Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Lennox Xu March 20, 2023

Effects of early life heavy metal exposure on the metabolome of mouse lungs

By

Lennox Xu

Dean Jones

Adviser

Chemistry

Dean Jones

Adviser

Khalid Salaita

Committee Member

James Kindt

Committee Member

2023

Effects of early life heavy metal exposure on the metabolome of mouse lungs

By

Lennox Xu

Dean Jones
Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Chemistry

2023

Abstract

Effects of early life heavy metal exposure on the metabolome of mouse lungs

By Lennox Xu

Heavy metal exposure from metals like cadmium and arsenic in fetal development and early life hurts the development of lungs and leads to decreased lung functioning. Previous work has shown that cadmium (Cd) and arsenic (As) at levels similar to non-occupational human exposures via everyday exposures, such as drinking water, can lead to postnatal lung function alterations and susceptibility to lung and airway disease. Additionally, these effects are greatly worsened by a combination of both low-level Cd and As. We aim to determine the exact mechanistic effects of the early life exposure to heavy metals by examining changes in the metabolome. We performed mass spectrometry and metabolomic analysis on early postnatal pups after they were exposed to Cd and As *in utero*. The dams were provided with low levels and high levels of Cd and As (50 and 250 ppb, respectively, each) in drinking water and purified AIN-93G diet ad libitum for 10 weeks prior to breeding and continued until 3 weeks postpartum, exposing the pups to heavy metals during pregnancy and lactation. The exposed postnatal pup lungs were processed for metabolite quantification with an Orbitrap mass spectrometer. High-resolution metabolomics was used to study metabolic changes and discover disease biomarkers. Our results have shown significant changes in various pathways vital for the development of lungs, as well as changes in the inflammation pathway. The regulation of the arachidonic acid inflammation metabolism pathway is especially essential in maintaining a balance between beneficial and harmful inflammation. The significant changes we have seen in this path, as well as the prevalence of heavy metal exposure in the everyday life of the world population, make this field of research a very important area of study. Ultimately, this research presents potential metabolites to focus on in the future for anti-inflammatory drugs to target and regulate the production for during the early life. There is hope that this would then lessen the risk of long-term negative health consequences of heavy metals.

Effects of early life heavy metal exposure on the metabolome of mouse lungs

By

Lennox Xu

Dean Jones

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences

of Emory University in partial fulfillment

of the requirements of the degree of

Bachelor of Science with Honors

Chemistry

2023

Acknowledgements

My thanks to Dr. Dean Jones, my adviser, and to my other committee members, Dr. Khalid Salaita and Dr. James Kindt, who were all very supportive throughout this project and have helped me throughout my time at Emory. A special thanks to Dr. Xin Hu, who worked actively with me through every step of this process and has been one of my biggest supporters throughout the year. Despite her heavy workload, she has been extremely willing to help me with my work and has remained patient with me, no matter the adversities. She is an amazing mentor, and I look forward to continuing to work with her. I would also like to thank other current and former lab members: Dr. Choon Lee, Grant Singer, Jaclyn Weinberg, and Dr. Ken Liu. They have taught me so many things before and during this process and have been pivotal in my development as a researcher and student. Thank you to Dr. Edward Morgan as well, for working me for the past four years before retiring in 2022. I could not have completed this paper without everyone here, so thank you again for everything.

Tables of Contents

Introduction

The prevalence of Chronic Respiratory Diseases (CRDs) has been highlighted more than ever in recent years due to the COVID-19 pandemic. Back in 2017, CRDs, which include asthma, chronic obstructive pulmonary disease (COPD), lung disease, and any other pulmonary related disease, remained as the leading cause of death worldwide with the total number of cases increasing by 39.5% from 1990 to 2017.[1] Since the pandemic started in 2019, there have been findings that these CRDs, most frequently COPD and asthma, can lead to much more severe effects and symptoms from the COVID-19 disease.[2] Importantly, while COVID-19 is rarely severe with children, it has a much more significant risk when infecting those with pre-existing CRDs.[3] The life-long consequences of chronic respiratory diseases and the potential vulnerability to other respiratory diseases throughout people's lives make them vital conditions to study and understand.

A possible culprit of the increasing prevalence of CRDs can be found in the massive industries being built and the pollution it brings to our environment. Heavy metals are spread through this pollution, eventually finding their way into food and drinking water. Exposure to these contaminants not only have severe effects on adults and children, but it also has been seen to affect developing fetuses of pregnant women. Arsenic and cadmium specifically (ranked number 1 and 7 respectively on the priority list of health hazards[4]), have been detected and found to accumulate in excess in placental or fetal issues of pregnant women.[5-7] In a study ranging from 2003-2012, in 4123 children (aged 6-18) were studied and found to have Arsenic at a level 98.7% above detectable limit and Cadmium at a level 73.9% above the detectable limit.[8] The heightened levels of these heavy metals in children have a very high potential of disrupting lung development in the early years and leading to greater health complications in later life stages. Considering these findings of heavy metal exposure *in utero,* there is a need to examine the mechanisms in which the exposure directly affects the fetuses of pregnant women in order to develop preventative measures that can directly oppose the progression of disease.

Previous research by the Dean Jones lab has shown decrease in lung function in mouse pups of dams exposed to heavy metals during pregnancy.[9] With this knowledge in mind, there is still a gap in knowledge regarding the exact effects of heavy metal exposure *in utero* to the offspring and the specific impacts on the metabolic pathways in these mice. Studying metabolism provides multi-level insight on how the body reacts to disease onset. At the mechanistic level, the regulation of the chemical reactions in various metabolic pathways is vital for maintaining homeostasis and preventing disease. At the molecular level, metabolism involves the transformation of molecules within the body into energy and other developmental products, which is critical for sustaining life and maintaining health. Investigating changes in metabolic pathways can also reflect changes in organisms that may indicate disease or other pathological conditions.[10] Importantly, metabolism can also serve as a biomarker for inflammation [11], which can be used to identify and monitor the progression of inflammatory diseases and may be useful for developing new treatments and therapies.

