

ARSENIC SPECIATION METHODS REVEAL SIGNIFICANT SHIFTING
IN MURINE MODELS OF HUMAN ARSENIC EXPOSURE

by
Lu Wang

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Microbiology & Immunology

MONTANA STATE UNIVERSITY
Bozeman, Montana

January 2025

©COPYRIGHT

by

Lu Wang

2025

All Rights Reserved

DEDICATION

I dedicate this work to Dr. Seth Walk and Dr. Jovanka Voyich-Kane, whose unwavering support and guidance made this journey possible. In moments of uncertainty, you extended a helping hand, offering not only academic advice but also heartfelt encouragement. The belief in my potential from Dr. Seth Walk and Dr. Jovanka Voyich-Kane gave me the strength to pursue this degree despite the challenges of transitioning between departments. I am forever grateful for the opportunity to learn and grow under your mentorship. Thank you for making this achievement possible.

To my advisor, Dr. Seth Walk, whose guidance, encouragement, and expertise were invaluable throughout this project. Your dedication to scholarship has inspired me to pursue knowledge with rigor and passion.

To my committee members, Dr. Jovanka Voyich-Kane, Dr. Douglas Kominsky, and Dr. Masafumi Yoshinaga, whose kindness, unwavering support and professional scientific spirit were the power to keep me finishing this dissertation.

To my wife, Dr. Qian Wang, who never stopped believing in me and supporting me.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Seth Walk, for his unwavering support, guidance, and mentorship throughout my doctoral journey. His expertise, encouragement, and constant belief in my abilities have been instrumental in my growth as a researcher and as an individual. I am incredibly fortunate to have had the opportunity to work with such a dedicated and inspiring mentor.

I would also like to thank my dissertation committee members, Dr. Jovanka Voyich-Kane, Dr. Douglas Kominsky, and Dr. Masafumi Yoshinaga, for their valuable **support**, feedback, constructive criticism, and thoughtful suggestions. Their collective expertise has significantly shaped the direction and quality of my research, and I am deeply grateful for their time and support.

A heartfelt thank you to my lab members, past and present, who have made this journey not only intellectually stimulating but also enjoyable. To Trenton Wolfe, Nick Pinkham, Karlin Balckwell, and Reece Erickson, thank you for your collaboration, friendship, and for creating a positive and motivating environment in the lab. Especially thanks for your support when I hit the wall during my PhD progress.

Finally, I owe my deepest thanks and love to my wife, Dr. Qian Wang, whose unwavering support and understanding have been my greatest source of strength. Throughout the long days and nights, your patience, encouragement, and belief in me never faltered. You have made countless sacrifices to help me achieve this milestone, and I could not have completed this journey without you by my side. Your love and dedication have been my anchor, and for that, I am eternally grateful.

TABLE OF CONTENTS

1. BACKGROUND	1
Arsenic	1
Different Toxicity of Different Arsenical Species	2
Arsenic Biotransformation in Microbes.....	2
Arsenite Oxidation	3
Arsenate Reduction.....	4
Arsenic Methylation.....	5
Organic Arsenical Demethylation.....	6
Arsenic Metabolism in Human	6
Arsenic Transformation in Gut Microbiome.....	10
Reduction of iAs^V to iAs^{III} in Gut Microbiome	10
Gut Microbial Methylation of iAs	10
Thiolation of Methylated Arsenic Species in Gut Microbiome	11
Gut Microbial As Resistance Mechanisms	12
Influence of Gut Microbiome on Arsenic Bioavailability and Toxicity	12
Implications for Health and Future Research Directions.....	13
Metabolomics.....	13
Different Tools to Study Metabolomics.....	16
Nuclear Magnetic Resonance (NMR) Spectroscopy	16
Mass Spectrometry (MS)	17
Other Analytical Tools in Metabolomics	18
NMR, A Toolset of Strategies for Studying Metabolism	22
High-Resolution One-Dimensional (1D) NMR.....	22
Two-Dimensional (2D) NMR Techniques	24
Stable Isotope Labeling for Metabolic Flux Analysis.....	25
Dynamic Nuclear Polarization (DNP) for Sensitivity Enhancement.....	26
Automated Quantification and Identification Algorithms	26
In Vivo NMR and Magnetic Resonance Imaging (MRI) for Metabolic Studies.....	27
Prospects and Challenges for NMR in Metabolism Studies	27
The Significance of the Study	28
The Significance of Novel Methodology of Arsenic Speciation	28
The Significance of Arsenic Diffusion in Different Mouse Tissues at Early-stage Exposure	29
The Significance of Metabolomics Analyses.....	30

TABLE OF CONTENTS CONTINUED

2. ARSENIC SPECIATION METHODS REVEAL INTER-INDIVIDUAL VARIABILITY IN MURINE MODELS OF HUMAN ARSENIC EXPOSURE.....	32
Introduction.....	32
Methods.....	34
Chemical Reagents.....	34
Experimental Animals.....	35
Murine Arsenic Exposures with Fasting and Diet Manipulation.....	35
Enzyme Comparisons	36
Papain Proteomics.....	37
Tissue Maceration	38
Processing of Mouse Tissues	39
HPLC-ICPMS	39
Statistical Analyses	40
Results.....	41
Enzyme-Assisted Extraction of Arsenic Species	41
Tissue Maceration Techniques for Arsenic Speciation	46
The Influence of Diet Type on Arsenic Burden and Inter-Individual Variability.....	52
The Influence of Fasting on Arsenic Burden and Inter-Individual Variability.....	56
The Influence of the Microbiome of Arsenic Burden and Inter-Individual Variability	57
Generalizability of HPLC-ICPMS-Resolved Arsenic Species	59
Discussion	61
Technical Sources of Variability	61
Biological Sources of Variability	62
Inter-Individual Variability	64
3. TIME-COURSE ARSENIC ACCUMULATION IN MURINE TISSUES DURING EARLY-STAGE OF ARSENIC EXPOSURE	66
Introduction.....	66
Methods.....	68
Chemical Reagents.....	68
Experimental Animals.....	68
Murine Arsenic Exposures and Tissue Collection	69
Processing of Mouse Tissues	69
HPLC-ICPMS	70
Results and Discussion	70

TABLE OF CONTENTS CONTINUED

Conclusion	84
4. PREVIOUS WORK DURING PHD PROGRAM - NMR METABOLOMICS ANALYSES.....	85
Introduction.....	85
Methods and Materials.....	89
Bacterial Strains, Genes, and Genetic Modifications	89
Quantification of Growth and Arsenic Biotransformation.....	90
Sample Collection.....	90
Metabolite Extraction for ¹ H NMR Analysis.....	91
Sample Preparation and NMR Acquisition.....	92
Spectral Processing and Metabolite Quantification.....	93
Statistical Analysis.....	93
Results and Discussion	94
Compare Metabolomics of <i>E. coli</i> W3110 and AW3110.....	94
Growth and iAs ^V Reduction Curve.....	94
NMR Metabolomic Differences in <i>E. coli</i> Strains With or Without iAs ^V – Whole Map	96
NMR Metabolomic Differences in <i>E. coli</i> Strains Without iAs ^V	98
NMR Metabolomic Differences in <i>E. coli</i> Strains With iAs ^V	105
The Possible As-C Bond Mechanism	112
Construction of AroA and PepM Expression Strains.....	112
Metabolic Analysis of Extracellular Samples	115
Metabolic Analysis of Intracellular Samples	117
Conclusion	122
The Compare Metabolomics Between <i>E. coli</i> Strains W3110 and AW3110	122
The Possible As-C Bond Mechanism	123
5. FUTURE DIRECTION	125
Arsenobetaine Metabolism in Host and Gut Microbiome	125
Time-Course Arsenic Diffusion in Human	126
Brain-Gut Microbiome Axis under Arsenic Exposure.....	127
REFERENCES CITED.....	128
APPENDICES	145
APPENDICES TABLES.....	146

TABLE OF CONTENTS CONTINUED

PUBLICATIONS DURING PHD PROGRAM.....	157
--------------------------------------	-----

LIST OF TABLES

Table	Page
1. Table 2.1. Proteins identified in the commercial papain protease product.	45
2. Table 2.2 Summary of statistical analysis.....	55
3. Table A1. Primers were used in this project.	147
4. Table A2. The summary table of the metabolites detected using ^1H NMR.	147
5. Table A3. The summary table of the extracellular metabolites from the <i>E. coli</i> strains AW3110 (pET-28a), AW3110 (pET- <i>aroA</i>), AW3110 (pET- <i>pepM</i>), and AW3110 (pET- <i>aroA-pepM</i>) with or without the addition of iAs ^V detected using ^1H NMR.	152
6. Table A4. The summary fold change table of the intracellular metabolites from the <i>E. coli</i> strains AW3110 (pET-28a), AW3110 (pET- <i>aroA</i>), AW3110 (pET- <i>pepM</i>), and AW3110 (pET- <i>aroA-pepM</i>) with the addition of iAs ^V detected using ^1H NMR.	153
7. Table A5. The summary fold change table of the intracellular metabolites from the <i>E. coli</i> strains AW3110 (pET-28a), AW3110 (pET- <i>aroA</i>), AW3110 (pET- <i>pepM</i>), and AW3110 (pET- <i>aroA-pepM</i>) without the addition of iAs ^V detected using ^1H NMR.	155

LIST OF FIGURES

Figure	Page
1. Figure 1.1 The model of arsenic transformation in bacterial cells, modified from Wang <i>et al.</i> , 2023.....	3
2. Figure 2.1 Enzyme-assisted extraction of arsenic from murine tissues. Enzyme digestion of murine tissues was performed using papain and pepsin. Background iAs ^{III} in commercial papain and pepsin preparations (A , unpaired t test with Welch's correction). Levels of arsenic species standards following incubation with papain and pepsin (B). Recovery efficiencies of papain (C) and pepsin (D) tested against the theoretical threshold of 100% (one sample two tailed t test). Recovery of arsenic species standards with and without incubation in murine liver homogenate (E , multiple Mann Whitney tests and FDR correction for multiple comparisons). Recovery of arsenic species from the liver of an exposed mouse using either papain and pepsin digestion (F , multiple Mann Whitney tests and FDR correction for multiple comparisons). For all panels containing circles, open and closed circles represent papain and pepsin results, respectively. In panel E, open and closed squares represent results with and without 12-hours of pepsin incubation.....	44
3. Figure 2.2 Background arsenic levels in commercial bead beat tubes. Four commercial bead beat tubes were compared before and after washing with strong acid (multiple unpaired t test with Welch's correction and FDR correction for multiple comparisons).....	48
4. Figure 2.3 Evaluation of bead beating for murine tissue maceration and arsenic speciation. Background iAs ^V in bead beat tubes containing lysing matrix B before and after washing with water and strong acid (A , Kruskal-Wallis test with Dunn's p value adjustment for multiple comparisons). Recovery efficiency of arsenic species standards following incubation with lysing matrix B (B , one sample two tailed t test; Bonferroni adjustment for multiple comparisons; alpha = 0.01). Arsenic species in stool of an unexposed (C) and exposed (D) mouse before and after acid washing (multiple Mann Whitney tests and FDR correction for multiple comparisons). Arsenic species in liver of an unexposed (E) and exposed (F) mouse before and after acid washing (multiple Mann Whitney tests in both panels; Bonferroni adjustment for multiple comparisons; alpha = 0.025 in panel E; FDR correction for multiple comparisons in panel F). Graph labels for unwashed and acid washed bead beat tubes are identical.	49

LIST OF FIGURES CONTINUED

Figure	Page
5. Figure 2.4 Comparison of bead beating and mechanical homogenization for murine tissue maceration and arsenic speciation. Background levels of arsenic species were compared in tubes mimicking bead beating or mechanical homogenization in the absence of tissue (A). Combined arsenic species (i.e., levels quantified by both bead beating and homogenization) in liver samples from three mice following controlled exposure to iAs^{III} (B ; multiple Kruskal-Wallis tests with FDR correction for multiple comparisons). No statistical tests were performed on undetectable arsenic species.	50
6. Figure 2.5 Arsenic species were compared in liver samples prepared by bead beating or homogenizer in three different mice following controlled exposure to iAs^{III} (circles represent $n=3$ technical replicates from the same liver). Graph label is the same for all panels as shown in the top panel.	51
7. Figure 2.6 Dietary impact on arsenic body burden. Background levels of arsenic species in three commercial mouse diets (A ; multiple Kruskal-Wallis tests with FDR correction for multiple comparisons). Background levels of AsB (B) and iAs^V (C) in different batches of the standard and germ-free mouse diets (Kruskal-Wallis test with Dunn's p value adjustment for multiple comparisons). Total arsenic levels in stool (D), liver (E), and bladder (F) of iAs^{III} -exposed mice eating either a standard or purified diet (t test with Welch's correction for unequal variance. Test performed on log-transformed stool data). Relationship between total stool and liver arsenic (G). Linear regression line (solid line) shown with 95% confidence (likelihood) interval (dotted lines) and equation. Arsenic species in stool (H), liver (I), and bladder (J) of iAs^{III} -exposed mice eating either a standard or purified diet (multiple Mann Whitney tests with FDR correction for multiple comparisons used for stool and bladder data; multiple t testing with Welch's correction for unequal variance and FDR correction for multiple comparisons used for square-root transformed liver data). No statistical tests were performed on undetectable arsenic species.	54
8. Figure 2.7 The impact of fasting on arsenic body burden. Total arsenic in stool (A), liver (B), and bladder (C) in unfasted and fasted mice (t test with Welch's correction for unequal variance on square root transformed data for stool and log transformed data for liver and bladder). Arsenic species in stool (D), liver (E), and bladder (F) in unfasted and fasted mice.	57

LIST OF FIGURES CONTINUED

Figure	Page
9. Figure 2.8 The impact of microbiome status on arsenic body burden. Total arsenic in stool (A), liver (B), and bladder (C) in conventional and germ-free mice (t test with Welch's correction for unequal variance on log transformed data for stool and bladder). Arsenic species in stool (D), liver (E), and bladder (F) in conventional and germ-free mice (multiple Mann Whitney tests with FDR correction for multiple comparisons used for stool and bladder data; two-way ANOVA with Sidak's test for multiple comparisons for liver data). Symbols in panels E-F represent groups labeled in A-C (closed diamonds = conventional mice, open diamonds = germ-free mice).	59
10. Figure 2.9 Inter-lab comparison of HPLC-ICPMS for arsenic speciation. Absolute abundance of arsenic species (AsB, iAs ^{III} , DMA ^V , MMA ^V , and iAs ^V) quantified in the same samples in labs A and B (A , t test with Welch's correction for unequal variance on log transformed data). Relationship between arsenic levels quantified in labs A and B (B). Origin constrained linear regression line (solid line) shown with 95% confidence (likelihood) interval (dotted lines) and equation. An origin constrained straight line with slope = 1 shown for comparison.	60
11. Figure 3.1 The process of murine arsenic exposure and tissue collection.....	69
12. Figure 3.2 The time-course AsB concentration in mouse livers, when mice were exposed to iAs ^{III}	71
13. Figure 3.3 The time-course iAs ^{III} concentration in mouse livers, when mice were exposed to iAs ^{III}	72
14. Figure 3.4 The time-course MMA concentration in mouse livers, when mice were exposed to iAs ^{III}	73
15. Figure 3.5 The time-course DMA concentration in mouse livers, when mice were exposed to iAs ^{III}	73
16. Figure 3.6 The time-course iAs ^V concentration in mouse livers, when mice were exposed to iAs ^{III}	74

LIST OF FIGURES CONTINUED

Figure	Page
17. Figure 3.7 The time-course AsB concentration in urinary bladder, when mice were exposed to iAs ^{III}	75
18. Figure 3.8 The time-course iAs ^{III} concentration in urinary bladder, when mice were exposed to iAs ^{III}	76
19. Figure 3.9 The time-course MMA concentration in urinary bladder, when mice were exposed to iAs ^{III}	76
20. Figure 3.10 The time-course DMA concentration in urinary bladder, when mice were exposed to iAs ^{III}	77
21. Figure 3.11 The time-course iAs ^V concentration in urinary bladder, when mice were exposed to iAs ^{III}	77
22. Figure 3.12 The time-course AsB concentration in mouse stool, when mice were exposed to iAs ^{III}	79
23. Figure 3.13 The time-course iAs ^{III} concentration in mouse stool, when mice were exposed to iAs ^{III}	79
24. Figure 3.14 The time-course MMA concentration in mouse stool, when mice were exposed to iAs ^{III}	80
25. Figure 3.15 The time-course DMA concentration in mouse stool, when mice were exposed to iAs ^{III}	80
26. Figure 3.16 The time-course iAs ^V concentration in mouse stool, when mice were exposed to iAs ^{III}	81
27. Figure 4.1 The growth curve of <i>E. coli</i> strains W3110 and AW3110 in M9 medium (100 μ M Pi), with or without the addition of 50 μ M iAs ^V . The red squares showed the sample collection time for different strains.	95
28. Figure 4.2 The iAs ^V reduction curve of <i>E. coli</i> strains W3110 and AW3110 in M9 medium (100 μ M Pi), with the addition of 50 μ M iAs ^V	95
29. Figure 4.3 Representative ¹ H NMR spectrum of intracellular metabolites obtained from AW3110 cells after 8h cultivation without iAs ^V	96

LIST OF FIGURES CONTINUED

Figure	Page
30. Figure. 4.4 Multivariate principal component analyses (PCA) of the metabolic profiles of <i>E. coli</i> W3110 (blue) and AW3110 (red) indicate mutant AW3110 separated from the wild type strain W3110 with or without the addition of iAs ^V . For WT (wild-type strain), 4 h and 8 h indicate early-log phase and mid-log phase, respectively. For Mutant (mutant strain AW3110), 8 h and 16 h indicate early-log phase and mid-log phase, respectively.	97
31. Figure 4.5 Hierarchical clustering analysis (HCA) and heatmap visualizations of <i>E. coli</i> strains W3110 (WT) and AW3110 (Mutant) intracellular samples. For WT, 4 h and 8 h indicate early-log phase and mid-log phase, respectively. For Mutant, 8 h and 16 h indicate early-log phase and mid-log phase, respectively. The red-blue color spectrum represents the scaled abundance of each metabolite, with dark red and blue indicating high and low abundance, respectively. The distinct clustering based on metabolite levels and genotype were observed in both wild type and mutant strains.	98
32. Figure. 4.6 Multivariate principal component analyses (PCA) of the metabolic profiles of early-log (green) and mid-log (blue) phases of <i>E. coli</i> W3110 cultured without iAs ^V , indicating the separation of metabolites in different growth phases of <i>E. coli</i> W3110. Additionally, 4 h and 8 h indicate early-log phase and mid-log phase, respectively.	99
33. Figure. 4.7 Volcano plot of metabolites of mid-log phase wild-type <i>E. coli</i> W3110, compared to the ones at early-log phase. 4 h and 8 h indicate early-log phase and mid-log phase, respectively. Y-axis shows statistical significance as the -Log ₁₀ P values while the x-axis the Log ₂ fold change in gene expression (positive values represent overexpression in the NMR). The default cut-off for log ₂ FC is > 1.5 and the default cut-off for P value is 10 ⁻⁶ , denoted by the dotted lines.....	100
34. Figure. 4.8 Multivariate principal component analyses (PCA) of the metabolic profiles of early-log (yellow) and mid-log (orange) phases of <i>E. coli</i> AW3110 cultured without iAs ^V , indicating the separation of metabolites in different growth phases of <i>E. coli</i> AW3110. 8 h and 16 h indicate early-log phase and mid-log phase, respectively.....	101

LIST OF FIGURES CONTINUED

Figure	Page
35. Figure. 4.9 Volcano plot of metabolites of mid-log phase mutant <i>E. coli</i> AW3110, compared to the ones at early-log phase. 8 h and 16 h indicate early-log phase and mid-log phase, respectively. Y-axis shows statistical significance as the $-\text{Log}_{10}P$ values while the X-axis the Log_2 fold change in gene expression (positive values represent overexpression in the NMR). The default cut-off for $\log_2$ FC is > 1.5 and the default cut-off for P value is 10^{-6} , denoted by the dotted lines.....	101
36. Figure. 4.10 Multivariate principal component analyses (PCA) and partial least squares discriminant analysis (PLSDA) of the metabolic profiles of early-log (yellow for mutant and green for WT) and mid-log (orange for mutant and blue for WT) phases of <i>E. coli</i> strains W3110 and AW3110 cultured without iAs^V , indicating the separation of metabolites in different <i>E. coli</i> strains with or without <i>ars</i> operon. For WT (wild-type strain), 4 h and 8 h indicate early-log phase and mid-log phase, respectively. For Mutant (mutant strain AW3110), 8 h and 16 h indicate early-log phase and mid-log phase, respectively.	103
37. Figure. 4.11 A VIP score of a variable calculated as a weighted sum of the squared correlations between the PLSDA components and the original variable of the top 20 metabolites different between <i>E. coli</i> strains W3110 and AW3110 cultured without iAs^V	104
38. Figure. 4.12 The two-sample <i>t</i> -test of signature metabolomic profiles identified in mutant strain that significantly discriminated against the wild-type strain. The <i>t</i> -test parameters assumed equal variance and the threshold <i>p</i> value < 0.5	104
39. Figure. 4.13 An illustration of how the glycolysis and NAD^+ regeneration cycle both affected by the deletion of <i>ars</i> operon.....	105
40. Figure. 4.14 Multivariate principal component analyses (PCA) of the metabolic profiles of early-log (blue) and mid-log (dark blue) phases of <i>E. coli</i> W3110 cultured with iAs^V , indicating the separation of metabolites in different growth phases of <i>E. coli</i> W3110. 4 h and 8 h indicate early-log phase and mid-log phase, respectively.....	106

LIST OF FIGURES CONTINUED

Figure	Page
41. Figure. 4.15 Volcano plot of metabolites of mid-log phase wild-type <i>E. coli</i> W3110 cultured with iAs ^V , compared to the ones at early-log phase. 4 h and 8 h indicate early-log phase and mid-log phase, respectively. Y-axis shows statistical significance as the -Log ₁₀ P values while the x-axis the Log ₂ fold change in gene expression (positive values represent overexpression in the NMR). The default cut-off for log ₂ FC is > 1.5 and the default cut-off for P value is 10 ⁻⁶ , denoted by the dotted lines.	106
42. Figure. 4.16 Multivariate principal component analyses (PCA) of the metabolic profiles of early-log (pink) and mid-log (brown) phases of <i>E. coli</i> AW3110 cultured with iAs ^V , indicating the separation of metabolites in different growth phases of <i>E. coli</i> AW3110. 8 h and 16 h indicate early-log phase and mid-log phase, respectively.....	107
43. Figure. 4.17 Volcano plot of metabolites of mid-log phase mutant <i>E. coli</i> AW3110, compared to the ones at early-log phase. 8 h and 16 h indicate early-log phase and mid-log phase, respectively. Y-axis shows statistical significance as the -Log ₁₀ P values while the X-axis the Log ₂ fold change in gene expression (positive values represent overexpression in the NMR). The default cut-off for log ₂ FC is > 1.5 and the default cut-off for P value is 10 ⁻⁶ , denoted by the dotted lines.....	108
44. Figure. 4.18 Multivariate principal component analyses (PCA) of the metabolic profiles of early-log (yellow for mutant and green for WT) and mid-log (orange for mutant and blue for WT) phases of <i>E. coli</i> strains W3110 and AW3110 cultured with iAs ^V , indicating the separation of metabolites in different <i>E. coli</i> strains with or without <i>ars</i> operon. For WT (wild-type strain), 4 h and 8 h indicate early-log phase and mid-log phase, respectively. For Mutant (mutant strain AW3110), 8 h and 16 h indicate early-log phase and mid-log phase, respectively.	110
45. Figure. 4.19 A VIP score of a variable calculated as a weighted sum of the squared correlations between the PLSDA components and the original variable of the top 20 metabolites different between <i>E. coli</i> strains W3110 and AW3110 cultured with iAs ^V	111

LIST OF FIGURES CONTINUED

Figure	Page
46. Figure. 4.20 The two-sample <i>t</i> -test of signature metabolomic profiles identified in mutant strain that significantly discriminated against the wild-type strain, culture with iAs ^V . The <i>t</i> -test parameters assumed equal variance and the threshold <i>p</i> value < 0.5.	111
47. Figure. 4.21 An illustration of how the TCA cycle and amino acid metabolism both affected by the deletion of <i>ars</i> operon with the present of iAs ^V	112
48. Figure. 4.22 Growth curve of <i>E. coli</i> strains AW3110 (pET-28a), AW3110 (pET- <i>aroA</i>), AW3110 (pET- <i>pepM</i>), and AW3110 (pET- <i>aroA-pepM</i>) with or without the addition of iAs ^V	114
49. Figure. 4.23 SDS-PAGE to confirm the expression of AroA and PepM in the constructions of <i>E. coli</i> strains AW3110 (pET-28a) (lane 1), AW3110 (pET- <i>aroA</i>) (lane 2), AW3110 (pET- <i>pepM</i>) (lane 3), and AW3110 (pET- <i>aroA-pepM</i>) (lane 4), with the induction of IPTG.	114
50. Figure. 4.24 The total As accumulation in the constructions of <i>E. coli</i> strains AW3110 (pET-28a), AW3110 (pET- <i>aroA</i>), AW3110 (pET- <i>pepM</i>), and AW3110 (pET- <i>aroA-pepM</i>). The cells were cultured in M9 medium with the addition of 50 μ M iAs ^V	115
51. Figure. 4.25 Two-dimensional (2D, left) and three-dimensional (3D, right) multivariate principal component analyses (PCA) of the extracellular metabolic profiles indicating the separation of metabolites in different <i>E. coli</i> constructs with or without the addition of iAs ^V	116
52. Figure. 4.26 Left: Three-dimensional (3D) multivariate principal component analyses (PCA) of the extracellular metabolic profiles indicating the separation of metabolites in different <i>E. coli</i> constructs with the addition of iAs ^V . Right: Plot of the log <i>p</i> -values obtained in a one-way ANOVA showing significant (red, <i>n</i> = 8) and non-significant (green, <i>n</i> = 7) NMR spectral bins (features) across the four <i>E. coli</i> constructs.	116

LIST OF FIGURES CONTINUED

Figure	Page
53. Figure. 4.27 Left: Three-dimensional (3D) multivariate principal component analyses (PCA) of the extracellular metabolic profiles indicating the separation of metabolites in different <i>E. coli</i> constructs without the addition of iAs ^V . Right: Plot of the log p-values obtained in a one-way ANOVA showing significant (red, n = 8) and non-significant (green, n = 7) NMR spectral bins (features) across the four <i>E. coli</i> constructs.....	117
54. Figure. 4.28 Two-dimensional (2D, left) and three-dimensional (3D, right) multivariate principal component analyses (PCA) of the intracellular metabolic profiles indicating the separation of metabolites in different <i>E. coli</i> constructs with or without the addition of iAs ^V	118
55. Figure. 4.29 Left: Two-dimensional (2D) multivariate principal component analyses (PCA) of the intracellular metabolic profiles indicating the separation of metabolites in different <i>E. coli</i> constructs with the addition of iAs ^V . Right: Plot of the log p-values obtained in a one-way ANOVA showing significant (red, n = 51) and non-significant (green, n = 14) NMR spectral bins (features) across the four <i>E. coli</i> constructs.....	119
56. Figure. 4.30 Left: Two-dimensional (2D) multivariate principal component analyses (PCA) of the intracellular metabolic profiles indicating the separation of metabolites in different <i>E. coli</i> constructs without the addition of iAs ^V . Right: Plot of the log p-values obtained in a one-way ANOVA showing significant (red, n = 58) and non-significant (green, n = 11) NMR spectral bins (features) across the four <i>E. coli</i> constructs.....	119
57. Figure. 4.31 The overlap intracellular metabolites significantly changed in the four <i>E. coli</i> constructs with or without the addition of iAs ^V	121
58. Figure. 4.32 The concentration of trehalose in four different <i>E. coli</i> constructions with or without the addition of iAs ^V	121
59. Figure. 4.33 The concentration of pyruvate in four different <i>E. coli</i> constructions with or without the addition of iAs ^V (Left), and the proposed As-C bond mechanism based on the metabolomics in <i>A. tumefaciens</i> 5A (Right).	122

ABSTRACT

Background: Arsenic is an environmental toxicant associated with adverse health effects in humans. Its toxicity depends largely on its chemical form, which requires advanced biochemical techniques to resolve. To date, there is little consensus on speciation methods or murine model design for human-equivalent arsenic exposure.

Objectives: The study objectives were to generate a robust methodological framework for arsenic speciation of mouse tissues and address current limitations and challenges associated with murine models of human arsenic exposure.

Methods: Arsenic species (arsenobetaine, inorganic arsenite, inorganic arsenate, dimethylarsinate, and monomethylarsonate) were quantified using HPLC-ICPMS against commercial standards. Enzyme-assisted digestion of tissue by papain vs. pepsin and tissue maceration by bead beating vs. mechanical homogenization were evaluated using liver samples spiked with standards. Between-lab HPLC-ICPMS data reproducibility was evaluated. Live animal exposures were conducted using C57BL/6 mice orally exposed to 1 µg of inorganic arsenite per gram body weight. Arsenic species in stool, liver, and bladder were compared between mice eating a standard vs. purified diet, between fasted vs. unfasted mice, and between conventional vs. germ-free mice. Background levels of arsenic in food, water, and processing reagents were also quantified.

Results: Papain digestion, although more efficient, exhibited significant conversion of inorganic arsenite to arsenate. Bead beating and mechanical homogenization of tissue performed similarly, but the lysing matrix of the beads contributed a detectable amount of inorganic arsenate background. HPLC-ICPMS results between laboratories were highly correlated but differed slightly in absolute abundance. Significant levels of arsenobetaine were observed in typical mouse chow samples, highlighting the need to consider dietary impacts on speciation analysis. Large inter-individual variability was observed within and between cages of the same treatment group, which was minimally affected by diet type, fasting, or microbiome status. Microbiome status had a significant effect on arsenical speciation of all sample types.

Conclusion: Improving the accuracy and reliability of murine models of human arsenic exposure is critical for understanding factors that influence disease risk. Our results provide a benchmark for evaluating and interpreting data from murine models. Importantly, experiments considering low sample sizes may not adequately account for inter-individual variability, leading to low power and erroneous conclusions.

CHAPTER ONE

BACKGROUND

Arsenic

Beginning with the early metallurgical developments of the Bronze Age, arsenic (As) has held a notable place in human history and culture (Wang *et al.*, 2023). The discovery of As is often credited to the German scholar Albertus Magnus around 1250 AD (Bowell *et al.*, 2014). Known for its notorious use as a poison throughout history, As has claimed many notable figures, including Hephaestus, the god of fire and metallurgy, and Napoleon Bonaparte during his exile on Saint Helena. This deadly reputation has earned arsenic the title of the “king of poisons” (Wang *et al.*, 2023).

Arsenic is a ubiquitous metalloid in the environment, belonging to group 15 in the periodic table, positioned directly below phosphorus (Rosen *et al.*, 2011; Strawn, 2018). The toxicity of As results in its classification as a class I human carcinogen, on the list of harmful agents released by the International Agency of Research on Cancer (IARC) (Li *et al.*, 2014; Minatel *et al.*, 2018). Arsenic is released into the environment by both natural processes, including volcanic eruptions and geomaterial weathering, and by the consumption of As-containing products, including the use of chemotherapeutic drugs, insecticides and wood preservatives (Zhu *et al.*, 2014; Bhowmick *et al.*, 2018). To date, more than 200 million people are affected by high groundwater As pollution in several areas, including Bangladesh, Brazil, Canada, China, Hungary, India, Indonesia, Mexico, Pakistan, USA and Vietnam (Chakraborti *et al.*, 2009; Bhowmick *et al.*, 2018; de Souza *et al.*, 2018; Rodríguez-Lado *et al.*, 2013). Even in the U.S., there are more than 2.1 million populations

are exposed to high level As in their private well (USGS, 2017). Chronic exposure to As has been documented to give rise to cardiovascular, diabetic, reproductive, hematological, hepatic, neurological and respiratory diseases, as well as to cause bladder, lung and skin cancers (Rahman *et al.*, 2009; Carlin *et al.*, 2016).

Different Toxicity of Different Arsenical Species

For decades, chronic exposure to inorganic arsenic (iAs) through contaminated drinking water has been a critical public health concern with global implications. As a well-established human carcinogen, arsenic is associated with cancers of the skin, bladder, lung, and liver, as well as cardiovascular disorders and diabetes. However, the toxicity of arsenic is highly dependent on its chemical form, which would be transformed among its chemical forms once absorbed into organisms. Four oxidation states of arsenical compounds are found in natural environments ($-III$, 0 , $+III$ and $+V$), of which, trivalent arsenite (iAs^{III}) and pentavalent arsenate (iAs^V) are the most prevalent forms (Zhu *et al.*, 2014). The toxicity level of different arsenic species is, in order, arsenite [iAs^{III}] > arsenate [iAs^V] > monomethylarsonate (MMA) > dimethylarsinate (DMA) (Jain and Ali, 2000). Thus, the As species biotransformation is critical to understand the bio-toxicity in organisms.

Arsenic Biotransformation in Microbes

In nature, microorganisms play an important role in As bioavailability, mobility, and speciation, generating a robust biogeochemical cycle in nature (Zhu *et al.*, 2014). The main microbial As biotransformations include: iAs^{III} oxidation, iAs^V reduction, iAs^{III} methylation and the following demethylation (Fig. 1.1).

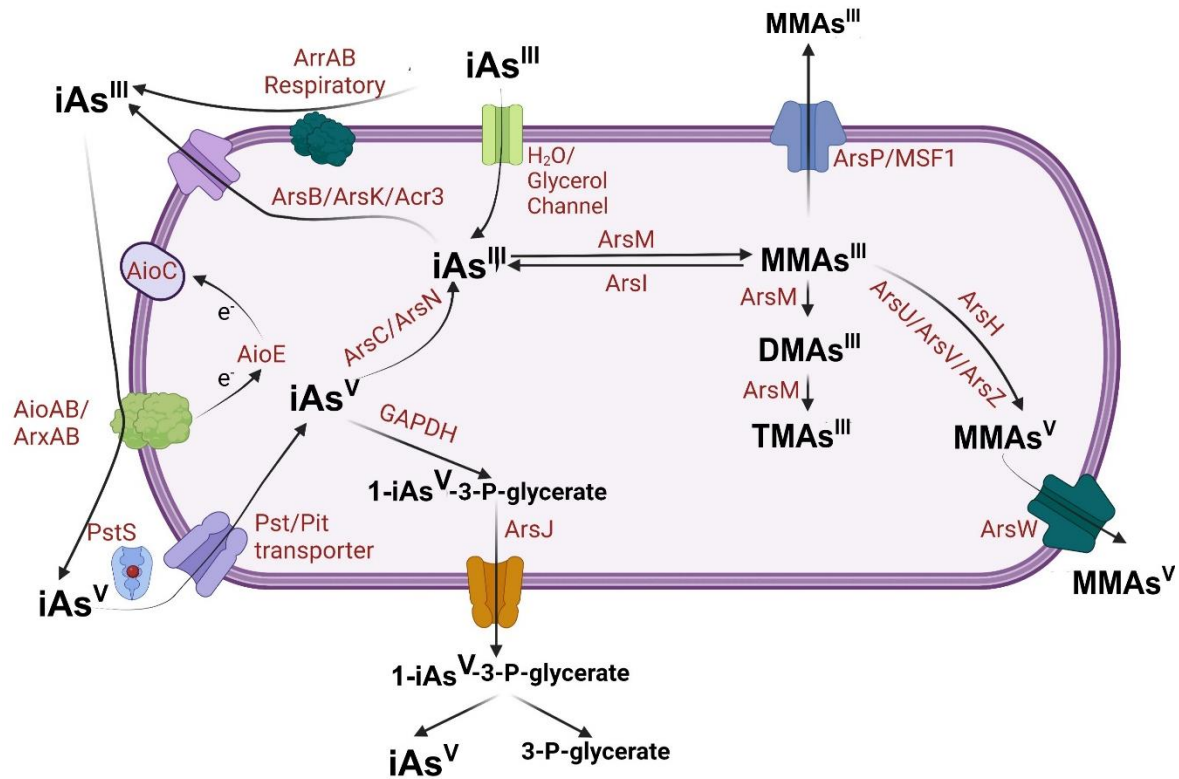


Figure 1.1 The model of arsenic transformation in bacterial cells, modified from Wang *et al.*, 2023.

Arsenite Oxidation. Arsenite oxidation, which transforms iAs^{III} to iAs^V , is considered a potential mechanism for environmental bioremediation because iAs^V is less toxic and less mobile than iAs^{III} (Xiong *et al.*, 2010; Wang *et al.*, 2015). In geothermal spring waters and water sediments, microbial iAs^{III} oxidation has been widely observed and is very efficient (Mukhopadhyay & Rosen, 2002, Oremland & Stolz 2003; Sun *et al.*, 2008), revealing the high efficiency of microbial iAs^{III} oxidation due to the availability of different electron acceptors. Currently, different genera of iAs^{III} -oxidizing bacteria and archaea have been reported, including α -, β -, γ -Proteobacteria, *Thermus*, green sulfur bacteria, filamentous green nonsulfur bacteria, Crenarchaeota and Euryarchaeota (Stolz *et al.*, 2006). Moreover, some iAs^{III} -oxidizing bacteria

use iAs^{III} oxidation as an energy source (Oremland & Stolz, 2003; Stolz et al., 2006; Cai et al., 2009; Wang et al., 2015). Chemolithoautotrophic iAs^{III}-oxidizing bacteria, such as *Rhizobium* sp. NT-26, use iAs^{III} as the electron donor, oxygen as the electron acceptor, and carbon dioxide-bicarbonate as the carbon source to maintain bacterial growth (Santini et al., 2000). More interestingly, heterotrophic iAs^{III}-oxidizing bacteria, such as *Hydrogenophaga* sp. N14 and *Agrobacterium tumefaciens* GW4, can also employ iAs^{III} oxidation as an energy resource to gain a growth advantage, and the subsequent product of iAs^{III} oxidation, iAs^V, can be used to substitute phosphate in lipids under phosphate starvation condition (vanden Hoven & Santini, 2004; Wang et al., 2015).

Arsenate Reduction. The ubiquity of As in the environment has led to the evolution of a bacterial As detoxification system, iAs^V reduction, which transforms iAs^V back to iAs^{III}, and is considered the initial step for bacterial As metabolism (Mukhopadhyay & Rosen, 2002). In prokaryotes, there are two different iAs^V reduction systems: respiratory iAs^V reductase ArrAB and cytoplasmic iAs^V reductase ArsC. ArrAB is a heterodimer consisting of a large subunit ArrA and a small subunit ArrB, using iAs^V as electron donor and acetate as electron acceptor to conserve energy for bacterial growth maintenance (Silver & Phung, 2005; Krafft & Macy, 1998). Besides energy generation using iAs^V reduction and since iAs^V entered the bacterial cell via the phosphate transport system, a soluble cytoplasmic arsenate reductase ArsC can reduce iAs^V to iAs^{III} in cytoplasm, and then the iAs^{III} is effluxed outside of the bacterial cell via an ATP-dependent iAs^{III} transporter ArsAB or Acr3 (Silver & Phung, 2005). The iAs^V reduction via ArsC does not appear to be involved in energy conservation (Rosen *et al.*, 1995; Krafft & Macy, 1998), which is also the biggest difference between these two iAs^V reduction systems.

Arsenic Methylation. Additionally, there exists a parallel pathway for synthesis and degradation of methylated arsenicals in microorganisms (Liu *et al.*, 2016). The bacterial iAs^{III} S-adenosylmethionine (SAM) methyltransferase ArsM, which was first identified from a soil bacterium *Rhodopseudomonas palustris* (Qin *et al.*, 2006), catalyzes the transfer of a methyl group from SAM to iAs^{III} , and subsequently methylates inorganic iAs^{III} in three sequential steps to monomethylarsenite [$MMAs^{III}$], dimethylarsenite [DMA^{III}] and trimethylarsenite [TMA^{III}]. ArsM enzymes characterized to date typically have three domains, an N-terminal SAM binding domain, a middle iAs^{III} binding domain, and a C-terminal domain of unknown function (Liu *et al.*, 2016; Huang *et al.*, 2018), yet some of them are composed of only the first two domains, missing the C-terminal domain. However, the iAs^{III} binding domain varies with iAs^{III} binding sites in different organisms. In mammalian and algal ArsM amino acid sequences, four conserved cysteine residues have always been identified (Liu *et al.*, 2016). Among the four conserved cysteine residues, the two closest to the N-terminus are disulfide bonded, and the other two bind iAs^{III} (Marapakala *et al.*, 2015). Mutating any one of the four cysteine residues surprisingly results in loss of nearly all methylation activity. Furthermore, mutating the two iAs^{III} binding cysteine residues results in the inability to methylate inorganic iAs^{III} to $MMAs^{III}$, but the enzyme is still able to methylate MA^{III} to DMA^{III} (Marapakala *et al.*, 2015). Recently, a novel ArsM has been found in *Aspergillus fumigatus* which only contains three conserved cysteine residues (Chen *et al.*, 2017), while another ArsM discovered in *Bacillus* sp. CX-1 requires only two cysteine residues (Huang *et al.*, 2018). All of these research findings indicate that i) the iAs^{III} binding cysteine residues are the most important elements; and ii) the iAs^{III} binding domain is

variable and performs different iAs^{III} bind mechanisms and different efficiency to methylate inorganic iAs^{III} .

Organic Arsenical Demethylation. Because MMAs^{III} is more toxic than iAs^{III} , the MMAs^{III} species has been proposed to be a primordial antibiotic (Yoshinaga *et al.*, 2011; Liu *et al.*, 2016). Microorganisms have evolved the ability to oxidize MMAs^{III} to less toxic MMAs^{V} using ArsH (Chen *et al.*, 2015b), and/or to actively extrude MMAs^{III} from bacterial cells via ArsP (Chen *et al.*, 2016), and/or even to demethylate MMAs^{III} back to less toxic iAs^{III} via action of a C-As bond lyase ArsI (Yoshinaga & Rosen, 2014).

Arsenic Metabolism in Human

Once inside the human body, iAs undergoes biotransformation through a methylation process aimed at detoxification, though toxicity and the efficiency of detoxification vary among individuals. This methylation process is primarily catalyzed in the liver, which is considered the main organ for arsenic metabolism (Khairul *et al.*, 2017). The methyltransferase enzyme AS3MT, similar to ArsM in microbes, with the help of a methyl donor (S-adenosylmethionine, SAM) and a cofactor (glutathione, GSH), facilitates the conversion of inorganic arsenic to mono- and dimethylated metabolites, such as MMAs^{III} and DMA^{III} . These metabolites are further oxidized to monomethylarsonic acid (MMAs^{V}) and dimethylarsinic acid (DMAs^{V}), which are ultimately excreted in urine (Mersaoui *et al.*, 2022). This transformation reduces As toxicity and enables its more efficient removal from the body.

The classical methylation pathway in human, originally proposed by Frederick Challenger (Challenger, 1951; Cullen *et al.*, 2008), involves a series of reduction and methylation steps, where

iAs^V is first reduced to the more toxic iAs^{III} , then methylated and oxidized into $MMAs^V$, subsequently reduced to $MMAs^{III}$, and methylated again to $DMAs^V$. Although $DMAs^V$ is considered the terminal metabolite in humans, certain species convert it further into trimethylarsine oxide ($TMAO^V$) (Drobná *et al.*, 2005), underscoring the interspecies differences in arsenic metabolism. Notably, these methylation steps result in a shift from toxic trivalent intermediates to relatively less reactive pentavalent forms, aiding in detoxification.

Alternative pathways to Challenger's model have also been proposed. One such alternative by Hayakawa *et al.* (2005) posits that iAs^{III} species persist during methylation, and that oxidation to pentavalent As species occurs post-methylation. This pathway proposes that arsenic-glutathione complexes, such as $iAs^{III}(GS)_3$, are key intermediates. These complexes can be sequentially methylated to form monomethylated and dimethylated arsenic-glutathione derivatives in the trivalent state. William Cullen (2014) further refined this theory, suggesting that thiol-containing proteins may stabilize arsenic by neutralizing its charge, allowing for additional methylation and subsequent detoxification.

The oxidative state of the methylated species directly impacts arsenical toxicity. Pentavalent methylated species, especially $DMAs^V$, are the primary urinary excretion forms in humans and exhibit lower reactivity and toxicity, making them easier to excrete. In contrast, trivalent methylated species are more reactive, binding to intracellular thiols like GSH and specific proteins, and are more likely to cause cellular damage (Khairul *et al.*, 2017). Consequently, As detoxification in humans largely relies on the efficient methylation of these trivalent intermediates to pentavalent species, thereby facilitating safer excretion.

Consistent with the above conclusion, recent epidemiological studies have revealed variability in how individuals metabolize and excrete As (González-Martínez *et al.*, 2024). Although people may be exposed to similar As levels through drinking water, urinary profiles of As metabolites vary significantly, highlighting differences in methylation capacities. These individual differences influence susceptibility to arsenic-related diseases. Factors such as As species, availability of thiol (-SH) containing molecules (like GSH), and enzyme activity levels play crucial roles in determining the distribution and retention of arsenic within tissues (Khairul *et al.*, 2017). The liver's role is central to this process, yet other tissues like the testes, kidneys, and lungs also exhibit methylation activity, suggesting the systemic nature of arsenic processing in humans.

Detoxification efficiency in humans is influenced by the rate at which iAs are methylated (Laue *et al.*, 2022). In this process, trivalent methylated forms such as MMAs^{III} and DMAs^{III}, though highly toxic, are transient intermediates that quickly convert into more stable pentavalent species like DMAs^V. Research indicates that increased proportions of pentavalent methylated arsenicals in urine are associated with a higher detoxification capacity, while elevated levels of trivalent intermediates, particularly MMAs^{III}, suggest a reduced ability to fully methylate iAs (Pace *et al.*, 2018). Individuals with higher proportions of urinary trivalent arsenicals may have an increased risk of health issues such as skin and bladder cancers.

Historically, iAs^{III} methylation was believed to serve only as a detoxification process. However, the detection of highly toxic trivalent methylated intermediates, such as MMAs^{III} and DMAs^{III}, has challenged this view, suggesting that under certain conditions, methylation may actually contribute to bioactivation (McDermott *et al.*, 2020). For example, in cases of chronic

exposure, trivalent intermediates accumulate, overwhelming the body's detoxification capacity and posing greater toxicity risks. In cases of lower or acute arsenic exposure, however, the methylation pathway primarily aids in detoxification, converting arsenic into less toxic and more readily excretable pentavalent forms.

An emerging area of interest is the effect of As on DNA methylation patterns (Reichard & Puga, 2010). Arsenic biotransformation mediated by AS3MT produces S-adenosylhomocysteine (SAH), which may interfere with DNA methylation, impacting the regulation of tumor suppressor genes and potentially playing a role in cancer prevention. Clinical studies in patients treated with arsenic trioxide (As_2O_3) for acute promyelocytic leukemia (APL) have shown methylated arsenic species in various biological samples (Chin *et al.*, 2022), further suggesting the significance of arsenic metabolism in therapeutic outcomes.

In conclusion, As detoxification in humans is a complex process marked by methylation of iAs to less toxic pentavalent species. This metabolic pathway transforms arsenic into a form that is efficiently cleared from the body, with DMAs^{V} being the primary excreted metabolite. The presence of toxic intermediates like MMAs^{III} and DMAs^{III} , however, highlights the potential for methylation to act as both a detoxification and bioactivation pathway, depending on exposure levels and individual metabolic capacity. Understanding this intricate balance between detoxification and bioactivation is essential to fully grasp arsenic's impact on human health and its therapeutic potential.

Arsenic Transformation in Gut Microbiome

The gut microbiome is now considered the biggest tissue in the human body. Because it is composed of microbes, the transformation of arsenic by the gut microbiome is a multi-step process involving microbial reduction, methylation, and sometimes thiolation of arsenical species (Coryell *et al.*, 2019). This transformation not only influences the bioavailability and toxicity of arsenic but also contributes to individual variability in arsenic-related host health outcomes. Key insights into these microbial transformations are grounded in studies examining the roles of specific gut bacteria and the biochemical pathways they employ.

Reduction of iAs^V to iAs^{III} in Gut Microbiome

Reduction of iAs^V to iAs^{III} is often the initial transformation step for ingested arsenic in the gut. In the gut, this process is mediated by the cytoplasmic iAs^V reductase ArsC, whose encoding gene was found in various gut bacteria such as *Bacteroides* and *Escherichia coli* (Wang *et al.*, 2022). McDermott *et al.* (2020) noted that these microbial redox transformations are a primary pathway in the gut and can lead to increased arsenic bioavailability in host tissues. The transformation of iAs^V to iAs^{III} serves as a bacterial detoxification mechanism, as iAs^{III} can be actively extruded from cells, reducing intracellular As levels. However, the resultant iAs^{III} is more readily absorbed by the human body, which could heighten toxicity, particularly in chronic exposure scenarios.

Gut Microbial Methylation of iAs

The next critical step in As transformation within the gut microbiome is methylation, which generally converts iAs into less toxic, more excretable forms. Methylation results in the production

of MMAs^V and DMAs^V. Studies have identified bacterial genera, such as *Clostridium* and *Bacteroides*, that exhibit significant methylation capabilities (Isokpehi *et al.* 2014; Wang *et al.*, 2017). This transformation is crucial for reducing the overall toxicity of iAs, as methylated arsenicals like DMAs^{III} and DMAs^V are easier to excrete and less toxic than their trivalent counterparts.

