids Res.* **2021**, *49*, W388–W396. [[CrossRef](#)] [[PubMed](#)]
66. Batista, B.L.; Nigar, M.; Mestrot, A.; Rocha, B.A.; Júnior, F.B.; Price, A.H.; Raab, A.; Feldmann, J. Identification and Quantification of Phytochelatins in Roots of Rice to Long-Term Exposure: Evidence of Individual Role on Arsenic Accumulation and Translocation. *J. Exp. Bot.* **2014**, *65*, 1467–1479. [[CrossRef](#)] [[PubMed](#)]
67. Zhang, D.; Zhao, L.; Wang, W.; Wang, Q.; Liu, J.; Wang, Y.; Liu, H.; Shang, B.; Duan, X.; Sun, H. Lipidomics Reveals the Changes in Non-Starch and Starch Lipids of Rice (*Oryza sativa* L.) during Storage. *J. Food Compos. Anal.* **2022**, *105*, 104205. [[CrossRef](#)]
68. Liu, L.; Waters, D.L.E.; Rose, T.J.; Bao, J.; King, G.J. Phospholipids in Rice: Significance in Grain Quality and Health Benefits: A Review. *Food Chem.* **2013**, *139*, 1133–1145. [[CrossRef](#)] [[PubMed](#)]
69. Rizov, I.; Doulis, A. Separation of Plant Membrane Lipids by Multiple Solid-Phase Extraction. *J. Chromatogr. A* **2001**, *922*, 347–354. [[CrossRef](#)]
70. Kobayashi, K. Role of Membrane Glycerolipids in Photosynthesis, Thylakoid Biogenesis and Chloroplast Development. *J. Plant Res.* **2016**, *129*, 565–580. [[CrossRef](#)] [[PubMed](#)]
71. Basnet, R.; Zhang, J.; Hussain, N.; Shu, Q. Characterization and Mutational Analysis of a Monogalactosyldiacylglycerol Synthase Gene OsMGD2 in Rice. *Front. Plant Sci.* **2019**, *10*, 992. [[CrossRef](#)] [[PubMed](#)]
72. Kadotani, N.; Akagi, A.; Takatsuji, H.; Miwa, T.; Igarashi, D. Exogenous Proteinogenic Amino Acids Induce Systemic Resistance in Rice. *BMC Plant Biol.* **2016**, *16*, 60. [[CrossRef](#)] [[PubMed](#)]
73. Guo, N.; Zhang, S.; Gu, M.; Xu, G. Function, Transport, and Regulation of Amino Acids: What Is Missing in Rice? *Crop J.* **2021**, *9*, 530–542. [[CrossRef](#)]
74. Tripathi, P.; Tripathi, R.D.; Singh, R.P.; Dwivedi, S.; Chakrabarty, D.; Trivedi, P.K.; Adhikari, B. Arsenite Tolerance in Rice (*Oryza sativa* L.) Involves Coordinated Role of Metabolic Pathways of Thiols and Amino Acids. *Environ. Sci. Pollut. Res.* **2013**, *20*, 884–896. [[CrossRef](#)]
75. de Juan, A.; Jaumot, J.; Tauler, R. Multivariate Curve Resolution (MCR). Solving the Mixture Analysis Problem. *Anal. Methods* **2014**, *6*, 4964–4976. [[CrossRef](#)]
76. Windig, W.; Stephenson, D.A. Self-Modeling Mixture Analysis of Second-Derivative Near-Infrared Spectral Data using the Simplisima Approach. *Anal. Chem.* **1992**, *64*, 2735–2742.