The aim of this research is to bridge the gap by examining the metabolome and discovering how inflammation pathways in the body are affected by heavy metal exposure. By conducting pathway analysis within metabolomics, it makes it possible to search for exact points in the pathway where metabolites are changed as a direct result of exposure.

Background

Cadmium and Arsenic

Prevalence and Danger of Heavy Metals

Exposure of heavy metals during pregnancy has been shown to have detrimental effects on offspring, even in the early stages of their lives. Heavy metals, specifically arsenic and cadmium, are a major environmental health concern, as it is prominent in many parts of the world and have led to deteriorating effects on lung health in humans and general health.[12] Aquifers throughout the world have been discovered to be contaminated with arsenic with average levels above 0.05 mg/L, 0.04mg/L above the Maximum Contaminant Level (MCL) recommended from the U.S. Environmental Protection Agency (EPA).[13, 14] Chronic Arsenic Toxicity (CAT) from this toxic drinking water has already become a major environmental health concern, causing chronic lung diseases like chronic obstructive pulmonary disease, liver disease like non-cirrhotic portal fibrosis, and other diseases like heart disease and hypertension.[15] The prominence of arsenic in groundwater and freshwater sources of drinking water worldwide makes it a vital pollutant to examine the effects on human metabolism. Additionally, cadmium is a naturally occurring environmental toxicant that could affect mitochondria function at l exposure levels.[16] It is a common heavy metal, most commonly spread through cigarette smoke, welding, and contaminated food and beverages.[17] This is the more common type of everyday exposure that occurs at low levels in most populations. This non-occupational exposure and the health effects contains less literature written on it but impacts more people. Research has suggested that excessive exposure to Cd can lead to impaired lung function and increased risk of lung cancer, and even chronic low exposure could also leave to negative effects on the kidney and bones of the general population.[18] The prevalence of these elements in everyday human consumption, along with the destructive impacts of them on human body systems, makes them an important contaminant to research and discover the metabolic pathways it has an effect on.

Harmful Effects of Heavy Metals in utero

Previous studies have displayed the ability for heavy metal exposure during pregnancy to have the potential to result in harmful effects for their offspring as well. These studies have displayed that maternal exposure to Cd has consequences on fetal levels of essential trace elements.[19] Most importantly, it leads to the deficiency in iron in the offspring, which is required for cardiovascular system development, oxygen homeostasis, and cellular metabolism.[19] Arsenic contains similar negative effects on offspring, as studies have shown that exposure to this metal *in utero* increases the risk of lower respiratory tract infections during infancy and mortality from bronchiectasis continuing into adulthood.[20]

Exposure to heavy metals such as cadmium (Cd) and arsenic (As) have been proven a major public health concern, with alarmingly high levels in pregnant women and children. Associations of metal levels and respiratory disease onset in children have been repeatedly found. Therefore, the common exposure that humans have to the heavy metals, along with the destructive impact it can have on offspring from the prenatal exposure, makes it important to study the exact mechanism of action of the harmful effects that it can result in.

The Mouse Model

Due to the invasiveness in measuring lung functions, a mouse model must be used in place of the human model for bioethical and technological considerations. This model contains many advantages with the ability to replicate critical features of human pathophysiology, most importantly in this research being human lung disease.

Differences and Similarities with the Human Model

Similar to the structure of the human organ, the mouse lung is subdivided into lobes of lung parenchyma containing a branching bronchial with very vascular pulmonary circulation.[21] Figure 1 displays how the five stages of lung development between humans and mice are similar as well.

There are some subtle variations in this general structure between species, but the largest differences between the mouse and human lung model are related to their vastly different sizes and respiratory rate.[21] This naturally leads the mice to have much smaller airway diameters and alveolar size. These differences lead to other differences in blood supply as well, which could have consequences when using mouse lungs to model human lung function.

As the particular focus for this study is the inflammatory metabolic pathways that heavy metals influence, it is important to specifically acknowledge the physical differences in the inflammatory pathways between the mouse model and human lungs. Analyzing the cell biology aspect of the lungs is the first step in determining the possibility of metabolomic differences between the two models. Mast cells are a key factor of human inflammatory lung disease, as they are a leukocyte that is mostly found in connective tissues as well as the lungs, regulating vasodilation, vascular homeostasis, innate and adaptive immune responses, and angiogenesis, all important aspects of maintaining healthy lung function.[22] The mast cells in each of the airways of these models differ in granule composition and localization: Mice mostly lack mast cells in the peripheral lung, while humans have many in this area; Mice lack an extensive pulmonary circulation due to a smaller size, displaying possible significant effects on leukocyte adhesion and migration.[21] The activation of these adhesion and migration molecules on the leukocytes are the two steps are initiated by inflammation in tissues.

Understanding the Advantages and Considering the Differences

It is important to take these differences into consideration when examining mouse models in the discovery of the effects of human ailments. The general advantages of the mouse model are their small size, fast breeding, ease of handling and extensive availability of biomedical research tools, which makes them ideal for practical reasons. Additionally, recent research has also shown similarities between the Human species genome and *Mus musculus* genome, providing greater of using mice as a model for human disease and metabolic discovery.[23] Even though mice do not naturally develop allergic asthma,

the potential for the manipulation of mice to develop similar phenotypes allow the utilization of the mouse model to study the pathobiology of allergic asthma.[21] Previous studies have also displayed metabolic similarities between the mouse and human model, which provides validation to the method of using a mouse model on this research.[24] In our study, with the differences in size, the dosages of heavy metal exposure to the mice are made relative to their size in comparison to the normal human heavy metal exposure.