The work by Wang *et al.* (2022) highlights the role of specific gut bacteria in transforming iAs through methylation pathways. In this study, *Lactobacillus* strains showed high iAs biosorption and methylation abilities, potentially reducing the host's iAs burden. This microbial methylation parallels the detoxification processes occurring in the liver, and the methylated arsenicals (MMA and DMA) found in urine indicate the significant role gut bacteria play in systemic arsenic detoxification.

Thiolation of Methylated Arsenic Species in Gut Microbiome

Apart from methylation, certain gut bacteria can thiolate methylated arsenic species, adding sulfur atoms to form thiolated derivatives (Potera, 2014). These thiolated arsenicals are less toxic and are thought to enhance As clearance from the host. Thiolation pathways were first examined in a controlled gut simulation model by Van de Wiele *et al.* (2010), who observed that gut microbes could convert methylated arsenic species into their thiolated forms under anaerobic conditions. The production of thiolated species, such as monomethylmonothioarsonic acid (MMMTA^V) and dimethylmonothioarsinic acid (DMMTA^V), represents a potential pathway for reducing the As toxicity in the human body. This transformation may be less common but is a critical aspect of the gut microbiome's As detoxification capabilities.

Gut Microbial As Resistance Mechanisms

Consistent with the environmental As-resistant microbes, As resistance in gut bacteria is supported by the same specific genetic operons, such as the *ars* operon, which includes genes like *arsC* for arsenate reduction and *arsB* and *acr3* for iAs^{III} efflux. The *ars* operon enables bacteria to transform iAs^V to iAs^{III} and subsequently pump it out of the cell to maintain As resistance. In a study on the microbial impacts of As exposure, Coryell *et al.* (2019) noted that these resistance mechanisms are critical for microbial survival in the As-rich environments and influence the availability of As species that the host absorbs. Additionally, in a study with mice, exposure to iAs^{III} induced changes in gut microbiota, favoring bacteria with high iAs^{III} resistance and transformation capabilities, demonstrating the microbiome's adaptive response to As exposure.

Influence of Gut Microbiome on Arsenic Bioavailability and Toxicity

The composition and functional capacity of the gut microbiome vary significantly among individuals, influencing the degree to which As species is transformed and its toxic potential. A study by Laue *et al.* (2022) in infants found that the gut microbiome's ability to transform As affected arsenical species in urine, with higher levels of methylated arsenicals (MMA and DMA) observed in infants with specific microbial As-transforming genes. This relationship between microbial composition and urinary As profile underscores the microbiome's role in modifying As toxicity and emphasizes the potential for using gut microbiota profiles as indicators of As detoxification efficiency.

Furthermore, Gokulan *et al.* (2018) demonstrated that arsenic exposure not only increases the abundance of As-resistant microbes in the gut but also affects host immune responses. In As-

exposed mice, the presence of arsenic-resistant bacteria was associated with elevated levels of gut-associated chemokines and cytokines, indicating that As exposure impacts both microbial and host immune systems. This immune modulation, influenced by the microbiome, may play a role in how the host manages As toxicity over time.

Implications for Health and Future Research Directions

The findings from these studies highlight the critical role of gut microbiome in modulating As bioavailability and toxicity through reduction, methylation, and resistance mechanisms. The microbial As transformation represents both a protective detoxification pathway and a potential source of toxicity, depending on the composition of gut microbiome and As exposure level. With the bidirectional relationship between As and the gut microbiome, high exposure levels can lead to microbial adaptation and potentially harmful shifts in microbiome composition, further influencing health outcomes.

Future research should focus on identifying key bacterial species and strains that contribute to As methylation and thiolation, as well as exploring probiotic interventions that enhance arsenic detoxification (Wang *et al.*, 2023). Understanding the complex interplay between microbiome composition, host As exposure, and individual health could enable targeted microbiome-based therapies for mitigating As toxicity, particularly in vulnerable populations with high exposure levels.

Metabolomics

Metabolomics is an advanced field that focuses on the comprehensive analysis of metabolites—small molecules generated through metabolic processes within biological systems (Keum *et al.*,

2010). These metabolites, typically weighing less than 1,500 Da, are crucial indicators of cellular function, offering a direct window into the biochemical state of an organism (Trifonova *et al.*, 2023; Qiu *et al.* 2023). Unlike genomics, transcriptomics, and proteomics, which provide information on genetic codes, RNA transcripts, and proteins, metabolomics uniquely captures the dynamic and real-time changes that reflect a physiological state in organisms. This makes it a powerful tool for understanding the biochemical underpinnings of various biological processes, as well as the influences of diet, environment, and genetics.

Metabolites serve as both substrates and products in cellular pathways, and they are directly involved in essential functions like energy production, signal transduction, and regulatory mechanisms (Wang & Lei, 2018; Milanesi *et al.*, 2020). Because of their immediate response to changes within a biological system, metabolites provide a sensitive and direct measure of the metabolic flux. This allows metabolomics to offer unparalleled insights into the functional state of cells, tissues, and whole organisms (Tang *et al.*, 2011; Burgess *et al.*, 2014). For instance, while genomics might reveal the potential for certain metabolic pathways, metabolomics provides a direct snapshot of their activity and the impact of external factors like environmental toxins or nutritional changes. This ability to observe the immediate biochemical consequences of cellular processes gives metabolomics a unique advantage over other omics fields.

There are two main approaches in metabolomics: targeted and untargeted analyses (Roberts *et al.*, 2012). Targeted metabolomics focuses on quantifying specific known metabolites within defined pathways, providing precise information on selected biochemical processes (Kim *et al.*, 2016; Qiu *et al.* 2023). This method is highly sensitive, allowing for accurate measurement of metabolites involved in particular physiological functions. In contrast, untargeted metabolomics

aims to profile all measurable metabolites in a sample without prior knowledge, making it a powerful approach for discovery and hypothesis generation (Roberts *et al.*, 2012; Tebani *et al.*, 2018). This untargeted approach allows researchers to reveal unexpected metabolic changes, leading to new insights into complex biological phenomena. However, untargeted metabolomics comes with challenges, such as complex data interpretation and the need for advanced informatics tools to identify unknown metabolites (Roberts *et al.*, 2012; Tebani *et al.*, 2018).

To address the vast chemical diversity of the metabolome, metabolomics often relies on sophisticated analytical techniques, primarily mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy (Emwas *et al.*, 2019; Marshall and Powers, 2017). Mass spectrometry, often paired with chromatographic separation like liquid or gas chromatography (LC-MS or GC-MS), excels in detecting low-abundance metabolites with high sensitivity. NMR spectroscopy, while less sensitive than MS, is highly reproducible and requires minimal sample preparation, making it ideal for quantifying abundant metabolites. The combination of MS and NMR in a multiplatform approach significantly enhances metabolite coverage, providing a more comprehensive understanding of the metabolome (Emwas *et al.*, 2019). This integrated approach is essential, as each technique captures different classes of metabolites, helping to overcome the limitations of individual methods.

Despite the progress in analytical techniques, metabolomics faces ongoing challenges in achieving complete metabolome coverage (Danzi *et al.*, 2023). The metabolome is chemically complex, with metabolites varying greatly in polarity, solubility, and abundance. As a result, even the most advanced technologies can detect only a fraction of the total metabolites present in a biological sample (Hajjar *et al.*, 2023). This limitation drives the development of innovative

strategies, such as combining ion mobility spectrometry with mass spectrometry, creating more robust computational tools, and expanding metabolite databases. These efforts aim to improve the accuracy and depth of metabolomic analyses, allowing for better characterization of metabolic pathways and the discovery of new biomarkers.

Different Tools to Study Metabolomics

Metabolomics, a field dedicated to the comprehensive analysis of small molecules within biological systems, has made substantial strides due to advancements in analytical technologies, particularly Nuclear Magnetic Resonance (NMR) spectroscopy and Mass Spectrometry (MS). Each method brings unique strengths to metabolomics, enhancing our capacity to study metabolites with high sensitivity, resolution, and accuracy (Emwas *et al.*, 2019; Marshall & Powers, 2017).

Nuclear Magnetic Resonance (NMR) Spectroscopy. NMR spectroscopy is a cornerstone of metabolomics, known for its quantitative precision, reproducibility, and non-destructive nature (Emwas *et al.*, 2019). This technique requires minimal sample preparation and does not depend on ionization, which makes it especially effective for detecting a broad range of metabolites (Nagana & Raftery, 2021), including those challenging for MS, such as polar molecules like sugars and organic acids. The ability of NMR to analyze samples without altering their structure is particularly beneficial in studies requiring high-throughput or longitudinal analysis, as the same sample can be analyzed multiple times under various conditions.

Technological advancements have propelled NMR capabilities significantly. High-field magnets and cryoprobes enhance signal sensitivity and resolution, allowing the detection of metabolites at low concentrations, though still not reaching the sensitivity levels of MS.

Additionally, two-dimensional (2D) NMR techniques, such as 2D ^1H - ^{13}C HSQC (Heteronuclear Single Quantum Coherence), are now frequently employed to improve spectral resolution and reduce peak overlap, which is a limitation in traditional one-dimensional (1D) NMR (Wu *et al.* 2022; Weitzel *et al.*, 2020). Furthermore, automated data processing tools, such as Chenomx, AQuA, and Bayesil, have emerged to streamline NMR data analysis, enabling accurate quantification and identification of metabolites with minimal intervention from person.

Despite its advantages, there are limitations from NMR metabolomics analysis. The sensitivity of NMR remains lower than MS, restricting its ability to detect metabolites in trace amounts (Emwas *et al.*, 2019). However, ongoing developments in hardware, such as microprobes and dynamic nuclear polarization, aim at enhancing detection limits (Letertre *et al.*, 2021). Additionally, future directions for NMR include integrating it more seamlessly with other platforms (such as MS) to provide a holistic analysis of the metabolome.

Mass Spectrometry (MS). Mass Spectrometry, often coupled with Liquid Chromatography (LC-MS) or Gas Chromatography (GC-MS), is celebrated for its sensitivity and broad metabolome coverage. MS can detect metabolites at femtomolar concentrations, a feature that makes it invaluable for untargeted metabolomics, where capturing as many metabolites as possible is essential (Rakusanova *et al.*, 2023). Chromatographic separation before MS analysis helps resolve complex mixtures by separating metabolites based on their physicochemical properties, reducing matrix effects and enhancing the accuracy of mass spectrometric analysis.

Recent innovations in MS technology, such as high-resolution Orbitrap and time-of-flight (TOF) detectors, as well as ion-mobility spectrometry (IMS), have further extended the capacity of MS to separate isomeric and polar compounds that are difficult to distinguish in traditional

setups. Mixed-mode chromatographic methods, which combine different stationary phases, are increasingly utilized to broaden metabolome coverage, enabling the analysis of highly polar and ionic metabolites that are essential in studying primary metabolic pathways.

However, MS also has its challenges. Ion suppression is a common issue in complex biological matrices, where some metabolites interfere with the ionization of others, resulting in reduced sensitivity for certain compounds (Furey *et al.*, 2013). Furthermore, MS data analysis requires extensive computational support to manage and interpret large datasets, especially in untargeted studies where unknown compounds are frequently encountered (Gorrochategui *et al.*, 2016; Ebbels *et al.*, 2023). Innovations in data-independent acquisition (DIA) methods and integration of advanced data processing software are essential to address these challenges and support the expanding role of MS in metabolomics.

Other Analytical Tools in Metabolomics. While Nuclear Magnetic Resonance (NMR) and Mass Spectrometry (MS) are the mainstays of metabolomics, several other tools and technologies play crucial roles in expanding metabolome coverage, enhancing sensitivity, and improving data quality. Tools such as Gas Chromatography (GC), Liquid Chromatography (LC), Capillary Electrophoresis (CE), and emerging methods like Ion Mobility Spectrometry (IMS) and Infrared Spectroscopy (IR) contribute to a more comprehensive approach to metabolite analysis (Patel *et al.*, 2021). Each tool offers unique strengths in resolving specific types of metabolites, often complementing NMR and MS in metabolomics studies.

1. Gas chromatography (GC)

Gas Chromatography (GC) is a widely used separation technique in metabolomics, particularly effective for analyzing volatile and semi-volatile metabolites. Coupled with Mass

Spectrometry (GC-MS), it provides high sensitivity and resolution for compounds like fatty acids, organic acids, and alcohols (Fiehn, 2016). One significant advantage of GC is the ability to separate isomers that might co-elute in LC, making it indispensable for metabolomics studies that focus on volatile or thermally stable metabolites. Additionally, derivatization procedures, which modify metabolites to increase volatility, extend GC's utility by enabling the analysis of otherwise non-volatile compounds, expanding its application in the metabolic profiling of biofluids and tissues (Fiehn, 2008; Dettmer *et al.* 2017).

Despite these strengths, GC is less versatile than LC due to the sample derivatization requirement and limited applicability to non-volatile compounds. However, for targeted analyses or studies focused on volatile metabolites, GC-MS remains a valuable tool that complements LC-MS and NMR by providing detailed insights into specific metabolite classes.

2. Liquid Chromatography (LC)

Liquid Chromatography (LC), especially when paired with MS, is one of the most flexible and powerful separation tools in metabolomics. LC can handle a wide range of polarities and molecular weights, allowing it to profile a broad spectrum of metabolites, from lipids and peptides to amino acids and nucleotides (Xiao *et al.*, 2012; Rakusanova & Cajka, 2024). High-Performance Liquid Chromatography (HPLC) and Ultra-High-Performance Liquid Chromatography (UHPLC) are frequently used in metabolomics for their high resolution and speed, improving both throughput and sensitivity in complex sample analyses.

Recent advances in mixed-mode chromatography, which combines different stationary phases, have expanded the ability of LC to separate structurally similar metabolites, such as isomers. This is particularly important for untargeted metabolomics studies, where complex

mixtures of metabolites must be resolved with high accuracy. Moreover, the flexibility of LC-MS in coupling with various detection methods (e.g., UV, fluorescence) makes it adaptable to a variety of experimental needs in metabolomics.

3. Capillary Electrophoresis (CE)

Capillary Electrophoresis (CE) has emerged as a powerful tool in metabolomics for separating charged or highly polar metabolites, such as amino acids, nucleotides, and small organic acids (Soga, 2023). CE separates metabolites based on their charge-to-mass ratio (Soga, 2023), allowing it to resolve ionic compounds that may be challenging for LC or GC. When paired with MS (CE-MS), it provides high sensitivity and is particularly effective for analyzing low-volume samples (Lageveen-Kammeijer *et al.*, 2019), making it suitable for single-cell metabolomics or studies with limited sample availability.

One of the main advantages of CE is its high efficiency and resolution, especially for ionic species. However, its utility is limited by the requirement that metabolites carry a net charge for optimal separation (Soga, 2023), which restricts its application for neutral molecules. Despite this limitation, CE is an important tool in metabolomics, especially for targeted studies involving charged metabolites, where it complements LC and GC-based methods.

4. Ion Mobility Spectrometry (IMS)

Ion Mobility Spectrometry (IMS) is an emerging technique in metabolomics, used to separate ions in the gas phase based on their size, shape, and charge (Burnum-Johnson *et al.*, 2019; Dodds *et al.*, 2019). When coupled with MS (IMS-MS), it provides an additional separation dimension, enhancing the identification and characterization of complex metabolite mixtures. IMS is

particularly useful for distinguishing structural isomers, which are often challenging to resolve using traditional MS alone.

IMS-MS has shown potential for applications in untargeted metabolomics, as it allows for the rapid screening of complex biological samples with improved resolution (Burnum-Johnson *et al.*, 2019; Dodds *et al.*, 2019). This technique is gaining traction in metabolomics as a complementary tool to LC-MS, as it can improve coverage of the metabolome by separating isomers that might otherwise be indistinguishable. However, IMS is still relatively new in metabolomics and may benefit from further optimization and integration with other platforms.

5. Infrared (IR) and Raman spectroscopy

Infrared (IR) and Raman spectroscopy are non-invasive analytical techniques that have gained attention in metabolomics for their ability to provide molecular information based on vibrational frequencies. IR spectroscopy measures the absorption of infrared light by molecular bonds (Maiti *et al.*, 2023), while Raman spectroscopy detects scattered light (Lima *et al.*, 2021). Both techniques offer unique spectral fingerprints that can be used to identify and quantify metabolites without the need for extensive sample preparation.

While less commonly used than NMR and MS, IR and Raman spectroscopy are valuable in situations where minimal sample processing is required or when studying complex biological matrices like tissues (Maiti *et al.*, 2023). They have shown potential in rapid diagnostic applications, as spectral data can indicate the presence of specific biomolecular features. However, their sensitivity is generally lower than NMR and MS, and they are best suited for targeted metabolomics applications where specific metabolites or classes of metabolites need to be monitored.

Each of these tools contributes to the field of metabolomics by offering unique advantages for specific types of metabolites or analytical needs. Together with NMR and MS, these techniques provide a comprehensive toolkit for exploring the metabolome, enhancing our understanding of biological processes, and supporting advances in biomedical research.

NMR, A Toolset of Strategies for Studying Metabolism

Nuclear Magnetic Resonance (NMR) spectroscopy has long been a foundational tool in studying metabolism, offering unique capabilities that set it apart from other analytical techniques (Moco, 2022). The non-destructive nature, quantitative accuracy, and structural elucidation capabilities make NMR a versatile tool in metabolomics research. Over the years, various NMR strategies have been developed to overcome its inherent limitations, such as sensitivity, and to enhance its utility in metabolic studies. Different NMR strategies will be summarized to advance our understanding of metabolism, with insights from recent advancements in NMR-based metabolomics.

High-Resolution One-Dimensional (1D) NMR. One-dimensional proton (^1H) NMR spectroscopy is a widely used technique in metabolomics due to its simplicity, non-destructive nature, quantitative reliability, and broad metabolite coverage (Emwas *et al.*, 2019; Wishart *et al.*, 2022). As one of the most accessible NMR applications, ^1H NMR allows for the rapid analysis of a range of metabolite types in biological samples: biofluids, tissues, and even intact organisms. One of the primary advantages of ^1H NMR is its high reproducibility and quantitative accuracy, making it ideal for longitudinal studies where precise metabolite concentration measurements are essential (Rolin *et al.*, 2013). Unlike other techniques, ^1H NMR provides direct, absolute quantification based on signal integrals directly proportional to metabolite

concentrations. This accuracy and reproducibility make ^1H NMR particularly useful for clinical studies that require reliable comparisons across multiple samples collected over time.

Another significant advantage of ^1H NMR is the non-destructive nature. Samples analyzed by ^1H NMR can be preserved for further analysis using other techniques, like mass spectrometry (MS), or for re-analysis under different conditions. This is especially beneficial in studies with limited sample availability: such as rare biofluids or samples from small experimental groups, where sample preservation is critical (Nagana Gowda & Raftery, 2023). The non-destructive ^1H NMR analysis capability is also valuable for experiments that require repeated measurements when monitoring metabolic changes over time in response to treatment or environmental changes.

The ^1H NMR typically requires minimal sample preparation, which allows for a streamlined workflow and increases the potential for high-throughput analysis. Unlike techniques requiring derivatization or complex extractions, ^1H NMR can be applied directly to various biological samples, from blood plasma to urine, with minimal intervention. This ease of preparation is advantageous in large-scale studies, where many samples need to be processed efficiently under consistent conditions. Additionally, the broad detection range of ^1H NMR, which encompasses various hydrogen-containing compounds, enables comprehensive metabolite profiling across key metabolic pathways, including those involving organic acids, amino acids, and sugars.

The ability of ^1H NMR to analyze complex biological mixtures without the need for prior separation makes it especially suitable for untargeted metabolomics studies. Biofluids, for instance, contain a complex array of metabolites that can be detected and differentiated based on distinct chemical shifts and signal patterns in ^1H NMR spectra. This capability is critical for clinical studies where the goal is to capture a global snapshot of metabolic profiles without extensive

sample alteration, thereby preserving the native state of metabolites. The robust, untargeted analysis capability has made ^1H NMR a popular choice in biomarker discovery for a wide range of diseases, including cancers, cardiovascular diseases, and neurological disorders, where it can reveal metabolic shifts associated with pathological conditions (Silva *et al.*, 2020; Nagana Gowda & Raftery, 2021; 2023).

In addition to its analytical strengths, ^1H NMR is relatively straightforward to operate compared to more complex NMR techniques. This accessibility makes it feasible for use in various research settings, from academic labs to clinical facilities (Silva *et al.*, 2020; Serkova & Niemann, 2006). Furthermore, since hydrogen nuclei are highly abundant, ^1H NMR does not require isotopic enrichment, which is necessary for other nuclei like ^{13}C or ^{15}N , thereby reducing costs and simplifying sample preparation (Rowlinson *et al.*, 2022; Nagana Gowda & Raftery, 2021; 2023). While ^1H NMR does have limitations, such as lower sensitivity compared to MS and potential peak overlap in complex mixtures, advancements in NMR hardware, including high-field magnets and cryoprobes, have significantly enhanced its sensitivity and resolution.

Overall, ^1H NMR provides a unique combination of quantitative accuracy, reproducibility, minimal sample preparation, and broad metabolite coverage, making it indispensable in metabolomics. Its strengths are particularly valuable for high-throughput and clinical studies, where sample preservation and quantitative stability are paramount. As NMR technology continues to advance, ^1H NMR will likely remain a foundational tool in metabolomics, playing a key role in unraveling the complex biochemical networks that drive health and disease.

Two-Dimensional (2D) NMR Techniques. To enhance the capabilities of 1D NMR, two-dimensional (2D) NMR techniques, such as ^1H - ^{13}C Heteronuclear Single Quantum Coherence

(HSQC) and Total Correlation Spectroscopy (TOCSY), have been developed. These 2D methods add an additional dimension of separation, greatly improving spectral resolution and facilitating more precise identification and quantification of metabolites in complex mixtures (Marchand *et al.*, 2017; Xia *et al.*, 2008; Emwas *et al.*, 2019). For example, HSQC provides correlation between protons and directly bonded carbons, increasing resolution and enabling the detailed characterization of complex metabolites.

The advantage of 2D NMR techniques is their ability to capture structural information about metabolites that may be difficult to resolve in 1D spectra (Dona *et al.*, 2016; Mahrous & Farag, 2015). While these methods are more time-consuming, non-uniform sampling (NUS) and parameter optimization have been developed to reduce acquisition times without compromising resolution. These advancements are making 2D NMR increasingly viable for high-throughput studies in metabolomics.

Stable Isotope Labeling for Metabolic Flux Analysis. Stable isotope labeling, particularly with ^{13}C and ^{15}N , is a powerful NMR strategy for tracing metabolic pathways and quantifying flux through different biochemical routes. By incorporating isotopically labeled substrates (like ^{13}C -glucose), researchers can track the incorporation of labeled atoms into metabolites across pathways, revealing metabolic flux and compartmentalization within cells and tissues (Walther *et al.*, 2012; Antoniewicz, 2018; Allen & Young, 2020). This approach is especially useful for studying metabolic dynamics and understanding how different pathways contribute to energy production, biosynthesis, and other cellular functions.

Stable isotope labeling combined with NMR provides unparalleled insight into metabolic flux without invasive sampling or cell disruption. This technique has been instrumental in metabolic

studies related to cancer, obesity, and other metabolic diseases, where tracing the flux of nutrients through specific pathways can reveal dysregulated metabolic processes.

Dynamic Nuclear Polarization (DNP) for Sensitivity Enhancement. A significant challenge in NMR is its lower sensitivity compared to Mass Spectrometry (MS), which limits the detection of low-abundance metabolites. Dynamic Nuclear Polarization (DNP) has emerged as an effective strategy to enhance NMR sensitivity by transferring polarization from unpaired electrons to nearby nuclei, thereby amplifying signal intensity by several orders of magnitude (Menzildjian *et al.*, 2023; Abhyankar & Szalai, 2021). This technique is beneficial for analyzing low-concentration metabolites in complex biological samples, expanding the range of detectable compounds in NMR-based metabolomics.

DNP-enhanced NMR is opening new avenues for studying metabolic processes that involve minor metabolites or metabolites with low natural abundance, such as those involved in signaling or secondary metabolism (Hu, 2011; Chaudhari *et al.*, 2017). While DNP setups are still relatively expensive and complex, ongoing advancements are expected to make this technology more accessible and integrative in routine metabolic studies.

Automated Quantification and Identification Algorithms. The development of automated software tools, such as AQuA, Chenomx, and Bayesil, has been pivotal in accelerating NMR-based metabolomics by reducing manual intervention in data processing. These algorithms use probabilistic models and spectral fittings to quantify metabolites in complex mixtures, providing faster and more accurate analysis (Röhnisch *et al.*, 2021; Crook *et al.*, 2020; Lipfert *et al.*, 2019). Bayesil applies Bayesian statistical models to identify and quantify metabolites from ^1D NMR spectra with high accuracy, making it particularly useful in high-throughput clinical studies.

Automated tools are essential for achieving consistency in inter-laboratory studies and for scaling up NMR-based metabolomics to larger cohorts. By minimizing human error and optimizing data interpretation, these tools facilitate the reproducible identification and quantification of metabolites across various studies, thereby enhancing the reliability of NMR as a quantitative tool in metabolomics.

In Vivo NMR and Magnetic Resonance Imaging (MRI) for Metabolic Studies. *In vivo* NMR and Magnetic Resonance Imaging (MRI) are non-invasive strategies allowing real-time metabolic monitoring within living organisms. MRI can provide spatial and temporal information on metabolic changes in tissues, making it ideal for studying metabolism in intact biological systems, including animal models and human subjects (Hwang & Choi, 2015; Liu *et al.*, 2017). By coupling MRI with ^{13}C -labeled substrates, researchers can observe the metabolic activity in specific tissues or organs, offering valuable insights into disease states, drug efficacy, and metabolic adaptations in response to environmental changes (Anderson *et al.*, 2021; Stewart *et al.*, 2022). This approach is particularly relevant for clinical applications, where *in vivo* NMR can be used for diagnostic purposes or to monitor treatment efficacy over time. MRI-based metabolomics has shown promise in studying cancer metabolism, neurological disorders, and cardiovascular health, where metabolic changes play a significant role in disease progression.

Prospects and Challenges for NMR in Metabolism Studies. As NMR technology advances, its potential for metabolic studies continues to grow. Nevertheless, challenges such as the need for improved sensitivity, high-cost infrastructure, and complex data processing persist. The integration of advanced hardware, such as DNP setups, and the development of more sophisticated computational algorithms are expected to address some of these limitations,

making NMR a more accessible and powerful tool in metabolomics (Emwas *et al.*, 2019; Wishart *et al.*, 2022). Additionally, expanding metabolite databases and adopting standardized protocols will further enhance NMR's reproducibility and integration with other analytical platforms like MS.

In summary, NMR provides a versatile toolkit for studying metabolism, from quantifying individual metabolites to tracing dynamic metabolic pathways. Its range of strategies—1D and 2D NMR, isotope labeling, DNP, automated software, and *in vivo* applications—enable comprehensive exploration of the metabolome. As these techniques evolve, NMR will continue to play a critical role in elucidating the complex networks that underlie metabolic processes, advancing both fundamental research and clinical applications in metabolomics.

The Significance of the Study

The Significance of Novel Methodology of Arsenic Speciation

This study advances our understanding of arsenic exposure in murine models by providing a comprehensive assessment of both technical and biological sources of variability in arsenic quantification. By refining methodologies for arsenic extraction and speciation analysis, this research establishes a robust experimental framework for high-performance liquid chromatography-inductively coupled plasma mass spectrometry (HPLC-ICPMS), enabling more accurate comparisons across studies. Additionally, the study reveals critical influences of diet composition and gut microbiome on arsenic body burden, highlighting the need to account for these factors in murine models to better approximate human arsenic exposure scenarios. Notably, this research identifies the microbiome as a significant source of inter-individual variability,

suggesting that microbiome composition may be a modifiable factor in human disease susceptibility. Collectively, these findings provide essential insights that can improve study designs for arsenic research, inform risk assessment for arsenic-induced diseases, and guide future epidemiological studies aimed at linking specific microbiome taxa to arsenic toxicity.

The Significance of Arsenic Diffusion in Different Mouse Tissues at Early-stage Exposure

This study addresses critical gaps in our understanding of arsenic metabolism and tissue-specific diffusion in mammals, focusing on the early stages of arsenic exposure. With over 200 million people globally and 2.1 million population in the United States chronically exposed to unsafe arsenic levels in drinking water, the need for clear insights into arsenic metabolic pathways and detoxification mechanisms is urgent. While the liver is recognized as the primary organ for arsenic detoxification, research suggests that other tissues, including the kidney, lungs, and gut microbiome, play significant roles in arsenic methylation and biotransformation. This study provides an essential view into the early dynamics of arsenic species distribution by examining arsenic distribution in the liver, urinary bladder, and fecal pellets within the first 12 hours of exposure in a murine model. After intravenous injection, iAs^{III} is rapidly taken up by the liver, which undergoes reduction, methylation, and conjugation with glutathione (GSH). A significant portion of iAs^{III} is conjugated to form GSH-conjugated complexes, such as $iAs^{III}(GS)_3$, which are subsequently excreted into bile and transferred to the gastrointestinal tract. The data will offer valuable insights for medical interventions during acute arsenic exposure and improve understanding of factors influencing arsenic toxicity and carcinogenicity in humans. This research

ultimately contributes to developing strategies for timely detoxification and risk mitigation in populations exposed to arsenic.

The Significance of Metabolomics Analyses

This study provides valuable insights into the metabolic impact of the *ars* operon on *E. coli* metabolism, offering a detailed comparison of the wild-type W3110 strain with the *arsRBC*-deleted mutant AW3110. By analyzing metabolomic differences under both standard and arsenic-stressed conditions, this research demonstrates the far-reaching effects of arsenic resistance genes on bacterial metabolism, including nucleotide, energy, carbohydrate, amino acid metabolism, and oxidative stress responses. The findings highlight that genetic differences in the *ars* operon have a more pronounced influence on metabolite profiles than arsenic stress itself, underscoring the operon's regulatory role in bacterial adaptation to environmental stressors.

The other metabolomics research provides critical insights a potential mechanism for arsenic-carbon (As-C) bond formation in the absence of *arsM*, supported by NMR metabolomics data showing the critical role of intracellular metabolites, especially pyruvate, in facilitating As-C bond generation. By identifying and characterizing novel genes and pathways potentially responsible for this process, this study sheds light on how arsenate (iAs^{V}) is incorporated into organic molecules within microbial systems. The integration of bacterial genetic engineering with advanced analytical techniques, such as HPLC-ICPMS, allows for a detailed examination of As-C bond formation and arsenic metabolism in model bacterial strains. This work not only deepens our understanding of the biochemical pathways for arsenic detoxification and assimilation but also contributes valuable knowledge about arsenic biogeochemical cycles.

These metabolomics findings contribute to our understanding of arsenic metabolism in bacteria. The results suggest new avenues for exploring arsenic detoxification and resistance in environmental and pathogenic bacteria, with implications for bioremediation and public health.

CHAPTER TWO

ARSENIC SPECIATION METHODS REVEAL INTER-
INDIVIDUAL VARIABILITY IN MURINE MODELS OF
HUMAN ARSENIC EXPOSUREIntroduction

Arsenic is a toxic metalloid occurring naturally and widely distributed across the environment (Wang *et al.*, 2023). According to the World Health Organization (WHO), over 200 million people worldwide are chronically exposed to arsenic in drinking water at levels exceeding the established maximum contaminant level of 10 µg/L (World Health Organization, 2022). The inorganic forms of arsenic—specifically arsenite (iAs^{III}) and arsenate (iAs^V)—are the most common species found in drinking water and are recognized carcinogens linked to skin, bladder, lung, and liver cancers (Faita *et al.*, 2013; Hong *et al.*, 2014). Prolonged exposure to inorganic arsenic (iAs) also increases the risk of cardiovascular diseases and diabetes (Faita *et al.*, 2013; Hong *et al.*, 2014). Once ingested, arsenic toxicity varies significantly depending on its chemical form, with trivalent species being more toxic than pentavalent species; generally, inorganic species are more toxic than their organic (methylated) counterparts, with limited exceptions (Cohen *et al.*, 2013). Both mammalian and microbial cells can metabolize inorganic arsenic into other forms, significantly impacting toxicity.

Rodent models, particularly mice, are widely used in toxicological research, yet there is a lack of standardized methods for processing and assessing arsenic species in mouse tissues. Research has involved exposing lab mice to various concentrations of iAs in drinking water, from

1 part per billion (ppb) to over 100 parts per million (ppm) (Lu *et al.*, 2014; Barai *et al.*, 2017; Coryell *et al.*, 2018; Chi *et al.*, 2019). Commonly overlooked in murine studies are details such as the methods used for sample collection/processing and the specific arsenic species produced in vivo. Additionally, arsenic present in mouse chow and local water can serve as considerable and variable background sources in experiments. Furthermore, both humans and mice demonstrate significant individual variability in arsenic body burden and toxicity responses (Coryell *et al.*, 2018), highlighting the importance of appropriate sample sizes for rigorous statistical analysis. Evaluating these experimental factors is essential for generating more accurate disease risk assessments and understanding contributory variables.

The combination of high-performance liquid chromatography (HPLC) with inductively coupled plasma mass spectrometry (ICPMS) is a sensitive and widely referenced technique for detecting arsenic species in arsenic exposure studies. A standardized HPLC-ICPMS protocol was initially developed to assess the safety of retail chicken exposed to the organoarsenical Roxarsone (Peng *et al.*, 2014). This protocol enables the quantification of Roxarsone, its organoarsenical metabolites, arsenobetaine (AsB), iAs^{III}, iAs^V, and the primary arsenic metabolites found in human urine: monomethylarsonic acid (MMA^V) and dimethylarsinic acid (DMA^V). A subsequent study refined this protocol by incorporating proteolytic enzymes to enhance arsenic extraction efficiency from tuna fish reference material, with the papain enzyme increasing efficiency from 28% to 55% (Liu *et al.*, 2015).

Combining enzymatic proteolysis with HPLC-ICPMS represents the current state-of-the-art protocol for arsenic speciation in animal tissue. In this study, we adapted this protocol for mouse samples, evaluating essential methodological steps such as enzyme selection for digestion and

common tissue maceration techniques. We assessed background arsenic levels in drinking water and commercial lab mouse diets and explored the effects of diet type, food restriction (fasting), and microbiome status on individual variability in arsenic species in different tissues (stool, liver, and bladder) in mice receiving controlled, human-relevant oral doses of iAs^{III}. Finally, we assessed the inter-laboratory portability of the HPLC-ICPMS technique. Our findings aim to standardize arsenic speciation methodologies for murine models of human arsenic exposure and underscore the importance of sufficient sample sizes to address the notable individual variability in arsenic exposure (i.e., body burden).

Methods

Chemical Reagents

Stock solutions (10 mg/mL) were prepared for the following compounds: arsenobetaine (AsB, 95% purity, Sigma-Aldrich, St. Louis, MO), arsenic (III) oxide (iAs^{III}, 99.995% trace metals basis, Sigma-Aldrich, St. Louis, MO), sodium arsenate dibasic heptahydrate (iAs^V, ≥98% purity, Sigma-Aldrich, St. Louis, MO), methylarsonic acid (monomethylarsonic acid, MMA^V, 95%, Wako Chemicals, Richmond, VA), and cacodylic acid (dimethylarsinic acid, DMA^V, 99%, Sigma-Aldrich, St. Louis, MO). Fresh working solutions of arsenic species at concentrations of 0.5, 1, 5, 10, 50, and 100 µg/L were prepared daily through serial dilutions from the stock solutions. Papain was obtained from Thermo Scientific (Waltham, MA), and pepsin was sourced from Sigma-Aldrich (St. Louis, MO). The mobile phase was prepared using HPLC-grade ammonium phosphate dibasic (Sigma-Aldrich, St. Louis, MO), with pH adjustments made using trace metal-grade ammonium hydroxide (VWR, West Chester, PA).

Experimental Animals

All mouse experiments were performed in compliance with protocols pre-approved by the Montana State University Institutional Animal Care and Use Committee (IACUC). Experimental mice were housed at the Montana State University Animal Resource Center, an American Association for the Accreditation of Laboratory Animal Care (AAALAC)-accredited facility. C57BL/6J mice were sourced from The Jackson Laboratory and housed in individually ventilated cages under specific pathogen-free conditions, including protection from murine norovirus, with sterilized bedding. Germ-free C57BL/6J mice were maintained within hermetically sealed, HEPA-filtered vinyl isolators and provided with autoclaved water and LabDiet mouse chow (5013 Autoclavable Rodent Breeder Diet, LabDiet, St. Louis, MO) to ensure sterility. All food and water supplied to the germ-free mice were autoclaved, quarantined, and tested for contaminants prior to use. Germ-free status was routinely confirmed using both aerobic and anaerobic culturing techniques on rich media (Mueller-Hinton broth and agar plates), supplemented by periodic PCR assays targeting the bacterial 16S rRNA gene, with DNA extracted from stool samples used as the template.

Murine Arsenic Exposures with Fasting and Diet Manipulation

All arsenic exposures were conducted using iAs^{III} with age-matched mice (7–13 weeks old) of both sexes. For enzyme digestion comparisons, a C57BL/6J mouse unexposed to arsenic was humanely euthanized by isoflurane overdose, and the liver was collected. Another C57BL/6J mouse was exposed to 10 ppm iAs^{III} in drinking water ad libitum for 12 hours, after which humane

euthanasia was performed via isoflurane overdose, and the liver was collected for arsenic speciation analysis.

To assess tissue maceration methods and examine the effects of fasting, diet, and microbiome status, conventional and germ-free C57BL/6J mice were exposed to iAs^{III} by oral gavage at a dose of 1 μg iAs^{III} per gram of body weight and euthanized after 24 hours. Stool samples were collected just before euthanasia, and liver and/or bladder tissues were collected post-euthanasia. For fasting experiments, food and water were removed for three hours before mice were exposed to iAs^{III} by oral gavage (1 μg iAs^{III} per gram of body weight). After an additional three-hour fasting period, food and water were reintroduced, and the mice were euthanized 24 hours post-gavage.

For diet evaluations, mice initially fed a standard lab chow (5053 PicoLab Rodent Diet 20, LabDiet, St. Louis, MO) were switched to a purified powdered diet (110700 AIN-93G Purified Rodent Diet, Dyets, Bethlehem, PA) one week before iAs^{III} exposure.

Enzyme Comparisons

Enzyme-assisted extraction methods were adapted based on protocols from Liu *et al.* (2015), who tested various protease digestions on chicken breast meat, and Peng *et al.* (2014), who modified prior pepsin digestion techniques for chicken liver. A 20 g/L papain solution was prepared in ultrapure water (MilliQ) and incubated overnight with shaking in a 60°C water bath. Similarly, a 20 g/L pepsin solution containing 0.5% HCl (trace metal grade) was prepared and incubated overnight with shaking at 37°C. For spiking experiments, arsenic stock solutions were added to the papain or pepsin mixtures to achieve a final concentration of 5 $\mu g/L$ (ppm), and incubated as previously described.

For extraction, three replicate 0.1 g liver samples from an unexposed mouse were combined with 400 μ L of the pepsin solution, macerated with a mechanical tissue homogenizer (Omni International, Kennesaw, GA) at 20,000 rpm for one minute, and incubated overnight with shaking as described. Additionally, six replicate 0.1 g liver samples from a mouse exposed to iAs^{III} (10 μ g per g bodyweight) were mixed with 400 μ L of either papain or pepsin solutions (n=3 per enzyme), homogenized using the same mechanical tissue homogenizer, and incubated overnight as described. After incubation, samples were centrifuged at 12,000 rpm for five minutes, and the supernatant was filtered with a reusable 3KDa molecular weight cutoff membrane filter (MilliporeSigma, Burlington, MA) by centrifugation at 12,000 rpm for 30 minutes at 4°C. Filtered samples were stored at -80°C until HPLC-ICPMS analysis.

To ensure the filters were free of humectants, which can interfere with arsenic detection using HPLC-ICP-MS, all 3K membrane filters were washed before use. Filters were placed in 50 mL conical tubes containing 30 mL of ultrapure water and incubated with shaking at room temperature for five minutes. This washing step was repeated three times, after which the filters were stored in fresh ultrapure water at 4°C.

Papain Proteomics

A sample of the commercial papain utilized in this study was sent to the IDeA National Resource for Quantitative Proteomics (Little Rock, AR) for proteomic analysis. The workflow included digestion of the papain suspension with trypsin, followed by protein identification using an Orbitrap Fusion LC-MS system.

Tissue Maceration

For bead beating, four types of 2 mL lysing matrix tubes were sourced from commercial suppliers, including lysing matrix types A, B, and E from MP Biomedicals (Santa Ana, CA) and Qiagen PowerBead Pro. Each lysing matrix tube received a 400 μ L aliquot of ultrapure water. Tubes were processed in a mini-BeadBeater 24 (BioSpec Products Inc., Bartlesville, OK, USA) for one minute, then centrifuged at 12,000 rpm for five minutes. The supernatant was filtered through a 3 KDa membrane filter by centrifugation at 12,000 rpm for 30 minutes at 4°C. To remove background arsenic, an acid wash was performed on the commercial beads: each tube was treated with 200 μ L of 50% nitric acid and 500 μ L of 20% hydrochloric acid, mixed by inversion, and incubated at 95°C for 30 minutes. After cooling to room temperature, the acid solution was discarded, and the matrix was rinsed with 1 mL of ultrapure water, followed by vortexing for 10 seconds. Tubes were then centrifuged at 12,000 rpm for one minute to remove water. This washing step with ultrapure water was repeated four times, and tubes were dried overnight in a chemical hood at 70°C using a heat block.

For liver maceration by bead beating, 0.1 g of fresh liver tissue was combined with 400 μ L of ultrapure water and transferred into an acid-washed lysing matrix tube. Tubes were processed for one minute, centrifuged at 12,000 rpm for five minutes, and the supernatant was filtered using a 3 KDa membrane filter by centrifugation at 12,000 rpm for 30 minutes at 4°C. Liver and stool samples from an unexposed germ-free C57BL/6J mouse and from a germ-free C57BL/6J mouse exposed to 1 μ g iAs^{III} per gram of body weight were used for tissue comparison studies.

For mechanical homogenization, an additional 0.1 g fresh liver sample was mixed with 400 μ L of ultrapure water and homogenized with an Omni THq Homogenizer (Omni International,

Kennesaw, GA, USA) at 20,000 rpm for one minute. Samples were then centrifuged at 12,000 rpm for five minutes, and the supernatants were collected and filtered using 3 KDa membrane filters with centrifugation at 12,000 rpm for 30 minutes at 4°C.

Processing of Mouse Tissues

Unless otherwise specified, 0.1 gram samples of liver and stool or 0.03 gram samples of urinary bladder tissue were combined with 400 μ L of pepsin solution for digestion. The tissues were then processed using a mechanical tissue homogenizer and filtered as previously described.

HPLC-ICPMS

High-performance liquid chromatography (HPLC) was conducted at Montana State University using an Agilent 1260 HPLC system (Agilent Technologies), which was equipped with a quaternary pump, an online degasser, and a system control unit. Arsenic species were separated using a Hamilton PRP-X100 anion-exchange column (250 $\times$ 4.1 mm, 10 μ m particle size). Both samples and standard solutions were introduced through a 50 μ L sample loop. Isocratic elution was applied to separate the arsenic compounds, using 10 mM (NH₄)₂HPO₄ at pH 8.25 (adjusted with trace metal-grade NH₄OH) as the mobile phase. The flow rate was set to 1.5 mL/min, achieving complete separation of arsenic species within 10 minutes. Inductively coupled plasma mass spectrometry (ICPMS) was performed on an Agilent 7800 ICPMS system, with arsenic compounds identified by matching retention times with single standards on the PRP-X100 column. Arsenic ions were detected by the ICPMS, with operating conditions optimized daily using a tuning solution to ensure precision. To evaluate HPLC-ICPMS consistency, the same operator processed all liver samples at Montana State University (Bozeman, MT, USA). Samples were then

split into two sets and aliquoted into separate mass spectrometry vials: half were shipped on dry ice to Florida International University (Miami, FL) for HPLC-ICPMS analysis, while the other half were analyzed at Montana State University.

At Florida International University, HPLC-ICPMS operations followed an established protocol (Dong *et al.*, 2022). Briefly, the HPLC system for arsenic speciation included a degasser, an HPLC column oven, a binary pump, and an autosampler (NexSar HPLC, PerkinElmer), coupled to an ICPMS (NexION 1000, PerkinElmer) for monitoring ^{75}As ($m/z = 75$).

Standard solutions containing 0, 1, 2, 5, 10, 20, and 50 ppb of iAs^{III} , iAs^{V} , MMA^{V} , and DMA^{V} were freshly prepared and analyzed at the beginning of each batch of sample analyses by the C18 column. For the cation exchange column assays, standard solutions containing 0, 1, 2, 5, 10, and 20 ppb of AsB, TMAAsO , and arsenic choline (AsC) were used. iAs^{V} , MMA^{V} , DMA^{V} , and AsB, TMAAsO , and AsC in the samples were quantified using Clarity software (version 8.4.0.47, DataApex Ltd., Prague, Czech Republic), based on the calibration curves prepared from the results of the above-mentioned standard solutions. The peaks of iAs^{III} and AsB overlap on the C^{18} column due to their close retention time, while these two species were clearly separated by the cation exchange column method. Therefore, iAs^{III} was quantified by subtracting AsB levels determined by the cation exchange column method from the sum of As^{III} and AsB determined by the C18 column method.

Statistical Analyses

All statistical analyses in this study were conducted using GraphPad Prism (La Jolla, CA). For comparisons involving enzyme and tissue maceration, small sample sizes (typically $n=3$) necessitated the use of unpaired Welch's t-tests, which do not assume equal standard deviations

between groups. P-values from multiple tests were adjusted using the two-stage step-up False Discovery Rate (FDR), with significant findings reported at $q=5\%$. One-sample tests were used to determine whether the percent recovery of arsenic species deviated from the theoretical value of 100%, with the significance level (α) set at 0.01 to account for multiple ($n=5$) one-sample tests, following a traditional Bonferroni adjustment.

For larger sample sizes ($n > 5$) and assessments involving multiple batches or cages, parametric tests were applied when normality was confirmed via the D'Agostino & Pearson test. Data transformations (square root or log base 10) were attempted when normality was not confirmed. If these transformations normalized the data, parametric statistics were applied to the transformed data; if not, non-parametric equivalents were used. For instances without non-parametric equivalents (e.g., two-way designs), multiple Mann-Whitney or Kruskal-Wallis tests with FDR correction ($q=5\%$) were performed, with q -values presented in the figures. For clarity and comparison, untransformed data are shown in all figures.

To assess variance, an F test was used for comparisons between two groups, while Levene's median test was used for three or more groups. The figure legends provide details of all tests and transformations.

Results

Enzyme-Assisted Extraction of Arsenic Species

Previous studies have shown that both papain and pepsin enhance arsenic extraction efficiency from chicken breast meat (Liu *et al.*, 2015). To assess these enzymes for arsenic extraction from mouse tissues, we first measured background arsenic levels in the commercial

preparations. Detectable levels of iAs^{III} were present in both papain and pepsin products, with papain containing approximately 3.5 times more iAs^{III} per gram than pepsin (Fig. 2.1A), and no other arsenic species were detected in either product. To determine enzyme-specific effects on arsenic recovery, new enzyme preparations were spiked with arsenic standards and incubated for 12 hours, after which arsenic species were quantified (see Methods). The recoveries of iAs^{III} , MMA^{V} , and iAs^{V} varied between the two enzymes (Fig. 2.1B). When recovery rates were plotted and compared to a theoretical 100% recovery, papain incubation showed an increase of 31% in iAs^{III} while reducing MMA^{V} and iAs^{V} by 45% and 81%, respectively. In contrast, pepsin incubation provided more consistent recoveries (closer to 100%), reducing iAs^{V} by only 10% (Fig. 2.1C and 2.1D).

Given that mammalian liver contains proteases, we next examined whether incubating arsenic standards with untreated (arsenic-free) murine liver influenced recovery, beyond what was observed with pepsin alone. Minor changes in iAs^{III} and DMA^{V} recovery were noted after a 12-hour incubation with untreated liver homogenate (Fig. 2.1E). However, these differences did not remain statistically significant after adjusting for multiple comparisons (FDR correction). Finally, we compared arsenic recovery in liver samples from a mouse exposed to 10 ppm iAs^{III} in drinking water (for 12 hours) using papain versus pepsin digestion (see Methods). Consistent with prior findings (Liu *et al.*, 2015), papain yielded higher levels for each arsenic species compared to pepsin (Fig. 2.1F). However, papain's higher iAs^{III} background (Fig. 2.1A) and differential recovery patterns with arsenic standards (Fig. 2.1C and 2.1D) indicated potential interconversion among species, such as the reduction of iAs^{V} to iAs^{III} .

To investigate further, we conducted a proteomic analysis of the commercial papain, which is a crude plant extract, to identify enzymes potentially involved in arsenic species conversion. The analysis identified chymopapain as the dominant protease, along with other proteolytic enzymes (Table 2.1). Notably, glyceraldehyde-3-phosphate dehydrogenase, a papaya enzyme, was detected with high probability (100%). This enzyme has been shown to catalyze the reduction of iAs^V in human red blood cells, hemolysate, and rat liver cytosol (Gregus & Nemeti, 2005; Nemeti *et al.*, 2006). Together, these results suggest that papain may interconvert arsenic species, which could counterbalance its higher extraction efficiency. As a result, all further experiments in this study employed pepsin digestion exclusively.