Mass Spectrometry

A Fusion Tribrid Orbitrap Mass Spectrometer and an Ultra High Performance Liquid Chromatography (uHPLC) instrument were the primary analytical tools used in this experiment. Specifically for metabolomics, MS has generally been coupled with HPLC for higher sensitivity and better efficiency of analysis.[25] The HPLC column effectively separates the compounds for maximum detection by the mass spectrometry instrument.

Ultra High-Performance Liquid Chromatography

The process begins with the sample injection into the HPLC-MS system, where it is run in a solvent under high pressure through the HPLC column.[26] They work in cooperation, as HPLC separates the components of a sample based on their physicochemical properties such as size, charge, hydrophobicity, or affinity, while MS provides information on the mass and structure of individual components after.[27] Additionally, the high pressure of HPLC, makes it much preferred over liquid chromatography, which operates on gravity.

The compounds interact with the stationary phase of the column, eluting from the column at different time points, known as retention time, depending on their biochemical properties. In the Normal-phase (NP) used in this experiment, the stationary phase is polar and the mobile phase is nonpolar, so the

nonpolar molecules elute faster. The Reverse phase (RP) is opposite, where the stationary phase is nonpolar and the mobile phase is polar. Other properties that might affect retention time are size, charge and volatility.[27] Smaller compounds will generally move faster and low volatility compounds, with a lower boiling point, will also generally move faster.

The LC system we used in this project operates two columns, HILIC and C18, in parallel to obtain orthogonal chemical separations to maximize chemical detection. The HILIC column is used to separate and retain polar compounds, while the C18 column is used to separate and retain nonpolar compounds.

Specifically, hydrophilic interaction liquid chromatography (HILIC) utilizes a variation of normal-phase chromatography method with a polar stationary phase, such as silica gel, but with water-miscible less polar organic solvent as its mobile phase, such as acetonitrile and water. The advantages of HILIC are that it can analyze molecules that elute too quickly in RP and polar samples are always soluble, a problem in NP.[28] C18 in parallel utilizes reverse-phrase chromatography, with the stationary phase being made up of a C18 carbon molecules, or octyldecylsilane. The large surface area and number of carbons makes it an ideal nonpolar solvent for the stationary phase, allowing nonpolar and slightly polar compounds to pass through. [29]

As the samples elute from the columns, they are ionized using an Electrospray Ionization (ESI) source as an ionization source. ESI produces gaseous ions directly from an aqueous or aqueous/organic solvent system by creating a spray of highly charged droplets in an electric field. The sample in the acetonitrile/water solvent is emitted as a large spray within an electrostatic field, the solvent evaporates leaving behind only the charged droplets with the sample molecules, and then the ionization process continues. A counterflow of dry gas removes any remaining solvent and further reduces the size of the droplets, increasing charge density increases and concentrating the sample molecules. The high charge density leads sample molecules to undergo ionization through the transfer of a proton or other charge.[30]

Previous metabolomic studies have indicated the most metabolite database matches with positive ionization with HILIC columns and with negative ionization with C18 columns,[31] so this is what was utilized in the experiment.

Fusion Tribrid Orbitrap Mass Spectrometer

The Orbitrap is a Mass spectrometry (MS) instrument that contains many advantages over other instruments in the industry. In history, MS has been widely used for protein analysis, but it has also found major usage recently in the field of metabolomics.[32] The Orbitrap instrument includes these specific steps for MS: ionization, ion trapping, orbital motion, detection.

The ions are initially produced in the ion source before being subsequently transferred into the mass analyzer. The mass analyzer then sorts the ions, in space or time, based on the mass-to-charge ratios (m/z). Rather than the traditional method of scanning, where the separated ions are detected by an ion detector in the space or time domain, an Orbitrap instead utilizes ion trapping and orbital motion. The advantage of trapping resides in its ability to isolate compounds from mixtures, with high mass resolution, sensitivity, and accuracy.[33] This is very important for metabolomics studies, as detection of low-abundance analytes is critical to measure thousands of metabolites within a short chromatography, and the high resolution makes it possible for the distinction of compounds with very close masses.[34]

The trapping tool of the Orbitrap operates with central and outer electrodes, with a radical electric field. Ions are injected into the volume between the electrodes in a special slot with a deflector electrode to control the overall trajectory of the ions. When voltage is applied between the central and outer electrodes, a radial electric field bends the ion trajectory toward the central electrode, creating a centripetal force, while the tangential velocity creates an opposing centrifugal force with the ion's inertia. This creates the orbital motion of the Orbitrap.[33]

While in orbit, a radio frequency (RF) voltage is applied by the conical electrodes, exciting the ions and initiating harmonic axial oscillations that will determine the m/z of each ion.[33] The outer electrodes are

used as receiver plates for image current detection of these axial oscillations resulting from free induction decay (ions dropping back down to ground state), which have frequencies that are proportional to the mass-to-charge ratio of each compound.[35] The ions with the same m/z group together and the digitized image current goes through a mathematic Fourier-transformation feature into the frequency domain and converted into a mass spectrum.[36]

The Orbitrap contains many scan modes, however the one chosen for this search is the Full MS mode, without Higher-energy collisional dissociation (HCD).[37] The quadrupole in front of the Orbitrap mass analyzer is operated as a wide-pass mass filter[37], which prevents filling the C-Trap with ions outside the defined mass range of 85–1275 m/z. This provides full detection of all compounds ionized in the electrospray (ESI) ion source with m/z values in this range. The Full MS mode is the most frequently used mode since it contains comprehensive information about individual masses and provides strong quantitative data.[37]

Overall, the Fusion Tribrid Orbitrap Mass Spectrometer is a powerful tool with its strong ability to analyze narrow mass ranges in a single composite MS spectrum and its rapid fragmentation of masses in succession.[38] During analysis it continuously stores and transmits ions from an ion source into the mass analyzer with a high rate of efficiency. Overall, it holds advantages with its fast speed, high mass resolution, high mass accuracy, and its ability to measure a complex mixture of ions.[35]

Metabolomics

Metabolomics is the study of small molecules, known as metabolites, that are produced by living organisms as they go through various metabolic processes. Metabolomics aims to identify and quantify all the metabolites within a biological sample, such as a cell or tissue, and to understand their biological functions and interactions. It involves highly precise, detailed characterizations of metabolic phenotypes,

allowing for the potential to discover new therapeutic targets biomarkers that may be used to diagnose disease and monitor effects of therapeutics.[39]

Metabolites are the end products of many pathways within cellular metabolism, and their levels can be influenced by a variety of factors such as diet, lifestyle, genetics, and disease. Studying metabolism provides multi-level insight on how the body reacts to disease onset. Metabolism involves the transformation of molecules within the body into energy and other developmental products through chemical reactions, which is critical for sustaining life and maintaining health. Changes in these metabolic pathways can also reflect changes in organisms that may indicate disease or other pathological conditions.[10] Importantly, metabolism serves as biomarkers for inflammation [11], which can be used to identify and monitor the progression of inflammatory diseases. Knowing the negative consequences of some of these factors, one could study the metabolite levels in response to these stimuli, gaining insight into the metabolic pathways that are involved in various physiological and pathological processes. New technologies in this field provide detailed characterizations of metabolic phenotypes and contain potential applications in a wide range of areas, including disease diagnosis and treatment, drug development, and personalized medicine.[39]

Metabolomics is the focus of this research, as there is still a lack of knowledge regarding the mechanisms of early life heavy metal exposure on offspring lung function, especially through the disruption of metabolic pathways.

Computational workflow of metabolomics

After the production of the individual mass spectrums from the LC-MS instrument, this extracted raw data must be processed for interpretation of the metabolites using statistical association methods. Multiple measures during this process are also taken to enhance metabolite identity prediction in high and low abundance metabolites. The major steps in this data process are as follows: peak detection and alignment,

parameter optimization, and quality assessment and data Correction.[40] apLCMS [41] and xMSanalyzer [42] are coupled together for efficient data extraction in this process.

They are both R packages that automatically processes the metabolomics data to enhance the aforementioned workflow. They improved this process by providing sensitive feature detection, accurate feature quantification, and reliable alignment.

apLCMS [41] performs peak detection, noise removal, peak quantification, peak alignment, and recovery of weak signals. The output of apLCMS includes retention times, m/z features that appeared in the LCMS run, and the ion intensities of each of these m/z features for every sample. Peak detection and alignment specifically detect peaks for individual mass-to-charge ratios and only retains those that meet the signal-to-noise threshold and peak shape criteria. apLCMS assists this by grouping features together based on m/z cutoff, splitting each group of m/z features based on the differing retention times, and using a run filter to identify true peaks.[41] The run filter considers the minimum retention time, as well as the proportion of the signal's intensities at the specific retention time. If the signal does pass through the filter, it is considered a feature, and if it does not, then it is considered noise. It constructs Extracted Ion Chromatography (ETC), which displays the intensities of individual mass-to-charge ratios at different times.

In this analysis, apLCMS generates two feature tables for each of the m/z values using two sets of parameters of min.press and min.run, which define the peaks by the minimal percentage of presence of data points and the minimal time for continuous data points. xMSanalyzer then picks the best quality features from each of the two features tables with different parameters and merges the tables together. Higher correlation of feature intensities between replicates of samples indicates good quality peaks with low noise and is kept.[42] This also accounts for systemic errors, as it only keeps results with low technical variation.

The quality of individual features is evaluated based on coefficient of variation (CV) within technical replicates, variability across pooled reference/quality control (QC) samples, percent missing values, signal-to-noise ratio, principal component analysis to identify outliers, and pairwise correlation within technical replicates to evaluate analytical reproducibility.[40, 42]

After the data is extracted into feature tables, it is then further processed and analyzed. xmsPanda is another R package that normalizes the data with a log2 transform quantile normalization and runs statistical tests. The Linear Models for Microarray Data (*limma*) test run is run on the data, which is a powerful tool for time course experiences, multi-variable comparisons, and significance testing.[43] Similar to t-tests and one-way ANOVAs, the process generates a p-value and test statistic for each feature. It also adjusts false discovery rates for multiple tests with the False Discovery Rate (FDR) Benjamin-Hochberg correction method. With a significance threshold of $q < 0.05$, it can then determine if a specific feature is significantly different in the mice exposed to heavy metals *in utero* compared to the control mice. xmsPANDA also provides visualizations, such as a heatmap and boxplot for individual mass-to-charge ratios.

Pathway analysis and metabolite annotations

Annotations for metabolite feature table data has been optimized over the years, with recent technology attempting to further automate the workflow. *Mummichog* [44] is a program that aims to predicts metabolic pathway data directly from feature table data, allowing for more efficient metabolite identification.[44] It requires the mass-to-charge ratio and retention time, as well as the p-value and test statistic resulting from limma test run from xmsPanda.

For each individual feature, *mummichog* searches for all possibly matched metabolites, and then cross references three reference metabolic networks, Kyoto Encyclopedia of Genes and Genomes (KEGG) [38],

University of California San Diego (UCSD) Recon1 [15], and Edinburgh human metabolic network [16]. By doing this, the program searches for every possible pathway that the features could be involved in. A null distribution is made between the pathways and the m/z's, and significant features are calculated and marked. The program produces a list of the predicted metabolites in the significant pathways and forms visualizations, including an "activity network."[44] Understanding this allows for the understanding of which specific metabolites in individual pathways change as a result of heavy metal exposure *in utero*.