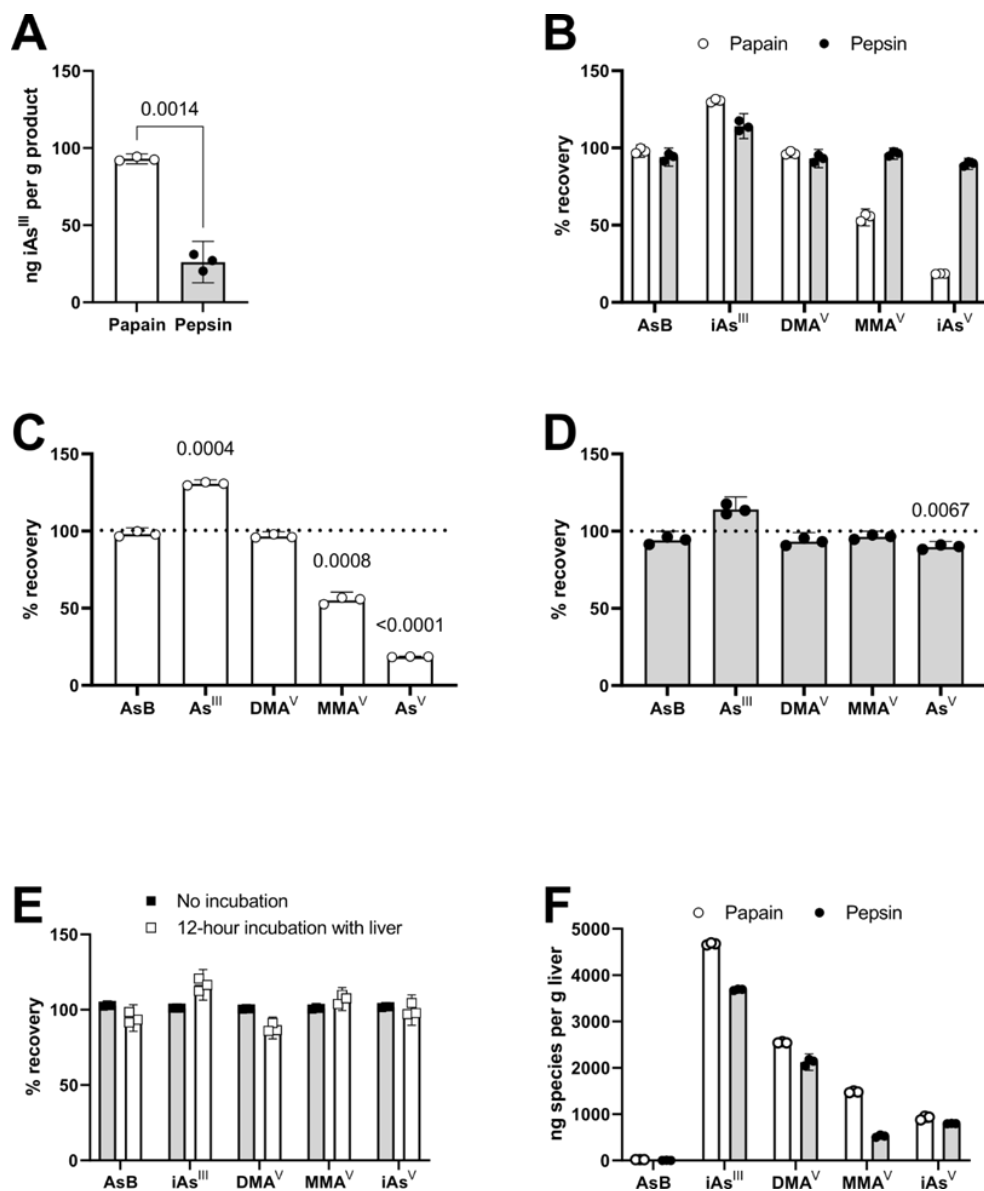


Figure 2.1 Enzyme-assisted extraction of arsenic from murine tissues. Enzyme digestion of murine tissues was performed using papain and pepsin. Background iAs^{III} in commercial papain and pepsin preparations (**A**, unpaired t test with Welch's correction). Levels of arsenic species standards following incubation with papain and pepsin (**B**). Recovery efficiencies of papain (**C**) and pepsin (**D**) tested against the theoretical threshold of 100% (one sample two tailed t test). Recovery of arsenic species standards with and without incubation in murine liver homogenate (**E**, multiple Mann Whitney tests and FDR correction for multiple comparisons). Recovery of arsenic species from the liver of an exposed mouse using either papain and pepsin digestion (**F**, multiple Mann Whitney tests and FDR correction for multiple comparisons). For all panels containing circles, open and closed circles represent papain and pepsin results, respectively. In panel E, open and closed squares represent results with and without 12-hours of pepsin incubation.

Table 2.1. Proteins identified in the commercial papain protease product.

Identified Proteins	UniProt Accession Number	Molecular Weight	Total Spectrum Count	Protein Identification Probability
Chymopapain (Fragment) OS=Carica papaya OX=3649 PE=3 SV=1	F6KSW9	39 kDa	286	100
Chymopapain OS=Carica papaya OX=3649 PE=1 SV=2	P14080	39 kDa	270	100
Papaya proteinase 4 OS=Carica papaya OX=3649 PE=1 SV=3	P05994	39 kDa	74	100
Papaya proteinase omega OS=Carica papaya OX=3649 GN=omegaii PE=2 SV=1	Q42673	41 kDa	71	100
Caricain OS=Carica papaya OX=3649 PE=1 SV=2	P10056	39 kDa	68	99
Glutamine cyclotransferase OS=Carica papaya OX=3649 PE=1 SV=1	O81226	33 kDa	64	100
Papain OS=Carica papaya OX=3649 PE=1 SV=1	P00784	39 kDa	67	100
Endochitinase OS=Carica papaya OX=3649 PE=1 SV=1	P85084	27 kDa	35	100
Cysteine proteinase 3 (Fragment) OS=Vasconcellea cundinamarcensis OX=35926 PE=1 SV=1	P32956	5 kDa	24	100
Papaya barwin-like protein OS=Carica papaya OX=3649 PE=1 SV=1	U5HK42	14 kDa	13	100
Cysteine proteinase 2 (Fragment) OS=Vasconcellea cundinamarcensis OX=35926 PE=1 SV=1	P32955	5 kDa	9	100
Latex serine proteinase inhibitor OS=Carica papaya OX=3649 PE=1 SV=1	P80691	21 kDa	8	100
Cysteine proteinase inhibitor OS=Carica papaya OX=3649 PE=2 SV=1	Q39561	11 kDa	8	100
GDSL esterase/lipase OS=Carica papaya OX=3649 PE=1 SV=1	P86276	38 kDa	5	100
Glyceraldehyde-3-phosphate dehydrogenase (Fragment) OS=Carica papaya OX=3649 GN=GAPDH PE=2 SV=1	M4I2N9	29 kDa	3	100
Protein TIC 214 OS=Jacaratia spinosa OX=1222208 GN=YCF1 PE=3 SV=1	A0A7L8UVM6 (+1)	222 kDa	2	98
CST complex subunit CTC1 (Fragment) OS=Carica papaya OX=3649 PE=2 SV=1	J7GLC1	45 kDa	3	92
Dicer-like 4 (Fragment) OS=Carica papaya OX=3649 PE=2 SV=1	A0A343J651	119 kDa	2	69
Heat shock protein 70 OS=Carica papaya OX=3649 PE=2 SV=1	A0A1I9RH42	71 kDa	2	98
Elongation factor 1-alpha (Fragment) OS=Carica papaya OX=3649 GN=EF1 PE=2 SV=1	M4I309	37 kDa	2	95

Tissue Maceration Techniques for Arsenic Speciation

Two commonly used techniques for processing large numbers of tissue samples are bead beating and tube-sized tissue homogenizers. To compare arsenic extraction and speciation between these methods, we first quantified background arsenic levels across four types of commercial bead-beating tubes (lysing matrices; see Methods). Detectable levels of iAs^V were present in all four tube types, with lysing matrix B showing the lowest level (Fig. 2.2). We attempted to reduce background iAs^V by acid washing, which effectively decreased background levels in lysing matrices A and B (Fig. 2.2). Further evaluation of two batches of lysing matrix B tubes confirmed that acid washing significantly reduced background iAs^V (Fig. 2.3A). Spiking experiments conducted with washed lysing matrix B tubes and arsenic standards showed complete recovery for DMA^V and MMA^V , with a slightly lower recovery for AsB (97%) (Fig. 2.3B). Inorganic arsenic species were somewhat affected by bead beating, with recoveries of 84% for iAs^{III} and 117% for iAs^V , suggesting that this process may oxidize some iAs^{III} to iAs^V .

To assess the impact of washing on arsenic recovery from mouse tissue, stool and liver samples were collected from an unexposed germ-free mouse and a germ-free mouse 24 hours after a controlled dose of $1 \mu g iAs^{III}$ per gram of body weight (see Methods). Background iAs^V levels in the stool from the unexposed mouse were influenced by the type of lysing matrix used for stool processing, with washed matrices showing lower levels (Fig. 2.3C); however, this difference was not statistically significant after correction for multiple comparisons. No significant differences were observed in arsenic species levels in the stool of the exposed mouse (Fig. 2.3D). Acid washing reduced levels of AsB and iAs^V in liver samples from exposed mice (Fig. 2.3E) and lowered the levels of all arsenic species in liver samples from the exposed mouse (Fig. 2.3F). Identical p-values

in panels E and F reflect the application of a rank sum test, given uniformly lower values in acid-washed versus unwashed tubes. These findings collectively indicate that acid washing effectively reduces background levels of arsenic species, particularly iAs^V . The presence of AsB in liver samples from an unexposed mouse further suggests a separate, background exposure source.

Due to the effect of bead beating on arsenic species recovery, we hypothesized that maceration using a mechanical tissue homogenizer, followed by pepsin digestion, would result in minimal iAs^{III} background and little to no iAs^V . To test this, we simulated maceration by using washed bead-beating tubes or in-tube homogenization (without tissue; see Methods), incubating with pepsin digestion solution and analyzing the results. Consistent with our hypothesis, tubes simulating mechanical homogenization showed lower iAs^{III} levels and no iAs^V , whereas both iAs^{III} and iAs^V were detected in washed bead-beating tubes (Fig. 2.4A).

We next compared arsenic species recovery in liver samples (three technical replicates per liver) from three C57BL/6J mice, 24 hours after controlled iAs^{III} exposure (1 μg per g body weight) by oral gavage. Species recovery did not differ significantly by maceration type (bead beating versus mechanical homogenization). However, there was notable inter-mouse variability (Fig. 2.5). Since the recovery was unaffected by the maceration method, we combined these data. We tested for inter-mouse differences in liver arsenic species levels, independent of maceration type. This analysis (Fig. 2.4B) confirmed significant inter-individual variability, suggesting that factors beyond the exposure dose impact arsenic body burden. These findings indicate that both bead beating and mechanical homogenization provide comparable results, and all further experiments employed mechanical tissue homogenization exclusively.

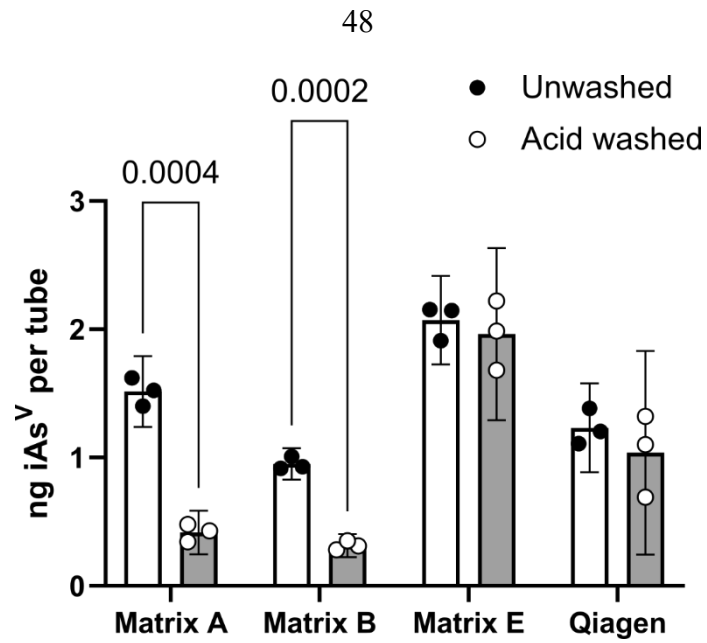


Figure 2.2 Background arsenic levels in commercial bead beat tubes. Four commercial bead beat tubes were compared before and after washing with strong acid (multiple unpaired t test with Welch's correction and FDR correction for multiple comparisons).

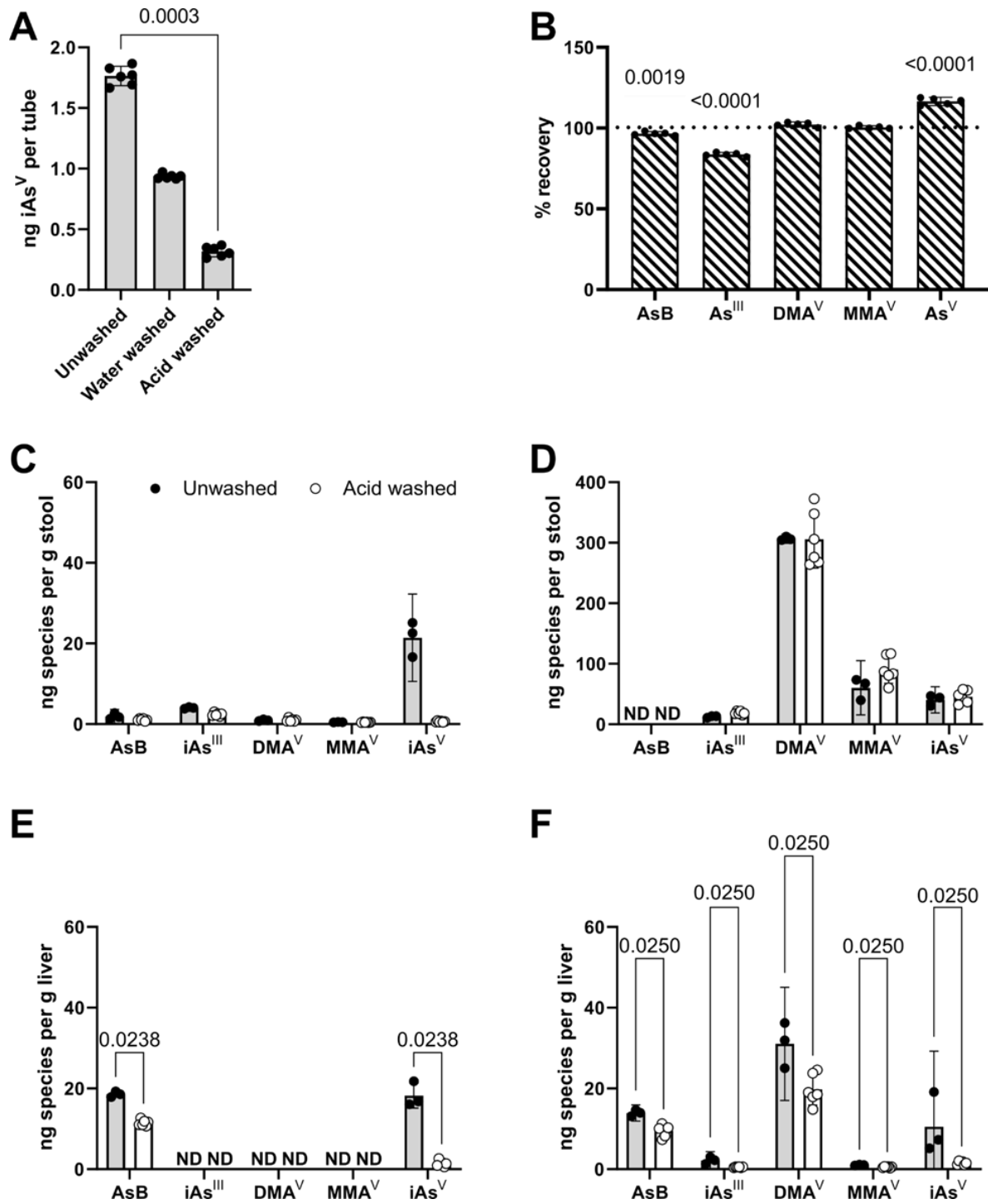


Figure 2.3 Evaluation of bead beating for murine tissue maceration and arsenic speciation. Background iAs^V in bead beat tubes containing lysing matrix B before and after washing with water and strong acid (A, Kruskal-Wallis test with Dunn's p value adjustment for multiple comparisons). Recovery efficiency of arsenic species standards following incubation with lysing

matrix B (B, one sample two tailed t test; Bonferroni adjustment for multiple comparisons; $\alpha = 0.01$). Arsenic species in stool of an unexposed (C) and exposed (D) mouse before and after acid washing (multiple Mann Whitney tests and FDR correction for multiple comparisons). Arsenic species in liver of an unexposed (E) and exposed (F) mouse before and after acid washing (multiple Mann Whitney tests in both panels; Bonferroni adjustment for multiple comparisons; $\alpha = 0.025$ in panel E; FDR correction for multiple comparisons in panel F). Graph labels for unwashed and acid washed bead beat tubes are identical for panels C-D and shown in panel C. No statistical tests were performed on undetectable arsenic species.

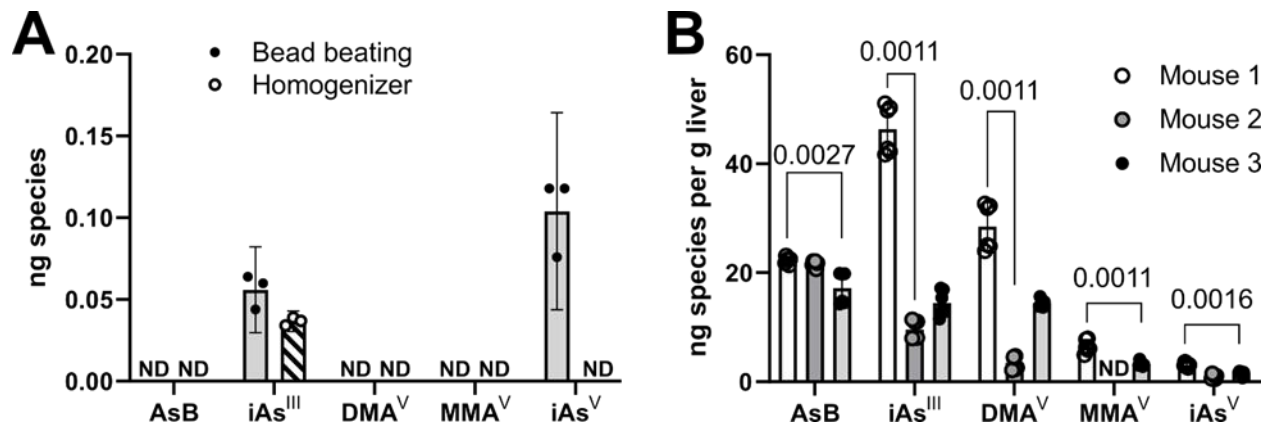


Figure 2.4 Comparison of bead beating and mechanical homogenization for murine tissue maceration and arsenic speciation. Background levels of arsenic species were compared in tubes mimicking bead beating or mechanical homogenization in the absence of tissue (A). Combined arsenic species (i.e., levels quantified by both bead beating and homogenization) in liver samples from three mice following controlled exposure to iAs^{III} (B; multiple Kruskal-Wallis tests with FDR correction for multiple comparisons). No statistical tests were performed on undetectable arsenic species.

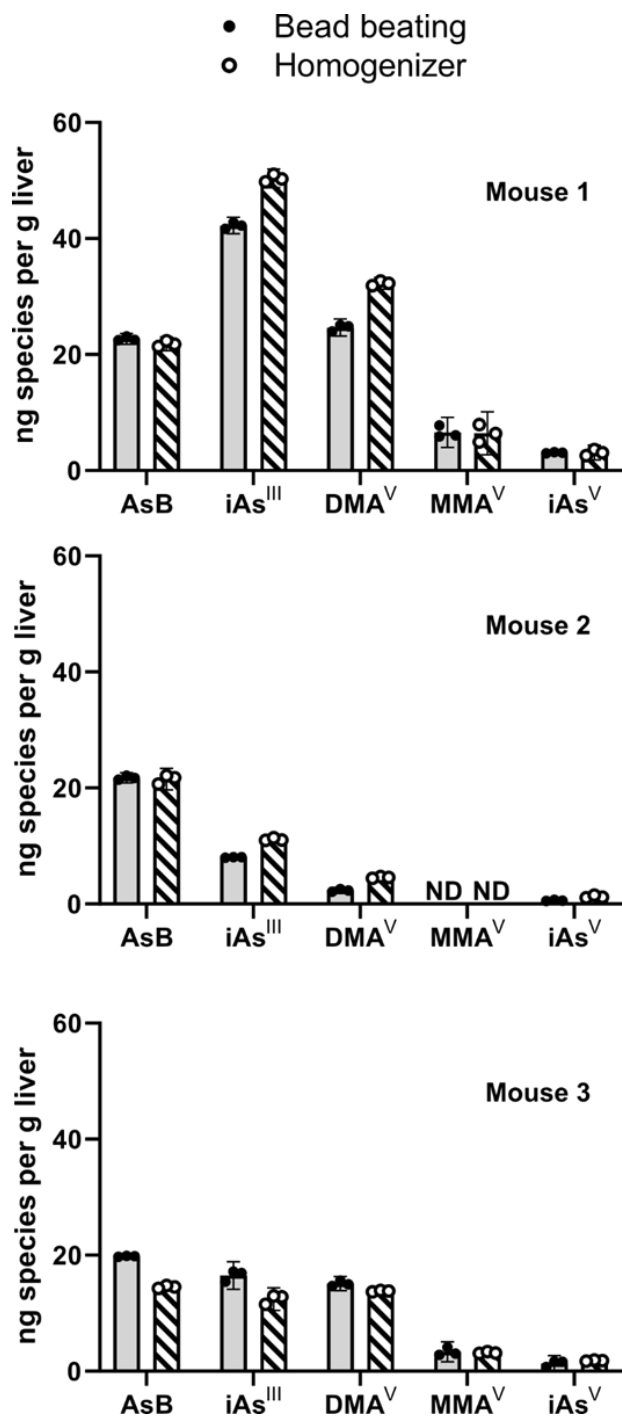


Figure 2.5 Arsenic species were compared in liver samples prepared by bead beating or homogenizer in three different mice following controlled exposure to iAs^{III} (circles represent n=3 technical replicates from the same liver). Graph label is the same for all panels as shown in the top panel.

The Influence of Diet Type on Arsenic Burden and Inter-Individual Variability

Human studies have indicated that diet composition influences individual variability in arsenic body burden and related disease risk. To explore the impact of different diets on arsenic body burden in mice, we first measured background arsenic species levels in three commercial lab mouse diets, including samples from multiple batches or shipments from the same suppliers. Both the standard lab mouse diet and an equivalent autoclavable diet for germ-free mouse husbandry contained considerable levels of AsB, DMA^V, and iAs^V (Fig. 2.6A). In contrast, a purified diet commonly used for controlled nutrition studies had undetectable levels of AsB, DMA^V, MMA^V, and significantly lower levels of iAs^V (Fig. 2.6A). All three diets showed undetectable or minimal amounts (<2.0 ng per g) of iAs^{III}. Batch-to-batch variability was significant for AsB in the germ-free diet (Fig. 2.6B) and iAs^V in the standard diet (Fig. 2.6C), while DMA^V variability between batches was not significant for either diet.

We then conducted controlled iAs^{III} exposures in groups of mice fed either the standard or purified diet and measured arsenic levels in stool, liver, and bladder samples after 24 hours. Total arsenic levels differed in stool (Fig. 2.6D) and liver (Fig. 2.6E) depending on diet, while no significant difference was noted in bladder samples (Fig. 2.6F). Mice on the purified diet excreted more arsenic in stool and accumulated less in the liver compared to mice on the standard diet. Interestingly, total arsenic levels in stool and liver showed a correlation (Pearson $r = -0.5187$, $R^2 = 0.2690$, $p=0.0079$), with linear regression indicating that for every 1% increase in total stool arsenic, total liver arsenic decreased by 0.06% (Fig. 2.6G).

Regarding specific arsenic species, AsB levels were significantly lower across all sample types (stool, liver, and bladder) in mice on the purified diet (Fig. 2.6H–J). In stool and bladder,

median AsB levels for mice on the purified diet were close to the detection limit (as confirmed by one-sample Wilcoxon tests), and only trace amounts of AsB were detected in the liver of these mice (range: 0.17–0.48 ng per g tissue; LOD: 0.16 ng per g tissue). Additionally, mice on the purified diet excreted more iAs^{III} and iAs^V in stool (Fig. 2.6H) and accumulated less AsB and iAs^{III} in the liver (Fig. 2.6I) and bladder (Fig. 2.6J) compared to those on the standard diet.

Examining inter-individual variability, the variance in total stool arsenic was higher in mice on the purified diet (Table 2.2), while variances in total liver and bladder arsenic were not significantly different between diets. At the species level, variance in both DMA^V and iAs^V was greater in the stool of mice on the purified diet, though no other species-level differences reached statistical significance (Table 2.2). Given that AsB levels in the stool of mice on the purified diet were at or below the detection limit, variance in AsB was not assessed for this group. Overall, these findings indicate that the purified diet did not reduce inter-individual variability compared to the standard diet.

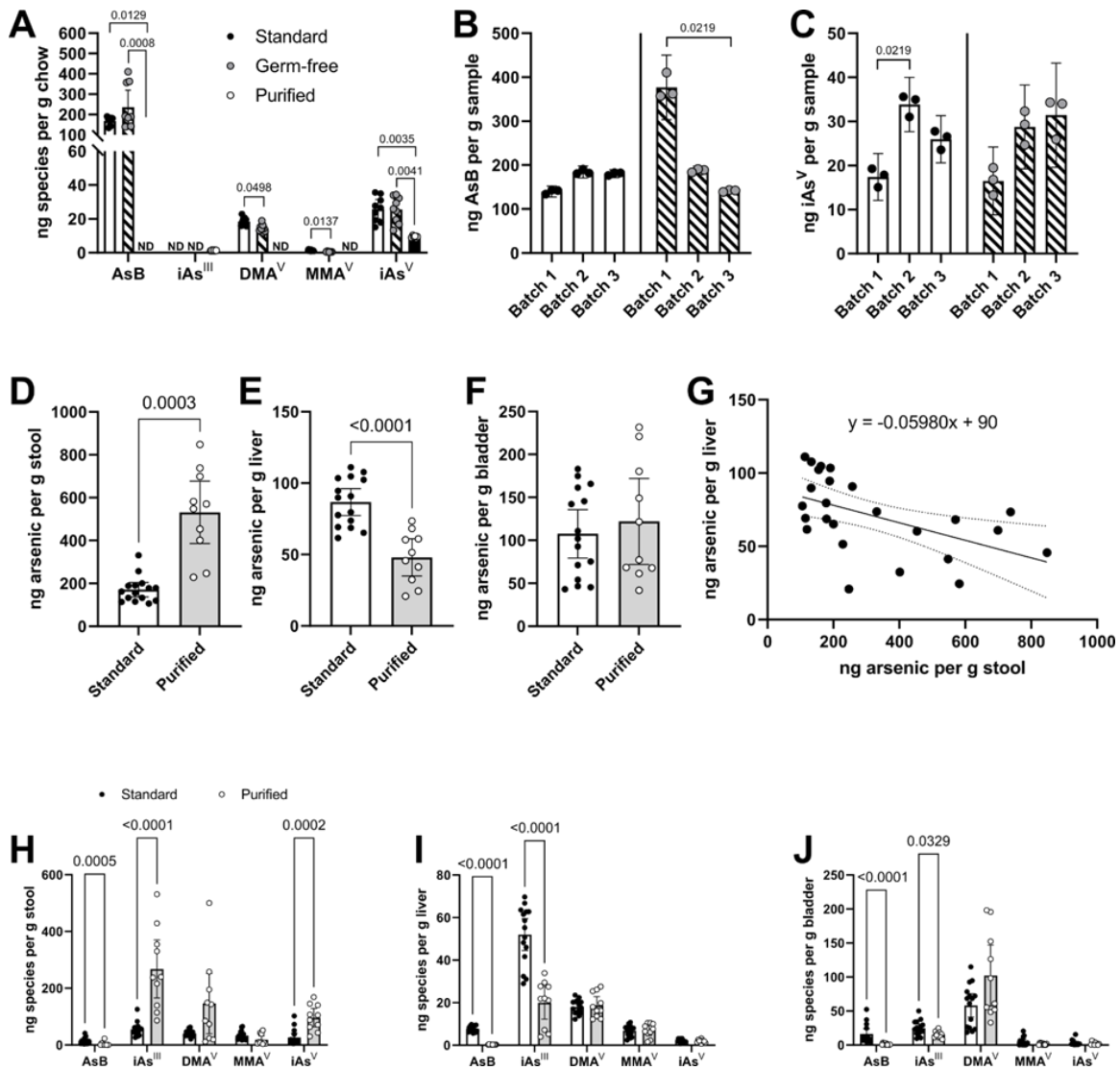


Figure 2.6 Dietary impact on arsenic body burden. Background levels of arsenic species in three commercial mouse diets (A; multiple Kruskal-Wallis tests with FDR correction for multiple comparisons). Background levels of AsB (B) and iAs^V (C) in different batches of the standard and germ-free mouse diets (Kruskal-Wallis test with Dunn's p value adjustment for multiple comparisons). Total arsenic levels in stool (D), liver (E), and bladder (F) of iAs^{III}-exposed mice eating either a standard or purified diet (t test with Welch's correction for unequal variance. Test performed on log-transformed stool data). Relationship between total stool and liver arsenic (G). Linear regression line (solid line) shown with 95% confidence (likelihood) interval (dotted lines) and equation. Arsenic species in stool (H), liver (I), and bladder (J) of iAs^{III}-exposed mice eating either a standard or purified diet (multiple Mann Whitney tests with FDR correction for multiple comparisons used for stool and bladder data; multiple t testing with Welch's correction for unequal variance and FDR correction for multiple comparisons used for square-root transformed liver data). No statistical tests were performed on undetectable arsenic species.

Table 2.2 Summary of statistical analysis

Total As (F test)					
Diet (standard vs. purified)	F	DFn	DFd	p-value	Greater variance
Stool	11.24	9	14	0.0001	purified diet
Liver	1.157	9	14	0.7788	
Bladder	1.875	9	14	0.2816	
Fasting (unfasted vs. fasted)					
Stool	1.156	19	14	0.7541	
Liver	2.087	19	14	0.1649	
Bladder	18.56	19	14	<0.0001	
Microbiome (conventional vs. germ-free)					
Stool	7.197	14	7	0.0139	conventional
Liver	2.006	14	7	0.2536	germ-free
Bladder	29.83	14	7	<0.0001	
Arsenic species (Levene's test)					
Diet (standard vs. purified)	Test ¹	Mean difference ²	Adjusted p-value	Greater variance	
Stool					
AsB			NA ³		
iAsIII	Kruskal-Wallis	-31.570	0.1641		
DMAV	Kruskal-Wallis	-55.370	0.0009	purified	
MMAV	Kruskal-Wallis	12.870	>0.9999		
iAsV	Kruskal-Wallis	-41.070	0.0275	purified	
Liver					
AsB			NA ³		
iAsIII	Kruskal-Wallis	16.870	>0.9999		
DMAV	Kruskal-Wallis	-10.030	>0.9999		
MMAV	Kruskal-Wallis	-6.467	>0.9999		
iAsV	Kruskal-Wallis	-0.033	>0.9999		
Bladder					
AsB			NA ³		
iAsIII	Kruskal-Wallis	14.630	>0.9999		
DMAV	Kruskal-Wallis	-23.770	0.5402		
MMAV	Kruskal-Wallis	20.270	0.8528		
iAsV	Kruskal-Wallis	-6.900	>0.9999		
Fasting (unfasted vs. fasted)					
Stool					
AsB	Kruskal-Wallis	-5.150	>0.9999		
iAsIII	Kruskal-Wallis	2.450	>0.9999		
DMAV	Kruskal-Wallis	2.183	>0.9999		
MMAV	Kruskal-Wallis	28.530	0.4958		
iAsV	Kruskal-Wallis	-1.533	>0.9999		
Liver					
AsB	Kruskal-Wallis	-9.267	>0.9999		
iAsIII	Kruskal-Wallis	9.842	>0.9999		
DMAV	Kruskal-Wallis	1.925	>0.9999		
MMAV	Kruskal-Wallis	8.558	>0.9999		
iAsV	Kruskal-Wallis	-10.010	>0.9999		
Bladder					
AsB	Kruskal-Wallis	-35.940	0.1890		
iAsIII	Kruskal-Wallis	5.175	>0.9999		
DMAV	Kruskal-Wallis	-37.350	0.1545		
MMAV	Kruskal-Wallis	10.300	>0.9999		
iAsV	Kruskal-Wallis	-17.380	>0.9999		
Microbiome (conventional vs. germ-free)					
Stool					
AsB	Kruskal-Wallis	48.080	0.0049	conventional	
iAsIII	Kruskal-Wallis	44.650	0.0111	conventional	
DMAV	Kruskal-Wallis	-2.200	>0.9999		
MMAV	Kruskal-Wallis	37.790	0.0481	conventional	
iAsV	Kruskal-Wallis	38.810	0.0392	conventional	
Liver					
AsB	one-way ANOVA ⁴	-8.035	0.0002	germ-free	
iAsIII	one-way ANOVA ⁴	-6.202	0.0053	germ-free	
DMAV	one-way ANOVA ⁴	-0.664	0.9983		
MMAV	one-way ANOVA ⁴	0.346	>0.9999		
iAsV	one-way ANOVA ⁴	-0.203	>0.9999		
Bladder					
AsB	Kruskal-Wallis	-35.940	0.0688		
iAsIII	Kruskal-Wallis	0.700	>0.9999		
DMAV	Kruskal-Wallis	-25.640	0.3942		
MMAV	Kruskal-Wallis	34.740	0.0863		
iAsV	Kruskal-Wallis	5.058	>0.9999		

¹Kruskal-Wallis = Nonparametric Kruskal-Wallis with Dunn's multiple comparisons test; one-way ANOVA = ordinary one-way ANOVA with Sidak's multiple comparisons test²Mean difference for Kruskal-Wallis is the mean rank difference.³NA

The Influence of Fasting on Arsenic Burden and Inter-Individual Variability

In the above experiments, mice had continuous access to food and water, and we hypothesized that the timing of arsenic exposure relative to eating or drinking might contribute to variability in arsenic levels. To test this hypothesis, we compared arsenic levels between mice exposed to arsenic under ad libitum (constant access) conditions and a group of mice undergoing food and water restriction (fasting). For the fasting group, food and water were removed three hours prior to exposure and continued to be withheld for an additional three hours post-exposure. Fasting did not impact total arsenic levels in stool (Fig. 2.7A) or liver (Fig. 2.7B) samples but did increase total arsenic in bladder samples (Fig. 2.7C). Similarly, fasting had no effect on individual arsenic species levels in stool (Fig. 2.7D) or liver (Fig. 2.7E) but increased both AsB and DMA^V in bladder samples (Fig. 2.7F). In terms of variability, bladder samples from fasted mice showed increased variance in total arsenic (Table 2.2), though the observed differences in variance for individual arsenic species across all sample types were not statistically significant (Table 2.2). Overall, fasting did not reduce inter-individual variability compared to mice with unrestricted access to food and water.

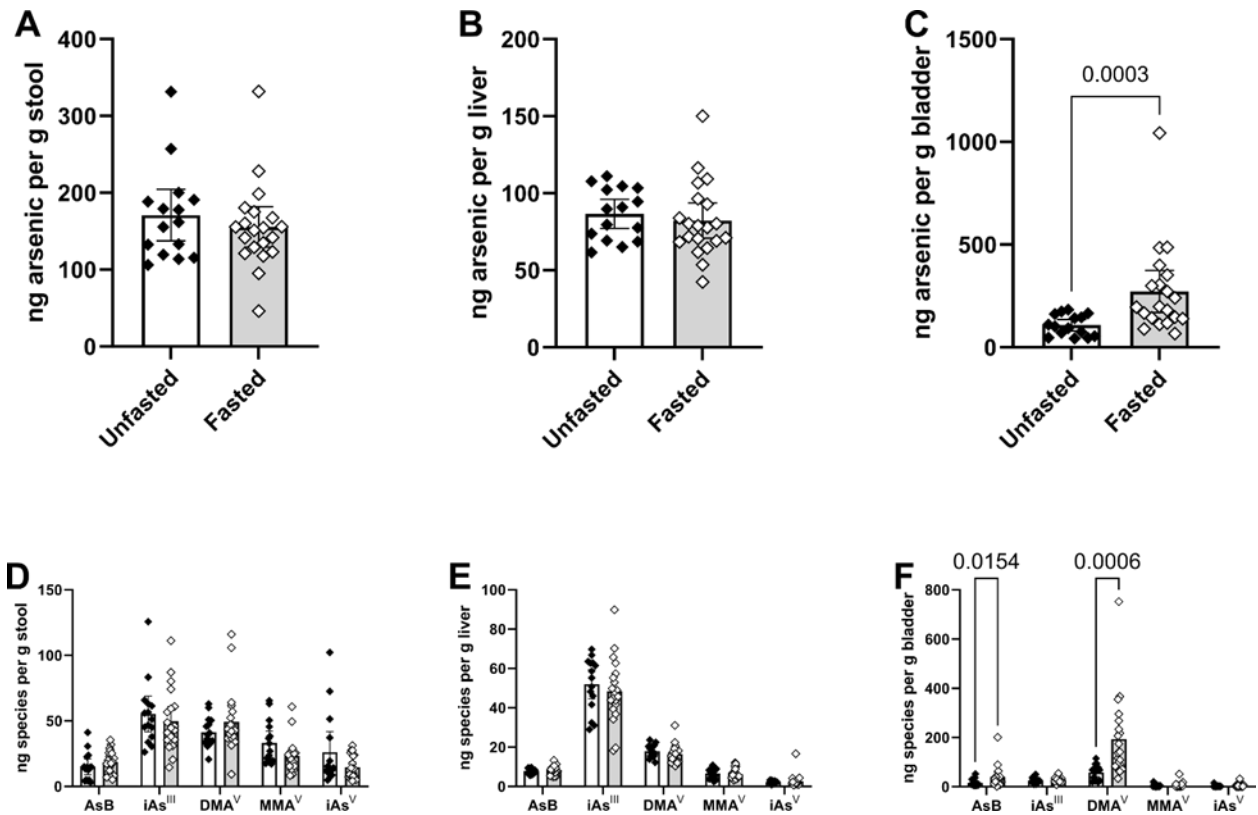


Figure 2.7 The impact of fasting on arsenic body burden. Total arsenic in stool (A), liver (B), and bladder (C) in unfasted and fasted mice (t test with Welch's correction for unequal variance on square root transformed data for stool and log transformed data for liver and bladder). Arsenic species in stool (D), liver (E), and bladder (F) in unfasted and fasted mice (multiple Mann Whitney tests with FDR correction for multiple comparisons used for stool, liver, and bladder). Symbols in panels E-F represent groups labeled in A-C (closed diamonds = fasted, open diamonds = unfasted).

The Influence of the Microbiome of Arsenic Burden and Inter-Individual Variability

Numerous studies, including our own, have shown that the microbiome plays a critical role in arsenic excretion, metabolism, and associated disease risk. In this study, we exposed groups of conventional and germ-free mice to arsenic and evaluated arsenic levels directly. Germ-free mice excreted less total arsenic in their stool (Fig. 2.8A), although differences in total arsenic levels in the liver (Fig. 2.8B) and bladder (Fig. 2.8C) were not statistically significant. Stool levels of AsB,

iAs^{III}, MMA^V, and iAs^V were lower in germ-free mice compared to conventional mice (Fig. 2.8D). In liver samples, AsB levels were higher in germ-free mice, while iAs^{III} and DMA^V levels were lower compared to conventional mice (Fig. 2.8E). There were no significant differences in arsenic species levels in bladder samples between the two groups (Fig. 2.8F).

The presence of a microbiome also affected the variability of total arsenic levels depending on the tissue type. Total arsenic levels in stool samples showed greater variance in conventional mice than in germ-free mice, no significant variance difference was found in liver samples, and total bladder arsenic exhibited less variance in conventional mice compared to germ-free mice (Table 2.2). For individual arsenic species, variance was greater for all species in stool samples from conventional mice, except for DMA^V (Table 2.2). Liver samples from germ-free mice showed greater variance in AsB and iAs^{III} compared to conventional mice, while no significant variance differences were observed in bladder samples (Table 2.2). In summary, our findings indicate that the microbiome substantially affects inter-individual differences in arsenic burden following exposure, with the direction and magnitude of this effect varying by tissue type and arsenic species.

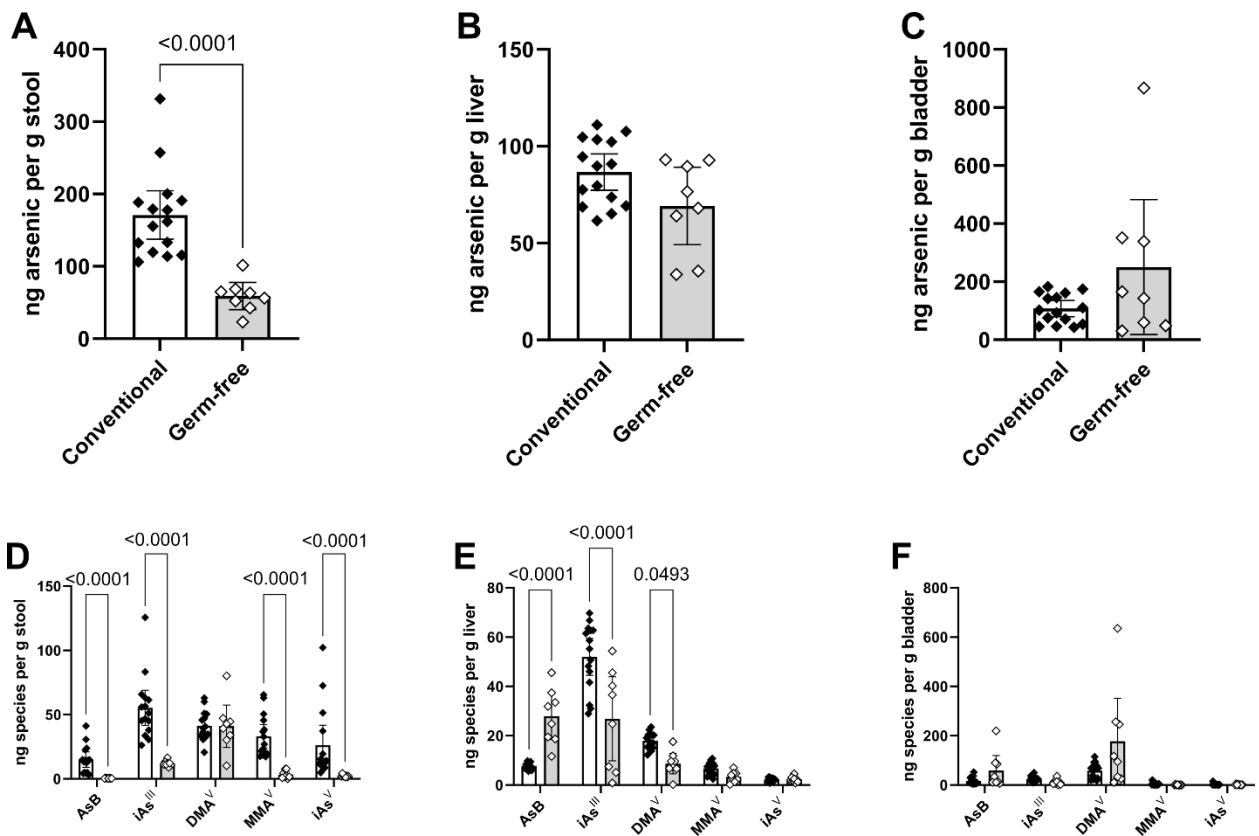


Figure 2.8 The impact of microbiome status on arsenic body burden. Total arsenic in stool (A), liver (B), and bladder (C) in conventional and germ-free mice (t test with Welch's correction for unequal variance on log transformed data for stool and bladder). Arsenic species in stool (D), liver (E), and bladder (F) in conventional and germ-free mice (multiple Mann Whitney tests with FDR correction for multiple comparisons used for stool and bladder data; two-way ANOVA with Sidak's test for multiple comparisons for liver data). Symbols in panels E-F represent groups labeled in A-C (closed diamonds = conventional mice, open diamonds = germ-free mice).

Generalizability of HPLC-ICPMS-Resolved Arsenic Species

HPLC-ICPMS offers a sensitive, accurate, and precise approach for arsenic speciation; however, technical variations in instrument setup—such as mobile phases, chromatography columns, and instrumentation—may impact the generalizability of results across different laboratories. To assess the potential effects of these technical differences, a single researcher

processed and prepared 15 samples, dividing each sample into two identical aliquots before freezing. One aliquot from each sample was then thawed and analyzed in one laboratory (Lab A), while the remaining aliquots were shipped on ice to a collaborating laboratory (Lab B) where they were also thawed and analyzed. The instruments were not standardized between the labs (see Discussion), which likely contributed to differences in absolute arsenic levels observed between the two laboratories (Fig. 2.9A). Nevertheless, the results from the two labs were highly correlated (Pearson's $r = 0.9847$, $R^2 = 0.9697$, $p < 0.0001$) and aligned well with a linear model ($R^2 = 0.9664$) with a slope of 0.6867 (Fig. 2.9B). These findings indicate that, despite variations in instrument setup, data from one laboratory can be reliably compared to data from another.

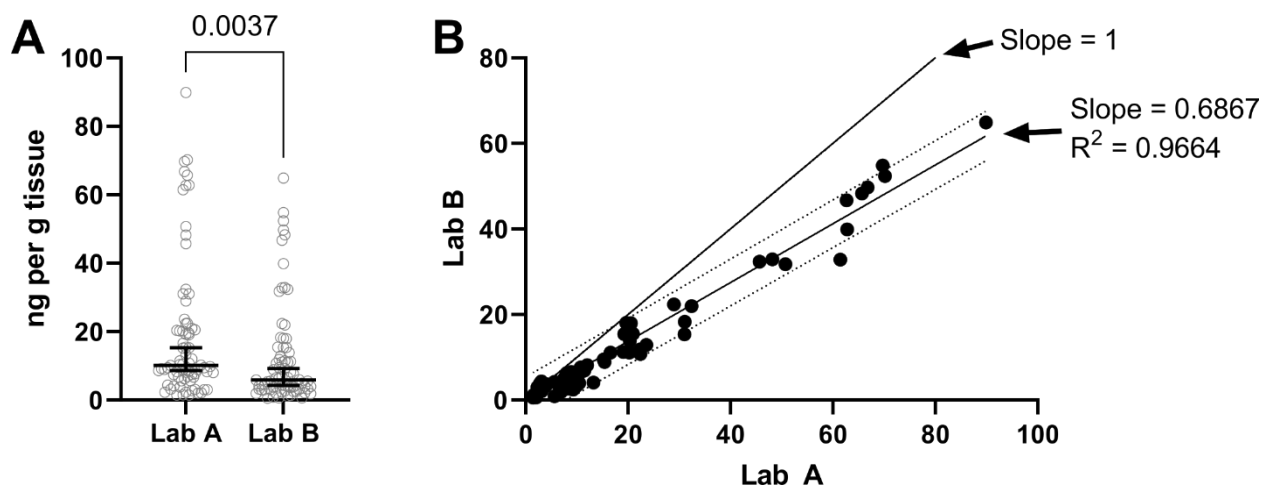


Figure 2.9 Inter-lab comparison of HPLC-ICPMS for arsenic speciation. Absolute abundance of arsenic species (AsB , iAs^{III} , DMA^{V} , MMA^{V} , and iAs^{V}) quantified in the same samples in labs A and B (A, t test with Welch's correction for unequal variance on log transformed data). Relationship between arsenic levels quantified in labs A and B (B). Origin constrained linear regression line (solid line) shown with 95% confidence (likelihood) interval (dotted lines) and equation. An origin constrained straight line with slope = 1 shown for comparison.

Discussion

Laboratory animal models of human arsenic exposure are valuable tools for identifying and assessing key factors that influence disease risk (Cohen *et al.*, 2013; Tsuji *et al.*, 2019). Many of these risk factors, including genetics (Xenakis *et al.*, 2022), can be controlled in laboratory animals, allowing researchers to untangle complex interactions among various host factors and covariates. However, both humans and lab animals display considerable inter-individual variability in their responses to arsenic exposure, a factor often overlooked in research studies. Some of this variability likely stems from technical aspects of sample processing and arsenic measurement, while other sources are biological and have yet to be fully understood. Clarifying the impacts of both technical and biological sources of variability is essential for accurately identifying modifiable risk factors and understanding their relative contributions to disease risk.

Technical Sources of Variability

This study establishes a comprehensive experimental framework for HPLC-ICPMS arsenic speciation in mouse tissues, assessing variability introduced by enzymatic digestion, tissue maceration techniques, and inter-laboratory differences in instrument setup. Previous studies have shown that arsenic extraction from protein-rich substrates, such as murine tissues, can be enhanced beyond simple aqueous extraction through the use of enzymatic digestion with papain, pepsin, and other proteases (Peng *et al.*, 2014; Liu *et al.*, 2015). Our findings align with these studies but indicate that, although papain provides better digestion efficiency, commercial papain preparations may inadvertently biotransform arsenic species, resulting in elevated background levels of iAs^{III}.

In terms of biotransformation, our results suggest that pepsin digestion offers a more consistent approach.

Arsenic extraction from animal tissues also requires tissue maceration, though there is limited consensus on the best methodology and the potential impact of different techniques. Mechanical methods, such as blade or pestle grinding, are commonly used for fresh or freeze-dried tissues from animal models or humans (Peng *et al.*, 2014; Liu *et al.*, 2015; Douillet *et al.*, 2023). However, the small size of mouse tissues restricts the use of certain devices, such as commercial food blenders (e.g., KitchenAid (Peng *et al.*, 2014; Liu *et al.*, 2015)). Mortar and pestle grinding of large sample batches can also be time-consuming and introduces potential experimental error due to human handling. Our results indicate that using a small mechanical homogenizer or bead beating yields comparable outcomes. Notably, mechanical homogenization introduced lower background arsenic levels than bead beating, suggesting that bead beating may not be ideal for studies involving low-level arsenic exposures (e.g., in the ppb range).

Finally, our study demonstrated that different laboratories with varying HPLC-ICPMS setups can produce highly comparable results in terms of relative, though not absolute, arsenic species abundances. This implies that cross-study comparisons should focus on reporting relative increases or decreases rather than absolute values of total or individual arsenic species.

Biological Sources of Variability

This study also demonstrates that diet composition and the gut microbiome are critical factors influencing the body burden of total and individual arsenic species. Consistent with previous findings showing varied background arsenic levels in mouse diets (Murko *et al.*, 2018), our results provide strong evidence that the purified mouse diet AIN-93G has a significantly lower arsenic

background than unpurified diets, particularly with regard to arsenobetaine (AsB). The presence of substantial AsB in unpurified diets may be relevant, as *in vitro* studies suggest that AsB can be transformed into other arsenic species (e.g., inorganic arsenic, iAs) (Men *et al.*, 2022; Liu *et al.*, 2023). Given this, it is advisable that murine studies investigating the *in vivo* transformation of AsB consider using a purified diet. Mice on different diets also showed significantly distinct levels of total and individual arsenic species in their tissues, an important finding given that current murine models of human arsenic exposure do not account for dietary variations, nor have they explored the impacts of specific dietary components on arsenic exposure outcomes. It remains unclear whether the differences in tissue arsenic accumulation are due to direct arsenic-diet interactions or more complex interactions, such as those involving diet, microbiome, and host metabolism. Future research will be essential to disentangle these multifaceted interactions.

Previous studies by our group (Coryell *et al.*, 2018; Roggenbeck *et al.*, 2021; Wang *et al.*, 2021) and others (Lu *et al.*, 2013; Lu *et al.*, 2014a, 2014b; Chi *et al.*, 2019) have identified the gut microbiome as a significant factor in arsenic exposure. Our findings further support this body of evidence and, to our knowledge, provide the first direct comparison between conventional and germ-free mice. Our results reaffirm our prior conclusion that the microbiome plays a key role in arsenic excretion through stool (Coryell *et al.*, 2018). Additionally, the observed negative correlation between arsenic levels in stool and liver (Coryell *et al.*, 2018) reinforces the concept that higher arsenic excretion in stool correlates with reduced arsenic exposure in the liver. As expected, the presence of a microbiome had the most substantial effect on arsenic levels in stool, followed by the liver; by the time arsenic reached the bladder, the microbiome's influence was no longer statistically significant. This pattern suggests that the microbiome plays a pivotal role early

in the exposure process, particularly in regulating absorption and liver exposure, consistent with a “facilitation by excretion” model.

Inter-Individual Variability

Many published studies on murine arsenic exposure rely on small sample sizes or data from a single group or cage of mice. In our experience, such study designs have limited statistical power, and findings may vary if repeated with larger samples. Previously, we demonstrated that antibiotic treatment in conventional mice resulted in inter-individual variability in arsenic-induced mortality (Roggenbeck *et al.*, 2021), but this outcome required testing large numbers ($n \geq 20$) of mice across different cages. The findings presented here further support the notion that, like humans, mice naturally exhibit variation in arsenic body burden. Based on the observed variance among conventional mice on a standard diet, we estimate a coefficient of variation ranging from 18% to 114% for arsenic species levels across murine tissues (stool, liver, and bladder). With standard statistical parameters ($\alpha = 0.05$ and power = 0.8), the required sample size will depend on the effect size of interest (e.g., differences between group means). Our data suggest that in most cases, sample sizes of more than five mice per group are needed; more critically, mean values varied between cages within the same treatment group. This indicates that achieving sufficient power in murine arsenic exposure studies will require at least two cages per group, with a minimum of 4–5 mice per cage. These guidelines are general, and studies should consider statistical power requirements relative to the effect sizes they aim to detect.

Among the three factors evaluated in this study (diet, fasting, and the microbiome), only the microbiome significantly affected inter-individual variability in tissue arsenic levels. Given the

complexities in controlling the microbiome compared to diet and other aspects of animal care, it is likely that results will vary between mice housed in different facilities due to differences in gut microbial composition. This underscores the importance of investigating specific human microbiome taxa that may influence or reduce arsenic variability between individuals. Identifying such taxa could help link specific microbiome compositions to arsenic exposure risk, providing insights that could be further validated through epidemiological studies in exposed human populations.

CHAPTER THREE

TIME-COURSE ARSENIC ACCUMULATION IN MURINE
TISSUES DURING EARLY-STAGE OF ARSENIC EXPOSUREIntroduction

So far, over 200 million people worldwide are chronically exposed to arsenic in drinking water above the current maximum contaminant level of 10 $\mu\text{g/L}$ (WHO, 2022), while about 2.1 million population in the U.S. exposed to arsenic in drinking water from private domestic wells, according to the U.S. Geological Survey and Centers for Disease Control and Prevention (2019). However, the data of the arsenicals time-course diffusion in mammals' system is understudied.

Epidemiological studies reveal that the methylation of iAs^{III} is the main detoxification mechanism in human, however it is often incomplete, leading to variability in As metabolite compositions in urine among individuals, even when exposed to the same levels of arsenicals in drinking water (Saint-Jacques *et al.*, 2014). The distribution of As species in humans depends on numerous factors, including the species of arsenic exposure, availability of glutathione (GSH) and other thiol ($-\text{SH}$) containing biomolecules, primary arsenic methylating enzymes (Harino *et al.*, 2004). The liver is generally considered the primary target organ for arsenic detoxification. In the liver, iAs is methylated by the enzyme arsenic methyltransferase (AS3MT) with the presence of the methyl donor S-adenosylmethionine (SAM) and the co-factor GSH, producing MMAs^{III} , MMAs^{V} , DMAs^{III} and DMAs^{V} , which will be mainly extruded through urine. However, in vitro studies suggest that other tissues may also have arsenic metabolic capacity, with the highest levels of arsenic methylation activity observed in the cytosol of testes, followed by the kidney, liver, and

lungs (Aposhian *et al.*, 1997, Healy *et al.*, 1998). The distribution and retention of arsenic in tissues, influenced by the bioconversion of methylated metabolites, which plays a critical role in determining its toxicity or carcinogenicity in humans (Cullen, 2014). Thus, it is important to study the diffusion of arsenicals in humans when exposed to iAs, especially in the early exposure stage for detoxification purpose.

Additionally, even though the iAs^V is the main As form in well water and drinking water, the transition from iAs^V to iAs^{III} happens quickly in the human body. In mammals, the reduction of iAs^V to iAs^{III} is related to liver enzyme iAs^V reductase, a type of protein similar to purine nucleoside phosphorylase, in the existence of dihydrolipoic acid with arsenolysis (Radabaugh *et al.*, 2002). Moreover, GSH and GSSG are also involved in the oxidation and reduction of iAs as non-enzymatic reactions (Khairul *et al.*, 2017). It seems like in the human body, the As metabolism is as complicated as in the bacteria. Nevertheless, the gut microbiome is a critical part involved in the As metabolism in humans.

In this part of the project, we used murine model to try to understand the diffusion of different species of arsenic in different tissues at the early-stage of As exposure. Based on the previous publications, liver is the most important tissue to detoxify As in mammals, and urine is the most important extrude method in mammals. Furthermore, the gut microbiome is now considered as the biggest tissue in humans. Thus, we focus on the arsenical species in liver, urinary bladder and fecal pellets, during the first 12 hours of As exposure. The data will lead us to the preview of the early-stage As metabolism in different mammalian tissues, and the diffusion of different arsenical species in different tissues at the early-stage of arsenical exposure. The results of this study will provide valuable data for medical detoxification during arsenical exposure.

Methods

Chemical Reagents

Stock solutions with a concentration of 10 mg/mL were prepared for various arsenic compounds: arsenobetaine (AsB, 95% purity, Sigma-Aldrich, St. Louis, MO), arsenic (III) oxide (iAs^{III} , 99.995% trace metals basis, Sigma-Aldrich, St. Louis, MO), sodium arsenate dibasic heptahydrate (iAs^V , $\geq 98\%$ purity, Sigma-Aldrich, St. Louis, MO), methylarsonic acid (monomethylarsonic acid, $MMAs^V$, 95%, Wako Chemicals, Richmond, VA), and cacodylic acid (dimethylarsinic acid, $DMAs^V$, 99%, Sigma-Aldrich, St. Louis, MO). Working solutions of arsenic species at concentrations of 0.5, 1, 5, 10, 50, and 100 $\mu\text{g/L}$ were freshly made on a daily basis through serial dilution of the stock solutions. Pepsin was purchased from Sigma-Aldrich (St. Louis, MO). The mobile phase was formulated using HPLC-grade ammonium phosphate dibasic (Sigma-Aldrich, St. Louis, MO), with pH adjustments made using trace metal grade ammonium hydroxide (VWR, West Chester, PA).

Experimental Animals

All protocols for mouse experiments were approved by the Institutional Animal Care and Use Committee (IACUC) at Montana State University. The mice were housed in the Montana State University Animal Resource Center, which is accredited by the American Association for the Accreditation of Laboratory Animal Care (AAALAC). C57BL/6J mice were obtained from The Jackson Laboratory and kept in individually ventilated cages under specific pathogen-free conditions, with sterilized bedding and protection from murine norovirus.

Murine Arsenic Exposures and Tissue Collection

The arsenic exposures utilized iAs^{III} and involved age-matched mice (7-13 weeks old) of both sexes. Euthanasia was carried out humanely via isoflurane overdose. As shown in Fig. 3.1, conventional C57BL/6J mice received $1 \mu g iAs^{III}$ per gram of body weight through oral gavage, and were euthanized after 3-, 6-, 9-, and 12-hours post-exposure ($n = 3$ for each time point). Control mice without arsenic exposure ($n = 4$) were euthanized to provide baseline data. Stool samples were collected prior to euthanasia, and tissues (liver and/or bladder) were collected afterward.

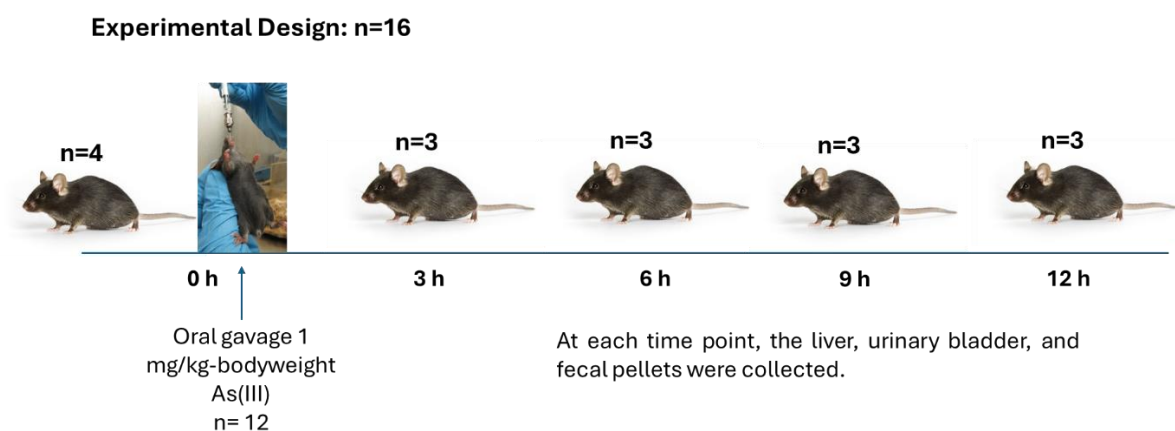


Figure 3.1 The process of murine arsenic exposure and tissue collection.

Processing of Mouse Tissues

Unless otherwise noted, 0.1 g samples of liver and stool, or 0.03 g samples of urinary bladder tissue, were digested in 400 μL of pepsin solution at 37°C overnight. The tissues were homogenized using an Omni THq Homogenizer (Omni International, Kennesaw, GA, USA) at 20,000 rpm for one minute. Samples then underwent centrifugation at 12,000 rpm for five minutes. Finally, the supernatants were collected and filtered through 3KDa membrane filters, followed by a second centrifugation at 12,000 rpm for 30 minutes at 4°C.

HPLC-ICPMS

HPLC analyses were performed at Montana State University with an Agilent 1260 HPLC system, which included a quaternary pump, an online degasser, and a system control unit. Separation of arsenic species was facilitated using a Hamilton PRP-X100 anion-exchange column (250 × 4.1 mm, 10 µm particle size). A 50 µL sample loop was used to introduce both sample and standard solutions. Arsenic compounds were separated through isocratic elution using 10 mM (NH₄)₂HPO₄ at pH 8.25 (adjusted with trace metal grade NH₄OH) as the mobile phase, with a consistent flow rate of 1.5 mL/min for a total separation time of 10 minutes. ICPMS measurements were conducted using an Agilent 7800 ICPMS system, where the identification of arsenic compounds relied on retention time matching with single standards on the PRP-X100 column, and arsenic ions were detected by the ICPMS. The ICPMS conditions were optimized daily using a tuning solution to ensure precise measurements.

Results and Discussion

In most As exposure experiments, mice are typically exposed to arsenic through drinking water, making it difficult to estimate the individual exposure levels precisely (Lu *et al.*, 2014; ref). Previous studies have demonstrated significant variability in arsenic metabolite profiles in human urine, even among individuals exposed to similar levels of arsenic through drinking water (ref). To address this, our study employed oral gavage of inorganic arsenic (iAs^{III}) in mice, with dosages calibrated according to body weight. This method standardizes arsenic exposure among subjects, thereby eliminating individual variability in exposure levels and providing a controlled murine model for studying arsenic metabolism and toxicity.

The liver is the primary organ responsible for detoxification following exposure to inorganic arsenic (iAs^{III}). Therefore, we initially tracked arsenic species in mouse livers over time (Figs. 3.2–3.7). Consistent with our earlier findings on arsenic background levels in mouse chow (Fig. 2.6), we observed the presence of arsenobetaine (AsB) in the liver, with no significant fluctuations during the period of iAs^{III} exposure. This result suggests that dietary AsB consistently contributes to the arsenic profile in conventional mice (Fig. 3.2).

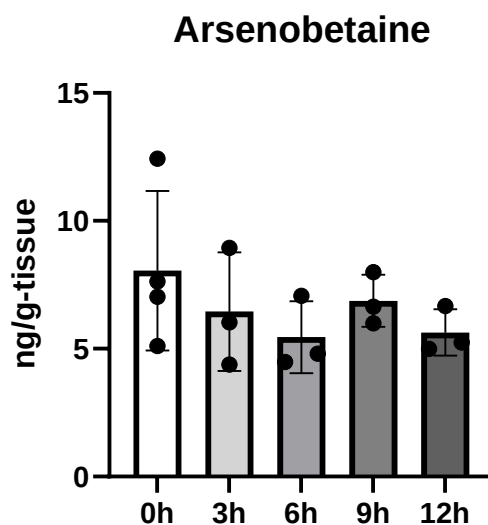


Figure 3.2 The time-course AsB concentration in mouse livers, when mice were exposed to iAs^{III} .

The time-course analysis of iAs^{III} levels in mouse livers revealed intriguing patterns (Fig. 3.3). In iAs^{III} -exposed mice ($n=3$), liver iAs^{III} concentrations were initially highest compared to control mice ($n=4$, without iAs^{III} oral gavage). Over time (from 3 to 12 hours), iAs^{III} levels in the liver decreased, suggesting active metabolism of iAs^{III} .

Similarly, MMA, DMA, and iAs^V followed a similar time-dependent trend in the liver (Figs. 3.3–3.6), indicating the progression of iAs^{III} detoxification. Consistent with previous studies, iAs^{III} was first methylated to MMA, then to DMA. Concurrently, iAs^{III} underwent oxidation to iAs^V , reflecting a detoxification process given the lower toxicity of iAs^V relative to iAs^{III} .

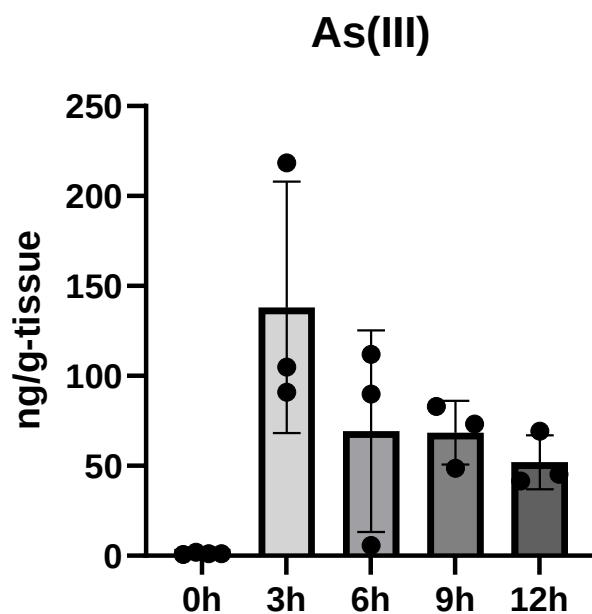


Figure 3.3 The time-course iAs^{III} concentration in mouse livers, when mice were exposed to iAs^{III} .

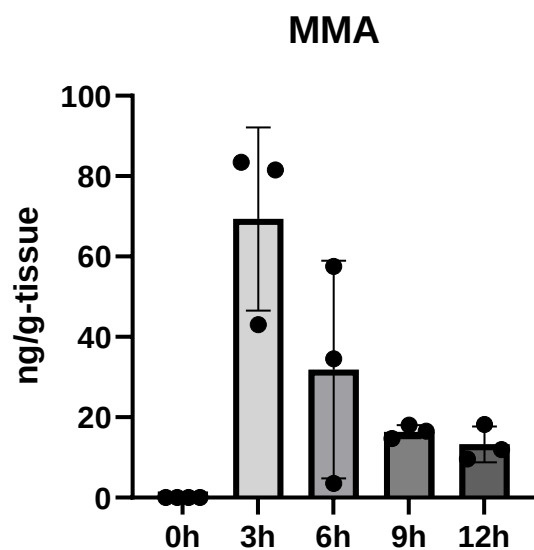


Figure 3.4 The time-course MMA concentration in mouse livers, when mice were exposed to iAs^{III} .

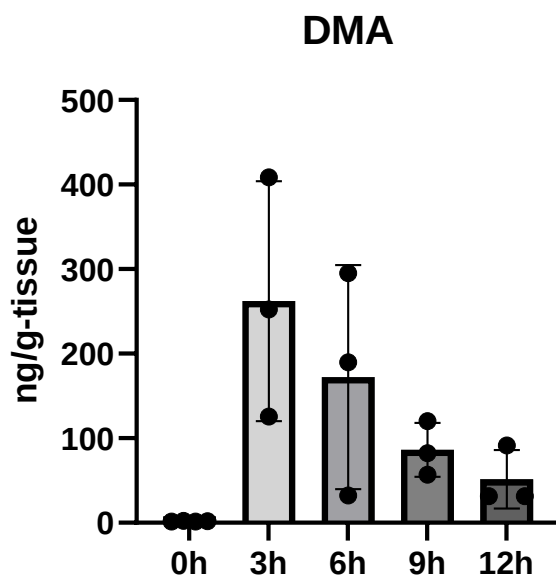


Figure 3.5 The time-course DMA concentration in mouse livers, when mice were exposed to iAs^{III} .

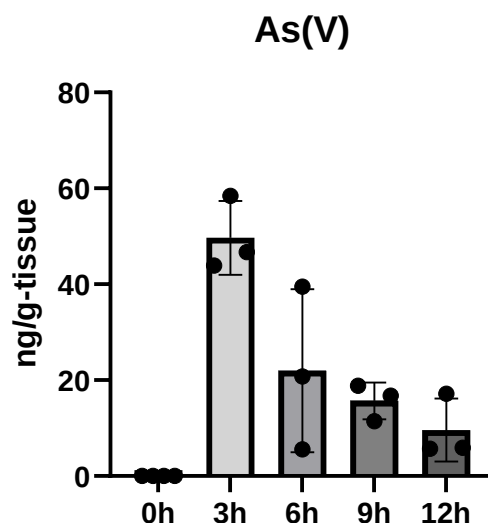


Figure 3.6 The time-course iAs^V concentration in mouse livers, when mice were exposed to iAs^{III}.

The similar decreasing trend observed for iAs^{III}, MMA, DMA, and iAs^V in liver suggests that these arsenicals transfer to other tissues via blood circulation and are subsequently excreted from the body. We monitored arsenic species in the urinary bladder (Figs. 3.7–3.11) and fecal pellets (Figs. 3.12–3.16) to trace this excretion process.

Due to the high background level of arsenobetaine (AsB) from dietary sources (Fig. 2.6), it remains unclear whether any AsB detected is a product of iAs^{III} metabolism. While AsB levels in the urinary bladder showed no significant changes (Fig. 3.7), its concentration in fecal pellets increased significantly after 6 hours (Fig. 3.12), with levels at 9–12 hours exceeding those at 3–6 hours. These results suggest that: (1) AsB may be involved in iAs^{III} metabolism post-exposure, and (2) fecal excretion could play a larger role in AsB elimination, especially under higher iAs^{III} exposure conditions. Given that AsB metabolism and excretion have not been thoroughly

investigated in prior studies, these findings may open a new research avenue in mammalian arsenic exposure studies.

In the urinary bladder, levels of iAs^{III} , MMA, DMA, and iAs^V increased significantly from 0 to 3 hours post-exposure (Figs. 3.8–3.11), followed by a decline from 6 to 12 hours. iAs^{III} levels decreased most markedly, followed sequentially by MMA, DMA, and iAs^V , likely reflecting the detoxification processes in the liver. These findings align with previous conclusions that urine is a primary excretion pathway following iAs^{III} exposure, as arsenic species are processed in the liver before being transported to the urinary bladder for elimination.

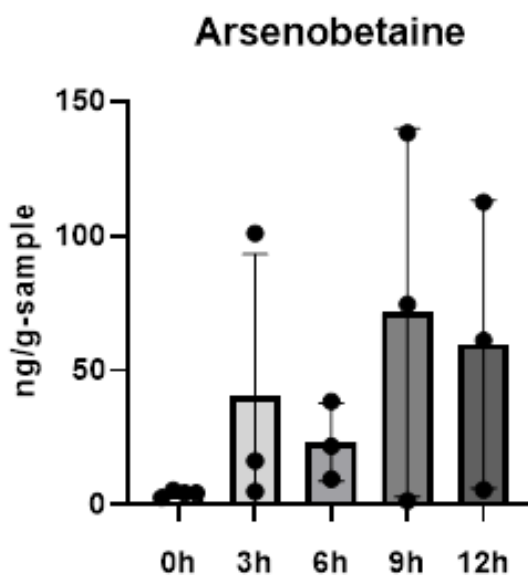


Figure 3.7 The time-course AsB concentration in urinary bladder, when mice were exposed to iAs^{III} .

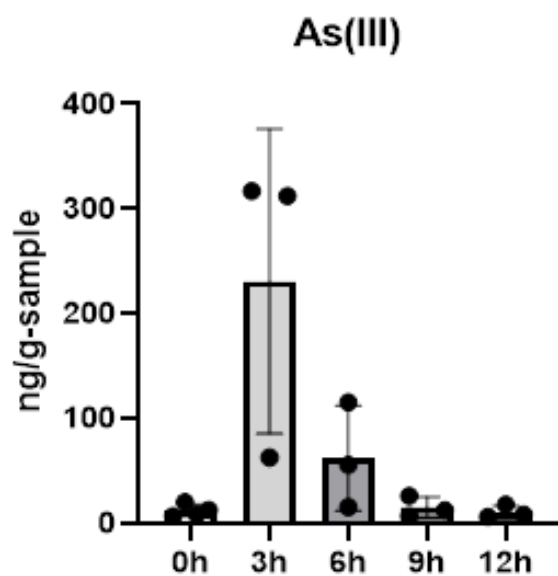


Figure 3.8 The time-course iAs^{III} concentration in urinary bladder, when mice were exposed to iAs^{III} .

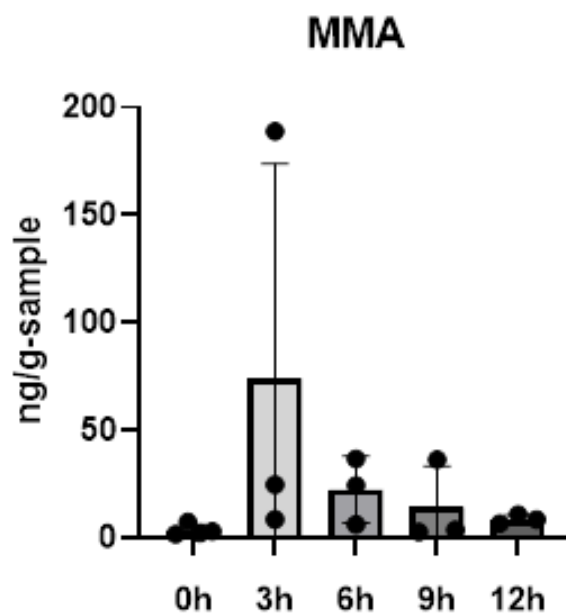


Figure 3.9 The time-course MMA concentration in urinary bladder, when mice were exposed to iAs^{III} .

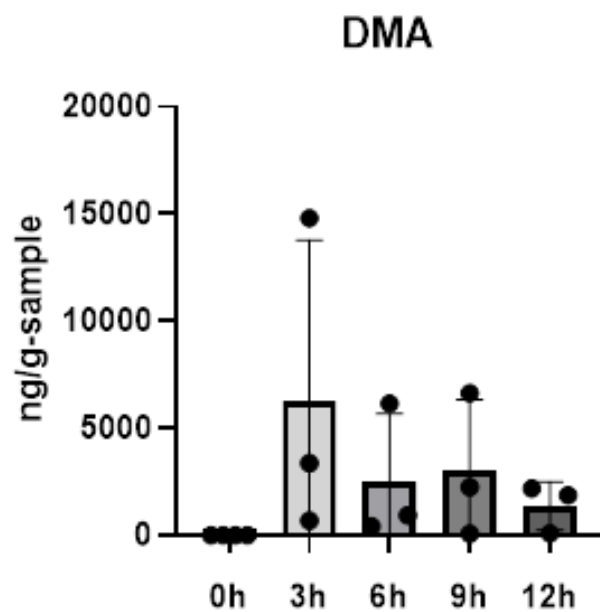


Figure 3.10 The time-course DMA concentration in urinary bladder, when mice were exposed to iAs^{III} .

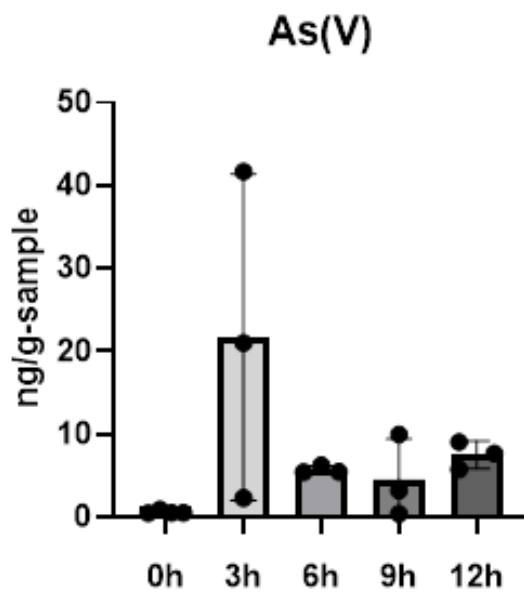


Figure 3.11 The time-course iAs^V concentration in urinary bladder, when mice were exposed to iAs^{III} .

Interestingly, fecal excretion also plays a significant role in arsenic elimination following iAs^{III} exposure. Although our group previously demonstrated fecal arsenic excretion (Coryell *et al.*, 2018), questions remain regarding the precise amounts excreted through feces. In this study, time-tracked analysis of arsenic species revealed that iAs^{III} , MMA, DMA, and iAs^V are all excreted via fecal matter post-exposure, albeit with a later onset and longer duration than urinary bladder (Figs. 3.12–3.16). While urinary bladder arsenic excretion peaked within 3 hours, fecal arsenic concentrations continued to rise from 0 to 9 hours before declining at 12 hours (Figs. 3.12–3.16). Additionally, arsenic concentrations in fecal samples were notably higher than those in the urinary bladder, suggesting that while urine bladder serves as the initial pathway for arsenic elimination, fecal excretion may remove residual arsenic left in the system after initial processing.

These findings highlight the potential role of the gut microbiome in arsenic detoxification, suggesting it may contribute to the capacity for arsenic excretion in mouse and potentially mitigate toxicity when the host is exposed to iAs^{III} .

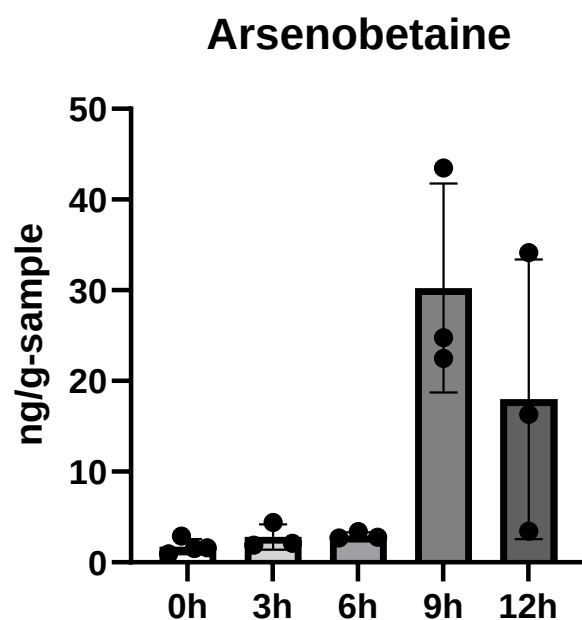


Figure 3.12 The time-course AsB concentration in mouse stool, when mice were exposed to iAs^{III} .

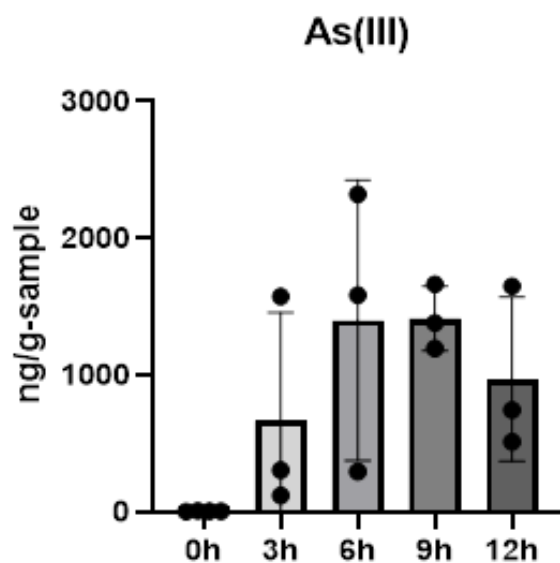


Figure 3.13 The time-course iAs^{III} concentration in mouse stool, when mice were exposed to iAs^{III} .

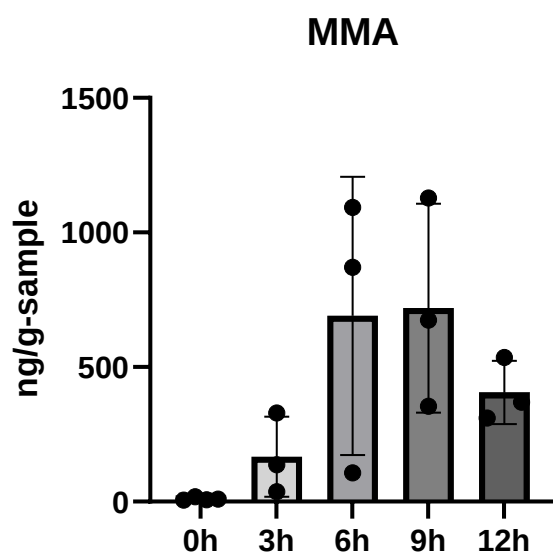


Figure 3.14 The time-course MMA concentration in mouse stool, when mice were exposed to iAs^{III} .

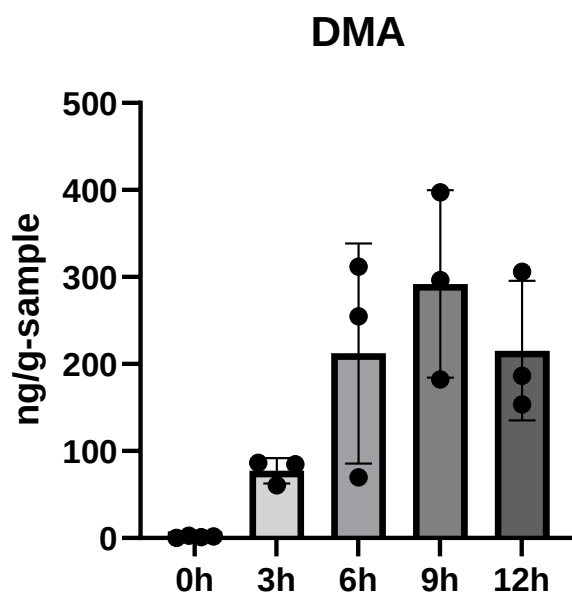


Figure 3.15 The time-course DMA concentration in mouse stool, when mice were exposed to iAs^{III} .

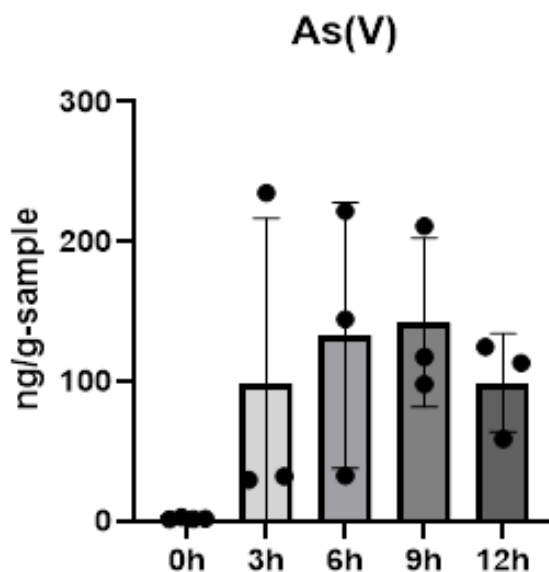


Figure 3.16 The time-course iAs^V concentration in mouse stool, when mice were exposed to iAs^{III} .

The hepatoenteric circulation may be the reason for the excretion mechanism we observed in mice. It is hypothesized to be a critical determinant in the metabolism, detoxification, and excretion of arsenic in mammals, mediating the distinctive species-specific differences observed in arsenic metabolism (Suzuki *et al.*, 2001). This pathway is proposed to facilitate arsenic excretion via stool, urinary bladder, and liver by orchestrating a complex series of biochemical processes that govern the transport, transformation, and elimination of arsenic species.

Upon intravenous injection, iAs^{III} is rapidly taken up by the liver, where it undergoes reduction, methylation, and conjugation with glutathione (GSH). A significant portion of iAs^{III} is conjugated to form GSH-conjugated complexes, such as $iAs^{III}(GS)_3$, which are subsequently excreted into bile and transferred to the gastrointestinal tract. This biliary excretion not only lowers

the concentration of reactive iAs^{III} in the liver but also initiates the recycling of arsenic within the hepatointestinal circulation, where iAs^{III} is deconjugated and reabsorbed for further transformation.

During hepatointestinal cycling, the liver metabolizes arsenic through methylation processes, producing methylated arsenic species such as MMA and DMA. Notably, MMA exists in bile in both trivalent conjugated ($MMA^{III}(GS)_2$) and pentavalent free forms- MMA^V , whereas DMA is absent from bile, suggesting a preferential pathway for its excretion into urine. The differential excretion patterns—biliary excretion for inorganic and monomethylated arsenic versus urinary excretion of DMA—highlight the unique functional roles of the hepatointestinal circulation and urinary elimination pathways.

As arsenic accumulates preferentially in red blood cells (RBCs), primarily in the form of DMA, the hepatointestinal circulation enables sequential reduction and methylation of arsenic during its passage through the liver. This process results in the gradual conversion of arsenic to DMA, the dominant urinary metabolite after the acute phase of exposure. The presence of iAs^{III} and iAs^V in nonmethylated forms, along with methylated species (MMA^V and DMA^V), in early-phase urine excretion underscores the temporally dynamic nature of arsenic metabolism.

This hypothesis posits that the hepatointestinal circulation is central to the observed species-specific arsenic metabolism in mammals (Suzuki *et al.*, 2001). It allows for efficient transformation and detoxification of arsenic while partitioning its excretion into distinct pathways - feces via bile for conjugated forms and urine for methylated forms. This mechanism well explained the long term (6-12h) fecal excretion of arsenic in this study.

Moreover, the bladder acts as a key excretory organ where arsenic metabolites concentrate as urine is stored. Reactive trivalent arsenicals, such as DMA^{III} , are particularly harmful due to

their ability to bind sulfhydryl groups in cellular proteins, leading to oxidative damage and disruption of cellular homeostasis. DMA^{III} is excreted into urine at concentrations known to be cytotoxic based on *in vitro* and *in vivo* studies. Prolonged contact with the urothelium during urine storage exacerbates cellular damage and contributes to sustained tissue exposure, a major risk factor for bladder cancer development (Arnold *et al.*, 2014). Since DMA^{III} induces superficial cell necrosis in the urothelium, primarily affecting the luminal superficial layer, which is the first point of contact with toxic urine. This cytotoxicity initiates compensatory regenerative proliferation, leading to hyperplasia - a preneoplastic condition characterized by increased urothelial cell layers.

Additionally, DMA^{III} binds with high affinity to cysteine residues in the hemoglobin of red blood cells (RBCs), forming stable complexes. This interaction prolongs the systemic circulation of DMA^{III} and facilitates its delivery to the bladder via renal filtration. The slow release of DMA^{III} from RBCs ensures a continuous supply of the reactive metabolite to the urinary system, exacerbating bladder exposure. This RBC-mediated transport mechanism further increases the bioavailability of DMA^{III} in urine, compounding its toxic effects on the bladder. Unlike the liver, which possesses robust enzymatic pathways to detoxify arsenic through methylation and conjugation, the bladder lacks such defenses. Consequently, the urothelium remains highly vulnerable to the cytotoxic effects of unmetabolized arsenicals, particularly DMA^{III}. The bladder functions as a urine reservoir exposes it to prolonged contact with toxic arsenicals. The chronic cycles of urothelial damage and subsequent regenerative proliferation establish a pro-carcinogenic environment. However, since there are more sulfhydryl groups in the fecal matter compared to the surface of the epithelium cells in the colon, the arsenicals probably mostly bind with the fecal

matters, which give the arsenicals less chance to toxify the epithelium layer in the colon, and lower the risk of colon cancer when exposed to arsenic.

Conclusion

In conclusion, this study provides important insights into the dynamics of arsenic metabolism and excretion in a controlled murine model, shedding light on the tissue-specific diffusion and detoxification processes following iAs^{III} exposure. Our findings reveal that the liver plays a primary role in detoxifying arsenic, with initial high concentrations of iAs^{III} , MMA, DMA, and iAs^V decreasing over time as these metabolites are processed and transported to excretory pathways. Both urinary and fecal routes contribute to arsenic elimination, though urine serves as the primary excretory pathway with rapid clearance, while fecal excretion persists over a longer timeframe. Notably, the increased fecal excretion of arsenobetaine and other arsenicals indicates that the gut microbiome may play an active role in arsenic detoxification. These results underscore the complex interplay of metabolic, microbiological, and excretory systems in managing arsenic exposure in mammals. Obviously, there are some more work need to be done for this study, however, this work not only enhances our understanding of arsenic biotransformation and elimination mechanisms but also highlights potential areas for further research, particularly regarding the role of the gut microbiome in modulating arsenic toxicity and supporting detoxification. The insights gained may inform strategies for managing arsenic exposure and mitigating associated health risks in exposed populations.

CHAPTER FOUR

PREVIOUS WORK DURING PHD PROGRAM - NMR

METABOLOMICS ANALYSES

Introduction

Among all of the As biotransformation [iAs^{III} oxidation, iAs^{V} reduction, As methylation and demethylation], the ArsR repressor regulates expression of genes, known as the *ars* operon, encoding functions associated with As detoxification (Andres and Bertin, 2016). While the bacterial iAs^{III} oxidase encoding genes *aioBA* is regulated by a three-component signal transduction system AioXSR, which consists of a periplasmic iAs^{III} binding protein AioX, a histidine sensor kinase AioS, and its cognate response regulator AioR (Kashyap *et al.*, 2006; Cai *et al.*, 2009; Liu *et al.*, 2012). Interestingly, this three-component signal transduction system is similar to the phosphate uptake two-component signal transduction system PhoBR; and a previous study showed that these two signal transduction systems crosstalk with each other (Wang *et al.*, 2018).

Phosphorus is an essential nutrient for microbial biological processes. However, phosphorus is typically scarce and normally present at nanomolar levels in natural water and soil environments (Vieira *et al.*, 2008). Thus, iAs^{V} substitution for phosphate is biologically significant in As-enriched phosphate starvation environments. Interestingly, genomic surveys have revealed that two sets of phosphate transport systems (Pst/Pho systems) often occur in iAs^{III} -oxidizing bacteria (Li *et al.*, 2013). The Pst/Pho system includes periplasmic phosphate binding protein (PstS), transmembrane proteins (PstA and PstC), ATP binding protein (PstB), inhibitor regulator (PhoU)

and PhoBR (Qiao *et al.*, 2019). Of those, the *pst1/pho1* system is adjacent to *aioBA* genes, and the *pst2/pho2* system is located elsewhere in the genome. Unsurprisingly, As resistant bacteria display a complete, logical and systematic regulation link between As regulation systems and phosphate two-component signal transduction system PhoBR as follows: 1) In some iAs^{III} oxidation bacterial genomes, there exists no AioXSR; in this case, PhoBR takes the place and regulates the expression of iAs^{III} oxidase AioBA under low-phosphate condition (Chen *et al.*, 2015a). 2) When iAs^{III} is present and environmental phosphate is depleted, the As resistant regulator ArsR1 can bind to iAs^{III} , and vacate the *phoB1* and *pstS1* promoters, which activate the expression of the phosphate uptake Pho/Pst system (Rawle *et al.*, 2019). 3) The ArsR1-vacated *PhoB1* is then phosphorylated by PhoR, which results in the induction of *aioXSR*, if the iAs^{III} oxidation three-component signal transduction system AioXSR is present (Shi *et al.*, 2018; Rawle *et al.*, 2019). 4) AioX can then sense iAs^{III} in the environment, and self-phosphorylates AioS, which subsequently phosphorylates AioR, which results in the activation of the *aioBA* gene expression. 5) Meanwhile, phosphorylated PhoR and AioS both phosphorylate *phoB1/phoB2* and *aioR*, which then activate the expression of both Pho/Pst- and Aio-encoding genes. All of the evidence indicates that these arsenic and phosphate regulatory systems are deeply integrated and make up the phosphate starvation response altogether (Wang *et al.*, 2018; Rawle *et al.*, 2019).

There are several studies revealed that iAs^V is able to substitute for its chemical analog, phosphate in bacteria (Wang *et al.*, 2015; 2018). Though As substitution in nucleic acids and ATP should not be occurring because of the instability of arseno-ester bonds (Rosen *et al.*, 2011), iAs^V is capable of being incorporated into arsenosugars and arsenolipids, via formation of a methyl iAs^V -C bond, which is very stable, and has been found in fungi, plants, lichens, marine organisms,

and in fish tissues (Dembitsky and Levitsky, 2004). Interestingly, the previous studies provided evidence of iAs^V substitution for phosphate in lipids in *Agrobacterium tumefaciens* (Wang *et al.*, 2015; 2018). Additionally, compare to *A. tumefaciens* grown in the absence of As, the compositions of five lipids were different with the addition of iAs^{III}, and only one lipid was increased in levels when *A. tumefaciens* was grown in the presence of iAs^{III}; and this lipid contains a carbon scaffold structure similar to 2-octyl-cyclopropaneoctanoic acid (Wang *et al.*, 2015). However, thus far, no recognizable ArsM homolog has been found in *A. tumefaciens* using the ArsM sequence from *R. palustris* as a template for sequence homology searches, implying that novel mechanism(s) may be involved and maybe encoding novel enzymes to catalyze As-C bond formation in *A. tumefaciens*.

Additionally, high-level resistance to arsenic compounds has been linked to the arsenical resistance (*ars*) operons (Carlin *et al.*, 1995), which were widespread in bacteria. The *ars* operon consists of five genes: *arsRDABC* (Rosen *et al.*, 1991). The *arsR* and *arsD* genes encode regulatory proteins, while the *arsA* and *arsB* genes encode subunits of an ATP-driven iAs^{III} pump. ArsA is the catalytic subunit, while ArsB forms the membrane component of the pump. Thus, in the absence of the *arsA* gene, ArsB alone provides partial iAs^{III} resistance, likely functioning as a secondary uniporter. iAs^V resistance is mediated by the ArsC, which reduces iAs^V to iAs^{III}, and the resulting iAs^{III} is then extruded by ArsAB.

E. coli strain W3110 is a plasmidless *E. coli*, which contains one copy of *arsRBC* as the *ars* operon in its genome. To investigate whether the chromosomal *arsRBC* operon contributes to this baseline arsenical resistance, the *arsRBC* genes were deleted, and *E. coli* strain AW3110 was generated (Rosen *et al.*, 1995). The resulting Δ *arsRBC* strain AW3110 displayed hypersensitivity

to iAs^{III}. However, this mutant strain AW3110 still show slightly iAs^V reduce activity with non-enzymatic reaction in the cell. This set of wild type strain W3110 and $\Delta arsRBC$ mutant strain AW3110 gave us a good chance to look into the metabolomics background generated by ArsRBC in model strain *E. coli*, which will provide a clear and clean strain background for our further research about the As-C bond mechanism.

Based on the background mentioned above, there exists limited knowledge about As-C bond formation in *A. tumefaciens*, and the molecular mechanism underlying the synthesis of the arsenolipids in *A. tumefaciens* is not well understood. Thus, in my PhD progress, I want to identify the unknown gene(s) potentially involved in the molecular mechanisms of As-C bond formation, and to further investigate and characterize the biosynthesis pathway responsible for arsenolipid synthesis in *A. tumefaciens*. This study will be modeled on the previous work (Wang *et al.*, 2015; 2018, Rawle *et al.*, 2019), using *A. tumefaciens* strain 5A as a model strain. Investigation of the genome of *A. tumefaciens* 5A suggests a potential As-C bond formation pathway: one catalyzed by two candidate genes: 3-phosphoshikimate 1-carboxyvinyltransferase (AroA), and phosphoenolpyruvate mutase (PepM), which may be involved in replacing phosphate with iAs^V, followed by formation of an As-C bond (Cruz-Morales *et al.*, 2016). By accomplishing this aim, this study will: i) characterize the As metabolism background of *E. coli* AW3110, which is the *arsRBC* deletion mutant from *E. coli* W3110; then ii) determine the initial chemical step of As assimilation in organic scaffolds as a first step to understand the mechanisms by which arsenolipids are synthesized in *A. tumefaciens* using the *E. coli* AW3110 as model strain.

Increasing evidence showed that *A. tumefaciens* strains are able to incorporate iAs^V into cellular biomass, such as lipids and proteins (Wang *et al.*, 2015; 2018). However, a major

unanswered question in the field is how are the As-C bonds formed in *A. tumefaciens*? This study used a highly interdisciplinary approach to enhance our knowledge of the biochemical mechanisms by which organoarsenicals can be formed in *A. tumefaciens*. With the integration of bacterial genetic engineering and HPLC-ICPMS analytical method, this study provided a foundational understanding of As-C bond formation, and identified the uncharacterized enzyme(s) catalyzed the As-C bond formation using either iAs^V or iAs^{III} in *A. tumefaciens* 5A. Overall, the study generated new knowledge about the diversity of protein functions involved in As biotransformation and explored the biogeochemical cycling of As in the natural environment.