Inflammation

Inflammation is the body's defense against disease and is the process where the immune system recognizes and removes harmful foreign substances and attempts to heal the body. The natural response includes the release of hormones to dilate the blood vessels and allow more blood flow, letting more immune system cells to reach that part of the body.[45]

The arachidonic acid pathway is an extensive metabolic pathway that has been known to produce metabolites that mediate inflammatory reactions. Going through 3 pathways, cyclooxygenases (COXs), lipoxygenases (LOXs), and cytochrome P450 (CYPs), many of the later metabolites in these pathways are considered potential targets for cardiovascular and inflammatory disease.[46] It is important to examine this pathway to validate these previous findings and ensure that these metabolites are still the ones to focus on with offspring exposed *in utero* to heavy metals.

Methods

Experiment

Animal Conditions and Sample Collection

Animal protocols were approved by the Institutional Animal Care and Use Committee at Emory University (Atlanta, GA), and experiments were performed in accordance with the relevant guidelines and regulations. 6 week old Female C57/BL6 mice were purchased from Jackson Laboratory and housed in clean facilities under conventional conditions of 21–24 °C and 12 h light/dark cycle with the purified AIN-93G diet ad libitum. Mice were then treated with low levels of cadmium (Cd) and arsenic (As) (50 and 250 ppb each) in the drinking water for 10 weeks before mating and continued to be under this treatment until the pups were 3 weeks old. The pups were not provided with any food outside of the lactation from the dams. For metabolomics, half of the original female mice were randomly divided (n= 6-8 per group) and assigned to vehicle control and two Cd/As treatment groups exposed to varying level of Cd and As in drinking water. The control group was given drinking water [provided with purified water by reverse osmosis system, no detectable Cd/As (< 0.0008 ppb)] for the entire study. At three weeks old, the pups were euthanized and intact mouse lungs (n >8 per group) from the pups were harvested and stored. The mouse lung samples were then weighed. Extraction solvent which is acetonitrile and water mixture (2:1 ratio) spiked with 2.5% (v/v) mixture of stable isotope-labeled internal standards. The acetonitrile:water extraction solution was added to a weighed tissue sample (approximately 20-40 mg tissue) at a ratio of 15 μL/mg tissue. The internal standards include: [13C6]-D-glucose, [15N,13C5]- L-methionine, [13C5]-L-glutamic acid, [15N]-L-tyrosine, [3,3- 13C2]-cystine, [trimethyl13C3]-caffeine, [U13C5, U15N2]-L-glutamine, [15N]-indole, and will be used to monitor the data quality during data acquisition and extraction.

Following homogenization and incubation on ice for 30 min, extracts were centrifuged to remove protein, randomized, and 10 μL aliquots were analyzed with three technical replicates using C18 chromatography

(C18) with negative electrospray ionization (ESI) or hydrophilic interaction liquid chromatography (HILIC) with positive ESI and mass spectrometry (85–1275 m/z) on Fusion Orbitrap mass spectrometer (Thermo Fisher) [29,30].

HILIC–positive chromatography method summary [47]

The HILIC column is a normal phase column used for simultaneous separation and flushing using a dual head HPLC pump equipped with switching valves. During HILIC separation, a 10 µL sample is injected onto the HILIC column while the reverse phase column is flushed with a wash solution. The flow rate is maintained at 0.35 mL/min for 1.5 min, increased to 0.4 mL/min at 4 min, and held for 1 min. The mobile phase consists of solvents A, B, and C, with a linear gradient from 22.5% A, 75% B, 2.5% C to 77.5% A, 20% B, 2.5% C over 4 min, followed by a 1 min hold, for a total run time of 5 min. The MS is operated in positive ion mode. During the flushing phase, the HILIC column is equilibrated with a wash solution of 77.5% A, 20% B, 2.5% C.

C18–negative chromatography method summary [47]

The C18 column is a reverse phase column that is used in parallel with the HILIC column for simultaneous separation and flushing using a dual head HPLC pump equipped with switching valves. During C18 separation, a 10 µL sample is injected onto the C18 column while the HILIC column is flushed with a wash solution. The flow rate is maintained at 0.4 mL/min for 1.5 min, increased to 0.5 mL/min at 2 min, and held for 3 min. The mobile phase consists of solvents A, B, and C, with a linear gradient from 60% A, 35% B, 5% C to 0% A, 95% B, 5% C over 1.5 min, followed by a 3.5 min hold, for a total run time of 5 min. The MS is operated in negative ion mode. During the flushing phase, the C18 column is equilibrated with a wash solution of 0% A, 95% B, 5% C until 2.5 min, followed by an equilibration solution of 60% A, 35% B, 5% C for 2.5 min.

Metabolomics

HRM: High Resolution Metabolomics

Mass spectrometry data were extracted with apLCMS [41] and xMSanalyzer [42] as metabolic features with high-resolution mass-to-charge ratios paired with retention times. Data were prefiltered to keep only features with nonzero values in with above 80% abundance in all samples, and data from triplicate analyses were averaged before statistical and bioinformatic analyses.