Methods and Materials

Bacterial Strains, Genes, and Genetic Modifications

The *E. coli* strain AW3110, along with its wild-type strain W3110 (*E. coli* Genetic Stock Center), does not naturally carry the *arsM* gene, and the complete *arsRBC* operon was experimentally deleted (Carlin *et al.*, 1995). As a result, AW3110 is incapable of iAs^{III} methylation via ArsM, iAs^V reduction to iAs^{III} via ArsC, or iAs^{III} extrusion via ArsB. The only iAs^V reduce activity will be chemical reduction in AW3110.

The genes *aroA* and *pepM* were identified based on the previous RNAseq study (Rawle *et al.*, 2019). The genes were PCR amplified and then the gene fragments were subcloned into pET-28a(+) as an BamHI-Hind^{III} restriction fragments, enabling convenient, inducible expression in *E. coli* strains. The *E. coli* AW3110 (pET-*aroA*), *E. coli* AW3110 (pET-*pepM*) and *E. coli* AW3110 (pET-*aroA-pepM*) were generated in this study. Among them, two same promoters (T7 promoter)

were added in front of the *aroA* gene and *pepM* gene, respectively, in *E. coli* AW3110 (pET-*aroA-pepM*).

Quantification of Growth and Arsenic Biotransformation

Characterizations of *E. coli* strains growth responses to inorganic iAs^{III} (as sodium arsenite) were conducted in M9 chemical defined media containing 50 μ M of iAs^V and 100 μ M Pi (phosphate) as needed. Growth was quantified as optical density (OD₆₀₀) using a SpectraMax microtiter plate reader (Molecular Devices, California). Viable cell counts were determined by serially diluting cultures, spreading the dilutions onto LB agar plates, and counting the colonies after overnight incubation at 37°C. The cell counts were then normalized to the volume of media (mL).

For identification of arsenic biotransformation, the cultures were taken during the culturing, after centrifugation ($8000 \times g$), the supernatant will be filtered and diluted. Then the arsenical concentrations would be detected using HPLC-ICPMS (Agilent Technologies). For quantification of arsenic bioaccumulation, cells were grown to either early-log or mid-log phase in desired M9 medium, collected via centrifugation ($8000 \times g$), washed $3\times$ in Tris-HCl buffer (20 mM, pH 7.2), and finally digested in 5% trace metal grade nitric acid using heat and pressure (115°C, 29.7 PSI for 30 min) (Coryell *et al.*, 2018). Then the samples were filtered and diluted before arsenical detection using HPLC-ICPMS (Agilent Technologies).

Sample Collection

For the compatible metabolomics of *E. coli* strains W3110 and AW3110, the cells were cultured in M9 medium containing 100 μ M Pi, with or without the addition of 50 μ M iAs^V. Then

1 L of W3110 cultures was collected at early-log phase (4 h), and mid-log phase (8 h), respectively, while 1 L of AW3110 cultures were collected at early-log phase (8 h), and mid-log phase (16 h), respectively. The cell pellets were washed three times using Tris-HCl (20 mM, pH 7.2) and then stored at -80 °C until metabolite extraction.

For the As-C bond metabolomics study, the strains *E. coli* AW3110 (pET-*aroA*), *E. coli* AW3110 (pET-*pepM*) and *E. coli* AW3110 (pET-*aroA-pepM*) were cultured in M9 medium containing 100 μ M Pi, with or without the addition of 50 μ M iAs^V. Then 500 mL the cultures were collected at 12 h with the addition of iAs^V (mid-log phase), and 8 h without the addition of iAs^V (mid-log phase). The cell pellets were washed three times using Tris-HCl (20 mM, pH 7.2) and then stored at -80 °C until metabolite extraction.

Metabolite Extraction for ¹H NMR Analysis

Cell pellets (100-200 mg) were retrieved from -80°C storage and prepared for extraction. Each sample was suspended in 1000 μ L of a 50:50 methanol-water solution and thoroughly mixed by vortexing or pipetting. The suspension was transferred into a 2 mL FastPrep lysis B-Matrix tube, and cells were lysed using a FastPrep-24 5g instrument set to *S. aureus*-specific parameters (6.0 m/s for 40 seconds, two cycles). Following lysis, the samples were centrifuged at 10,000g for 10 minutes.

A 10 μ L aliquot of each sample was collected and stored in a 1 mL tube for subsequent protein quantification via the Bradford assay. For phase separation, 600 μ L of chloroform was added to each sample in the FastPrep 2 mL tubes, which were then vortexed and frozen at -20°C for 1 hour. The samples were centrifuged again at 10,000g for 10 minutes to achieve layer separation.

The polar (upper) layer (600 μ L) was carefully removed and transferred to a fresh 2 mL tube. The polar layer was dried overnight in a SpeedVac under manual, non-heated conditions. Once dried, samples were stored at -80°C until further analysis.

For NMR analysis, the dried polar extracts were reconstituted in 600 μ L of NMR buffer, comprising 0.25 mM DSS, 0.4 mM imidazole, 25 mM phosphate buffer, and 10% D₂O. The reconstituted samples were then transferred to NMR tubes, and one-dimensional proton (¹H) NMR spectra were acquired on a 600 MHz NMR spectrometer using the Bruker pulse sequence "zgesgp".

Sample Preparation and NMR Acquisition

Metabolite extracts were resuspended in 600 μ L of a phosphate buffer solution (10 mM NaH₂PO₄/Na₂HPO₄ containing 0.25 mM DSS in 100% D₂O, pH 7.0) and transferred to 5 mm NMR tubes. ¹H NMR spectra were acquired on a 600 MHz Bruker AVANCE III solution NMR spectrometer (Bruker Daltonics, Billerica, MA) equipped with an automatic sample loading system (SampleJet™), a 5 mm triple-resonance (¹H, ¹³C, ¹⁵N) helium-cooled cryoprobe (Cryoprobe™), and TopSpin™ software (version 3.2). The probe temperature was maintained at 300 K, and data acquisition was performed using the gradient-based water suppression pulse sequence "zgesgp" over a spectral width of 12 ppm. Each spectrum was collected with 2048 scans and 65 K data points, with a dwell time of 69 μ s, resulting in a total free induction decay (FID) acquisition time of 4.5 seconds. The inter-scan recovery delay (D1) was set to 2 seconds, providing a total relaxation recovery delay of 6.5 seconds between scans.

TopSpin software (version 3.2, Bruker, Billerica, MA, USA) was employed for chemical shift referencing to the DSS peak at 0.0 ppm. Each spectrum was manually phased, baseline-corrected, and a 0.5 Hz line-broadening function was applied to enhance resolution.

Spectral Processing and Metabolite Quantification

Processed spectra were analyzed with Chenomx NMR Suite software (version 8.0; Chenomx Inc., Edmonton, AB, Canada). For metabolite identification, the Chenomx 600 MHz, small molecule library was utilized to independently fit spectral patterns to each sample. Metabolite concentrations were quantified using DSS as an internal reference standard, calibrated to 0.0 ppm. This standard enabled precise chemical shift referencing and quantification across samples.

Statistical Analysis

Metabolite concentrations were exported from Chenomx in micromolar units and analyzed statistically using MetaboAnalyst (version 5.0, Edmonton, AB, Canada). The data were log₁₀-transformed and normalized based on colony-forming unit (CFU) scaling, followed by both univariate and multivariate statistical analyses, including Principal Component Analysis (PCA), hierarchical clustering, t-tests, and volcano plot analysis for assessing fold changes and p-values. Partial Least Squares Discriminant Analysis (PLS-DA) was performed to highlight significant metabolite differences, with metabolites showing VIP scores greater than 1.1 considered critical for distinguishing between control and experimental groups.

Results and Discussion

Compare Metabolomics of *E. coli* W3110 and AW3110

Growth and iAs^V Reduction Curve. The growth and iAs^V reduction were characterized in *E. coli* strains W3110 and AW3110. With the inoculation in M9 medium (100 μ M Pi), wild-type strain *E. coli* W3110 grew better than mutant strain *E. coli* AW3110 (Fig. 4.1), indicating that the deletion of *arsRBC* may not only affect the arsenical resistance for *E. coli* strains, may also show effect on other metabolic pathways in *E. coli* strain. Based on the iAs^V reduction curve, *E. coli* W3110 reduced the total of 50 μ M iAs^V in 8 h, while *E. coli* AW3110 cannot finish the reduction of the same amount of iAs^V in 24 h (Fig. 4.2), indicating the deletion of *arsRBC* showed the effect on the iAs^V reduction efficiency. Based on the growth and iAs^V reduction results, we decided to collect the culture samples both at early-log and mid-log phases, but for wild type strain, cultures were collected at 4 h and 8 h, while the mutant strain cultures were collected at 8 and 16h.

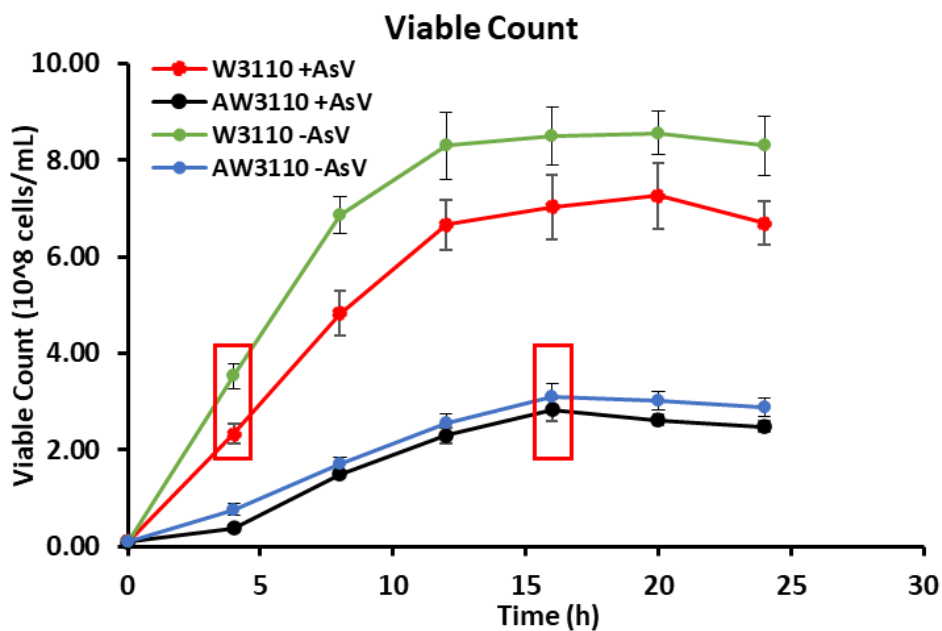


Figure 4.1 The growth curve of *E. coli* strains W3110 and AW3110 in M9 medium (100 μ M Pi), with or without the addition of 50 μ M iAs^V. The red squares showed the sample collection time for different strains.

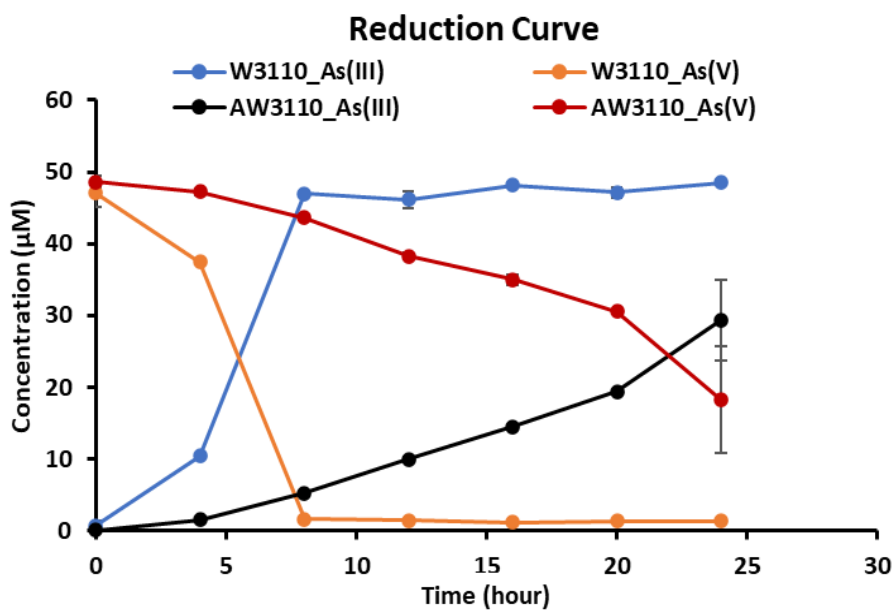


Figure 4.2 The iAs^V reduction curve of *E. coli* strains W3110 and AW3110 in M9 medium (100 μ M Pi), with the addition of 50 μ M iAs^V.

NMR Metabolomic Differences in *E. coli* Strains With or Without iAs^V – Whole Map. We speculated that metabolic pathways would be altered due to iAs^{III} stress between *E. coli* strains W3110 and AW3110. The representative ¹H NMR spectrum of intracellular metabolites from AW3110 cells after 8 h cultivation without the addition of iAs^V showed in Fig. 4.3. In total of 70 metabolites showed changes between *E. coli* strains W3110 and AW3110 have been listed in Table A2.

Differences in metabolic profiles between *E. coli* strains W3110 and AW3110 were first assessed using an unsupervised multivariate PCA analysis (Fig. 4.4). Two dimensional (2D) PCA scores plot shown in Fig. 4.4 revealed the clear separation of metabolites between the wild-type strain and mutant strain (Fig. 4.4).

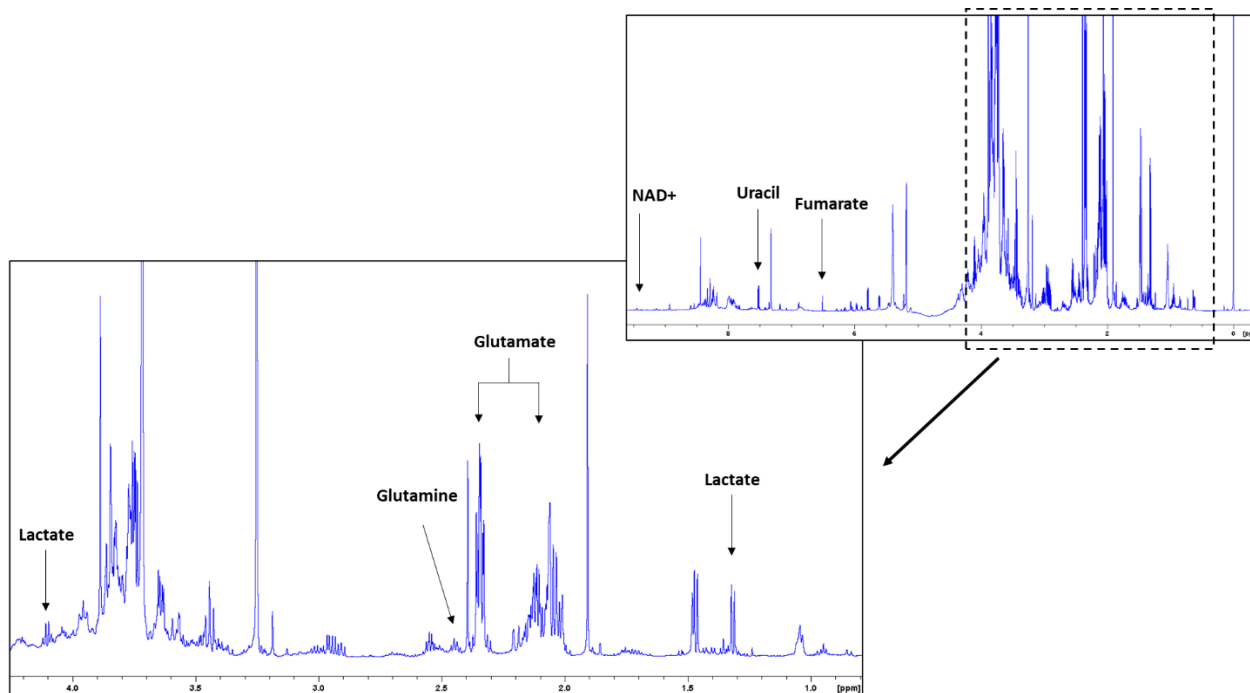


Figure 4.3 Representative ¹H NMR spectrum of intracellular metabolites obtained from AW3110 cells after 8h cultivation without iAs^V.

Differences in the metabolite profiles of *E. coli* strains were further assessed through a hierarchical cluster analysis (HCA) and visualized using heatmaps (Fig. 4.5). Based on the HCA, the metabolites from the same strain clustered more separately than the metabolites from different iAs^V stress (Fig. 4.5). Consistent with the PCA results, both of the results indicated the deletion of *ars* operon showing more effect on the metabolic changes than arsenical stress.

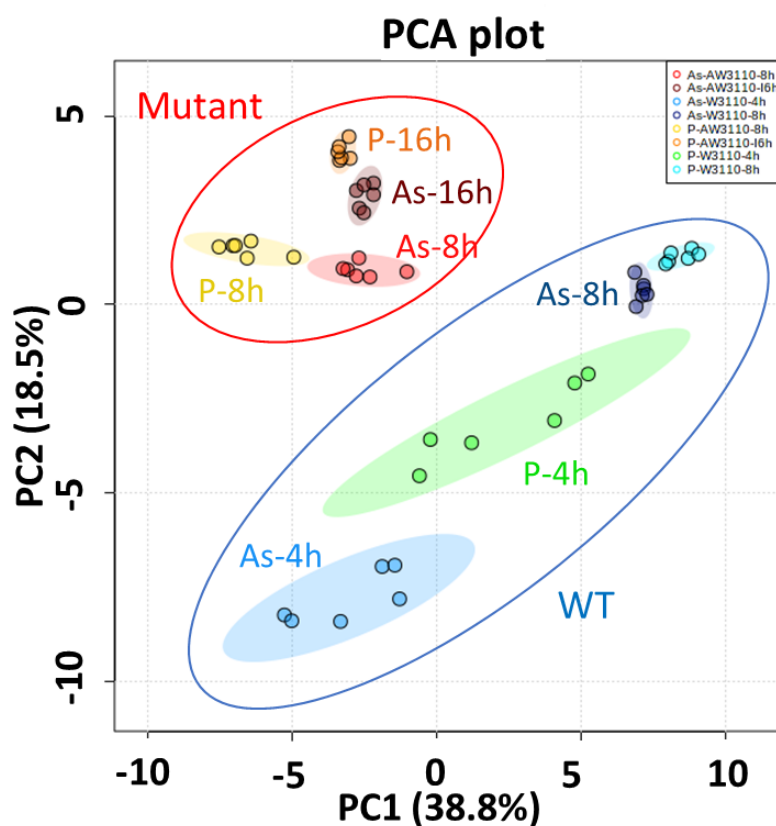


Figure. 4.4 Multivariate principal component analyses (PCA) of the metabolic profiles of *E. coli* W3110 (blue) and AW3110 (red) indicate mutant AW3110 separated from the wild type strain W3110 with or without the addition of iAs^V. For WT (wild-type strain), 4 h and 8 h indicate early-log phase and mid-log phase, respectively. For Mutant (mutant strain AW3110), 8 h and 16 h indicate early-log phase and mid-log phase, respectively.

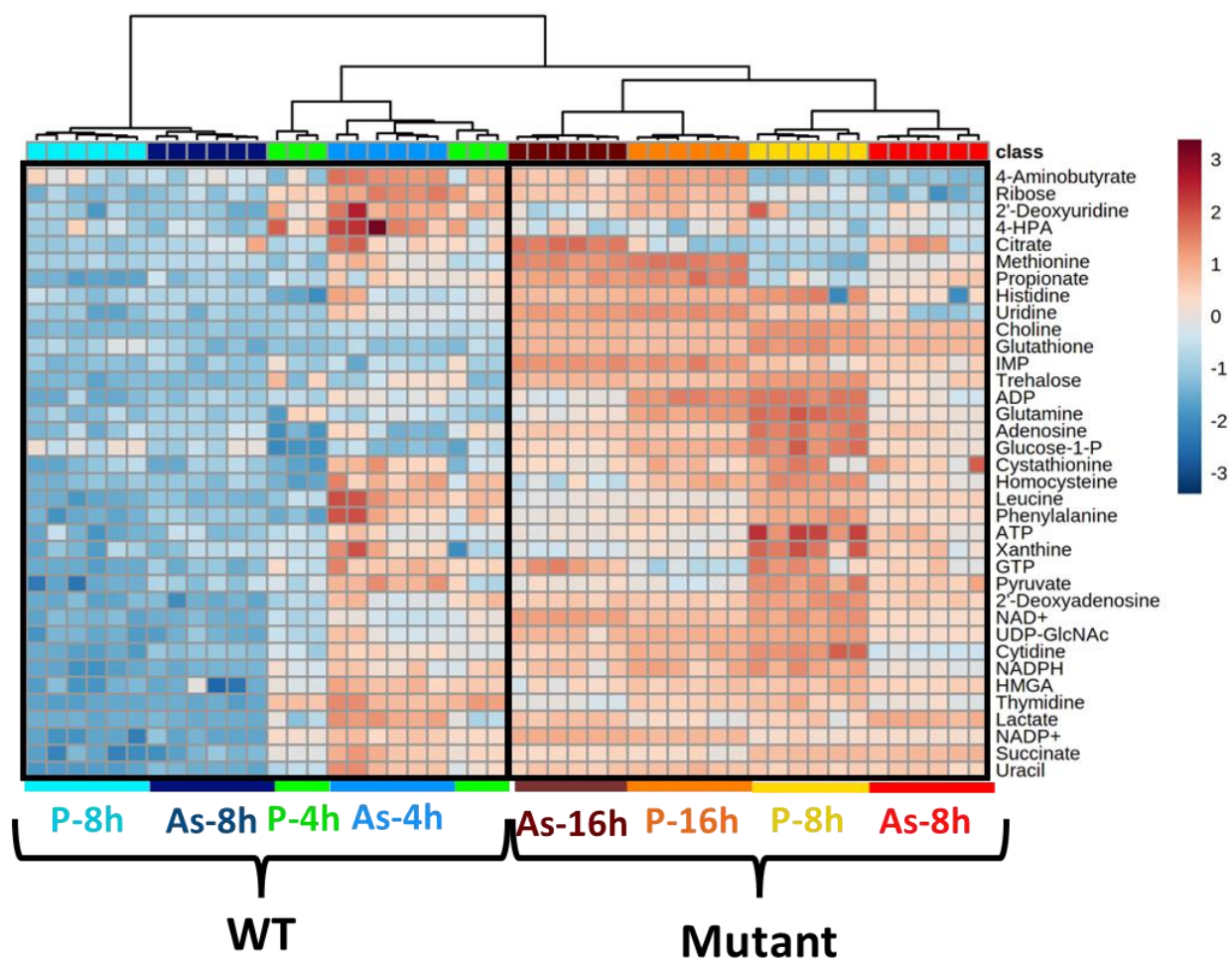


Figure 4.5 Hierarchical clustering analysis (HCA) and heatmap visualizations of *E. coli* strains W3110 (WT) and AW3110 (Mutant) intracellular samples. For WT, 4 h and 8 h indicate early-log phase and mid-log phase, respectively. For Mutant, 8 h and 16 h indicate early-log phase and mid-log phase, respectively. The red-blue color spectrum represents the scaled abundance of each metabolite, with dark red and blue indicating high and low abundance, respectively. The distinct clustering based on metabolite levels and genotype were observed in both wild type and mutant strains.

NMR Metabolomic Differences in *E. coli* Strains Without iAs^V. Comparing the metabolites from early-log phase and mid-log phase of wild-type strain W3110 cultured without iAs^V, the PCA plot showed significant separation between the two groups, indicating the difference of metabolic pathways between early-log phase and mid-log phase in *E. coli* W3110 (Fig. 4.6). Compared to

early-log phase sample without As, WT at mid-log phase showed 33 metabolites down and 16 metabolites up with at least 1.5-fold change ($p<0.05$) (Fig. 4.7).

When comparing the metabolites from early-log phase and mid-log phase of mutant strain AW3110, the same trend of results showed up (Figs. 4.8 and 4.9). The significant separation between the two groups indicating the difference of metabolic pathways between early-log phase and mid-log phase in *E. coli* AW3110 (Fig. 4.8). Compare to the early-log phase sample, the mutant strain at the mid-log phase showed 39 metabolites down and 10 metabolites up with at least 1.5-fold change ($p<0.05$) (Fig. 4.9).

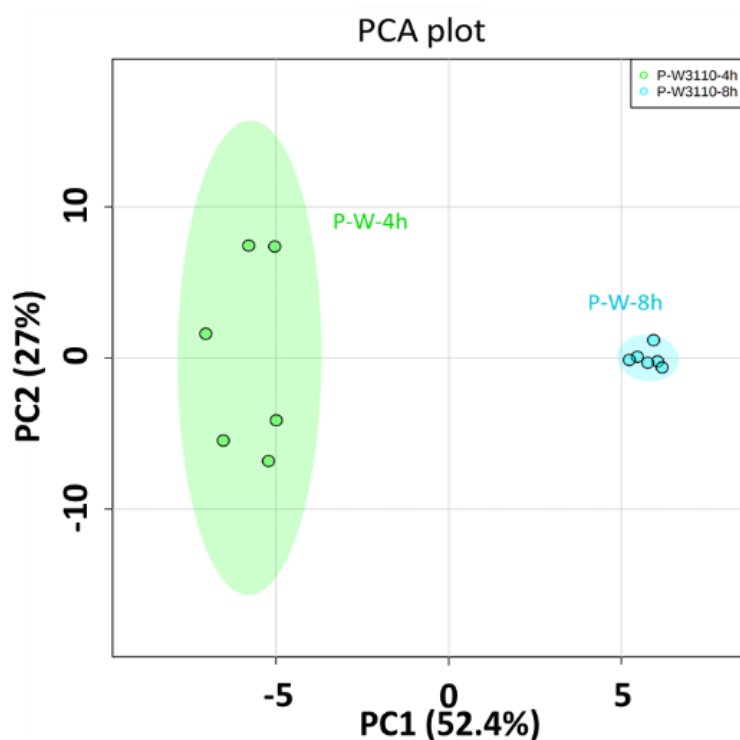


Figure. 4.6 Multivariate principal component analyses (PCA) of the metabolic profiles of early-log (green) and mid-log (blue) phases of *E. coli* W3110 cultured without iAs^V, indicating the separation of metabolites in different growth phases of *E. coli* W3110. Additionally, 4 h and 8 h indicate early-log phase and mid-log phase, respectively.

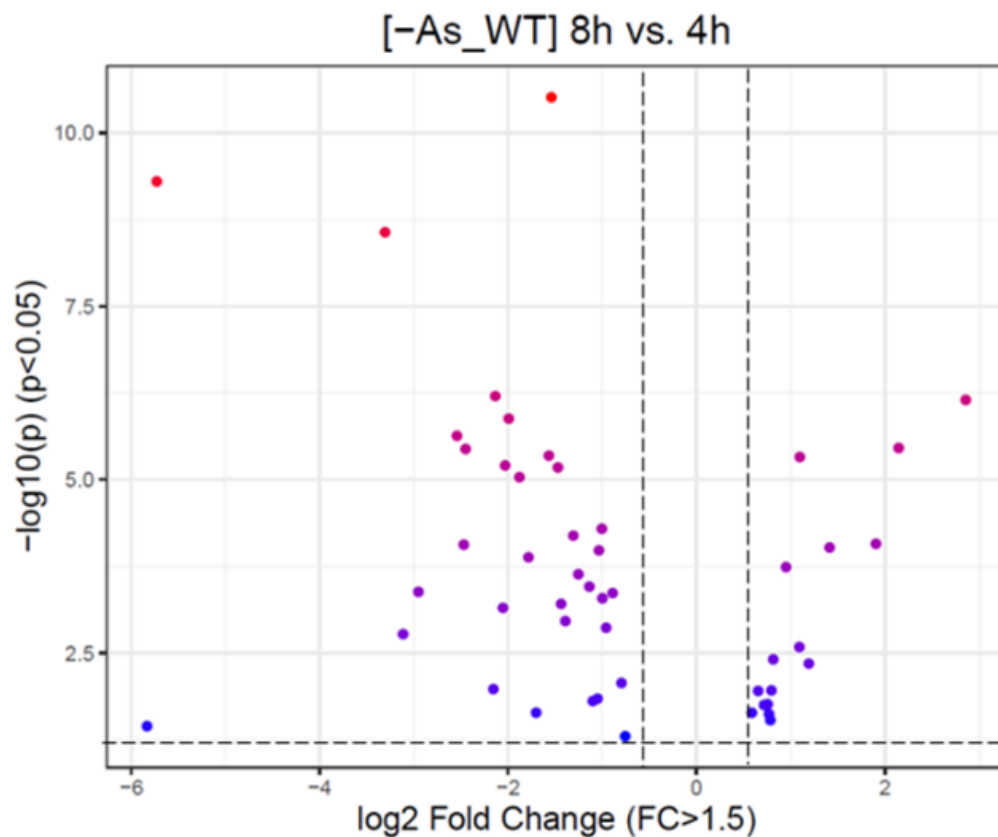


Figure. 4.7 Volcano plot of metabolites of mid-log phase wild-type *E. coli* W3110, compared to the ones at early-log phase. 4 h and 8 h indicate early-log phase and mid-log phase, respectively. Y-axis shows statistical significance as the $-\text{Log}_{10}P$ values while the x-axis the Log_2 fold change in gene expression (positive values represent overexpression in the NMR). The default cut-off for log_2 FC is > 1.5 and the default cut-off for P value is 10^{-6} , denoted by the dotted lines.

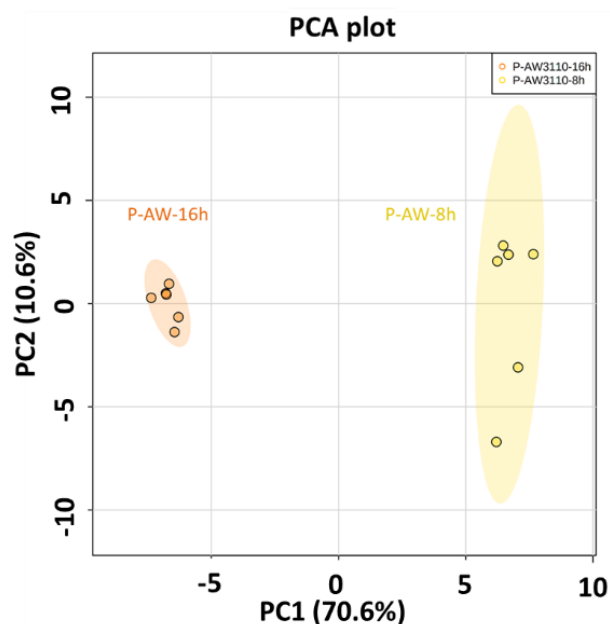


Figure. 4.8 Multivariate principal component analyses (PCA) of the metabolic profiles of early-log (yellow) and mid-log (orange) phases of *E. coli* AW3110 cultured without iAs^V , indicating the separation of metabolites in different growth phases of *E. coli* AW3110. 8 h and 16 h indicate early-log phase and mid-log phase, respectively.

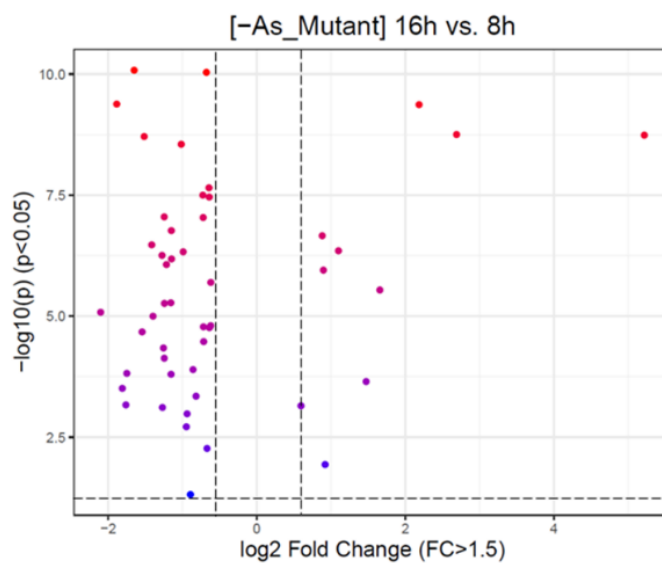


Figure. 4.9 Volcano plot of metabolites of mid-log phase mutant *E. coli* AW3110, compared to the ones at early-log phase. 8 h and 16 h indicate early-log phase and mid-log phase, respectively. Y-axis shows statistical significance as the $-\log_{10}P$ values while the X-axis the $\log_2$ fold change in

gene expression (positive values represent overexpression in the NMR). The default cut-off for $\log_2$ FC is > 1.5 and the default cut-off for P value is 10^{-6} , denoted by the dotted lines.

When compared the metabolites of mutant strain to the ones of wild type strain (Fig. 4.10), both PCA plot and PLSDA plot showed the separation of metabolites between the mutant strain and wild-type strain, no matter the growth phases (Fig. 4.10). Additionally, the metabolite change between early-log phase and mid-log phase were in different directions between the wild-type strain W3110 and mutant strain AW3110, implying the deletion of *ars* operon shifted the global metabolic pathway in *E. coli* strains. The top 20 metabolites drove the difference between the two *E. coli* strains were mainly belonging to nucleotide metabolism, oxidative stress, energy metabolism and carbohydrate metabolism (Fig. 4. 11).

With the analysis of signature metabolomic profiles identified in mutant strain that significantly discriminated against the wild-type strain (Fig. 4.12), the increased lactate in mutant strain, compared with WT, indicates elevated lactate fermentation of the NAD^+ regeneration process in mutant strain AW3110 (Fig. 4.13). While the carbohydrate metabolites pattern is consistent with the energy metabolism. (Figs. 4.12 and 4.13), both metabolites pattern of carbohydrate and energy metabolism consistent with the growth deficiency of the mutant strain AW3110. All the results revealed the deletion of *ars* operon did not only show effect on arsenical resistance, but also interrupt the key metabolism pathways related with the bacterial growth. This conclusion is consistent with the previous publications (Shi *et al.*, 2018, Rawle *et al.*, 2019), indicating that *ars* operon involved in the global metabolism in bacterial cells.

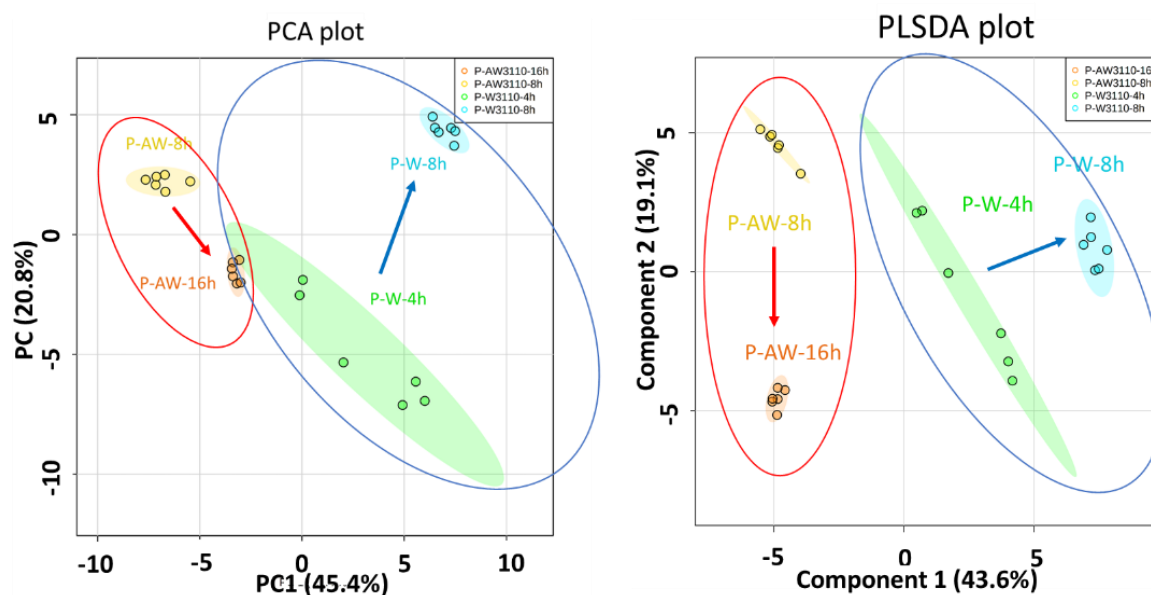


Figure. 4.10 Multivariate principal component analyses (PCA) and partial least squares discriminant analysis (PLSDA) of the metabolic profiles of early-log (yellow for mutant and green for WT) and mid-log (orange for mutant and blue for WT) phases of *E. coli* strains W3110 and AW3110 cultured without iAs^V, indicating the separation of metabolites in different *E. coli* strains with or without *ars* operon. For WT (wild-type strain), 4 h and 8 h indicate early-log phase and mid-log phase, respectively. For Mutant (mutant strain AW3110), 8 h and 16 h indicate early-log phase and mid-log phase, respectively.

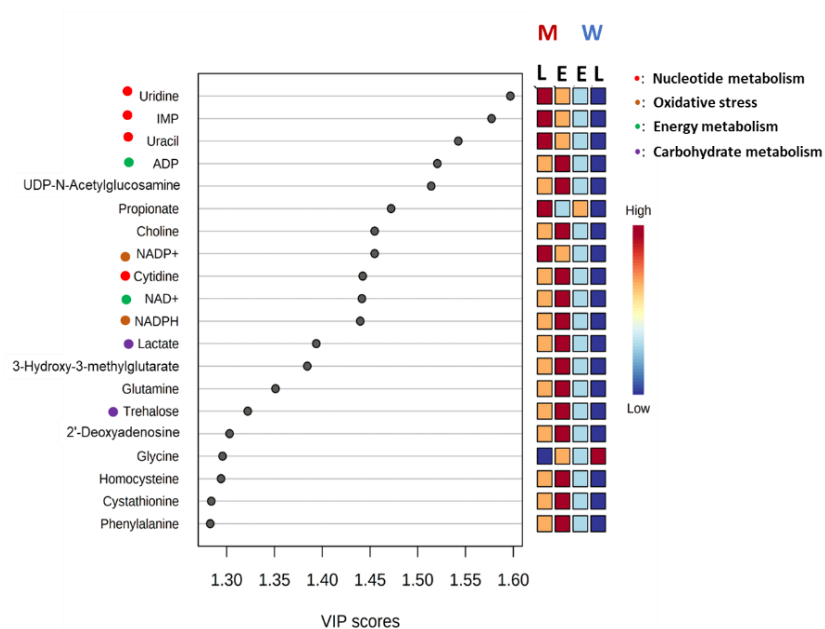


Figure. 4.11 A VIP score of a variable calculated as a weighted sum of the squared correlations between the PLSDA components and the original variable of the top 20 metabolites different between *E. coli* strains W3110 and AW3110 cultured without iAs^V.

•: Carbohydrate metabolism •: Energy metabolism

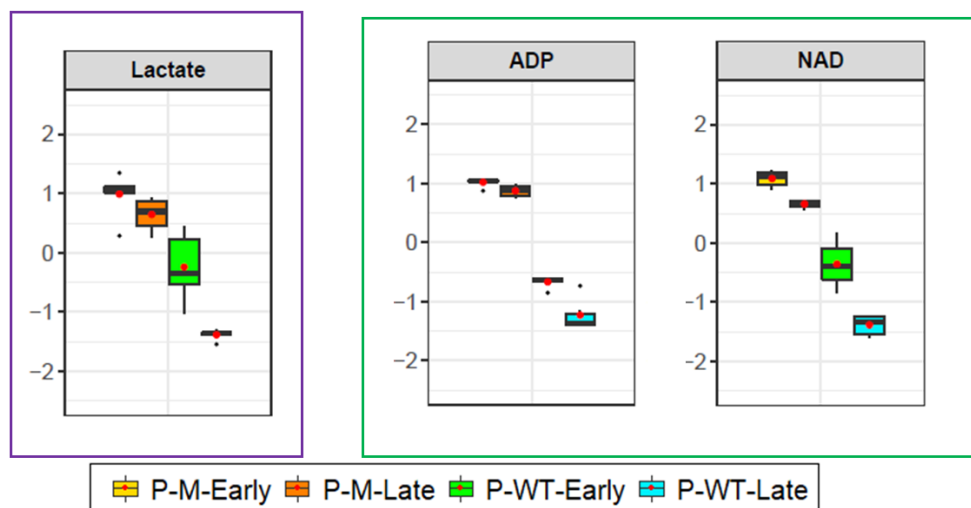


Figure. 4.12 The two-sample *t*-test of signature metabolomic profiles identified in mutant strain that significantly discriminated against the wild-type strain. The *t*-test parameters assumed equal variance and the threshold *p* value < 0.5.

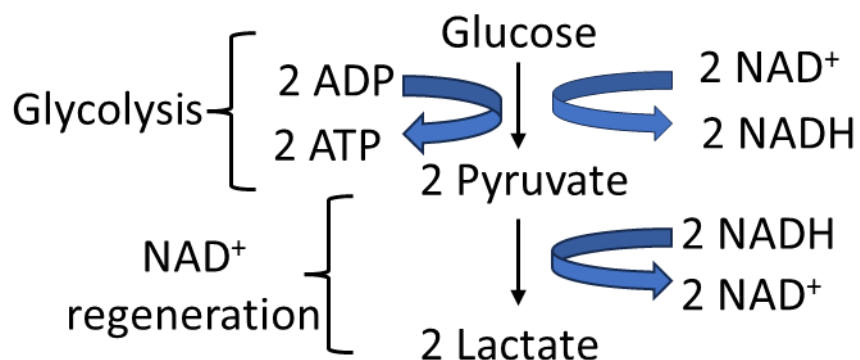


Figure. 4.13 An illustration of how the glycolysis and NAD⁺ regeneration cycle both affected by the deletion of *ars* operon.

NMR Metabolomic Differences in *E. coli* Strains With iAs^V. Comparing the metabolites from early-log phase and mid-log phase of wild-type strain W3110 cultured with iAs^V, the PCA plot also showed significant separation between the two groups, indicating the difference of metabolic pathways between early-log phase and mid-log phase in *E. coli* W3110 cultured with iAs^V (Fig. 4.14). Compared to early-log phase sample without As, WT at mid-log phase showed 43 metabolites down and 5 metabolites up with at least 1.5-fold change ($p < 0.05$) (Fig. 4.15).

When comparing the metabolites from early-log phase and mid-log phase of mutant strain AW3110 cultured with iAs^V, the same trend of results showed up (Figs. 4.16 and 4.17). The significant separation between the two groups indicates the difference of metabolic pathways between early-log phase and mid-log phase in *E. coli* AW3110 (Fig. 4.16). Compared to the early-log phase sample, the mutant strain at the mid-log phase showed 16 metabolites down and 15 metabolites up with at least 1.5-fold change ($p < 0.05$) (Fig. 4.17).

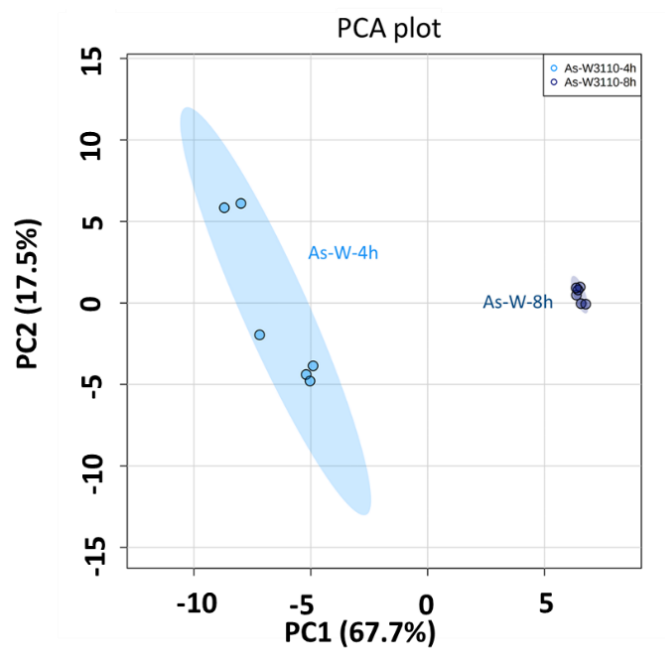


Figure. 4.14 Multivariate principal component analyses (PCA) of the metabolic profiles of early-log (blue) and mid-log (dark blue) phases of *E. coli* W3110 cultured with iAs^V, indicating the separation of metabolites in different growth phases of *E. coli* W3110. 4 h and 8 h indicate early-log phase and mid-log phase, respectively.

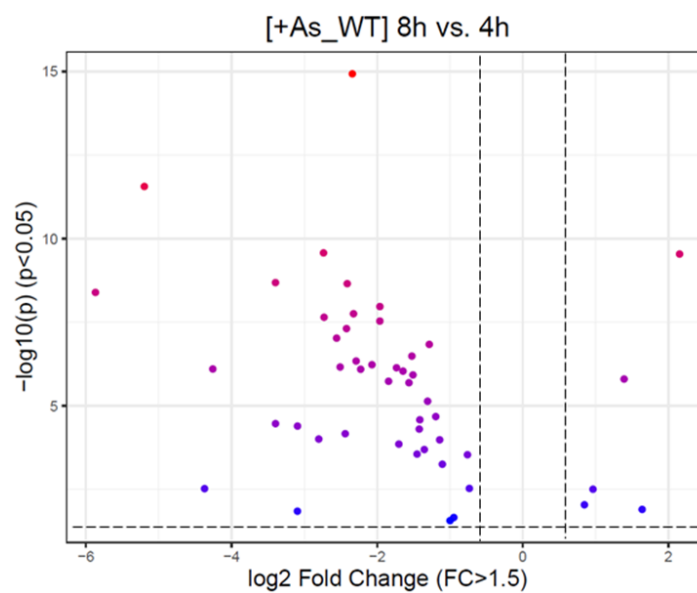


Figure. 4.15 Volcano plot of metabolites of mid-log phase wild-type *E. coli* W3110 cultured with iAs^V, compared to the ones at early-log phase. 4 h and 8 h indicate early-log phase and mid-log

phase, respectively. Y-axis shows statistical significance as the $-\text{Log}_{10}P$ values while the x-axis the Log_2 fold change in gene expression (positive values represent overexpression in the NMR). The default cut-off for log_2 FC is > 1.5 and the default cut-off for P value is 10^{-6} , denoted by the dotted lines.

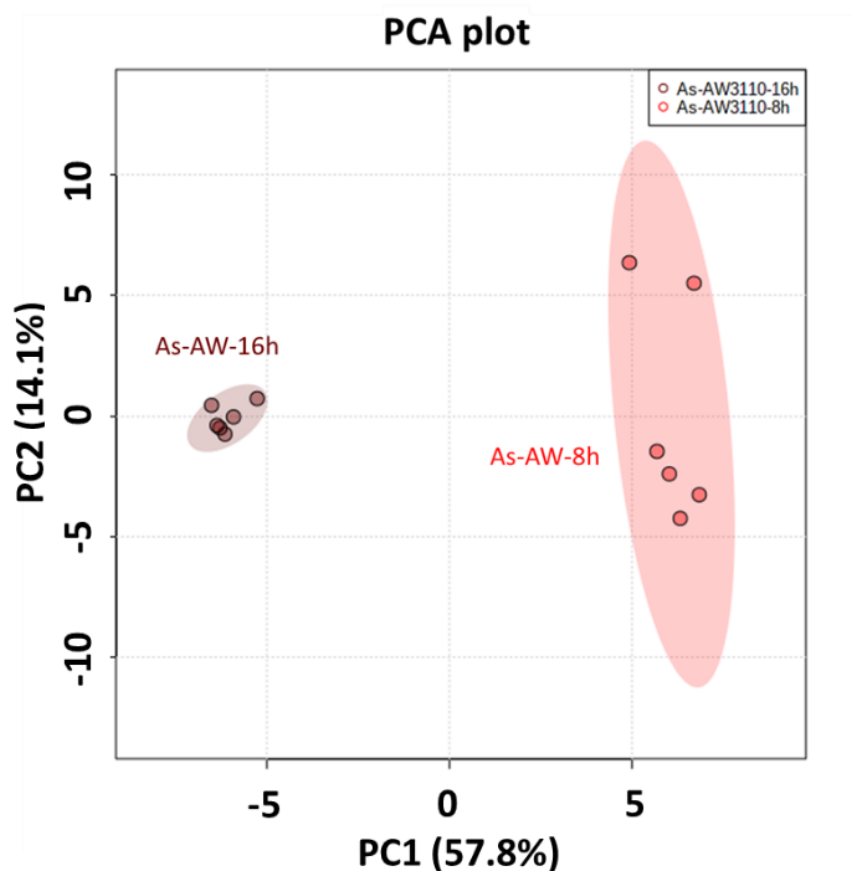


Figure. 4.16 Multivariate principal component analyses (PCA) of the metabolic profiles of early-log (pink) and mid-log (brown) phases of *E. coli* AW3110 cultured with iAs^V , indicating the separation of metabolites in different growth phases of *E. coli* AW3110. 8 h and 16 h indicate early-log phase and mid-log phase, respectively.

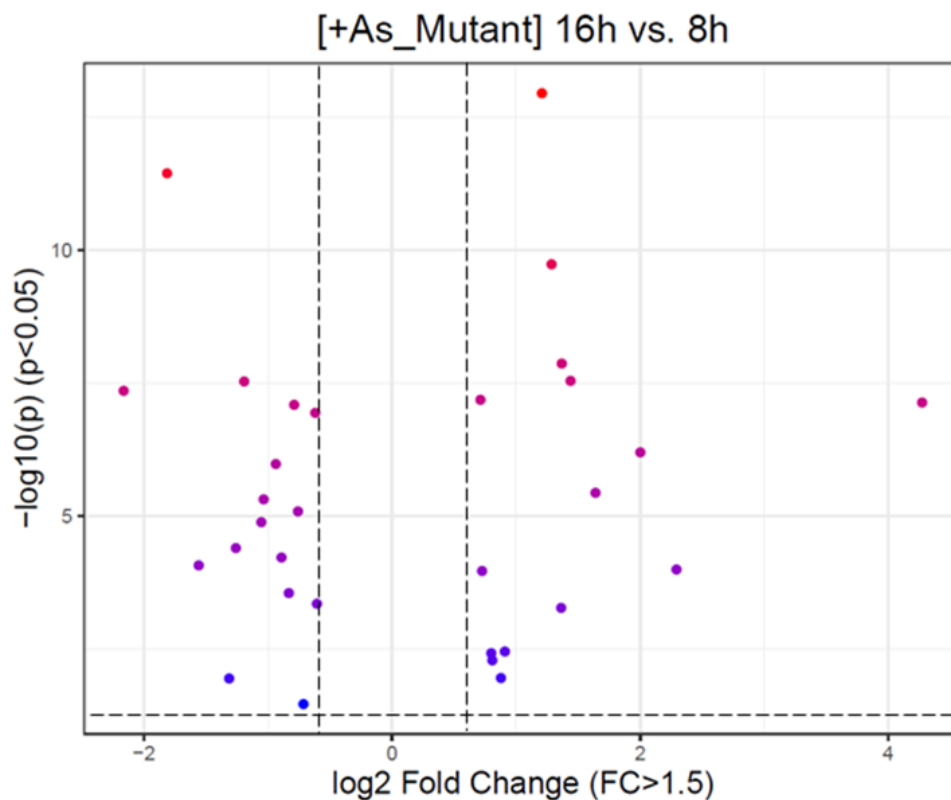


Figure. 4.17 Volcano plot of metabolites of mid-log phase mutant *E. coli* AW3110, compared to the ones at early-log phase. 8 h and 16 h indicate early-log phase and mid-log phase, respectively. Y-axis shows statistical significance as the $-\text{Log}_{10}P$ values while the X-axis the Log_2 fold change in gene expression (positive values represent overexpression in the NMR). The default cut-off for log_2 FC is > 1.5 and the default cut-off for P value is 10^{-6} , denoted by the dotted lines.

When compared the metabolites of mutant strain to the ones of wild type strain (Fig. 4.18), PCA plot showed the significant separation of metabolites between the mutant strain and wild-type strain, no matter the growth phases (Fig. 4.18). Consistently with the without iAs^V results, the metabolite change between early-log phase and mid-log phase were in different directions between the wild-type strain W3110 and mutant strain AW3110, implying the deletion of *ars* operon shifted the global metabolic pathway in *E. coli* strains. Additionally, the arsenical stress combined with the deletion of *ars* operon drive the metabolites even more significant to different directions during

the bacterial growth. The top 20 metabolites driven the difference between the two *E. coli* strains were mainly belonging to oxidative stress, carbohydrate metabolism and amino acid metabolism (Fig. 4. 19).

With the analysis of signature metabolomic profiles identified in mutant strain that significantly discriminated against the wild-type strain (Fig. 4.20), the decreased fumarate in mutant strain, compared with WT, indicates the block TCA cycle in the full growth period of mutant strain with the addition of arsenical (Fig. 4.21). Because of the iAs^{III} will block the process from α -ketoglutarate to succinyl-CoA by binding the enzyme catalyze this reaction (Rawle *et al.*, 2019), the accumulated α -ketoglutarate explained the increase of glutamine, while the decreased oxaloacetate because of the block TCA cycle resulted in the decrease of aspartate. While the carbohydrate metabolites pattern is consistent with the energy metabolism in the mutant strain (Figs. 4.20 and 4.21). All the results are consistent and indicated the iAs^V reduction in the mutant strain still showed block of TCA cycle. The TCA cycle blocked more significantly in the mutant strain, compare to the wild-type strain, could be explained by the ArsB – iAs^{III} efflux function – in the wild-type strain. Because the iAs^{III} generated from iAs^V reduction could be pump out from the bacterial cell, the TCA cycle showed less suppressed in the wild-type strain. Thus, the carbohydrate and amino acid metabolism showed less change in wild-type strain W3110.

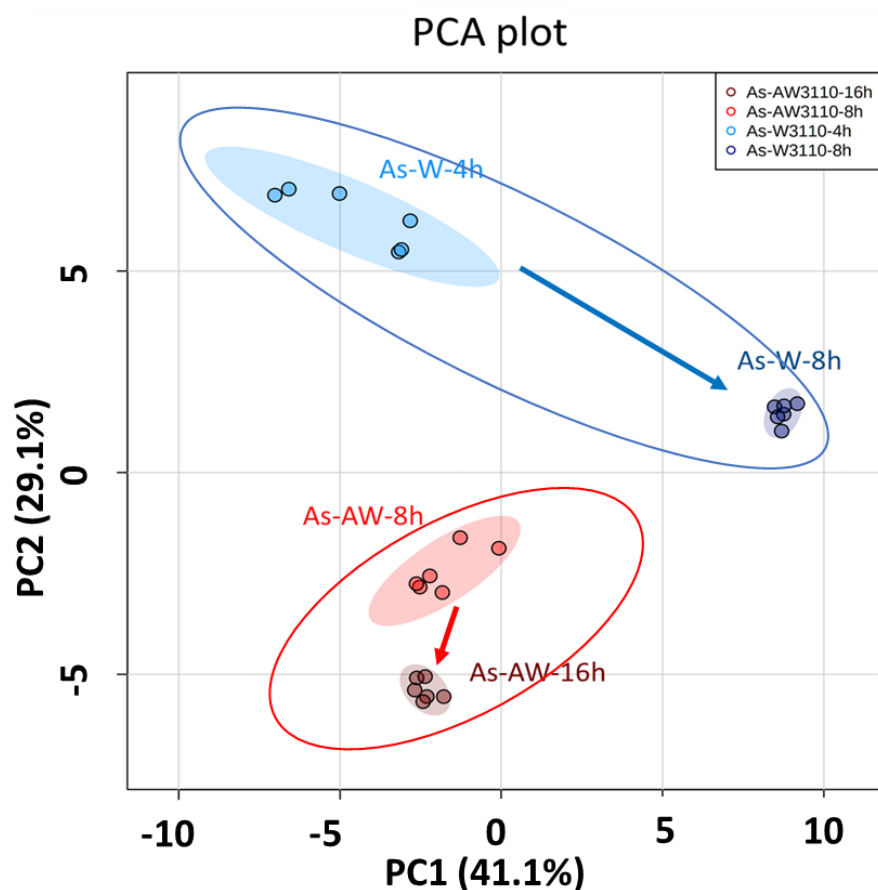


Figure. 4.18 Multivariate principal component analyses (PCA) of the metabolic profiles of early-log (yellow for mutant and green for WT) and mid-log (orange for mutant and blue for WT) phases of *E. coli* strains W3110 and AW3110 cultured with iAs^V, indicating the separation of metabolites in different *E. coli* strains with or without *ars* operon. For WT (wild-type strain), 4 h and 8 h indicate early-log phase and mid-log phase, respectively. For Mutant (mutant strain AW3110), 8 h and 16 h indicate early-log phase and mid-log phase, respectively.

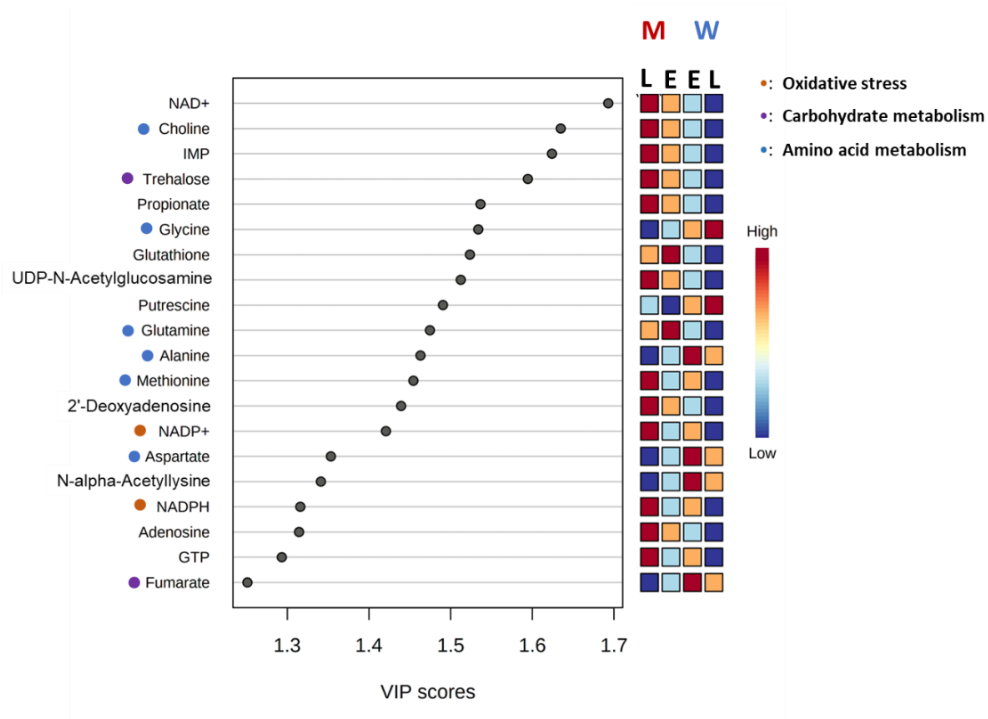


Figure. 4.19 A VIP score of a variable calculated as a weighted sum of the squared correlations between the PLSDA components and the original variable of the top 20 metabolites different between *E. coli* strains W3110 and AW3110 cultured with iAs^V.

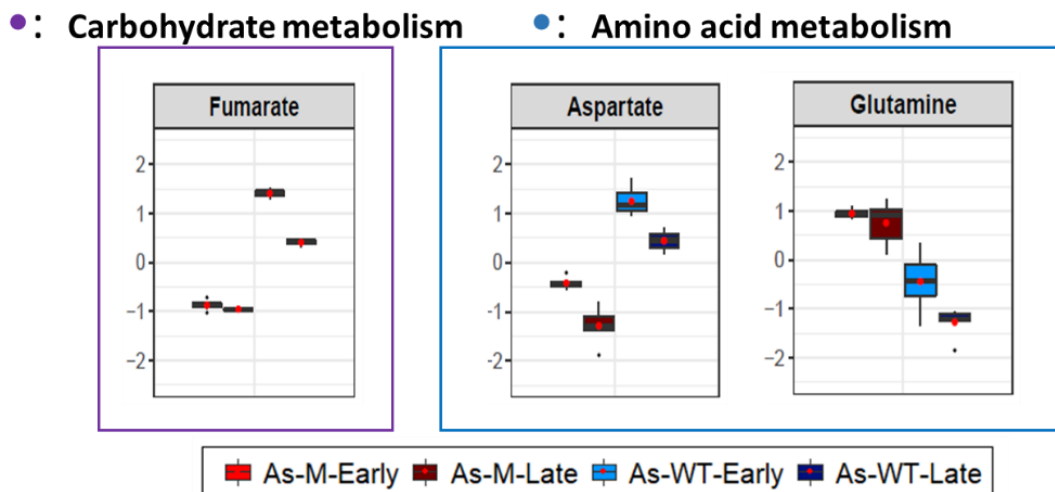


Figure. 4.20 The two-sample *t*-test of signature metabolomic profiles identified in mutant strain that significantly discriminated against the wild-type strain, culture with iAs^V. The *t*-test parameters assumed equal variance and the threshold *p* value < 0.5.

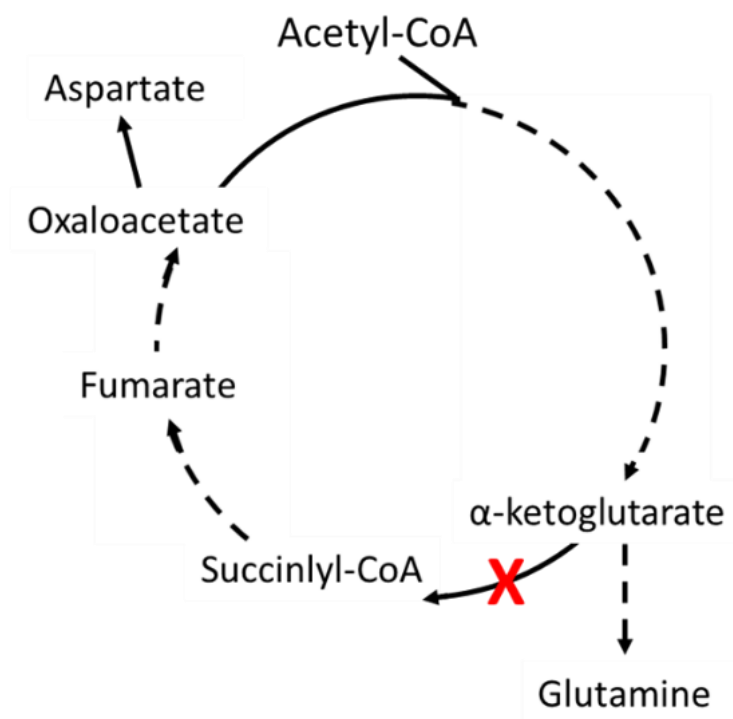


Figure. 4.21 An illustration of how the TCA cycle and amino acid metabolism both affected by the deletion of *ars* operon with the present of iAs^V

The Possible As-C Bond Mechanism

Construction of AroA and PepM Expression Strains. The gene fragments of *aroA* and *pepM* were amplified from the genome of *A. tumefaciens* 5A, and ligated on the plasmid pET-28a(+) using BamHI-HindIII enzymatic digestion sites. In the double-expressed construct, there are T7 promoters in front of both *aroA* and *pepM* gene fragments, to make sure their expression will be under the same induce strength. Finally, the constructed plasmid pET-28a, pET-*aroA*, pET-*pepM* and pET-*aroA-pepM* were transferred into *E. coli* strain AW3110, respectively, and build up the four constructs: *E. coli* strains AW3110 (pET-28a), AW3110 (pET-*aroA*), AW3110 (pET-*pepM*), and AW3110 (pET-*aroA-pepM*).

For identify the construction, the growth of the four constructions were tested in M9 medium containing only 100 μM Pi, with or without the addition of 50 μM iAs^V (Fig. 4.22). The results indicated that the growth of AW3110 (pET-*aroA*), AW3110 (pET-*pepM*), and AW3110 (pET-*aroA-pepM*) was better than AW3110 (pET-28a) either with or without iAs^V (Fig. 4.22). Additionally, the SDS-PAGE confirmed the growth differences shown in Fig. 4.22 caused by the expression of AroA and PepM (Fig. 4.23).

Lastly, the arsenical accumulation in bacterial cells of *E. coli* strains AW3110 (pET-28a), AW3110 (pET-*aroA*), AW3110 (pET-*pepM*), and AW3110 (pET-*aroA-pepM*) were characterized. As shown in Fig. 4.24, the total arsenical contents were decreased significantly, with the expression of AroA and PepM.

Based on the results, the expression of AroA and PepM in *E. coli* AW3110 showed the effect on the arsenical resistance and accumulation. Thus, we collected the cultures 8 h after the IPTG addition when iAs^V present, while 4 h after the IPTG addition when iAs^V absent. Then the metabolites from the cell pellets were considered as the metabolites from intracellular, while the metabolites from the supernatant after centrifugation were considered as the metabolites from the extracellular.

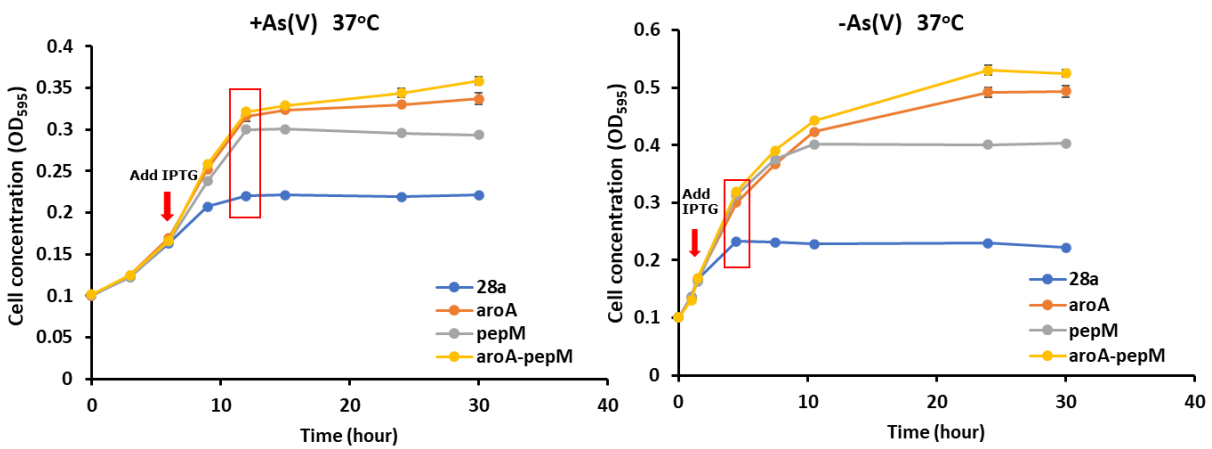


Figure. 4.22 Growth curve of *E. coli* strains AW3110 (pET-28a), AW3110 (pET-*aroA*), AW3110 (pET-*pepM*), and AW3110 (pET-*aroA-pepM*) with or without the addition of iAs^V.

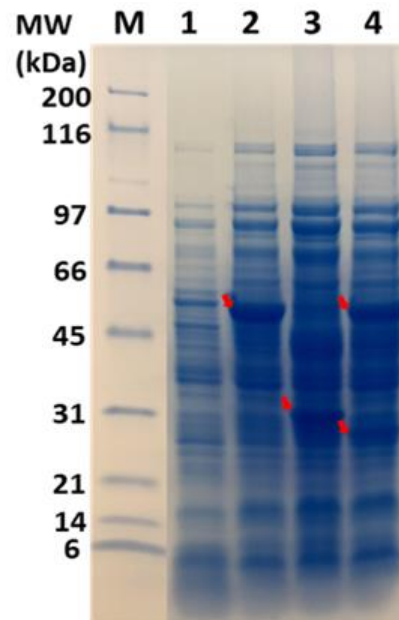


Figure. 4.23 SDS-PAGE to confirm the expression of AroA and PepM in the constructions of *E. coli* strains AW3110 (pET-28a) (lane 1), AW3110 (pET-*aroA*) (lane 2), AW3110 (pET-*pepM*) (lane 3), and AW3110 (pET-*aroA-pepM*) (lane 4), with the induction of IPTG.

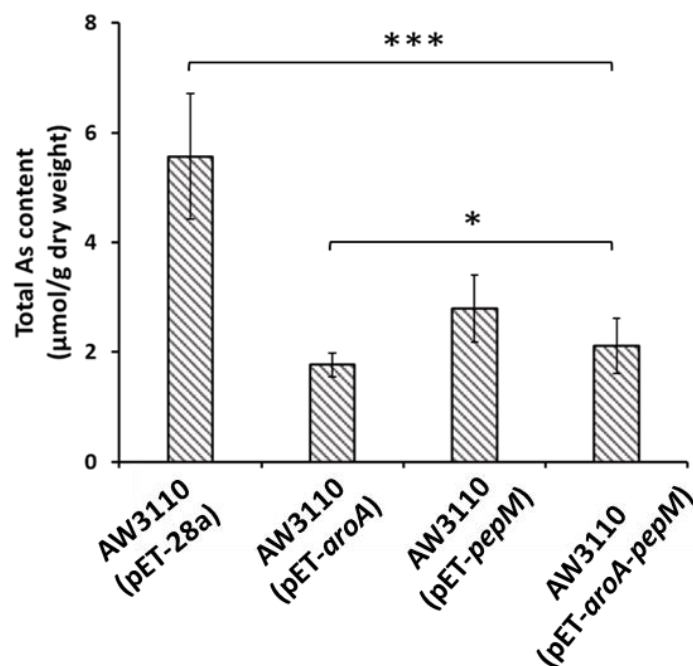


Figure. 4.24 The total As accumulation in the constructions of *E. coli* strains AW3110 (pET-28a), AW3110 (pET-*aroA*), AW3110 (pET-*pepM*), and AW3110 (pET-*aroA-pepM*). The cells were cultured in M9 medium with the addition of 50 μ M iAs^V.

Metabolic Analysis of Extracellular Samples. When compared the metabolites from extracellular samples of the four constructed *E. coli* strains (Fig. 4.25), PCA plot showed the significant separation of metabolites among the *E. coli* constructs, no matter the arsenical stress (Fig. 4.25), indicating the expression of AroA and PepM would change the metabolic pattern in *E. coli* AW3110. All of the 15 metabolites detected from the extracellular samples were listed in Table A3, while the eight significantly changed metabolites were underlined in Table A3.

The separation of the extracellular metabolites among the four *E. coli* constructs with (Fig. 4.26) or without (Fig. 4.27) the addition of iAs^V consistent with the results showed in Fig. 4.25, with

eight metabolites significant changed no matter the arsenical stress, just driven by the expression of AroA and PepM.

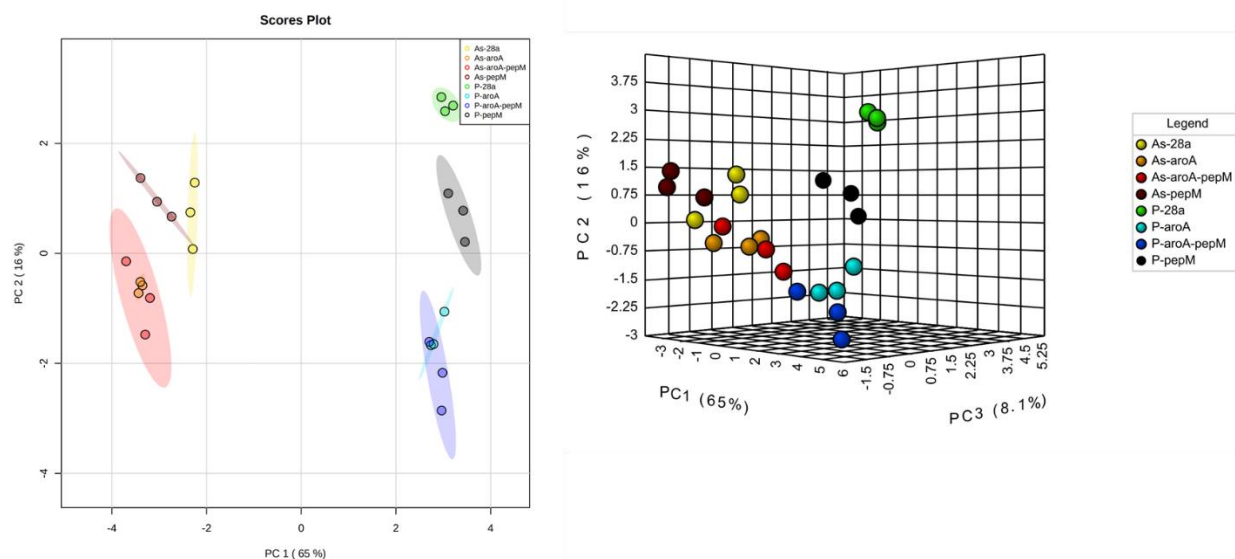


Figure. 4.25 Two-dimensional (2D, left) and three-dimensional (3D, right) multivariate principal component analyses (PCA) of the extracellular metabolic profiles indicating the separation of metabolites in different *E. coli* constructs with or without the addition of iAs^V .

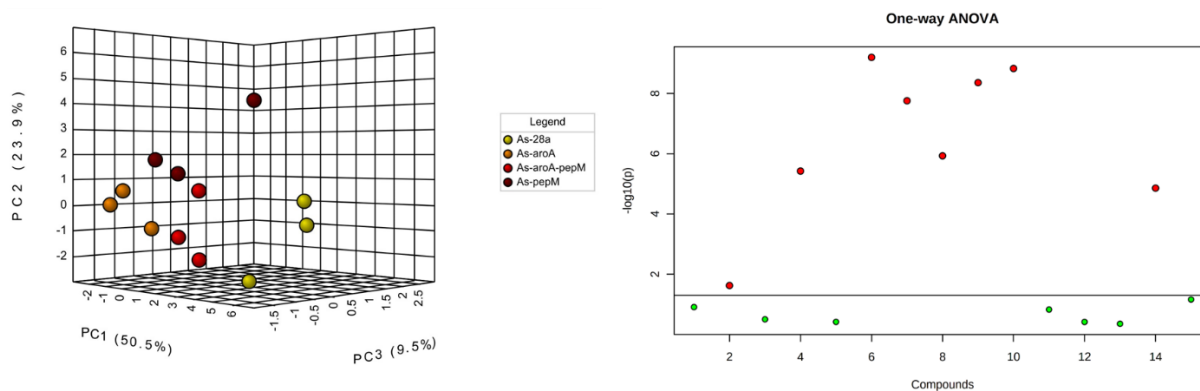


Figure. 4.26 Left: Three-dimensional (3D) multivariate principal component analyses (PCA) of the extracellular metabolic profiles indicating the separation of metabolites in different *E. coli* constructs with the addition of iAs^V . Right: Plot of the log p-values obtained in a one-way ANOVA showing significant (red, $n = 8$) and non-significant (green, $n = 7$) NMR spectral bins (features) across the four *E. coli* constructs.

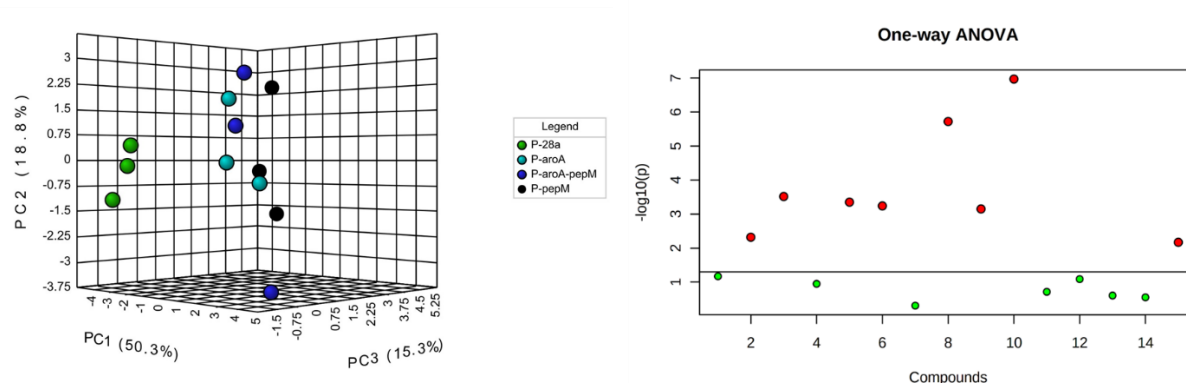


Figure. 4.27 Left: Three-dimensional (3D) multivariate principal component analyses (PCA) of the extracellular metabolic profiles indicating the separation of metabolites in different *E. coli* constructs without the addition of iAs^V . Right: Plot of the log p-values obtained in a one-way ANOVA showing significant (red, $n = 8$) and non-significant (green, $n = 7$) NMR spectral bins (features) across the four *E. coli* constructs.

Metabolic Analysis of Intracellular Sample. When compared the metabolites from intracellular samples of the four constructed *E. coli* strains (Fig. 4.28), PCA plot showed the significant separation of metabolites among the *E. coli* constructs, without the addition of iAs^V (Fig. 4.28, 2D PCA). However, the intracellular metabolites from the four constructs exposed to iAs^V could only show separation when 3D PCA was performed (Fig. 4.28, 3D PCA), indicating the expression of AroA and PepM would change the intracellular metabolites in *E. coli* AW3110 more significantly without the addition of iAs^V . However, with the addition of iAs^V , the expression of AroA and PepM drove the metabolites change to the same direction, the intracellular metabolites clustered either AroA or PepM expressed, indicating that the function of AroA and PepM may be related. All of the 69 metabolites detected from the intracellular samples were listed in Table A4, containing the 55 significantly changed metabolites.

Interestingly, on the PCA plot of the intracellular metabolites from the four constructs cultured with iAs^V (Fig. 4.29) or without iAs^V (Fig. 4.30), the expression of AroA and PepM

showed good separation, revealed that even their function may be related, but still different. However, with the addition of iAs^V , in total of 65 intracellular metabolites were detected, with 51 metabolites significantly changed (Fig. 4.29). While without the addition of iAs^V , in total of 69 intracellular metabolites were detected, with 55 metabolites significantly changed (Fig. 4.30).

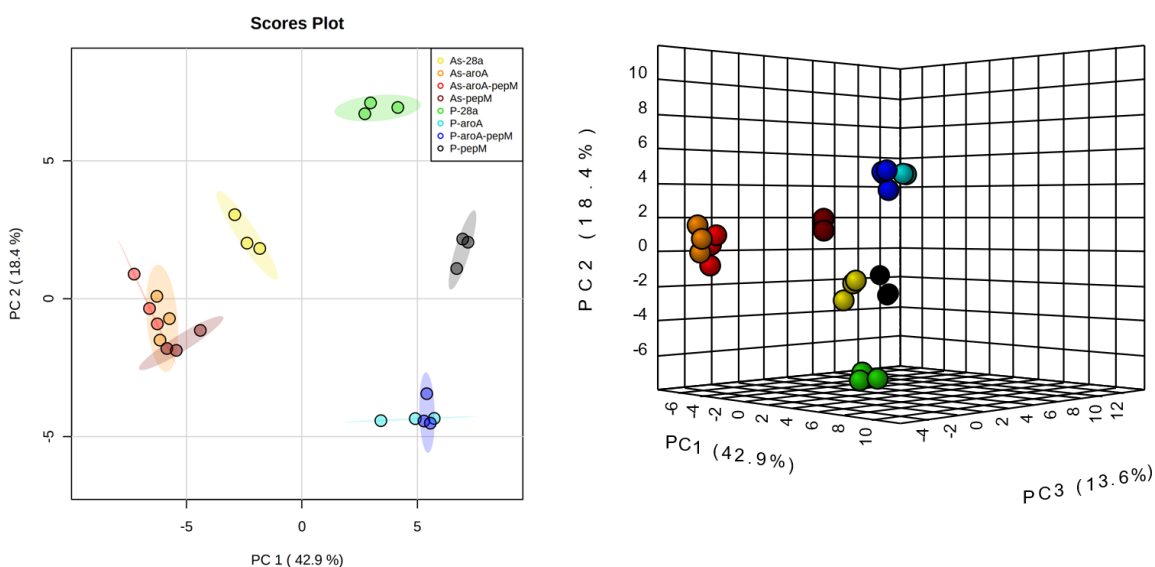


Figure. 4.28 Two-dimensional (2D, left) and three-dimensional (3D, right) multivariate principal component analyses (PCA) of the intracellular metabolic profiles indicating the separation of metabolites in different *E. coli* constructs with or without the addition of iAs^V .

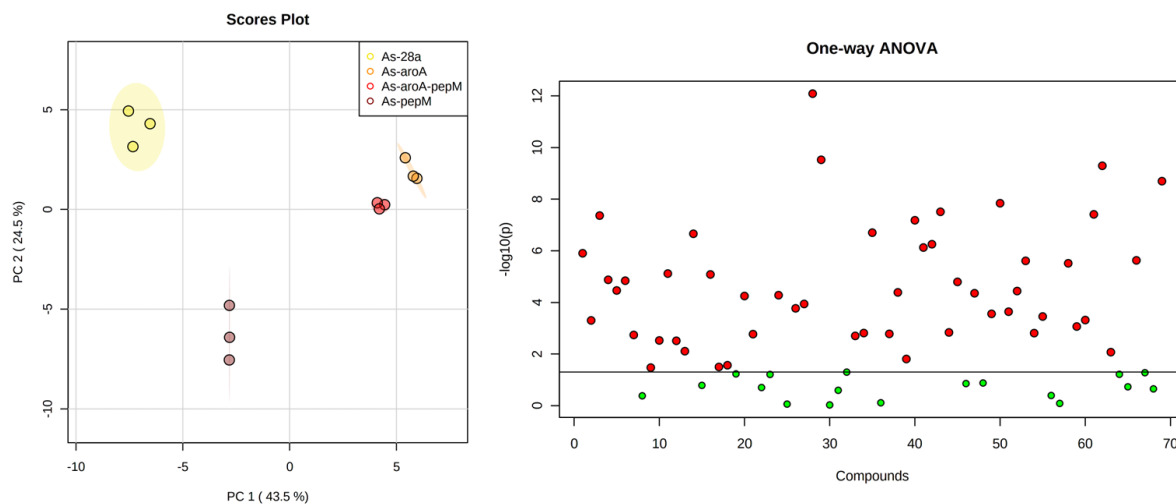


Figure. 4.29 Left: Two-dimensional (2D) multivariate principal component analyses (PCA) of the intracellular metabolic profiles indicating the separation of metabolites in different *E. coli* constructs with the addition of iAs^V. Right: Plot of the log p-values obtained in a one-way ANOVA showing significant (red, n = 51) and non-significant (green, n = 14) NMR spectral bins (features) across the four *E. coli* constructs.

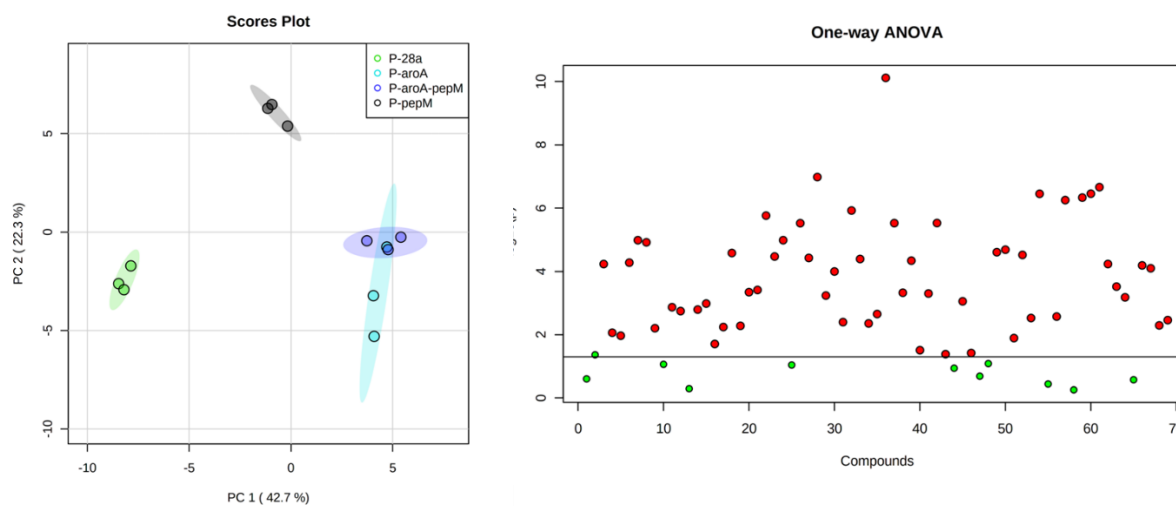


Figure. 4.30 Left: Two-dimensional (2D) multivariate principal component analyses (PCA) of the intracellular metabolic profiles indicating the separation of metabolites in different *E. coli* constructs without the addition of iAs^V. Right: Plot of the log p-values obtained in a one-way ANOVA showing significant (red, n = 58) and non-significant (green, n = 11) NMR spectral bins (features) across the four *E. coli* constructs.

Based on the 51 significantly changed intracellular metabolites with the addition of iAs^V and the 58 significantly changed intracellular metabolites without the addition of iAs^V , there are 8 metabolites showed up only when iAs^V present, and 15 metabolites showed up only when iAs^V absent. Additionally, 43 metabolites significantly changed no matter the arsenical stress, just driven by the expression of AroA and PepM (Fig. 4.31).

Compared to samples without iAs^V , AW3110 (pET-28a) showed high amount of trehalose with the addition of iAs^V , indicating the addition of iAs^V creates stress environment for cell growth (Fig. 4.32), consistent with the growth profile (Fig. 4.22) of the four constructs. Meanwhile, compared to pET-28a, with the expression of AroA, trehalose decreased significantly with iAs^V , however, PepM showed no significant difference to pET-28a, which consistent with the arsenical accumulation pattern (Fig. 4.24) of the four constructs.

Based on these metabolites data, we proposed the possible As-C bond mechanism as shown in the right panel of Fig. 4.33. AroA will catalyze the transformation from EPSP to AEP. If the PepM didn't co-exist, then the AEP will automatic hydrolyze to pyruvate and iAs^V , consistent with the arsenical accumulation pattern and pyruvate pattern in four constructs (Fig. 4.24 and 4.33). However, when PepM expressed, the AEP will transform into arsenopyruvate by PepM, which builds up the As-C bond, and have a chance to get into the whole-map metabolic pathways to generate more arseno-sugars and arsenolipids.

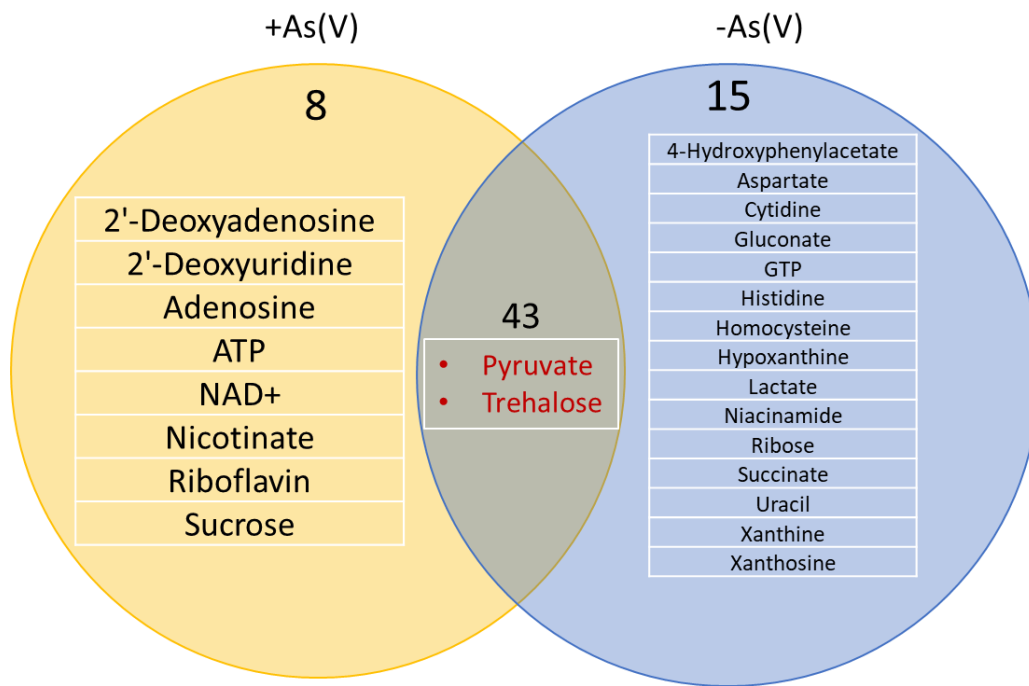


Figure. 4.31 The overlap intracellular metabolites significantly changed in the four *E. coli* constructs with or without the addition of iAs^V.

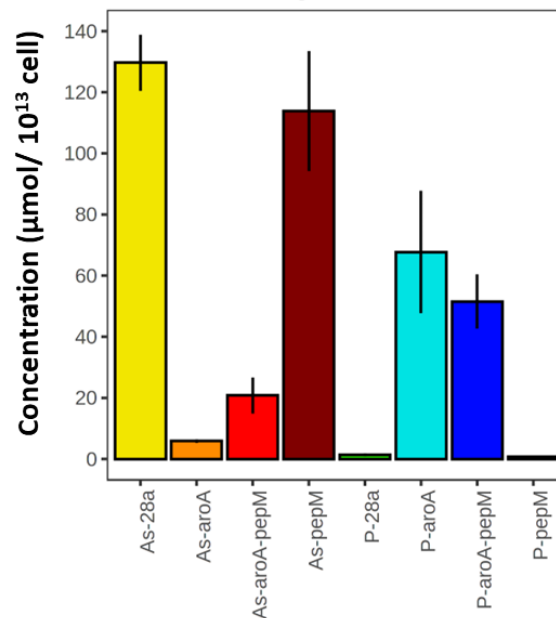


Figure. 4.32 The concentration of trehalose in four different *E. coli* constructions with or without the addition of iAs^V.

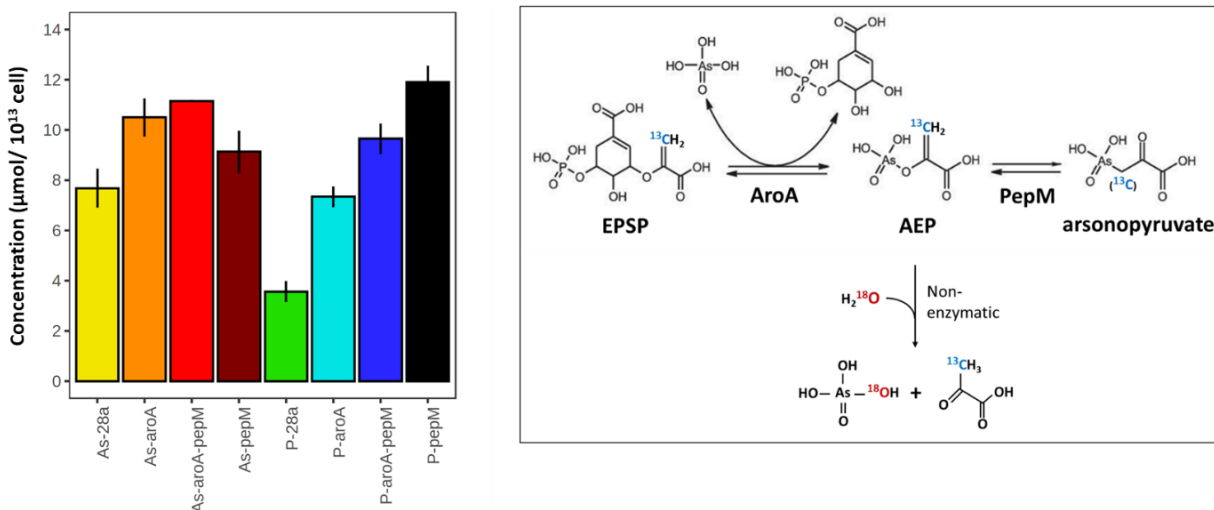


Figure. 4.33 The concentration of pyruvate in four different *E. coli* constructions with or without the addition of iAs^V (Left), and the proposed As-C bond mechanism based on the metabolomics in *A. tumefaciens* 5A (Right).

Conclusion

The Compare Metabolomics Between *E. coli* Strains W3110 and AW3110

Arsenic is classified as a Group-A carcinogen and tops the US EPA Priority List of Hazardous Substances. In nature, its mobility and toxicity are largely governed by bacterial redox transformations that are encoded by arsenic resistance (*ars*) operons which are widespread in both gram-negative and gram-positive bacteria. Generally, the simplest *ars* operon encodes the transcriptional regulator ArsR, the iAs^{III} efflux protein ArsB, and arsenate reductase, ArsC. The recent transcriptomics studies revealed the hierarchical relationship among multiple ArsRs in the genome of a single strain of *Agrobacterium tumefaciens*, showing how ArsR exerts global repression and activation effects on gene regulation that extends well beyond the current *ars* operon model (Rawle *et al.*, 2019). This study used nuclear magnetic resonance (NMR) to examine

metabolome variations. *Escherichia coli* strain W3110 is a wild-type strain which contains the *arsRBC* operon, whereas strain AW3110 is an *arsRBC* deletion mutant, making these strains ideal candidates to study how the *ars* operon influences metabolism under normal growth conditions as well as under arsenic stress. Growth of AW3110 was significantly weaker than strain W3110 regardless of arsenic, indicating that the loss of the *ars* genes results in significant cellular changes. NMR verified this to be the case; multivariate statistical analysis of profiled metabolites reveals:

- i) The driven power of the genetic difference is stronger than arsenical stress on metabolite level changes during growth; ii) Arsenical introduces whole different pattern of metabolism; and iii) The growth differences between *E. coli* strains can be explained by nucleotide, energy, carbohydrate, and amino acid metabolism and oxidative stress. This study provides insights into the metabolic adaptations of *E. coli* strains W3110 and AW3110 following arsenic stress and demonstrates the global effects of *ars* operon on bacterial metabolism.

The Possible As-C Bond Mechanism

Based on the NMR metabolomics study with the extracellular and intracellular samples, the extracellular metabolites did not provide much information for the metabolic pathways, which may be because the early to mid-log phase cells were robustly growing, and intent to keep more metabolites in the bacterial cells for growth. However, based on the intracellular samples, we got dozens of metabolites to work with. The growth profiles and arsenical accumulation patterns changes among the constructions with or without the expression of AroA and/or PepM could be explained by the metabolites' changes. Additionally, the most important metabolite, pyruvate, fits the predict As-C bond mechanism dramatically consistent, indicating the potential of the predicted

mechanism is important for the novel As-C bond generation when the *arsM* (known As-C bond enzyme encoding gene, iAs^{III} methylase) is absent in the bacterial genome.

CHAPTER FIVE

FUTURE DIRECTION

While significant progress has been made in this study to understand the arsenic species in human using murine as model, and to discover the time-course arsenic toxicity in mammals at the early-stage of arsenic exposure. However, there are several aspects that remain unexplored, presenting opportunities for future research.

Arsenobetaine Metabolism in Host and Gut Microbiome

In Chapter II, we evaluated arsenical contamination in commercial mouse chow and detected 200–400 ng of arsenobetaine (AsB) per gram in both standard and germ-free mouse chow, though no AsB was found in purified chow (Fig. 2.6). Interestingly, AsB appeared in the stool, liver, and urinary bladder of mice consuming both standard and purified chow, suggesting its presence in these tissues was independent of the chow type. In contrast, other arsenic species, particularly inorganic arsenic (iAs^{III} and iAs^V), differed significantly depending on the type of chow consumed (Fig. 2.6), suggesting a potential connection among AsB, iAs^{III} , and iAs^V in a shared metabolic pathway.

A key direction for future research will be to investigate AsB metabolism in either the host or the gut microbiome. To eliminate AsB contamination from dietary sources, all mice in this experiment will be fed purified mouse chow. Subsequently, AsB standards will be administered via oral gavage to both conventional (C57BL/6; 1 ng per g body weight) and germ-free mice (C57BL/6; 1 ng per g body weight). After 24 hours of exposure, samples from stool, liver, and

urinary bladder will be collected for analysis. Since no other arsenic species will be introduced in this study, comparing the arsenic species in conventional and germ-free mice should provide insight into AsB metabolism in both the host and gut microbiome. Given the complexities of gut microbiome-host interactions, further research could explore AsB metabolism in humans and the role of the gut microbiome in this process. Findings from such studies could uncover novel AsB metabolic pathways in the host and gut microbiome, opening a new avenue in arsenic toxicity research.

Time-Course Arsenic Diffusion in Human

In Chapter III, we tracked different arsenic species over 12 hours following precise iAs^{III} exposure (Fig. 3.2–3.16). While we uncovered several intriguing findings, more questions remain. It is evident that 12 hours is insufficient for the complete extrusion of arsenic from the host. The first question, therefore, is how much time is required to expel all arsenic or at least 80–90% of it. Additionally, after 6 hours, stool became the primary route of arsenic excretion, surpassing urinary excretion seen in the first 6 hours. This raises the question of which other tissue(s) may contribute to arsenic extrusion during this timeframe.

To address these questions, the experiment should be extended beyond 12 hours, possibly up to 24 hours, with the collection of additional tissues beyond stool, liver, and urinary bladder. With more time points and comprehensive arsenic speciation data across different tissues, we could develop a clearer understanding of arsenic distribution and elimination in the body. This enhanced insight into early-stage arsenic diffusion following acute exposure would aid in designing improved strategies for arsenic detoxification post-exposure.

Brain-Gut Microbiome Axis under Arsenic Exposure

Research indicates that arsenic exposure severely impairs learning, memory, and cognitive functions in animal models, highlighting the brain as a critical target for arsenic-induced toxicity (Lewchalermvong *et al.*, 2018). Both inorganic and organic arsenic species have been observed to accumulate in distinct brain regions (Baker *et al.*, 2018), further emphasizing the neurotoxic potential of arsenic compounds (Chen *et al.*, 2016). The ability of arsenic to cross the blood-brain barrier facilitates its distribution within various brain areas, contributing to neurological impairments and disorders (Uttara *et al.*, 2009). Arsenic exposure induces significant morphological and physiological alterations in neural cells, with the highly metabolic demand in brain rendering it particularly vulnerable to oxidative stress. There is hypothetical theory that arsenic-induced neurotoxicity is largely driven by elevated oxidative stress, which results from the depletion of antioxidant enzyme levels in brain tissue (Breijyeh *et al.*, 2020). However, does increased oxidative stress stand as a primary mechanism through arsenic damage on brain health?