Metabolomics Data Analysis

Limma test, which is similar to analysis of variance, using xmsPanda, was used to discover features that differed between the control and exposure groups. Significant features (with a $q < 0.05$ [$q < 0.1$ for high dose versus control group]) were further studied by pathway enrichment analyses using *mummichog version 2.0*. This approach protects against type 2 statistical error by including all features at $q < 0.1$ and protects against type 1 statistical error by permutation testing in pathway enrichment analysis (with a $p < 0.05$ for the pathways). Hierarchical clustering analysis and principal component analysis were used for untargeted comparison of the significant features differentiating treatment groups ($q < 0.05$ by limma test). Selected metabolic features from metabolic pathway enrichment analysis was done using *mummichog*. Metabolite annotations were made conservatively by searching accurate m/z at 5 ppm tolerance, with M+H adduct formed by positive ionization or M-H adduct formed by negative ionization. For each chromatography, whether multiple adducts could be detected at the same retention time were considered in annotation, a feature implemented by *mummichog 2.0*. Identities of selected metabolites were confirmed by co-elution of detected features relative to authentic chemical standards.

Statistical Analysis

Quantification data were analyzed using xmsPanda. Metabolomics data was analyzed as described above. $q < 0.1$ was considered statistically significant, and subject to *mummichog* pathway analysis. The significance level for the *mummichog* permutation test was $p < 0.05$. Due to the thousands of tests being run simultaneously, there is a high probability of false positives (Type I Error). The False Discovery Rate (FDR) Benjamin-Hochberg correction accounts for this error in the statistical analysis process by controlling the proportion of false discoveries among all the significant results. The FDR correction involves a stepwise procedure that ranks the p-values of all the tests from smallest to largest and then comparing them to a set threshold value. This threshold value is determined based on the desired false discovery rate (FDR), which was set at 0.1. The Benjamin-Hochberg method then calculates the critical value for each p-value, which is determined by multiplying the p-value by the number of tests being performed and then dividing by the rank of the p-value. All the tests with critical values below the FDR threshold value are then considered statistically significant.

Results

This study comprised of a Metabolome Wide Association Study (MWAS), where a collaboration of high-throughput metabolomic technologies was used to investigate the association between all the metabolites in a biological sample and health consequences of heavy metal exposure.[48] This was done using limma and performing a one-way Analysis of Variance (ANOVA) test.

apLCMS [41] and xMSanalyzer [42] were utilized to extract the feature tables from the mass spectrometry instrument, providing mass-to-charge ratios and retention times for each feature. xmsPanda ran a limma test for individual features, providing a p-value and test statistic for each feature as well. Importing these values into *mummichog* generated enriched pathways containing metabolites that were changed as a result of heavy metal dosage ($q < 0.05$). For the hilic positive HPLC columns, 2670 and 633 features, from male and female pups, respectively, were significantly changed. For the c18 negative columns, 1550 and 605 features, from male and female pups, respectively, were significantly changed. This is visualized in Figures 2 and 3. Additionally, these Manhattan plots illustrate the significance threshold and the fold change of features that were significantly different.

Results from this experiment have displayed significant change in metabolic pathways in mouse pup lungs. In unsupervised two-way hierarchical cluster analysis (HCA) (Figures 6 and 7), samples from the control, high dose exposure, and low dose exposure groups cluster together, showing samples from the same group have more similar metabolic profiles. On the y axis, the metabolic features that differed among three groups ($q < 0.05$ by limma test) were grouped into clusters. The clusters of metabolic features in samples from the high-dose group had different intensities compared to the control group, either increased or decreased, and samples from the low group showed intermediate difference that supported a dose-response relationship. The separated clustering validates the significant difference between the three groups. Further validation of the study method is proven with the Manhattan plots produced by *mummichog* software (Figures 2 and 3). There were two datasets that were inputted into this analysis: (1) all three groups, high dose, low dose, and control, or just (2) high dose and control. The

results of this study can be found in Figures 2 and 3. The similarities in shape for the graphs between these two analyses validate the metabolites found in the study by showing consistency and continuity between low and high dosages.

The pathway analysis produced by *mummichog* is illustrated in figures 4 and 5, illustrating trends in pathways that tend to change with the early life exposure to heavy metals. Besides the main inflammation pathways, other notable pathways found include those involved with development of the body and lipid metabolism. Pathways involved with lipid metabolism include glycerophospholipid metabolism, carnitine shuttle, and fatty acid biosynthesis. Pathways involved with development include the vitamin and gene pathways. Analysis of the *mummichog* results and manual annotation of the features display a variety of metabolites to examine within some of these pathways. Additionally, this study provides information on individual metabolites that were affected as a result of this experiment. Tables 1-4 display the mass-to charge ratios that were significantly changed from the control group and the potential compounds in these aforementioned pathways that they match with. The test statistics are included, with positive values meaning a significant increase and negative values meaning a significant decrease.

Pathways that are in high demand in early life development include: Pyrimidine metabolism, Purine metabolism, Vitamin A (retinol metabolism), and Vitamin E metabolism. Vitamin metabolism has been proven in previous to be important in the development of biological functions, with both vitamin A and E in particular being vital for lung development.[49]

There has been long standing literature on the importance of pyrimidine, purine, and hexose metabolism for genetic growth and development as well. Some potential metabolites associated with the pathways include deoxyuridine, inosine, and phosphatide. Overall, after examining the tests statistics of the significant metabolites, there were mixed responses for each path, with some metabolites increasing and others decreasing in the treatment groups. With the expectation for many of these features to be decreased due to the faltered lung development seen in previous models, the mixed responses are important to look more into for the future.

Lipid metabolism is largely connected to the arachidonic acid pathways (Figure 8), as the cell membranes and lipids are the main source of fatty acids for this pathway. Additionally, they play a large role in energy consumption has the main source during this newborn period.[50] The lipid and energy pathways include these pathways: Carnitine shuttle, glycerophospholipid metabolism, glycosphingolipid metabolism. Some of the potential metabolite associated with the pathways include: linoelaidyl carnitine, diacylglycerol, and sn-glycerol 3-phosphate. The test statistics of these features tended to show an increase of these features in the treatment group. The increase of these metabolites is understandable based on the contribution lipid metabolism has on energy needs of the fetus, as well as their role as the precursors of inflammation pathways.