Research on arsenic species within brain tissue remains in its early stages. Leveraging the techniques developed in this study, we are now able to precisely track arsenic accumulation in the brain using controlled arsenic exposure in mice models. This precise detection of arsenic species in brain tissue will significantly enhance our understanding of arsenic neurotoxic effects. By comparing conventional mice with germ-free mice, we aim to elucidate the role of the brain-gut microbiome axis in arsenic toxicity and detoxification, specifically examining its involvement in neurological regulation. This approach promises to shed light on the complex interactions between the microbiome and brain in the context of arsenic exposure.

REFERENCES CITED

- Abhyankar, N., & Szalai, V. (2021). Challenges and Advances in the Application of Dynamic Nuclear Polarization to Liquid-State NMR Spectroscopy. *The Journal of Physical Chemistry B*, 125(20), 5171-5190. doi: 10.1021/acs.jpcc.0c10937.
- Allen, D. K., & Young, J. D. (2020). Tracing Metabolic Flux Through Time and Space With Isotope Labeling Experiments. *Current Opinion in Biotechnology*, 64, 92-100. doi:10.1016/j.copbio.2019.11.003
- Anderson, S., Grist, J. T., Lewis, A., & Tyler, D. J. (2020). Hyperpolarized ^{13}C Magnetic Resonance Imaging for Noninvasive Assessment of Tissue Inflammation. *NMR in Biomedicine*. 34(3), e4460. doi: 10.1002/nbm.4460.
- Andres, J., & Bertin, P. N. (2016). The Microbial Genomics of Arsenic.” *FEMS Microbiology Review*, 40 (2), 299-322.
- Antoniewicz, M. R. (2018). A Guide to ^{13}C Metabolic Flux Analysis for the Cancer Biologist. *Experimental & Molecular Medicine*, 50(4), 1-13. doi: 10.1038/s12276-018-0060-y
- Arnold LL, Suzuki S, Yokohira M, Kakiuchi-Kiyota S, Pennington KL & Cohen SM (2014) Time Course of Urothelial Changes in Rats and Mice Orally Administered Arsenite. *Toxicologic Pathology* 42: 855-862.
- Baker, B. A., Cassano, V. A., & Murray, C. (2018). Arsenic Exposure, Assessment, Toxicity, Diagnosis, and Management: Guidance for Occupational and Environmental Physicians. *Journal of Occupational and Environmental Medicine*, 60, e634–e639. doi: 10.1097/JOM.0000000000001485
- Barai, M., Ahsan, N., Paul, N., Hossain, K., Abdur , R. M, Kato M, Ohgami N & Azim Akhand A (2017). Amelioration of Arsenic-Induced Toxic Effects in Mice by Dietary Supplementation of *Syzygium cumini* Leaf Extract. *Nagoya Journal of Medical Science*, 79, 167-177.
- Bhowmick, S., Pramanik, S., Singh, P., Mondal, P., Chatterjee, D., & Nriagu, J. (2018). Arsenic in groundwater of West Bengal, India: A review of human health risks and assessment of possible intervention options. *The Science of the Total Environment*, 612, 148-169.
- Bowell, R. J., Charles N., Alpers, H. E., Jamieson, D., Kirk, N., & Juraj, M. (2014). The Environmental Geochemistry of Arsenic — An Overview —. *Reviews in Mineralogy and Geochemistry*, 79 (1), 1–16. doi: [10.2138/rmg.2014.79.1](https://doi.org/10.2138/rmg.2014.79.1)
- Breijyeh, Z., & Karaman, R. (2020). Comprehensive Review on Alzheimer’s Disease: Causes and Treatment. *Molecules*, 25, 5789. doi: 10.3390/molecules25245789.
- Burgess, K., Rankin, N., & Weidt, S. (2014). Metabolomics. *Handbook of Pharmacogenomics and Stratified Medicine*, 181-205. doi:10.1016/B978-0-12-386882-4.00010-4

- Burnum-Johnson, K. E., Zheng, X., Dodds, J. N., Ash, J., Fourches, D., Nicora, C. D., Wendler, J. P., Metz, T. O., Waters, K. M., Jansson, J. K., Smith, R. D., & Baker, E. S. (2019). Ion Mobility Spectrometry and the Omics: Distinguishing Isomers, Molecular Classes and Contaminant Ions in Complex Samples. *Trends in Analytical Chemistry*, 116, 292-299. doi: 10.1016/j.trac.2019.04.022.
- Cai, L., Rensing, C., Li, X., & Wang, G. (2009). Novel gene Clusters Involved in Arsenite Oxidation and Resistance in Two Arsenite Oxidizers: *Achromobacter* sp. SY8 and *Pseudomonas* sp. TS44. *Applied Microbiology Biotechnology*, 83, 715-725.
- Carlin A., Shi W., Dey S., & Rosen B. P. (1995). The *ars* Operon of *Escherichia coli* Confers Arsenical and Antimonial Resistance. *Journal of Bacteriology*, 177(4), 981-986.
- Carlin, D. J., Naujokas, M. F., Bradham, K. D., Cowden, J., Heacock, M., Henry, H. F., Lee, J. S., Thomas, D. J., Thompson, C., Tokar, E. J., Waalkes, M. P., Birnbaum, L. S., & Suk, W. A. (2016). Arsenic and Environmental Health: State of the Science and Future Research Opportunities. *Environmental health*, 124(7), 890-899. doi: 10.1289/ehp.1510209.
- Chakraborti, D., Das, B., Rahman, M. M., Chowdhury, U. K., Biswas, B., Goswami, A. B., Nayak, B., Pal, A., Sengupta, M. K., Ahamed, S., Hossain, A., Basu, G., Roychowdhury, T., & Das, D. (2009). Status of Groundwater Arsenic Contamination in The State of West Bengal, India: A 20-year Study Report. *Molecular Nutrition & Food Research*, 53(5), 542-551. doi: 10.1002/mnfr.200700517.
- Challenger, F. (1951) Biological Methylation. *Advances in Enzymology and Related Areas of Molecular Biology*, 12, 429-491. doi: 10.1002/9780470122570.ch8
- Chaudhari, S. R., Wisser, D., Pinon, A. C., Berruyer, P., Gajan, D., Tordo, P., Ouari, O., Reiter, C., Engelke, F., Copéret, C., Lelli, M., Lesage, A., & Emsley, L. (2017). Dynamic Nuclear Polarization Efficiency Increased by Very Fast Magic Angle Spinning. *Journal of the American Chemical*, 139(31), 10609-10612. doi: 10.1021/jacs.7b05194. Society 139: 10609-10612.
- Chen, F., Cao, Y., Wei, S., Li, Y., Li, X., Wang, Q., & Wang, G. (2015a). Regulation of Arsenite Oxidation by the Phosphate Two-Component System PhoBR in *Halomonas* sp. HAL1. *Frontiers in microbiology*, 6, 923. doi: 10.3389/fmicb.2015.00923.
- Chen, J., Bhattacharjee, H., & Rosen, B. P. (2015b). ArsH is An Organoarsenical Oxidase that Confers Resistance to Trivalent Forms of The Herbicide Monosodium Methylarsenate and the Poultry Growth Promoter Roxarsone. *Molecular microbiology*, 96(5), 1042-1052. doi: 10.1111/mmi.12988.
- Chen, J., Madegowda, M., Bhattacharjee, H., & Rosen, B. P. (2015c). ArsP: A Methylarsenite Rfflux Permease. *Molecular microbiology*, 98(4), 625-635. doi: 10.1111/mmi.13145.

- Chen, J., Li, J., Jiang, X., & Rosen, B. P. (2017). Conserved Cysteine Residues Determine Substrate Specificity in A Novel As(III) S-adenosylmethionine Methyltransferase from *Aspergillus fumigatus*. *Molecular microbiology*, 104(2), 250-259. doi: 10.1111/mmi.13628.
- Chen, P., Miah, M. R., & Aschner, M. (2016). Metals and Neurodegeneration. *F1000Research*, 5, 366. doi: 10.12688/f1000research.7431.1.
- Chi, L., Xue, J., Tu, P., Lai, Y., Ru, H., & Lu, K. (2019). Gut Microbiome Disruption Altered the Biotransformation and Liver Toxicity of Arsenic in Mice. *Archives of Toxicology*, 93(1), 25-35. doi: 10.1007/s00204-018-2332-7.
- Chin, L., Kumana, C. R., Kwong, Y. L., & Gill, H. (2022). The Development and Clinical Applications of Oral Arsenic Trioxide for Acute Promyelocytic Leukaemia and Other Diseases. *Pharmaceutics*, 14(9), 1945. doi: 10.3390/pharmaceutics14091945.
- Cohen, S. M., Arnold, L. L., Beck, B. D., Lewis, A. S. & Eldan, M. (2013). Evaluation of the Carcinogenicity of Inorganic Arsenic. *Critical Reviews in Toxicology*, 43 (9), 711-752. doi: 10.3109/10408444.2013.827152.
- Coryell, M., McAlpine, M., Pinkham, N. V., McDermott, T. R. & Walk, S. T. (2018). The Gut Microbiome is Required for Full Protection Against Acute Arsenic Toxicity in Mouse Models. *Nature Communications*, 9, 5424. doi: 10.1038/s41467-018-07803-9.
- Coryell, M., Roggenbeck, B. A., & Walk, S. T. (2019). The Human Gut Microbiome's Influence on Arsenic Toxicity. *Current Pharmacology Reports*, 5(6), 491-504. doi: 10.1007/s40495-019-00206-4.
- Crook, A. A., & Powers, R. (2020). Quantitative NMR-Based Biomedical Metabolomics: Current Status and Applications. *Molecules*, 25(21), 5128. doi: 10.3390/molecules25215128.
- Cruz-Morales, P., Kopp, J. F., Martínez-Guerrero, C., Yáñez-Guerra, L. A., Selem-Mojica, N., Ramos-Aboites, H., Feldmann, J., & Barona-Gómez, F. (2016). Phylogenomic Analysis of Natural Products Biosynthetic Gene Clusters Allows Discovery of Arseno-Organic Metabolites in Model *Streptomyces*. *Genome Biology and Evolution*, 8(6), 1906-1916. doi: 10.1093/gbe/evw125.
- Cullen, W. R. (2014). Chemical Mechanism of Arsenic Biomethylation. *Chemical Research in Toxicology Journal*, 27(4), 457-461. doi: 10.1021/tx400441h. PMID: 24517124.
- Danzi, F., Pacchiana, R., Maffiini, A., Scupoli, M. T., Scarpa, A., Donadelli, M., & Fiore, A. (2023). To Metabolomics and Beyond: A Technological Portfolio to Investigate Cancer Metabolism. *Signal Transduction and Targeted Therapy*, 8(1), 1-22. doi:10.1038/s41392-023-01380-0.

- de Souza, A. C. M., de Almeida, M. G., Pestana, I. A., & de Souza, C. M. M. (2019). Arsenic Exposure and Effects in Humans: A Mini-Review in Brazil. *Archives of Environmental Contamination and Toxicology*, 76(3), 357-365. doi: 10.1007/s00244-018-00586-6. PMID: 30569210.
- Dembitsky, V. & Levitsky, D. (2004). Arsenolipids. *Progress in Lipid Research*, 43(5), 403-448. doi: 10.1016/j.plipres.2004.07.001. PMID: 15458814.
- Dettmer, K., Aronov, P. A., & Hammock, B. D. (2007) Mass Spectrometry-Based Metabolomics. *Mass Spectrometry Reviews*, 26(1), 51-78. doi: 10.1002/mas.20108. PMID: 16921475.
- Dona, A. C., Kyriakides, M., Scott, F., Shephard, E. A., Varshavi, D., Veselkov, K., & Everett, J. R. (2016). A Guide to the Identification of Metabolites in NMR-based Metabonomics/Metabolomics Experiments. *Computational and Structural Biotechnology Journal*, 14, 135-153. doi: 10.1016/j.csbj.2016.02.005.
- Dodds, J. N., & Baker, E. S. (2019). Ion Mobility Spectrometry: Fundamental Concepts, Instrumentation, Applications, and the Road Ahead. *Journal of the American Society for Mass Spectrometry*, 30(11), 2185-2195. doi: 10.1007/s13361-019-02288-2. PMID: 31493234; PMCID: PMC6832852.
- Dong, O., Powers, M., Liu, Z. & Yoshinaga, M. (2022). Arsenic Metabolism, Toxicity and Accumulation in the White Button Mushroom *Agaricus bisporus*. *Toxics*, 10(10), 554. doi: 10.3390/toxics10100554. PMID: 36287835; PMCID: PMC9609160.
- Douillet, C., Miller, M., Cable, P. H., Shi, Q., El-Masri, H., Matousek, T., Koller, B. H., Thomas, D. J. & Styblo, M. (2023). Fate of Arsenicals in Mice Carrying the Human AS3MT Gene Exposed to Environmentally Relevant Levels of Arsenite in Drinking Water. *Scientific Reports*, 13(1), 3660. doi: 10.1038/s41598-023-30723-8. PMID: 36871058; PMCID: PMC9985638.
- Drobná, Z., Waters, S. B., Devesa, V., Harmon, A. W., Thomas, D. J., & Stýblo, M. (2005). Metabolism and Toxicity of Arsenic in Human Urothelial Cells expressing Rat Arsenic (+3 oxidation state)-Methyltransferase. *Toxicology and Applied Pharmacology*, 207(2), 147-159. doi: 10.1016/j.taap.2004.12.007. PMID: 16102566; PMCID: PMC2366102.
- Ebbels, T. M. D., van der Hooft, J. J. J., Chatelaine, H., Broeckling, C., Zamboni, N., Hassoun, S., & Mathé, E.A. (2023). Recent advances in mass spectrometry-based computational metabolomics. *Current Opinion in Chemical Biology*, 74, 102288. doi: 10.1016/j.cbpa.2023.102288. PMID: 36966702; PMCID: PMC11075003.
- Emwas, A. H., Roy, R., McKay, R. T., Tenori, L., Saccenti, E., Gowda, G. A. N., Raftery, D., Alahmari, F., Jaremko, L., Jaremko, M., & Wishart, D. S. (2019). NMR Spectroscopy for Metabolomics Research. *Metabolites*. 9(7), 123. doi: 10.3390/metabo9070123. PMID: 31252628; PMCID: PMC6680826.

- Faita, F., Cori, L., Bianchi, F. & Andreassi, M. G. (2013). Arsenic-Induced Genotoxicity and Genetic Susceptibility to Arsenic-Related Pathologies. *International Journal of Environmental Research and Public Health*, 10(4), 1527-46. doi: 10.3390/ijerph10041527. PMID: 23583964; PMCID: PMC3709332.
- Fiehn, O. (2008). Extending the Breadth of Metabolite Profiling by Gas Chromatography Coupled to Mass Spectrometry. *TrAC Trends in Analytical Chemistry*, 27(3), 261-269. doi:10.1016/j.trac.2008.01.007.
- Fiehn, O. (2016). Metabolomics by Gas Chromatography-Mass Spectrometry: Combined Targeted and Untargeted Profiling. *Current Protocols in Molecular Biology*, 114, 30.4.1-30.4.32. doi: 10.1002/0471142727.mb3004s114. PMID: 27038389; PMCID: PMC4829120.
- Furey, A., Moriarty, M., Bane, V., Kinsella, B., & Lehane, M. (2013). Ion suppression; A critical review on causes, evaluation, prevention and applications. *Talanta*, 115, 104-122. doi: 10.1016/j.talanta.2013.03.048. Epub 2013 Apr 16. PMID: 24054567.
- Gokulan, K., Arnold, M. G., Jensen, J., Vanlandingham, M., Twaddle, N. C., Doerge, D. R., Cerniglia, C. E., & Khare, S (2018). Exposure to Arsenite in CD-1 Mice during Juvenile and Adult Stages: Effects on Intestinal Microbiota and Gut-Associated Immune Status. *mBio*, 9(4), e01418-18. doi: 10.1128/mBio.01418-18. PMID: 30108172; PMCID: PMC6094480.
- González-Martínez, F., Johnson-Restrepo, B., & Quiñones, L. A. (2024). Arsenic inorganic exposure, metabolism, genetic biomarkers and its impact on human health: A mini-review. *Toxicology Letters*, 398, 105-117. doi:10.1016/j.toxlet.2024.06.008.
- Gorrochategui, E., Jaumot, J., Lacorte, S., & Tauler, R. (2016). Data analysis strategies for targeted and untargeted LC-MS metabolomic studies: Overview and workflow. *TrAC Trends in Analytical Chemistry*, 82, 425-442. doi:10.1016/j.trac.2016.07.004
- Gregus, Z. & Nemeti, B. (2005). The Glycolytic Enzyme Glyceraldehyde-3-phosphate Dehydrogenase Works as An Arsenate Reductase in Human Red Blood Cells and Rat Liver Cytosol. *Toxicological Sciences*, 85(2), 859-69. doi: 10.1093/toxsci/kfi158. PMID: 15788719.
- Hajjar, G., Barros Santos, M. C., Bertrand-Michel, J., Canlet, C., Castelli, F., Creusot, N., Dechaumet, S., Diémé, B., Giacomoni, F., Giraudeau, P., Guillon, Y., Thévenot, E., Tremblay-Franco, M., Junot, C., Jourdan, F., Fenaille, F., Comte, B., Pétriaccq, P., & Pujos-Guillot, E. (2023). Scaling-up metabolomics: Current state and perspectives. *TrAC Trends in Analytical Chemistry*, 167, 117225. doi:10.1016/j.trac.2023.117225.
- Hayakawa, T., Kobayashi, Y., Cui, X., & Hirano, S. (2005). A New Metabolic Pathway of Arsenite: Arsenic-Glutathione Complexes Are substrates for Human Arsenic Methyltransferase

- Cyt19. Archives of Toxicology, 79(4), 183-191. doi: 10.1007/s00204-004-0620-x. Epub 2004 Nov 4. PMID: 15526190.
- Hong, Y. S., Song, K. H., & Chung, J. Y. (2014). Health Effects of Chronic Arsenic Exposure. Journal of Preventive Medicine and Public Health, 47(5), 245-52. doi: 10.3961/jpmph.14.035. PMID: 25284195; PMCID: PMC4186552.
- Hu, K. (2011). Polarizing Agents and Mechanisms for High-Field Dynamic Nuclear Polarization of Frozen Dielectric Solids. Solid State Nuclear Magnetic Resonance, 40(2), 31-41. doi:10.1016/j.ssnmr.2011.08.001.
- Huang, K., Xu, Y., Packianathan, C., Gao, F., Chen, C., Zhang, J., Shen, Q., Rosen, B. P., & Zhao, F. J. (2018). Arsenic Methylation by A Novel ArsM As(III) S-adenosylmethionine Methyltransferase that Requires Only Two Conserved Cysteine Residues. Molecular microbiology, 107(2), 265-276. doi: 10.1111/mmi.13882. PMID: 29134708; PMCID: PMC5760297.
- Hwang, J. H., & Choi, C. S. (2015). Use of In Vivo Magnetic Resonance Spectroscopy for Studying Metabolic Diseases. Experimental & Molecular Medicine, 47(2), e139. doi: 10.1038/emm.2014.101. PMID: 25656949; PMCID: PMC4346484.
- Isokpehi, R. D., Udensi, U. K., Simmons, S. S., Hollman, A. L., Cain, A. E., Olofinsae, S. A., Hassan, O. A., Kashim, Z. A., Enejoh, O. A., Fasesan, D. E., & Nashiru, O. (2014). Evaluative Profiling of Arsenic Sensing and Regulatory Systems in The Human Microbiome Project Genomes. Microbiology Insights, 7, 25-34. doi: 10.4137/MBI.S18076. PMID: 25452698; PMCID: PMC4230230.
- Jain, C., & Ali, I. (2000). Arsenic: Occurrence, Toxicity and Speciation Techniques. Water Research, 34(17), 4304-4312. doi:10.1016/S0043-1354(00)00182-2.
- Kashyap, D. R., Botero, L. M., Franck, W. L., Hassett, D. J., & McDermott, T. R. (2006). Complex Regulation of Arsenite Oxidation in *Agrobacterium tumefaciens*. Journal of Bacteriology, 188(3), 1081-1088. doi: 10.1128/JB.188.3.1081-1088.2006. PMID: 16428412; PMCID: PMC1347330.
- Keum, Y. S., Kim, J., & Li, Q. X. (2010). Metabolomics in Pesticide Toxicology. Hayes' Handbook of Pesticide Toxicology (Third Edition), 627-643. doi: 10.1016/B978-0-12-374367-1.00022-7.
- Khairul, I., Wang, Q. Q., Jiang, Y. H., Wang, C., & Naranmandura, H. (2017). Metabolism, Toxicity and Anticancer Activities of Arsenic Compounds. Oncotarget, 8, 23905–23926. doi: 10.18632/oncotarget.14733.

- Kim, S. J., Kim, S. H., Kim, J. H., Hwang, S., & Yoo, H. J. (2016). Understanding Metabolomics in Biomedical Research. *Endocrinology and Metabolism*, 31(1), 7-16. doi: 10.3803/EnM.2016.31.1.7. PMID: 26676338; PMCID: PMC4803564.
- Krafft, T. & Macy, J. M. (1998). Purification and characterization of the respiratory arsenate reductase of *Chrysiogenes arsenatis*." *European journal of biochemistry*, 255(3), 647-653. doi: 10.1046/j.1432-1327.1998.2550647.x. PMID: 9738904..
- Lageveen-Kammeijer, G. S. M., de Haan, N., Mohaupt, P., Wagt, S., Filius, M., Nouta, J., Falck, D., & Wuhrer, M. (2019). Highly Sensitive CE-ESI-MS Analysis of N-glycans From Complex Biological Samples. *Nature Communications*, 10(1), 2137. doi: 10.1038/s41467-019-09910-7. PMID: 31086181; PMCID: PMC6513864.
- Laue, H. E., Moroishi, Y., Palys, T. J., Jackson, B. P., Madan, J. C., & Karagas, M. R. (2022). Contribution of Gut Bacteria to Arsenic Metabolism in the First Year of Life in A Prospective Birth Cohort. *Environmental Research*, 214(Pt 4), 114099. doi: 10.1016/j.envres.2022.114099. PMID: 35998698; PMCID: PMC10319341.
- Letertre, M. P. M., Giraudeau, P., & de Tullio, P. (2021). Nuclear Magnetic Resonance Spectroscopy in Clinical Metabolomics and Personalized Medicine: Current Challenges and Perspectives. *Frontiers in Molecular Biosciences*, 8, 698337. doi: 10.3389/fmolb.2021.698337. PMID: 34616770; PMCID: PMC8488110.
- Lewchalermvong, K., Rangkadilok, N., Nookabkaew, S., Suriyo, T., & Satayavivad, J. (2018). Arsenic Speciation and Accumulation in Selected Organs after Oral Administration of Rice Extracts in Wistar Rats. *Journal of Agricultural and Food Chemistry*, 66, 3199–3209. doi: 10.1021/acs.jafc.7b05746.
- Li, H., Li, M., Huang, Y., Rensing, C., & Wang, G. (2013). In Silico Analysis of Bacterial Arsenic Islands Reveals Remarkable Synteny and Functional Relatedness Between Arsenate and Phosphate. *Frontiers in Microbiology*, 4, 347. doi: 10.3389/fmicb.2013.00347. PMID: 24312089; PMCID: PMC3834237.
- Li, J., Dong, F., Lu, Y., Yan, Q., & Shim, H. (2014). Mechanisms Controlling Arsenic Uptake in Rice Grown in Mining Impacted Regions in South China. *PLoS One*, 9(9), e108300. doi: 10.1371/journal.pone.0108300. PMID: 25251438; PMCID: PMC4177218.
- Lima, C., Muhamadali, H. & Goodacre, R. (2021). The role of Raman Spectroscopy Within Quantitative Metabolomics. *Annual Review of Analytical Chemistry*, 14(1), 323-345. doi: 10.1146/annurev-anchem-091420-092323. PMID: 33826853.
- Lipfert, M., Rout, M. K., Berjanskii, M., & Wishart, D. S. (2019). Automated Tools for the Analysis of 1D-NMR and 2D-NMR Spectra. In: Gowda, G., Raftery, D. (eds) *NMR-Based Metabolomics. Methods in Molecular Biology*, 2037. Humana, New York, NY. doi: 10.1007/978-1-4939-9690-2_24.

- Liu, G., Liu, M., Kim, E. H., Maaty, W. S., Bothner, B., Lei, B., Rensing, C., Wang, G., & McDermott, T. R. (2012). A Periplasmic Arsenite-Binding Protein Involved in Regulating Arsenite Oxidation. *Environmental Microbiology*, 14(7), 1624-1634. doi: 10.1111/j.1462-2920.2011.02672.x. Epub 2011 Dec 19. PMID: 22176720.
- Liu, Y., Gu, Y., & Yu, X. (2017). Assessing Tissue Metabolism by Phosphorous-31 Magnetic Resonance Spectroscopy and Imaging: A Methodology Review. *Quantitative Imaging in Medicine and Surgery*, 7(6), 707-726. doi: 10.21037/qims.2017.11.03. PMID: 29312876; PMCID: PMC5756783.
- Liu, B., Sui, J., Feng, R., Lin, H., Han, X., Sun, X. & Cao, L. (2023). Transformation of Arsenic Species from Seafood Consumption During in Vitro Digestion. *Frontiers in Nutrition*, 10, 1207732. doi: 10.3389/fnut.2023.1207732. PMID: 37899842; PMCID: PMC10602890.
- Liu, Q., Peng, H., Lu, X. & Le, X. C. (2015). Enzyme-assisted Extraction and Liquid Chromatography Mass Spectrometry for the Determination of Arsenic Species in Chicken Meat. *Analytica Chimica Acta*, 888, 1-9. doi: 10.1016/j.aca.2015.05.001. Epub 2015 May 7. PMID: 26320952.
- Lu, K., Cable, P. H., Abo, R. P., Ru, H., Graffam, M. E., Schlieper, K. A., Parry, N. M., Levine, S., Bodnar, W. M., Wishnok, J. S., Styblo, M., Swenberg, J. A., Fox, J. G., & Tannenbaum, S. (2013). Gut Microbiome Perturbations Induced by Bacterial Infection Affect Arsenic Biotransformation. *Chemical Research in Toxicology Journal*, 26(12), 1893-1903. doi: 10.1021/tx4002868. PMID: 24134150; PMCID: PMC3974266.
- Lu, K., Abo, R. P., Schlieper, K. A., Graffam, M. E., Levine, S., Wishnok, J. S., Swenberg, J. A., Tannenbaum, S. R., & Fox, J. G. (2014a). Arsenic exposure perturbs the gut microbiome and its metabolic profile in mice: an integrated metagenomics and metabolomics analysis. *Environmental Health Perspectives*, 122(3), 284-291. doi: 10.1289/ehp.1307429. PMID: 24413286; PMCID: PMC3948040.
- Lu, K., Mahbub, R., Cable, P. H., Ru, H., Parry, N. M., Bodnar, W. M., Wishnok, J. S., Styblo, M., Swenberg, J. A., Fox, J. G., & Tannenbaum, S. R. (2014b). Gut Microbiome Phenotypes Driven by Host Genetics Affect Arsenic Metabolism. *Chemical Research in Toxicology*, 27(2), 172-4. doi: 10.1021/tx400454z. PMID: 24490651; PMCID: PMC3997221.
- Mahrous, E. A., & Farag, M. A. (2015). Two dimensional NMR Spectroscopic Approaches for Exploring Plant Metabolome: A review. *Journal of Advanced Research*, 6(1), 3-15. doi: 10.1016/j.jare.2014.10.003.
- Maiti, K. S. (2023). Non-Invasive Disease Specific Biomarker Detection Using Infrared Spectroscopy: A Review. *Molecules*. 28(5), 2320. doi: 10.3390/molecules28052320. PMID: 36903576; PMCID: PMC10005715.

- Marapakala K, Packianathan C, Ajees AA, Dheeman DS, Sankaran B, Kandavelu P, Rosen BP. (2015). A Disulfide-Bond Cascade Mechanism for Arsenic(III) S-adenosylmethionine Methyltransferase.” *Acta Crystallographica, Section D, Biological Crystallography*, 71(Pt 3), 505-15. doi: 10.1107/S1399004714027552. PMID: 25760600; PMCID: PMC4356363.
- Marchand, J., Martineau, E., Guitton, Y., Dervilly-Pinel, G., & Giraudeau, P. (2017). Multidimensional NMR Approaches Towards Highly Resolved, Sensitive and High-throughput quantitative metabolomics. *Current Opinion in Biotechnology*, 43, 49-55. doi: 10.1016/j.copbio.2016.08.004.
- Marshall, D. D., & Powers, R. (2017). Beyond the Paradigm: Combining Mass Spectrometry and Nuclear Magnetic Resonance for Metabolomics. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 100, 1-16. doi: 10.1016/j.pnmrs.2017.01.001. PMID: 28552170; PMCID: PMC5448308.
- Men, X., Wu, C. X., Zhang, X., Wei, X., Ye, W. Q., Chen, M. L., Yang, T., Xu, Z. R. & Wang, J. H. (2022). Tracking Cellular Transformation of iAs^{III} in HepG2 Cells by Single-Cell Focusing/Capillary Electrophoresis Coupled to ICP-MS. *Analytica Chimica Acta*, 1226, 340268. doi: 10.1016/j.aca.2022.340268. PMID: 36068064.
- Menzildjian, G., Schlagnitweit, J., Casano, G., Ouari, O., Gajan, D., & Lesage, A. (2023). Polarizing Agents for Efficient High Field DNP Solid-State NMR Spectroscopy Under Magic-Angle Spinning: From Design Principles to Formulation Strategies. *Chemical Science*, 14(23), 6120. doi: 10.1039/d3sc01079a.
- Mersaoui, S. Y., Guilbert, C., Chou, H., Douillet, C., Bohle, D. S., Stýblo, M., Richard, S., & Mann, K. K. (2022). Arsenic 3 Methyltransferase (AS3MT) Automethylates on Cysteine Residues in Vitro. *Archives of Toxicology*, 96(5), 1371-1386. doi: 10.1007/s00204-022-03248-8.
- Mcdermott, T. R., Stolz, J. F., & Oremland, R. S., (2020). Arsenic and the Gastrointestinal Tract Microbiome. *Environmental Microbiology Reports*, 12, 136-159. doi:10.1111/1758-2229.12814.
- Milanesi, R., Coccetti, P., & Tripodi, F. (2020). The Regulatory Role of Key Metabolites in the Control of Cell Signaling. *Biomolecules*, 10(6), 862. doi:10.3390/biom10060862.
- Minatel, B. C., Sage, A. P., Anderson, C., Hubaux, R., Marshall, E. A., Lam, W. L., & Martinez, V. D. (2017). Environmental Arsenic Exposure: From Genetic Susceptibility to Pathogenesis. *Environment international*, 112, 183-197. doi: 10.1016/j.envint.2017.12.017. PMID: 29275244.
- Moco, S. (2022). Studying Metabolism by NMR-Based Metabolomics. *Frontiers in Molecular Biosciences*, 9, 882487. doi: 10.3389/fmolb.2022.882487. PMID: 35573745; PMCID: PMC9094115.

- Mukhopadhyay, R. & Rosen, B. P. (2002). Arsenate Reductases in Prokaryotes and Eukaryotes. *Environmental health perspectives*, 110(5), 745-748. doi: 10.1289/ehp.02110s5745. PMID: 12426124; PMCID: PMC1241237.
- Murko, M., Elek, B., Styblo, M., Thomas, D. J. & Francesconi, K. A. (2018). Dose and Diet - Sources of Arsenic Intake in Mouse in Utero Exposure Scenarios. *Chemical Research in Toxicology Journal*, 31 (2), 156-164. doi: 10.1021/acs.chemrestox.7b00309. PMID: 29244955; PMCID: PMC6611170.
- Nagana Gowda, G. A., & Raftery, D. (2021). NMR-Based Metabolomics. *Advances in Experimental Medicine and Biology*, 1280, 19-37. doi: 10.1007/978-3-030-51652-9_2. PMID: 33791972; PMCID: PMC8816450.
- Nagana Gowda, G. A., & Raftery, D. (2023). Quantitative NMR Methods in Metabolomics. *Handbook of Experimental Pharmacology*, 277, 143-164. doi: 10.1007/164_2022_612. PMID: 36308549; PMCID: PMC9992124.
- Nemeti, B., Csanaky, I. & Gregus, Z. (2006). Effect of An Inactivator of Glyceraldehyde-3-Phosphate Dehydrogenase, A Fortuitous Arsenate Reductase, on Disposition of Arsenate in Rats. *Toxicological Sciences*, 90, 49-60. doi: 10.1093/toxsci/kfj058. Epub 2005 Dec 1. PMID: 16322075.
- Oremland, R. S., and Stolz, J. F. (2003). The Ecology of Arsenic. *Science* (New York, N.Y.) vol. 300(5621), 939-44. doi: 10.1126/science.1081903. PMID: 12738852.
- Organization, W. H, (2022) Guidelines for Drinking-Water Quality: Incorporating the First and Second addenda. World Health Organization.
- Pace, C., Smith-Gagen, J., & Angermann, J. (2018). Arsenic Methylation Capacity and Metabolic Syndrome in the 2013-2014 U.S. National Health and Nutrition Examination Survey (NHANES). *International Journal of Environmental Research and Public Health*, 15(1), 168. doi: 10.3390/ijerph15010168. PMID: 29361794; PMCID: PMC5800267.
- Patel, M. K., Pandey, S., Kumar, M., Haque, M. I., Pal, S., & Yadav, N. S. (2021). Plants Metabolome Study: Emerging Tools and Techniques. *Plants* (Basel), 10(11), 2409. doi: 10.3390/plants10112409. PMID: 34834772; PMCID: PMC8621461.
- Peng, H., Hu, B., Liu, Q., Yang, Z., Lu, X., Huang, R., Li, X. F., Zuidhof, M. J. & Le, X. (2014). Liquid Chromatography Combined with Atomic and Molecular Mass Spectrometry for Speciation of Arsenic in Chicken Liver. *Journal of Chromatography A*, 1370: 40-49. doi: 10.1016/j.chroma.2014.10.012. PMID: 25459646.
- Potera, C. (2014). Fire in the Belly? Sulfur-Reducing Gut Microbes Fuel Arsenic Thiolation. *Environmental Health Perspectives*, 122(8), A222. doi: 10.1289/ehp.122-A222. PMID: 25229084; PMCID: PMC4123018.

- Qiao, Z., Huang, J., Cao, Y., Shi, K., & Wang, G. (2019). Genetics and Proteomics Analyses Reveal the Roles of PhoB1 and PhoB2 Regulators in Bacterial Responses to Arsenite and Phosphate.” *Research in Microbiology*, 170 (6-7), 263-271. doi: 10.1016/j.resmic.2019.06.006. Epub 2019 Jul 3. PMID: 31279088.
- Qin, J., Rosen, B. P., Zhang, Y., Wang, G., Franke, S., & Rensing, C. (2006). Arsenic Detoxification and Evolution of Trimethylarsine Gas by A Microbial Arsenite S-adenosylmethionine Methyltransferase. *Proceedings of the National Academy of Sciences of the United States of America*, 103(7), 2075-80. doi: 10.1073/pnas.0506836103. PMID: 16452170; PMCID: PMC1413689..
- Qiu, S., Cai, Y., Yao, H., Lin, C., Xie, Y., Tang, S., & Zhang, (2023). A. Small Molecule Metabolites: Discovery of Biomarkers and Therapeutic Targets. *Signal Transduction and Targeted Therapy*, 8(1), 132. doi: 10.1038/s41392-023-01399-3. PMID: 36941259; PMCID: PMC10026263.
- Rahman, M. M., Ng, J. C., Naidu, R. (2009), Chronic Exposure of Arsenic via Drinking Water and Its Adverse Health Impacts on Humans. *Environmental geochemistry and health*, Suppl 1:189-200. doi: 10.1007/s10653-008-9235-0. PMID: 19190988.
- Rakusanova, S., Fiehn, O., & Cajka, T. (2023). Toward Building Mass Spectrometry-Based Metabolomics and Lipidomics Atlases for Biological and Clinical Research. *TrAC Trends in Analytical Chemistry*, 158, 116825. doi: 10.1016/j.trac.2022.116825.
- Rakusanova, S., & Cajka, T. (2024). Tips and tricks for LC–MS-based metabolomics and lipidomics analysis. *TrAC Trends in Analytical Chemistry*, 180, 117940. doi: 10.1016/j.trac.2024.117940.
- Rawle, R. A., Kang, Y. S., Bothner, B., Wang, G., & McDermott, T. R. (2019). Transcriptomics Analysis Defines Global Cellular Response of *Agrobacterium tumefaciens* 5A to Arsenite Exposure Regulated Through the Histidine Kinases PhoR and AioS. *Environmental Microbiology*, 21(8), 2659-2676. doi: 10.1111/1462-2920.14577. PMID: 30815967.
- Reichard, J. F., & Puga, A. (2010). Effects of Arsenic Exposure on DNA Methylation and Epigenetic Gene Regulation. *Epigenomics*, 2(1), 87-104. doi: 10.2217/epi.09.45. PMID: 20514360; PMCID: PMC2877392.
- Roberts, L. D., Souza, A. L., Gerszten, R. E., & Clish, C. B. (2012). Targeted Metabolomics. *Current Protocols in Molecular Biology*, Chapter 30, Unit 30.2.1-24. doi: 10.1002/0471142727.mb3002s98. PMID: 22470063; PMCID: PMC3334318.
- Rodríguez-Lado, L., Sun, G., Berg, M., Zhang, Q., Xue, H., Zheng, Q., & Johnson, C. A. (2013). Groundwater Arsenic Contamination Throughout China. *Science*, 341(6148), 866-8. doi: 10.1126/science.1237484. PMID: 23970694.

- Roggenbeck, B. A., Bull Chief, L. K. & Walk, S. T. (2021). Antibiotic Perturbation of The Murine Gut Microbiome Introduces Inter-Individual Susceptibility to Arsenic. *Toxicology*, 456, 152798. doi: 10.1016/j.tox.2021.152798. PMID: 33901602; PMCID: PMC8204511.
- Röhnisch, H. E., Eriksson, J., Tran, L. V., Müllner, E., Sandström, C., & Moazzami, A. A. (2021). Improved Automated Quantification Algorithm (AQuA) and Its Application to NMR-Based Metabolomics of EDTA-Containing Plasma. *Analytical Chemistry*, 93, 8729–8738.. doi: 10.1021/acs.analchem.0c04233.
- Rolin, D., Deborde, C., Maucourt, M., Cabasson, C., Fauvelle, F., Jacob, D., Canlet, C., & Moing, A. (2013). High-Resolution ¹H-NMR Spectroscopy and Beyond to Explore Plant Metabolome. *Advances in Botanical Research*, 67, 1-66. doi: 10.1016/B978-0-12-397922-3.00001-0
- Rosen, B. P., Bhattacharjee, H., & Shi, W. (1995). Mechanisms of Metalloregulation of an Anion-Translocating ATPase. *Journal of Bioenergetics and Biomembranes*, 27(1), 85-91. doi: 10.1007/BF02110335. PMID: 7629056.
- Rosen, B. P., Ajees, A. A., & McDermott, T. R. (2011). Life and Death with Arsenic. *Arsenic Life: An Analysis of the Recent Report "A Bacterium That Can Grow by Using Arsenic instead of Phosphorus."* *Bioessays*, 33(5), 350-357. doi: 10.1002/bies.201100012. PMID: 21387349; PMCID: PMC3801090.
- Rowlinson, B., Crublet, E., Kerfah, R., & Plevin, M. J. (2022). Specific Isotopic Labelling and Reverse Labelling for Protein NMR Spectroscopy: Using Metabolic Precursors in Sample Preparation. *Biochemical Society Transactions*, 50(6), 1555-1567. doi: 10.1042/BST20210586. PMID: 36382942; PMCID: PMC9788560.
- Santini, J. M., Sly, L. I., Schnagl, R. D., & Macy, J. M. (2000). A New Chemolithoautotrophic Arsenite-Oxidizing Bacterium Isolated From A Gold Mine: Phylogenetic, Physiological, and Preliminary Biochemical Studies. *Applied and Environmental Microbiology*, 66(1), 92-97. doi: 10.1128/AEM.66.1.92-97.2000. PMID: 10618208; PMCID: PMC91790.
- Serkova, N. J. & Niemann, C. U. (2006). Pattern Recognition and Biomarker Validation Using Quantitative ¹H-NMR-Based Metabolomics. *Expert Review of Molecular Diagnostics*, 6(5), 717-731. doi: 10.1586/14737159.6.5.717. PMID: 17009906.
- Silva, R. A., Pereira, T. C., Souza, A. R., & Ribeiro, P. R. (2020). ¹H NMR-Based Metabolite Profiling for Biomarker Identification. *Clinica Chimica Acta*, 502, 269-279. doi:[10.1016/j.cca.2019.11.015](https://doi.org/10.1016/j.cca.2019.11.015).
- Silver, S., & Phung L. T. (2005). Genes and Enzymes Involved in Bacterial Oxidation and Reduction of Inorganic Arsenic. *Applied and Environmental Microbiology*, 71(2), 599-608. doi: 10.1128/AEM.71.2.599-608.2005. PMID: 15691908; PMCID: PMC546828.

- Shi, K., Fan, X., Qiao, Z., Han, Y., McDermott, T. R., Wang, Q., & Wang, G. (2017). Arsenite Oxidation Regulator AioR Regulates Bacterial Chemotaxis Towards Arsenite in *Agrobacterium tumefaciens* GW4. *Scientific Reports*, 7, 43252. doi: 10.1038/srep43252. PMID: 28256605; PMCID: PMC5335332.
- Shi, K., Wang, Q., Fan, X., & Wang, G. (2018). Proteomics and Genetic Analyses Reveal the Effects of Arsenite Oxidation on Metabolic Pathways and The Roles of AioR in *Agrobacterium tumefaciens* GW4. *Environmental Pollution*, 235, 700-709. doi: 10.1016/j.envpol.2018.01.006. Epub 2018 Jan 12. PMID: 29339339.
- Stolz, J. F., Basu, P., Santini, J. M., & Oremland, R. S. (2006). Arsenic and Selenium in Microbial Metabolism." *Annual review of microbiology*, 60, 107-30. doi: 10.1146/annurev.micro.60.080805.142053. PMID: 16704340.
- Strawn, D. G. (2018). Review of Interactions Between Phosphorus and Arsenic in Soils from Four Case Studies. *Geochemical transactions*, 19(1), 10. doi: 10.1186/s12932-018-0055-6. PMID: 29611006; PMCID: PMC5880798.
- Stewart, N. J., Sato, T., Takeda, N., Hirata, H., & Matsumoto, S. (2022). Hyperpolarized ^{13}C Magnetic Resonance Imaging as a Tool for Imaging Tissue Redox State, Oxidative Stress, Inflammation, and Cellular Metabolism. *Antioxid Redox Signal*, 36(1-3), 81-94. doi: 10.1089/ars.2021.0139. PMID: 34218688; PMCID: PMC8792501.
- Soga, T. (2023). Advances in Capillary Electrophoresis Mass Spectrometry for Metabolomics. *TrAC Trends in Analytical Chemistry*, 158, 116883. doi: 10.1016/j.trac.2022.116883.
- Sun, Y., Li, Z., Guo, B., Chu, G., Wei, C. & Liang, Y. (2008). Arsenic Mitigates Cadmium Toxicity in Rice Seedlings. *Environmental and experimental botany*, 64(3), 264-270. doi: 10.1016/j.envexbot.2008.05.009.
- Sun, W., Sierra-Alvarez, R., & Field, J. A. (2010). The Role of Denitrification on Arsenite Oxidation and Arsenic Mobility in an Anoxic Sediment Column Model with activated Alumina. *Biotechnology and Bioengineering*, 107(5), 786-94. doi: 10.1002/bit.22883. PMID: 20662039; PMCID: PMC4518441.
- Suzuki KT, Tomita T, Ogra Y & Ohmichi M (2001) Glutathione-conjugated Arsenics in the Potential Hepato-enteric Circulation in Rats. *Chemical Research in Toxicology* 14: 1604-1611.
- Tang, J. (2011). Microbial Metabolomics. *Current Genomics*, 12(6), 391. doi: 10.2174/138920211797248619.
- Tebani, A., Afonso, C., & Bekri, S. (2018). Advances in Metabolome Information Retrieval: Turning Chemistry into Biology. Part II: Biological Information Eecovery. *Journal of*

- Inherited Metabolic Disease, 41(3), 393-406. doi: 10.1007/s10545-017-0080-0. PMID: 28842777; PMCID: PMC5959951.
- Tsuji, J. S., Chang, E. T., Gentry, P. R., Clewell, H. J., Boffetta, P. & Cohen, S. M. (2019). Dose-Response for Assessing the Cancer Risk of Inorganic Arsenic in Drinking Water: The Scientific Basis for Use of A Threshold Approach. *Critical Reviews in Toxicology*, 49(1), 36-84. doi: 10.1080/10408444.2019.1573804. PMID: 30932726.
- Trifonova, O. P., Maslov, D. L., Balashova, E. E., & Lokhov, P. G. (2023). Current State and Future Perspectives on Personalized Metabolomics. *Metabolites*, 13(1), 67. doi: 10.3390/metabo13010067. PMID: 36676992; PMCID: PMC9863827.
- USGS. (2017). National News Release. <https://www.usgs.gov/news/national-news-release/study-estimates-about-21-million-people-using-wells-high-arsenic>
- Uttara, B., Singh, A., Zamboni, P., & Mahajan, R. (2009). Oxidative Stress and Neurodegenerative Diseases: A Review of Upstream and Downstream Antioxidant Therapeutic Options. *Current Neuropharmacology*, 7, 65-74. doi: 10.2174/157015909787602823.
- vanden Hoven, R. N., & Santini, J. M. (2004). Arsenite Oxidation by the Heterotroph *Hydrogenophaga* sp. str. NT-14: The Arsenite Oxidase and Its Physiological Electron Acceptor. *Biochimica et Biophysica Acta*, 1656(2-3), 148-55. doi: 10.1016/j.bbabo.2004.03.001. PMID: 15178476.
- Van de Wiele, T., Gallawa, C. M., Kubachka, K. M., Creed, J. T., Basta, N., Dayton, E. A., Whitacre, S., Du Laing, G., & Bradham, K. (2010). Arsenic Metabolism by Human Gut Microbiota upon In Vitro Digestion of Contaminated Soils. *Environmental Health Perspectives*, 118, 1004–1009. doi:10.1289/ehp.0901794.
- Vieira, R. P., Gonzalez, A. M., Cardoso, A. M., Oliveira, D. N., Albano, R. M., Clementino, M. M., Martins, O. B., & Paranhos, R. (2008). Relationships between Bacterial Diversity and Environmental Variables in A Tropical Marine Environment, Rio de Janeiro.” *Environmental Microbiology*, 10(1):189-99. doi: 10.1111/j.1462-2920.2007.01443.x. Epub 2007 Sep 24. PMID: 17892478.
- Walther, J. L., Metallo, C. M., Zhang, J., & Stephanopoulos, G. (2011). Optimization of ¹³C Isotopic Tracers for Metabolic Flux Analysis in Mammalian Cells. *Metabolic Engineering*, 14(2), 162-171. doi: 10.1016/j.ymben.2011.12.004. PMID: 22198197; PMCID: PMC3586220.
- Wang, Q., Qin, D., Zhang, S., Wang, L., Li, J., Rensing, C., McDermott, T. R., & Wang, G. (2015). Fate of Arsenate Following Arsenite Oxidation in *Agrobacterium tumefaciens* GW4. *Environmental Microbiology*, 17(6), 1926-40. doi: 10.1111/1462-2920.12465. Epub 2014 Apr 28. PMID: 24673976.

- Wang, Q., Kang, Y. S., Alowaifeer, A., Shi, K., Fan, X., Wang, L., Jetter, J., Bothner, B., Wang, G., & McDermott, T. R. (2018). Phosphate Starvation Response Controls Genes Required to Synthesize the Phosphate Analog Arsenate. *Environmental Microbiology*, 20(5), 1782-1793. doi: 10.1111/1462-2920.14108. PMID: 29575522.
- Wang, Q., McDermott, T. R. & Walk, S. T. (2021). A Single Microbiome Gene Alters Murine Susceptibility to Acute Arsenic Exposure. *Toxicological Sciences*, 181(1):105-114. doi: 10.1093/toxsci/kfab017. PMID: 33560341; PMCID: PMC8599829.
- Wang, Q., Pinkham, N. V. & Walk, S. T. (2023). The Gut Microbiome's Potential to Influence Arsenic Exposure. *Current Opinion in Toxicology*, 36, 100431. doi: 10.1016/j.cotox.2023.100431.
- Wang, P., Du, H., Fu, Y., Cai, X., Yin, N., & Cu,i Y. (2022). Role of Human Gut Bacteria in Arsenic Biosorption and Biotransformation. *Environment International*, 165, 107314. doi: 10.1016/j.envint.2022.107314. PMID: 35635965.
- Wang, N., Xue, X. M., Juhasz, A. L., Chang, Z. Z., & Li, H. B. (2017). Biochar Increases Arsenic Release from An Anaerobic Paddy Soil Due to Enhanced Microbial Reduction of Iron and Arsenic. *Environmental Pollution*, 220(Pt A), 514-522. doi: 10.1016/j.envpol.2016.09.095. PMID: 27720546.
- Wang, Y., & Lei, Q. (2018). Metabolite Sensing and Signaling in Cell Metabolism. *Signal Transduction and Targeted Therapy*, 3(1), 1-9. doi: 10.1038/s41392-018-0024-7.
- Wang, P. P., Sun, G. X., & Zhu, Y. G. (2014). Identification and Characterization of Arsenite Methyltransferase from An Archaeon, *Methanosarcina acetivorans* C2A. *Environmental Science & Technology*, 48(21), 12706-13. doi: 10.1021/es503869k. Epub 2014 Oct 22. PMID: 25295694.
- Weitzel, A., Samol, C., Oefner, P. J., & Gronwald, W. (2020). Robust Metabolite Quantification from J-Compensated 2D ^1H - ^{13}C -HSQC Experiments. *Metabolites*, 10(11), 449. doi: 10.3390/metabo10110449. PMID: 33171777; PMCID: PMC7695005.
- Wishart, D. S., Cheng, L. L., Copié, V., Edison, A. S., Eghbalian, H. R., Hoch, J. C., Gouveia, G. J., Pathmasiri, W., Powers, R., Schock, T. B., Sumner, L. W., & Uchimiy, M. (2022). NMR and Metabolomics-A Roadmap for the Future. *Metabolites*, 12(8), 678. doi: 10.3390/metabo12080678. PMID: 35893244; PMCID: PMC9394421.
- Wu, Y. S., Li, B. X., & Long, Y. Y. (2022). Rapid Quantitative ^1H - ^{13}C Two-Dimensional NMR with High Precision. *RSC Advances*, 12(9), 5349-5356. doi: 10.1039/d1ra08423b. PMID: 35425561; PMCID: PMC8981411.
- Xenakis, J. G., Douillet, C., Bell, T. A., Hock, P., Farrington, J., Liu, T., Murphy, C. E. Y., Saraswatula, A., Shaw, G. D., Nativio, G., Shi, Q., Venkatratnam, A., Zou, F., Fry, R. C.,

- Stýblo, M., & Pardo-Manuel de Villena, F. (2022). An Interaction of Inorganic Arsenic Exposure with Body Weight and Composition on Type 2 Diabetes Indicators in Diversity Outbred Mice. *Mammalian Genome*, 33(4), 575-589. doi: 10.1007/s00335-022-09957-w. PMID: 35819478; PMCID: PMC9761582.
- Xia, J., Bjorndahl, T. C., Tang, P., & Wishart, D. S. (2008). MetaboMiner-Semi-Automated Identification of Metabolites from 2D NMR Spectra of Complex Biofluids. *BMC Bioinformatics*, 9, 507. doi: 10.1186/1471-2105-9-507. PMID: 19040747; PMCID: PMC2612014.
- Xiao, J. F., Zhou, B., & Ransom, H. W. (2012). Metabolite Identification and Quantitation in LC-MS/MS-Based Metabolomics. *TrAC Trends in Analytical Chemistry*, 32, 1-14. doi: 10.1016/j.trac.2011.08.009. PMID: 22345829; PMCID: PMC3278153.
- Xiong, J., Wu, L., Tu, S., Van Nostrand, J. D., He, Z., Zhou, J., & Wang, G. (2010). Microbial communities and functional genes associated with soil arsenic contamination and the rhizosphere of the arsenic-hyperaccumulating plant *Pteris vittata* L. *Applied and Environmental Microbiology*, 76(21):7277-84. doi: 10.1128/AEM.00500-10. PMID: 20833780; PMCID: PMC2976218.
- Yoshinaga, M., Cai, Y., & Rosen, B. P. (2011). Demethylation of Methylarsonic Acid by A Microbial Community. *Environmental Microbiology*, 13(5):1205-15. doi: 10.1111/j.1462-2920.2010.02420.x. PMID: 21272184; PMCID: PMC3763196.
- Yoshinaga, M., & Rosen, B. P. (2014). A C-As Lyase for Degradation of Environmental Organoarsenical Herbicides and Animal Husbandry Growth Promoters. *Proceedings of the National Academy of Sciences of the United States of America*, 111(21), 7701-6. doi: 10.1073/pnas.1403057111. PMID: 24821808; PMCID: PMC4040596.
- Zhang, F., & Berti, P. J. (2006). Phosphate Analogues as Probes of the Catalytic Mechanisms of MurA and AroA, Two Carboxyvinyl Transferases. *Biochemistry*, 45(19):6027-37. doi: 10.1021/bi0601914. PMID: 16681375.
- Zhu, Y. G., Yoshinaga, M., Zhao, F. J., Rosen, B. P. (2014). Earth Abides Arsenic Biotransformations. *Annual Review of Earth and Planetary Sciences*, 42, 443-467. doi: 10.1146/annurev-earth-060313-054942. PMID: 26778863; PMCID: PMC4712701.

APPENDICES

APPENDIX A

APPENDICES TABLE

Table A1. Primers were used in this project.

Strain/Construct	Primer	Primer name/Digestion site
<i>E. coli</i> AW3110 (pET- <i>aroA</i>)	AAAGGATCCATGATCGAACTGACCA	aroA-F / BamHI
	AAACTCGAGGTTTCAGTGCTGCGGCT	aroA-R / XhoI
<i>E. coli</i> AW3110 (pET- <i>pepM</i>)	AAAGAGCTCATGAGCAATCAGATTGAAC	pepM-F / SacI
	AAACTCGAGAGTACAAGAGTTGCTTTC	pepM-R / XhoI
<i>E. coli</i> AW3110 (pET- <i>aroA-pepM</i>)	AAAGGATCCATGATCGAACTGACCA	aroA-F1 / BamHI
	AAAGAATTCGTTTCAGTGCTGCGGCT	aroA- R1 / EcoRI
	AAAGAGCTCAGAAGGAGATATACCATGAGCAATCAGATTGAAC	pepM-F1 / SacI
	AAACTCGAGAGTACAAGAGTTGCTTTC	pepM-R1 / XhoI

Table A2. The summary table of the metabolites detected using ^1H NMR. *WT is wild-type strain *E. coli* W3110, while MT is the mutant strain *E. coli* AW3110.

Metabolites	Early log phase				Middle log phase			
	P-WT	As-WT	P-MT	As-MT	P-WT	As-WT	P-MT	As-MT
2'-Deoxyadenosine	2.20 $\pm$ 1.89	2.51 $\pm$ 1.88	8.85 $\pm$ 1.36	3.52 $\pm$ 0.46	0.25 $\pm$ 0.06	0.24 $\pm$ 0.07	3.66 $\pm$ 0.41	3.70 $\pm$ 0.55
2'-Deoxyuridine	0.41 $\pm$ 0.08	0.57 $\pm$ 0.16	0.33 $\pm$ 0.18	0.28 $\pm$ 0.06	0.20 $\pm$ 0.03	0.19 $\pm$ 0.02	0.42 $\pm$ 0.04	0.30 $\pm$ 0.06
2-Aminobutyrate	0.98 $\pm$ 0.88	2.09 $\pm$ 2.11	1.39 $\pm$ 0.17	1.39 $\pm$ 0.10	0.51 $\pm$ 0.07	0.61 $\pm$ 0.10	0.52 $\pm$ 0.08	0.31 $\pm$ 0.07
2-Hydroxyisobutyrate	0.22 $\pm$ 0.08	0.29 $\pm$ 0.17	3.32 $\pm$ 1.08	2.42 $\pm$ 0.22	0.60 $\pm$ 0.04	0.91 $\pm$ 0.06	0.99 $\pm$ 0.06	1.17 $\pm$ 0.28
3-Hydroxy-3-methylglutarate	1.57 $\pm$ 0.54	2.65 $\pm$ 0.59	2.50 $\pm$ 0.37	1.92 $\pm$ 0.10	0.29 $\pm$ 0.07	0.38 $\pm$ 0.04	2.22 $\pm$ 0.21	1.56 $\pm$ 0.46
4-Aminobutyrate	51.32 $\pm$ 9.95	279.2 $\pm$ 89.5	4.33 $\pm$ 2.27	3.77 $\pm$ 1.44	23.19 $\pm$ 20.5	14.59 $\pm$ 7.00	161.08 $\pm$ 15.94	72.84 $\pm$ 25.49

4-Hydroxyphenylacetate	0.17 ± 0.08	0.36 ± 0.02	0.08 ± 0.02	0.09 ± 0.01	0.08 ± 0.03	0.07 ± 0.02	0.10 ± 0.04	0.10 ± 0.03
ADP	0.08 ± 0.01	0.15 ± 0.03	0.71 ± 0.06	0.15 ± 0.05	0.04 ± 0.01	0.06 ± 0.01	0.59 ± 0.08	0.13 ± 0.03
ATP	0.37 ± 0.06	0.61 ± 0.18	2.33 ± 0.78	0.75 ± 0.22	0.18 ± 0.08	0.24 ± 0.07	0.54 ± 0.08	0.53 ± 0.04
Acetate	80.00 ± 29.85	115.71 ± 9.1	173.98 ± 44.0	99.70 ± 26.4	140.68 ± 20.5	134.39 ± 12.7	157.96 ± 3.23	122.60 ± 10.93
Adenine	2.54 ± 1.17	4.42 ± 0.60	19.76 ± 1.74	16.53 ± 0.88	18.41 ± 1.00	19.64 ± 0.78	6.29 ± 0.30	4.70 ± 0.29
Adenosine	2.96 ± 1.53	3.14 ± 1.47	9.30 ± 0.77	5.19 ± 0.7	2.37 ± 0.55	2.80 ± 0.39	5.64 ± 0.19	5.67 ± 0.19
Alanine	38.90 ± 10.15	115.26 ± 10	81.47 ± 11.18	51.32 ± 5.09	75.14 ± 7.90	83.39 ± 4.87	34.35 ± 1.36	26.81 ± 2.99
Aspartate	1.94 ± 0.76	4.32 ± 0.57	5.63 ± 1.58	2.08 ± 0.11	3.06 ± 0.19	3.03 ± 0.27	1.93 ± 0.25	1.44 ± 0.22
Betaine	214.87 ± 17.6	431.5 ± 16.8	229.26 ± 24.7	220.39 ± 15	73.81 ± 1.23	85.12 ± 2.11	103.28 ± 11.62	107.56 ± 21.06
Butyrate	1.00 ± 0.28	1.61 ± 0.50	0.39 ± 0.04	0.44 ± 0.05	0.34 ± 0.02	0.38 ± 0.02	0.38 ± 0.02	0.52 ± 0.05
Choline	0.42 ± 0.14	0.55 ± 0.07	9.51 ± 0.81	4.89 ± 0.14	0.18 ± 0.01	0.23 ± 0.03	4.70 ± 0.12	5.30 ± 0.21
Citrate	0.86 ± 0.33	1.71 ± 0.88	0.57 ± 0.02	1.41 ± 0.67	0.50 ± 0.03	0.86 ± 0.47	0.71 ± 0.31	2.58 ± 0.18
Cystathionine	2.05 ± 0.72	5.01 ± 0.98	5.69 ± 1.89	5.46 ± 2.04	1.94 ± 0.43	2.02 ± 0.38	4.89 ± 0.78	3.97 ± 0.33
Cytidine	0.59 ± 0.18	1.27 ± 0.24	4.22 ± 1.37	0.56 ± 0.09	0.14 ± 0.02	0.24 ± 0.02	2.20 ± 0.30	1.51 ± 0.11
Formate	95.31 ± 31.33	173.8 ± 37.2	60.59 ± 5.87	63.75 ± 39.4	156.66 ± 33.8	181.89 ± 12.7	114.62 ± 9.59	81.29 ± 6.67
Fumarate	26.85 ± 11.10	88.08 ± 13.2	5.37 ± 0.90	1.20 ± 0.24	9.91 ± 0.97	13.21 ± 1.42	2.31 ± 0.19	1.02 ± 0.04
GTP	0.25 ± 0.08	0.44 ± 0.14	0.52 ± 0.22	0.29 ± 0.08	0.07 ± 0.004	0.09 ± 0.02	0.15 ± 0.04	0.52 ± 0.14
Gluconate	7.14 ± 2.56	17.60 ± 7.79	17.01 ± 9.31	11.66 ± 4.70	10.15 ± 7.33	13.98 ± 6.05	9.16 ± 1.44	7.11 ± 1.49
Glucose	215.0 ± 135.6	431.6 ± 298	79.7 ± 1.54	79.19 ± 6.14	48.18 ± 9.01	50.49 ± 4.27	51.02 ± 3.38	129.79 ± 4.45

Glucose-1-phosphate	2.24 ± 1.08	2.71 ± 0.69	14.50 ± 2.30	6.16 ± 0.38	4.77 ± 1.07	3.79 ± 1.19	9.75 ± 0.57	6.69 ± 0.35
Glutamate	176.5 ± 37.74	163.8 ± 38.8	755.9 ± 66.52	443.1 ± 39.6	265.11 ± 70.8	429.34 ± 13.5	264.26 ± 22.14	238.93 ± 51.44
Glutamine	8.92 ± 5.46	8.69 ± 1.64	38.41 ± 3.58	13.39 ± 0.40	5.78 ± 1.25	6.57 ± 0.58	26.42 ± 1.78	12.69 ± 1.76
Glutathione	10.03 ± 2.70	9.87 ± 2.05	63.82 ± 6.36	45.53 ± 2.24	13.97 ± 4.47	9.46 ± 2.81	41.54 ± 1.46	44.97 ± 1.24
Glycine	6.89 ± 3.84	21.16 ± 13.8	9.37 ± 0.61	10.44 ± 0.38	25.79 ± 2.83	30.04 ± 2.96	3.96 ± 0.96	6.80 ± 0.48
Histidine	0.54 ± 0.32	1.06 ± 0.71	2.13 ± 0.97	1.02 ± 0.40	0.51 ± 0.11	0.53 ± 0.02	1.80 ± 0.03	1.68 ± 0.07
Homocysteine	2.32 ± 1.60	4.81 ± 1.17	5.85 ± 0.56	2.95 ± 0.40	1.29 ± 0.13	1.44 ± 0.16	4.48 ± 0.57	2.24 ± 0.64
Hypoxanthine	4.62 ± 0.50	9.61 ± 0.12	5 ± 1.79	6.04 ± 0.39	2.25 ± 0.54	3.07 ± 0.35	2.07 ± 0.48	3.49 ± 0.26
IMP	0.21 ± 0.18	0.12 ± 0.04	0.74 ± 0.27	0.67 ± 0.16	0.07 ± 0.01	0.09 ± 0.03	2.06 ± 0.43	2.09 ± 0.17
Isocitrate	13.04 ± 6.32	22.52 ± 8.73	50.64 ± 3.17	37.60 ± 2.07	22.23 ± 4.93	15.83 ± 0.99	27.93 ± 6.64	21.13 ± 5.30
Isoleucine	0.62 ± 0.12	1.38 ± 0.56	1.99 ± 0.27	1.87 ± 0.078	1.32 ± 0.06	1.51 ± 0.06	1.28 ± 0.06	1.45 ± 0.14
Lactate	17.63 ± 10.97	160.3 ± 31.3	61.24 ± 19.36	121.3 ± 15.3	4.24 ± 0.35	4.37 ± 0.47	41.12 ± 11.59	84.78 ± 7.70
Leucine	5.55 ± 2.45	11.50 ± 3.98	8.83 ± 0.63	6.93 ± 0.26	2.11 ± 0.18	2.46 ± 0.17	5.34 ± 0.19	4.97 ± 0.36
Lysine	6.29 ± 4.34	11.55 ± 5.72	18.75 ± 0.83	7.71 ± 0.88	8.68 ± 0.37	7.32 ± 2.39	12.03 ± 0.69	6.37 ± 0.52
Malate	5.00 ± 3.78	4.61 ± 0.87	1.26 ± 0.20	1.21 ± 0.13	0.65 ± 0.08	0.70 ± 0.14	0.53 ± 0.14	1.05 ± 0.18
Maltose	9.99 ± 0.48	11.06 ± 3.81	8.91 ± 0.81	8.89 ± 0.77	17.20 ± 3.25	19.89 ± 4.78	40.49 ± 5.21	10.33 ± 0.88
Methionine	0.62 ± 0.15	1.43 ± 0.42	0.50 ± 0.10	1.08 ± 0.14	0.57 ± 0.03	0.65 ± 0.02	3.21 ± 0.19	2.80 ± 0.18
N-Acetyltyrosine	0.07 ± 0.03	0.11 ± 0.04	0.07 ± 0.02	0.07 ± 0.01	0.065 ± 0.01	0.08 ± 0.02	0.06 ± 0.007	0.06 ± 0.008
NAD ⁺	0.58 ± 0.25	0.60 ± 0.19	2.79 ± 0.42	1.10 ± 0.09	0.17 ± 0.03	0.16 ± 0.02	1.70 ± 0.11	2.68 ± 0.09

NADP ⁺	0.33 ± 0.03	0.49 ± 0.08	0.34 ± 0.05	0.32 ± 0.01	0.03 ± 0.007	0.05 ± 0.01	0.62 ± 0.03	0.73 ± 0.01
NADPH	0.70 ± 0.30	0.89 ± 0.30	1.78 ± 0.38	0.60 ± 0.09	0.12 ± 0.01	0.15 ± 0.01	1.44 ± 0.35	1.04 ± 0.27
Niacinamide	0.18 ± 0.10	0.28 ± 0.06	0.40 ± 0.18	0.40 ± 0.19	0.19 ± 0.03	0.16 ± 0.01	0.11 ± 0.008	0.16 ± 0.016
Nicotinate	6.98 ± 4.31	14.64 ± 7.58	10.24 ± 0.26	7.60 ± 0.27	6.57 ± 0.28	7.59 ± 0.25	6.39 ± 0.21	7.28 ± 0.27
Nicotinurate	0.13 ± 0.06	0.17 ± 0.08	0.28 ± 0.10	0.27 ± 0.08	0.23 ± 0.06	0.20 ± 0.04	0.14 ± 0.016	0.18 ± 0.02
N-alpha-Acetyllysine	3.03 ± 1.89	5.61 ± 2.08	3.24 ± 0.18	3.038 ± 0.66	2.15 ± 1.03	3.99 ± 0.39	1.36 ± 0.51	1.27 ± 0.28
Phenylalanine	1.47 ± 0.71	2.98 ± 1.04	2.84 ± 0.20	1.958 ± 0.04	0.87 ± 0.07	1.11 ± 0.04	2.15 ± 0.12	1.76 ± 0.07
Proline	2.37 ± 1.58	5.08 ± 0.98	3.28 ± 0.25	1.98 ± 1.16	4.11 ± 0.53	2.22 ± 0.44	4.87 ± 0.69	2.23 ± 0.60
Propionate	2.37 ± 0.73	3.04 ± 0.79	2.10 ± 0.70	3.26 ± 0.61	0.96 ± 0.08	1.42 ± 0.35	6.63 ± 0.79	5.39 ± 0.47
Putrescine	1.40 ± 1.03	1.97 ± 0.02	0.56 ± 0.11	0.57 ± 0.07	3.19 ± 0.25	3.83 ± 0.46	0.85 ± 0.07	0.73 ± 0.04
Pyridoxine	0.41 ± 0.11	0.79 ± 0.26	0.73 ± 0.22	0.95 ± 0.10	0.45 ± 0.03	0.54 ± 0.04	0.28 ± 0.02	0.42 ± 0.03
Pyruvate	7.62 ± 2.19	16.36 ± 2.46	19.09 ± 2.43	13.33 ± 2.00	3.46 ± 0.78	5.69 ± 0.92	8.61 ± 1.02	10.03 ± 0.84
Riboflavin	0.65 ± 0.42	1.18 ± 0.16	0.61 ± 0.16	0.59 ± 0.09	0.30 ± 0.15	0.30 ± 0.05	0.38 ± 0.05	0.39 ± 0.05
Ribose	33.373 ± 8.99	55.24 ± 8.73	18.01 ± 2.48	7.85 ± 2.44	8.38 ± 1.94	10.31 ± 2.52	38.59 ± 3.42	31.43 ± 1.47
Succinate	40.25 ± 7.81	70.78 ± 15.4	67.76 ± 2.40	67.83 ± 4.04	9.15 ± 2.55	12.02 ± 2.77	44.06 ± 5.71	48.87 ± 1.72
Thymidine	3.81 ± 1.56	4.78 ± 2.56	2.96 ± 1.03	0.97 ± 0.27	0.07 ± 0.01	0.08 ± 0.01	2.20 ± 0.20	0.78 ± 0.15
Thymine	3.20 ± 1.90	8.97 ± 2.12	7.38 ± 1.08	4.42 ± 0.55	2.88 ± 0.17	3.15 ± 0.14	4.48 ± 0.20	3.75 ± 0.24
Trehalose	11.33 ± 15.39	5.14 ± 0.60	115.46 ± 11.9	16.57 ± 7.25	0.20 ± 0.06	0.25 ± 0.04	31.25 ± 2.82	42.63 ± 2.80
Tyrosine	2.03 ± 0.59	4.38 ± 0.50	7.73 ± 1.01	3.77 ± 0.52	2.37 ± 0.24	3.10 ± 0.19	3.89 ± 0.17	2.23 ± 0.11

UDP-N-Acetylglucosamine	1.22 ± 0.65	2.18 ± 0.47	7.07 ± 0.75	2.51 ± 0.05	0.22 ± 0.06	0.25 ± 0.09	5.73 ± 0.20	4.72 ± 1.29
Uracil	21.70 ± 6.20	48.3 ± 11.42	36.44 ± 2.12	36.20 ± 3.10	7.82 ± 0.43	9.634 ± 0.87	41.69 ± 2.85	37.73 ± 2.33
Uridine	0.54 ± 0.24	1.09 ± 0.45	1.89 ± 0.20	0.61 ± 0.50	0.26 ± 0.07	0.33 ± 0.09	3.49 ± 0.21	3.00 ± 0.15
Valine	3.55 ± 0.91	7.76 ± 1.18	2.09 ± 0.32	1.65 ± 0.15	1.91 ± 0.23	1.99 ± 0.29	0.94 ± 0.14	1.56 ± 0.17
Xanthine	2.07 ± 0.54	5.40 ± 2.22	7.49 ± 1.86	3.90 ± 0.78	1.62 ± 0.37	1.97 ± 0.35	3.36 ± 0.40	2.73 ± 0.27
Xanthosine	0.11 ± 0.03	0.20 ± 0.01	0.16 ± 0.02	0.18 ± 0.07	0.46 ± 0.14	0.34 ± 0.22	0.09 ± 0.02	0.06 ± 0.01
gamma- Glutamylphenylalanine	0.21 ± 0.04	0.32 ± 0.05	0.18 ± 0.02	0.21 ± 0.03	0.16 ± 0.03	0.19 ± 0.05	0.15 ± 0.01	0.14 ± 0.02

Table A3. The summary table of the extracellular metabolites from the *E. coli* strains AW3110 (pET-28a), AW3110 (pET-*aroA*), AW3110 (pET-*pepM*), and AW3110 (pET-*aroA-pepM*) with or without the addition of iAs^V detected using ¹H NMR.

Metabolites	As				Non-As			
	pET28a	pET-aroA	pET-pepM	pET-aroA-pepM	pET28a	pET-aroA	pET-pepM	pET-aroA-pepM
<u>2-Hydroxyisobutyrate</u>	1665.33 ± 465.56	2390.85 ± 370.72	2777.5 ± 828.88	2016 ± 362.85	11.63 ± 1.80	6.64 ± 2.02	6.869565217	6.82 ± 0.86
<u>3-Aminoisobutyrate</u>	144.72 ± 10.72	229.04 ± 40.61	161.85 ± 9.04	216.54 ± 50.19	36.20 ± 9.66	20.81 ± 4.44	17.56521739	17.85 ± 0.56
Acetamide	266.03 ± 93.41	324.41 ± 22.31	357.93 ± 79.45	378.16 ± 88.19	540.11 ± 10.28	279.19 ± 44.60	295.6521739	225.83 ± 42.32
<u>Acetate</u>	18258.49 ± 1181.80	11954.01 ± 345.04	12968.48 ± 312.47	10472.86 ± 638.61	65575.22 ± 1166.80	54602.33 ± 5007.71	56163.65217	53346.37 ± 8097.17
<u>Formate</u>	163.82 ± 12.76	161.13 ± 23.29	173.15 ± 8.03	152.43 ± 1.23	5690.54 ± 402.37	3377.67 ± 282.19	4033.91 ± 239.74	2589.96 ± 549.99
Fumarate	214.57 ± 5.88	366.44 ± 4.11	270.87 ± 2.64	345.62 ± 3.76	181.74 ± 4.14	272.83 ± 7.28	234.96 ± 20.22	333.81 ± 56.77
Glucose	-172327.64 ± 1630.13	-202934.24 ± 690.30	-200480.33 ± 981.26	-190439.14 ± 1966.32	-117017.39 ± 3016.48	-114944.04 ± 1965.76	-107209.13 ± 9957.13	-112202.87 ± 12061.82
<u>Glutamate</u>	6535.38 ± 604.21	17731.07 ± 729.72	10426.52 ± 648.03	15244.11 ± 514.98	109.78 ± 5.87	857.58 ± 153.03	79.57 ± 37.08	1578.83 ± 348.34
<u>Lactate</u>	18859.30 ± 510.63	31771.07 ± 791.67	26171.74 ± 197.82	28213.08 ± 209.63	146.63 ± 18.23	157.13 ± 20.70	376.52 ± 85.90	119.19 ± 32.75
Maleate	762.81 ± 10.26	564.41 ± 4.75	707.17 ± 15.54	519.89 ± 1.62	722.17 ± 26.15	346.01 ± 11.31	395.48 ± 34.46	281.52 ± 7.81
N-Acetylglycine	273.47 ± 43.57	295.59 ± 97.26	155.33 ± 48.94	335.03 ± 56.40	352.28 ± 7.78	296.41 ± 37.99	297.65 ± 28.19	288.16 ± 46.34
Propionate	68.14 ± 6.79	81.02 ± 2.33	66.41 ± 19.96	77.51 ± 0.35	210.98 ± 18.50	181.70 ± 17.97	201.22 ± 2.42	183.41 ± 5.97
<u>Pyruvate</u>	3935.98 ± 1182.22	4571.30 ± 754.47	5092.28 ± 1186.37	5117.84 ± 812.36	11.20 ± 3.49	18.39 ± 6.49	16.52 ± 5.88	10.94 ± 3.83
<u>Succinate</u>	209.45 ± 11.71	88.25 ± 21.86	279.67 ± 60.11	61.19 ± 4.87	320.65 ± 20.98	364.22 ± 26.05	320.52 ± 23.08	376.05 ± 67.49
Trehalose	148.14 ± 8.36	141.47 ± 16.13	122.28 ± 9.06	181.30 ± 46.36	103.80 ± 15.38	241.43 ± 36.09	96.87 ± 33.66	200.54 ± 69.32

Table A4. The summary fold change table of the intracellular metabolites from the *E. coli* strains AW3110 (pET-28a), AW3110 (pET-*aroA*), AW3110 (pET-*pepM*), and AW3110 (pET-*aroA-pepM*) with the addition of iAs^V detected using ¹H NMR.

Metabolites	As-aroA / As-28a				As-pepM / As-28a				As-aroA-pepM / As-28a			
	Fold change	log2(fold change)	P value	Significance of P	Fold change	log2(fold change)	P value	Significance of P	Fold change	log2(fold change)	P value	Significance of P
2'-Deoxyadenosine	7.58	2.92	1.5E-03	**	5.19	2.38	1.7E-04	***	6.65	2.73	5.1E-04	***
2'-Deoxyuridine	0.34	-1.56	2.4E-02	*	0.68	-0.56	1.6E-01	NS	0.38	-1.38	3.0E-02	*
2-Aminobutyrate	0.88	-0.18	1.4E-02	*	0.77	-0.37	9.8E-04	***	0.47	-1.09	3.8E-05	****
2-Hydroxyglutarate	0.83	-0.27	7.0E-03	**	0.38	-1.41	8.0E-05	****	0.59	-0.77	4.2E-03	**
2-Hydroxyisobutyrate	1.28	0.35	5.4E-03	**	0.83	-0.27	2.2E-02	*	1.24	0.30	2.1E-02	*
3-Hydroxy-3-methylglutarate	1.12	0.16	1.5E-01	NS	0.41	-1.27	2.1E-04	***	1.31	0.39	7.0E-02	*
4-Aminobutyrate	0.84	-0.25	3.8E-01	NS	1.18	0.24	1.8E-01	NS	0.47	-1.09	1.0E-02	*
4-Hydroxyphenylacetate	0.92	-0.12	7.4E-01	NS	0.84	-0.26	2.4E-01	NS	0.71	-0.49	1.5E-02	*
ADP	0.71	-0.49	1.0E-01	NS	0.66	-0.59	5.3E-02	*	0.83	-0.28	2.7E-01	NS
ATP	0.81	-0.31	2.3E-02	*	1.00	-0.01	9.2E-01	NS	0.88	-0.18	4.1E-02	*
Acetate	0.62	-0.69	6.4E-04	***	0.68	-0.56	7.5E-04	***	0.55	-0.86	3.2E-04	***
Adenine	1.16	0.21	1.6E-01	NS	0.80	-0.31	1.0E-01	NS	1.12	0.16	2.5E-01	NS
Adenosine	1.30	0.38	3.4E-02	*	1.35	0.43	2.9E-03	**	1.10	0.14	2.8E-01	NS
Alanine	1.02	0.02	5.9E-01	NS	0.59	-0.75	1.7E-04	***	0.82	-0.29	4.2E-03	**
Aspartate	0.73	-0.45	1.2E-01	NS	0.64	-0.63	1.2E-02	*	0.75	-0.41	9.9E-02	*
Betaine	0.95	-0.07	5.2E-01	NS	0.43	-1.21	3.3E-04	***	0.85	-0.24	1.3E-01	NS
Citrate	0.54	-0.89	6.4E-02	NS	1.19	0.25	4.0E-01	NS	0.63	-0.67	4.4E-02	*
Cystathionine	0.79	-0.34	1.6E-01	NS	0.83	-0.26	4.8E-02	*	0.60	-0.73	7.8E-03	**
Cytidine	2.17	1.12	1.4E-02	*	0.78	-0.36	3.6E-01	NS	1.35	0.44	4.4E-01	NS
Formate	0.53	-0.92	1.2E-03	**	0.73	-0.45	5.7E-03	**	0.53	-0.91	1.6E-03	**
Fumarate	2.49	1.32	6.2E-05	****	0.80	-0.32	5.2E-01	NS	2.39	1.26	8.0E-03	**
GTP	0.71	-0.49	1.8E-01	NS	0.88	-0.19	5.5E-01	NS	0.94	-0.09	7.5E-01	NS
Gluconate	1.32	0.40	4.1E-04	***	1.02	0.02	8.8E-01	NS	1.12	0.16	3.4E-01	NS
Glucose	0.49	-1.02	1.1E-04	***	0.87	-0.20	1.6E-01	NS	0.68	-0.55	3.8E-04	***
Glucose-1-phosphate	1.02	0.03	8.2E-01	NS	0.96	-0.06	6.9E-01	NS	1.00	0.01	9.7E-01	NS
Glutamate	1.48	0.56	2.7E-02	*	0.57	-0.81	1.7E-02	*	1.17	0.22	1.7E-01	NS
Glutamine	1.92	0.94	4.0E-04	***	0.94	-0.09	7.0E-01	NS	2.21	1.15	3.2E-04	***
Glutathione	0.08	-3.57	2.5E-07	****	0.30	-1.73	7.7E-07	****	0.21	-2.25	4.7E-07	****
Glycine	0.23	-2.11	3.7E-06	****	0.38	-1.39	2.3E-05	****	0.15	-2.72	2.1E-06	****
Histidine	0.98	-0.03	9.1E-01	NS	0.93	-0.11	7.4E-01	NS	0.99	-0.02	9.5E-01	NS
Homocysteine	0.74	-0.44	3.0E-01	NS	1.13	0.18	6.0E-01	NS	0.81	-0.30	4.1E-01	NS
Hypoxanthine	0.73	-0.45	1.5E-01	NS	0.62	-0.69	5.3E-02	*	0.72	-0.48	1.2E-01	NS

IMP	0.35	-1.51	6.9E-02	NS	0.63	-0.67	2.4E-01	NS	0.25	-2.00	4.7E-02	*
Isocitrate	0.33	-1.61	1.8E-02	*	0.33	-1.61	1.4E-02	*	0.46	-1.12	3.0E-02	*
Isoleucine	0.16	-2.68	4.9E-06	****	0.23	-2.10	7.5E-06	****	0.12	-3.09	9.5E-06	****
Lactate	1.11	0.15	8.3E-01	NS	1.34	0.43	3.7E-01	NS	0.89	-0.16	8.0E-01	NS
Leucine	0.73	-0.46	1.3E-02	*	0.70	-0.51	5.0E-03	**	0.77	-0.39	8.4E-03	**
Lysine	0.65	-0.62	1.6E-02	*	0.34	-1.55	2.6E-03	**	0.42	-1.24	2.8E-03	**
Malate	0.71	-0.49	2.5E-01	NS	1.67	0.74	1.9E-02	*	0.90	-0.15	6.1E-01	NS
Maltose	0.27	-1.86	1.2E-04	***	0.75	-0.42	1.2E-02	*	0.39	-1.37	3.1E-04	***
Methionine	0.50	-1.01	9.8E-05	****	0.65	-0.62	1.0E-03	**	0.43	-1.22	2.1E-05	****
N-Acetyltyrosine	4.63	2.21	3.8E-03	**	0.43	-1.22	7.1E-03	**	7.87	2.98	2.2E-04	***
N6-Acetyllysine	0.15	-2.72	7.3E-07	****	0.61	-0.71	3.0E-04	***	0.09	-3.48	5.8E-08	****
NAD+	0.89	-0.16	2.2E-01	NS	0.65	-0.63	2.4E-03	**	0.75	-0.41	8.3E-03	**
NADPH	0.47	-1.09	1.0E-03	**	0.53	-0.92	1.3E-03	**	0.35	-1.50	3.7E-04	***
Niacinamide	0.95	-0.07	2.2E-01	NS	1.01	0.02	9.5E-01	NS	1.44	0.53	1.3E-01	NS
Nicotinate	0.80	-0.33	8.1E-03	**	0.67	-0.58	1.7E-03	**	0.74	-0.43	2.5E-03	**
Nicotinurate	0.96	-0.06	6.4E-01	NS	0.77	-0.38	2.5E-01	NS	0.65	-0.62	3.4E-02	*
N α -Acetyllysine	0.54	-0.90	1.8E-03	**	0.64	-0.65	4.2E-03	**	0.66	-0.60	1.3E-02	*
Phenylalanine	2.15	1.11	9.0E-06	****	0.53	-0.91	1.2E-04	***	0.96	-0.06	5.0E-01	NS
Propionate	0.71	-0.50	9.4E-04	***	0.71	-0.49	1.0E-02	*	1.00	-0.01	9.0E-01	NS
Putrescine	0.65	-0.62	3.5E-05	****	0.79	-0.33	1.2E-02	*	0.72	-0.48	8.4E-05	****
Pyridoxine	0.58	-0.78	1.4E-04	***	0.78	-0.36	2.3E-03	**	0.70	-0.51	5.5E-04	***
Pyruvate	1.23	0.30	3.2E-01	NS	1.19	0.25	9.2E-02	*	1.45	0.54	1.5E-03	**
Riboflavin	0.21	-2.22	3.9E-02	*	0.64	-0.63	2.4E-01	NS	0.36	-1.48	6.9E-02	*
Ribose	0.83	-0.26	3.7E-01	NS	0.79	-0.34	2.0E-01	NS	0.82	-0.29	2.6E-01	NS
Succinate	0.98	-0.04	7.0E-01	NS	1.00	0.00	9.5E-01	NS	1.05	0.07	6.0E-01	NS
Sucrose	0.29	-1.81	5.0E-04	***	0.76	-0.39	2.3E-02	*	0.41	-1.28	9.4E-04	***
Thymidine	0.19	-2.40	1.3E-04	***	0.41	-1.28	1.6E-04	***	0.36	-1.49	1.2E-03	**
Thymine	2.80	1.48	2.6E-03	**	2.00	1.00	5.8E-03	**	2.21	1.14	8.6E-03	**
Trehalose	0.05	-4.47	2.0E-05	****	0.88	-0.19	2.7E-01	NS	0.16	-2.64	6.6E-05	****
Tyrosine	1.49	0.57	2.6E-07	****	0.50	-1.01	2.0E-06	****	0.77	-0.38	3.9E-04	***
UDP-N-Acetylglucosamine	0.68	-0.56	2.5E-02	*	0.76	-0.39	4.7E-02	*	0.73	-0.44	2.8E-02	*
Uracil	0.98	-0.03	8.0E-01	NS	0.84	-0.25	1.4E-01	NS	1.02	0.03	8.0E-01	NS
Uridine	1.07	0.10	7.7E-01	NS	0.68	-0.55	1.4E-01	NS	0.99	-0.02	9.3E-01	NS
Valine	1.45	0.54	1.1E-04	***	0.85	-0.24	2.2E-02	*	1.17	0.23	3.8E-03	**
Xanthine	1.02	0.03	8.8E-01	NS	0.76	-0.39	1.4E-01	NS	0.96	-0.06	7.9E-01	NS
Xanthosine	0.48	-1.06	6.7E-03	**	0.57	-0.82	1.1E-01	NS	0.60	-0.74	8.8E-02	*
γ -Glutamylphenylalanine	3.06	1.61	1.2E-04	***	0.78	-0.35	2.8E-02	*	5.91	2.56	3.2E-05	****

Table A5. The summary fold change table of the intracellular metabolites from the *E. coli* strains AW3110 (pET-28a), AW3110 (pET-*aroA*), AW3110 (pET-*pepM*), and AW3110 (pET-*aroA-pepM*) without the addition of iAs^V detected using ¹H NMR.

Metabolites	P-aroA / P-28a				P-pepM / P-28a				P-aroA-pepM / P-28a			
	Fold change	log2(fold change)	P value	Significance of P	Fold change	log2(fold change)	P value	Significance of P	Fold change	log2(fold change)	P value	Significance of P
2'-Deoxyadenosine	0.89	-0.17	5.2E-01	NS	1.06	0.09	7.0E-01	NS	1.25	0.33	2.1E-01	NS
2'-Deoxyuridine	0.43	-1.21	1.5E-01	NS	0.36	-1.45	1.2E-01	NS	0.43	-1.23	1.5E-01	NS
2-Aminobutyrate	0.68	-0.55	8.4E-03	**	1.03	0.04	7.2E-01	NS	0.57	-0.81	1.9E-03	**
2-Hydroxyglutarate	0.42	-1.24	5.5E-02	*	0.61	-0.72	1.4E-01	NS	0.47	-1.09	7.7E-02	*
2-Hydroxyisobutyrate	0.89	-0.16	3.5E-01	NS	1.24	0.31	2.8E-03	**	1.20	0.26	1.7E-02	*
3-Hydroxy-3-methylglutarate	0.37	-1.44	2.3E-03	**	1.60	0.67	2.0E-03	**	0.57	-0.81	2.5E-02	*
4-Aminobutyrate	11.73	3.55	3.5E-03	**	2.25	1.17	6.0E-02	*	12.58	3.65	2.4E-05	****
4-Hydroxyphenylacetate	3.16	1.66	5.4E-03	**	2.06	1.04	4.0E-04	***	3.74	1.90	6.6E-05	****
ADP	1.40	0.48	7.8E-02	*	1.47	0.55	5.7E-03	**	1.54	0.62	1.8E-03	**
ATP	0.98	-0.03	8.5E-01	NS	1.31	0.39	2.6E-02	*	1.01	0.01	9.4E-01	NS
Acetate	1.89	0.92	3.9E-03	**	1.35	0.44	4.4E-02	*	1.31	0.39	5.6E-02	*
Adenine	1.44	0.53	9.6E-03	**	1.24	0.31	3.3E-02	*	1.12	0.16	1.7E-01	NS
Adenosine	1.05	0.06	7.9E-01	NS	0.90	-0.16	2.7E-01	NS	0.88	-0.18	2.1E-01	NS
Alanine	1.20	0.26	1.2E-01	NS	1.54	0.62	7.7E-04	***	1.17	0.23	6.9E-02	*
Aspartate	1.65	0.72	2.0E-02	*	1.64	0.71	1.7E-03	**	0.95	-0.07	6.9E-01	NS
Betaine	1.23	0.30	1.1E-01	NS	1.31	0.39	1.2E-02	*	1.29	0.37	1.8E-02	*
Citrate	1.15	0.20	3.1E-01	NS	1.47	0.55	5.3E-03	**	0.50	-1.01	5.6E-02	*
Cystathionine	2.22	1.15	2.7E-03	**	2.18	1.13	3.7E-04	***	2.13	1.09	5.0E-04	***
Cytidine	4.48	2.16	7.1E-03	**	2.02	1.02	1.0E-01	NS	1.97	0.98	1.3E-01	NS
Formate	1.67	0.74	4.5E-03	**	1.27	0.34	1.7E-02	*	1.33	0.41	3.5E-03	**
Fumarate	0.81	-0.30	1.7E-01	NS	1.14	0.19	1.7E-01	NS	1.84	0.88	8.8E-04	***
GTP	2.64	1.40	6.4E-04	***	1.10	0.13	2.3E-01	NS	1.48	0.56	1.2E-03	**
Gluconate	3.19	1.68	2.3E-04	***	1.76	0.81	1.9E-02	*	3.30	1.72	5.3E-04	***
Glucose	2.96	1.57	5.8E-04	***	0.92	-0.12	6.4E-01	NS	3.06	1.61	3.4E-07	****
Glucose-1-phosphate	0.95	-0.07	6.3E-01	NS	1.13	0.17	2.2E-01	NS	1.14	0.19	2.2E-01	NS
Glutamate	0.44	-1.18	6.7E-04	***	1.72	0.79	1.6E-04	***	0.63	-0.66	3.9E-04	***
Glutamine	0.49	-1.02	1.3E-03	**	1.08	0.11	3.1E-01	NS	0.66	-0.60	4.9E-03	**
Glutathione	0.31	-1.71	4.2E-04	***	0.83	-0.26	6.9E-02	*	0.13	-2.97	1.9E-04	***
Glycine	0.78	-0.37	2.7E-03	**	0.81	-0.31	1.1E-03	**	0.82	-0.29	5.7E-05	****
Histidine	1.51	0.59	2.6E-02	*	2.26	1.18	1.3E-04	***	1.41	0.50	9.6E-03	**
Homocysteine	1.66	0.73	5.5E-02	*	2.07	1.05	3.2E-03	**	1.83	0.87	4.1E-03	**
Hypoxanthine	0.72	-0.48	3.1E-02	*	2.26	1.18	1.2E-04	***	2.17	1.12	3.1E-05	****

IMP	0.50	-1.00	1.1E-03	**	1.14	0.19	1.2E-01	NS	1.29	0.37	2.3E-02	*
Isocitrate	0.72	-0.48	3.6E-02	*	1.32	0.40	1.5E-01	NS	0.76	-0.39	4.7E-02	*
Isoleucine	1.21	0.28	1.6E-01	NS	1.35	0.44	2.1E-03	**	0.84	-0.26	4.1E-02	*
Lactate	34.59	5.11	1.4E-04	***	12.67	3.66	5.5E-05	****	25.55	4.68	5.1E-06	****
Leucine	0.65	-0.63	1.6E-03	**	1.08	0.11	2.8E-01	NS	0.52	-0.94	2.5E-04	***
Lysine	1.21	0.27	1.3E-01	NS	1.35	0.44	1.5E-02	*	1.79	0.84	8.7E-06	****
Malate	1.74	0.80	4.4E-03	**	1.96	0.97	4.0E-04	***	1.20	0.26	8.4E-02	*
Maltose	1.02	0.02	9.2E-01	NS	1.16	0.22	3.1E-03	**	1.40	0.48	2.2E-04	***
Methionine	1.38	0.47	9.6E-02	*	2.01	1.01	7.9E-03	**	0.76	-0.40	2.7E-02	*
N-Acetyltyrosine	0.35	-1.51	2.2E-05	****	0.74	-0.43	2.9E-03	**	0.31	-1.70	2.3E-05	****
N6-Acetyllysine	0.75	-0.42	6.6E-02	*	0.84	-0.26	1.5E-01	NS	0.91	-0.13	4.3E-01	NS
NAD+	1.38	0.47	4.1E-01	NS	2.03	1.02	1.7E-03	**	1.80	0.85	5.4E-02	*
NADPH	0.76	-0.39	1.7E-01	NS	1.75	0.81	2.3E-02	*	1.13	0.17	4.3E-01	NS
Niacinamide	0.70	-0.51	8.4E-02	*	0.84	-0.26	3.0E-01	NS	1.14	0.18	4.6E-01	NS
Nicotinate	0.96	-0.06	5.4E-01	NS	0.99	-0.02	3.9E-01	NS	0.90	-0.16	3.8E-03	**
Nicotinurate	0.59	-0.77	5.0E-02	*	0.77	-0.37	2.4E-01	NS	0.77	-0.38	2.5E-01	NS
N α -Acetyllysine	0.47	-1.08	3.7E-03	**	0.99	-0.01	9.4E-01	NS	0.72	-0.47	3.6E-02	*
Phenylalanine	0.71	-0.50	1.9E-04	***	0.88	-0.19	3.1E-02	*	0.68	-0.55	1.3E-04	***
Propionate	0.77	-0.39	1.2E-01	NS	0.90	-0.16	3.3E-01	NS	0.63	-0.66	2.1E-02	*
Putrescine	0.54	-0.90	1.0E-03	**	0.73	-0.45	4.4E-03	**	0.63	-0.67	8.5E-04	***
Pyridoxine	0.79	-0.35	6.5E-02	*	0.69	-0.53	2.6E-02	*	0.89	-0.16	2.8E-01	NS
Pyruvate	2.06	1.04	3.7E-04	***	3.34	1.74	5.0E-05	****	2.71	1.44	1.4E-04	***
Riboflavin	0.63	-0.66	3.9E-01	NS	0.70	-0.52	4.7E-01	NS	0.52	-0.93	2.8E-01	NS
Ribose	0.74	-0.43	7.2E-02	*	0.56	-0.85	1.3E-02	*	0.60	-0.73	3.4E-02	*
Succinate	2.90	1.54	2.7E-04	***	2.06	1.04	1.7E-03	**	4.85	2.28	5.1E-05	****
Sucrose	1.13	0.18	5.7E-01	NS	1.22	0.29	2.0E-01	NS	1.19	0.25	2.6E-01	NS
Thymidine	0.05	-4.40	8.3E-06	****	0.05	-4.28	6.9E-06	****	0.07	-3.94	7.8E-06	****
Thymine	2.93	1.55	4.6E-04	***	2.08	1.06	9.7E-04	***	3.31	1.73	2.2E-06	****
Trehalose	48.25	5.59	4.6E-03	**	0.57	-0.82	1.2E-01	NS	36.74	5.20	6.1E-04	***
Tyrosine	0.68	-0.57	4.7E-03	**	1.08	0.11	1.4E-01	NS	0.82	-0.29	1.3E-02	*
UDP-N-Acetylglucosamine	1.15	0.20	2.0E-01	NS	1.42	0.50	1.2E-03	**	1.65	0.72	5.4E-04	***
Uracil	1.07	0.09	5.1E-01	NS	1.41	0.49	5.5E-03	**	1.61	0.68	5.6E-04	***
Uridine	1.30	0.38	2.4E-01	NS	1.19	0.25	3.8E-01	NS	1.25	0.32	2.7E-01	NS
Valine	0.93	-0.10	4.0E-01	NS	1.36	0.44	1.0E-03	**	0.79	-0.33	4.8E-03	**
Xanthine	0.52	-0.94	3.6E-03	**	0.50	-1.00	3.3E-03	**	0.56	-0.83	4.0E-03	**
Xanthosine	0.68	-0.55	2.0E-02	*	0.77	-0.37	1.0E-01	NS	0.36	-1.48	2.4E-03	**
γ -Glutamylphenylalanine	0.69	-0.54	4.9E-02	*	0.73	-0.46	2.3E-02	*	0.51	-0.97	3.3E-03	**

APPENDIX B

PUBLICATIONS DURING PHD PROGRAM

- Wang L**, Wang Q, Wolfe T, Pinkham NV, Erickson R, Hamner S, Yoshinaga M, McDermott TR, & Walk ST. Arsenic speciation methods reveal significant inter-individual variability in murine models of human arsenic exposure. In preparation.
- Larson J, Sather B, **Wang L**, Westrum J, Tokmina-Lukaszewska M, Pauley J, Copié V, McDermott TR, & Bothner B. (2024) Metalloproteomics Reveals Multi-Level Stress Response in *Escherichia coli* When Exposed to Arsenite. *Int J Mol Sci.* 25(17):9528. doi: 10.3390/ijms25179528. PMID: 39273475; PMCID: PMC11394912.
- Zhuang W, Balasubramanian N, **Wang L**, Wang Q, McDermott TR, Copié V, Wang G, & Bothner B. (2021). Arsenate-Induced Changes in Bacterial Metabolite and Lipid Pools during Phosphate Stress. *Appl Environ Microbiol* 87:e02261-20. Doi: 10.1128/AEM.02261-20.
- Wang Q, Kang YS, Alowaifeer A, Shi K, Fan X, **Wang L**, Jetter J, Bothner B, Wang G, & McDermott TR. (2018) Phosphate starvation response controls genes required to synthesize the phosphate analog arsenate. *Environ Microbiol.* 20(5):1782-1793. doi: 10.1111/1462-2920.14108. PMID: 29575522.
- Wang Q, Han Y, Shi K, Fan X, **Wang L**, Li M, & Wang G. (2017) An oxidoreductase AioE is responsible for bacterial arsenite oxidation and resistance. *Sci Rep.* doi: 10.1038/srep41536. PMID: 28128323; PMCID: PMC5270249.