The pathways mainly focused on in this study are the inflammation pathways discussed earlier. It includes arachidonic acid metabolism and its associated pathways, such as prostaglandin formation, leukotriene metabolism, and linoleate metabolism. Figure 7 displays the core arachidonic acid metabolism pathway with the downstream pathways and metabolites as well, such as the prostaglandins and leukotrienes. Linoleate metabolism is a precursor to arachidonic acid, suggesting that pro-inflammatory precursors were also broadly impacted by the metals. As discussed previously, the arachidonic pathway and the pathways associated with it have already been proven to be inflammation biomarkers. The inflammation pathways marked by *mummichog* (Figures 4-5) include these pathways: Prostaglandin formation, arachidonic acid metabolism. The potential metabolite list associate with the pathways include: Thrombaxane B2, prostaglandin B1, prostaglandin A2, prostaglandin G2, leukotriene A4, eicosatetranoic acid (arachidonic acid). The test statistics of these features showed a consistent increase in the treatment groups. The increase of these metabolites is understandable due to their nature as inflammation metabolites. The prostaglandin, thromboxane, and leukotriene pathways are inflammatory mediators produced simultaneously in the arachidonic pathway, and changes in one pathway will also lead to change in the others.[51] Prostaglandins are lipid metabolites derived from arachidonic acid through the COX pathway, which sustain homeostasis and mediate pathogenic mechanisms, including the inflammatory

response.[52] Additionally, there has been some debate in their potential function in both the promotion and resolution of inflammation.[52] Thromboxanes are produced by platelets and has been linked implicated in many cardiovascular conditions, such as asthma.[51] Leukotrienes produced by immune system cells that facilitate the production of immune system signaling molecules and are linked to many pathological conditions, such as bronchospasms and asthma.[51] Overall, this displays the specific metabolites that are targeted from the early life heavy metal exposure and illustrates how the inflammation pathway is affected at all levels. The boxplots representing some of these significant changes in metabolites within the inflammation pathway is visualized with Figures 10-12. These figures illustrate the large and significant changes in metabolite intensities in the body, as there is little to no overlap in the overall ranges of each of the groups.

Discussion

The results of this research successfully accomplished the original research aims at the beginning of this project. Studying metabolomics data has provided insight into potential metabolites that are affected as a result of early life arsenic and cadmium *in utero* exposure. Previous human studies have shown the negative consequences of heavy metal exposure on the fetal lungs, shown by increased asthma risks in children with higher Cd levels in cord blood at birth.[7] As this past work has largely been focused on the health effects of occupational exposure, the mechanisms are not fully understood with low levels of exposure.

Cd and As are both commonly found in natural environment. According to USDA soil survey, areas of the United States, including some southwestern and midwestern states, have high levels of Cd and As in the soil. At molecule level, Cd is known to causes oxidation of mitochondrial proteome by attacking mitochondrial anti-oxidant system and As is known to replace phosphate and interfere with enzyme activities (termed “arsenolysis”), most commonly found in cytosolic glycolysis. [53, 54] Therefore, mixed exposure to Cd and As may attack the multi-compartmentalized enzymatic reactions in cell metabolism and cause a synergistic impact. This was discovered in our preliminary studies of fetal lung branching morphogenesis (i.e. the process to generate airway branches in development).[9]

Additionally, we are more concentrated on the health effects of low level exposure, which we have less knowledge, but causes a much wider impact on worldwide populations.[55] In order to study these low-levels of exposure, the dosages to the mouse model are adjusted to be relative to the what would affect the human model. This takes into account the largest difference been the mice and human model, being the size differences.

The average Arsenic level was as high as 120 ppb in some county populations in community water systems, greatly exceeding the United States Environmental Protection Agency’s 10 ppb maximum.[56, 57] Therefore, based on this literature and previous experiments done on mice, for *in vivo* studies, female

mice were treated with 50 and 250 ppb Cd and As (as $CdCl_2$ reach a steady body homeostasis under this treatment, with continued exposure the entire time until euthanization. 50 ppb represents the lower dose, and 250 ppb represents the higher dose. The 50 ppb dose level (HED [human equivalent dose]: 0.72 µg/kg bodyweight [bw]/day [58]) is similar to the Reference Dose (RfD) for iAs (0.3 µg/kg bw/day) and estimated exposure scenarios for average Americans (total As: 0.36, iAs: 0.25 µg/kg bw/day).[59, 60] The doses of Cd (HED: 0.5 -2.8 µg/kg bw daily) are also similar to the average non-smoking human daily oral Cd intake of 0.14-0.5 µg/kg bw. [57] The diet is accompanied with the AIN-93G diet as well to avoid uncontrolled intake of Cd and As. Importantly, it has also been shown that some of the metabolites and pathways, such as prostaglandins, found in our study were compounds that had similar metabolism patterns in both the human and mouse models.[24]

Previous work done in the Dean Jones Lab has seen lowered lung function in the mouse model. The experiments done with this thesis was done on the same group of mice, treated with the same conditions. The mice were split in half for this metabolic research and the previous lung function testing research due to the possibility of the Flexivent machine affecting the metabolism and mass spectrometry analysis. In order for this machine to measure lung function parameters with forced oscillations through air flow, it requires anesthetic, paralyzing injection, and tracheostomizing. This procedure has the possibility of affecting metabolism, so this research split the sample of dams in half for each experiment. This work builds off the previous research by conducting a metabolomics study and thoroughly examining the metabolic pathways for points of interest. The significant pathways discovered as a result of this analysis aligns with the prior expectations of this project. As illustrated previously, the developmental, energetic, and inflammation pathways found to the most affected pathways by the heavy metal exposure. Due to the use of glycerophospholipids for general energy consumption, as well as being the main building blocks for the arachidonic inflammation pathways, the general increase of the metabolites in the pathway aligns with our previous knowledge. The effects on the vitamin and developmental pathways are also understandable considering their large role in early development of offspring.

The arachidonic pathway and the inflammation mediators they contain have been continuously mentioned throughout this paper due to their direct connections with pathological conditions. The biosynthesis of prostaglandins, thromboxanes, leukotrienes have long been considered pathological agents in the human body. These mediators play crucial roles in the body's response to injury and infection. However, excessive or prolonged inflammation can lead to tissue damage and chronic disease. The regulation of arachidonic acid metabolism is, therefore, essential in maintaining a balance between beneficial and harmful inflammation. This information, as well as the extreme prevalence of heavy metal exposure in the everyday life of most humans, illustrates why this field of research is vital for maintaining the well-being of our worldwide populations. Our research presents possible metabolites to focus on in the future for anti-inflammatory drugs to target and regulate the production. This has the potential to lessen the risk of long-term negative consequences of heavy metal exposure in the early life. Additionally, even at the lower dosages for this experiment, similar pathways were found to be significantly different from the control group of mice. This displays the potential for even commonplace exposure to have negative consequences on the lung function at the early stages of life.

Through examination of the data, there were large differences in the male and female mouse in the metabolomics data. There were much fewer pathways and metabolic annotations found for the female mice compared to the male mice. Previous studies in humans have stated that lipid metabolism in particular contains the most significant differences between the two sexes, with males having a higher abundance of various metabolites in these lipid pathways.[61] This is seen in our research, as more significant features were found in the c18 negative column (most effective in separating and retaining non-polar compounds) with the male pups versus the female pups (Figures 3 and 4). However, there is still a lack of research done on why there are the differences between the number of metabolites for each sex in the early stages and in mice. For future directions, it would be interesting to do a full scan MS with fragmentation in order to have an MS/MS spectrum of the significant features found in the data. This

fragmentation data would provide exact structural groups and functional groups of the features, allowing for the confirmation of the characterization of the mass-to-charge ratios and retention times. This would also help differentiate isomers from each other, which is especially important with the arachidonic pathway, as many of the prostaglandins and other downstream metabolites have the same masses. We could also run a transcriptomics study with RNA sequencing, as RNA transcripts would reflect enzymatic activity and provide more insight on the regulation of the metabolites in these pathways.

In conclusion, our study shows that early life cadmium and arsenic at levels that are comparable to human exposure found in the environment can cause issues with lung development and function that are also reflected in changes within the metabolome. We discovered changes in many metabolites, such as prostaglandins and different fatty acids, within various metabolic pathways, with the lipid, developmental, and inflammation pathways being the most prominent. These differences represent specific metabolites that were changed as a result of the heavy metal exposure and have the potential to be biomarkers for inflammation in the early development and function of lungs. In the future, we will look to confirm the exact identities of the proinflammatory metabolites and mark them as therapeutic targets for future drug discovery.

Figures

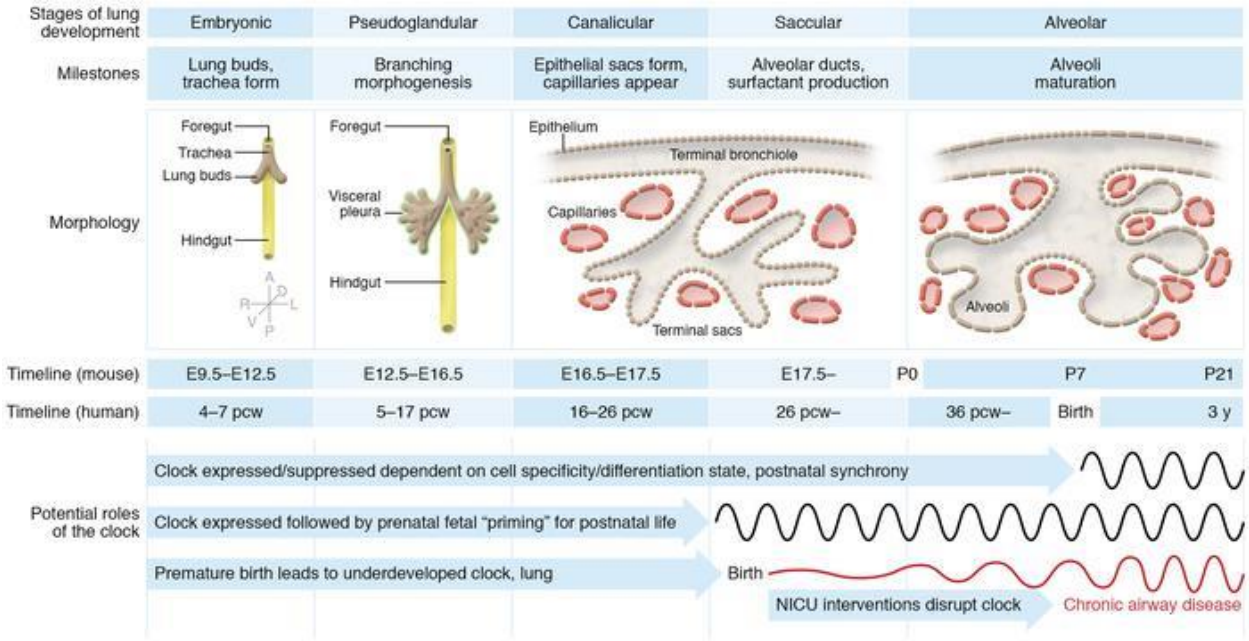

Figure 1. Five stages of lung development for humans and mice. The stages are as listed: embryonic (formation of lung buds), pseudoglandular (branching morphogenesis), canalicular (formation of epithelial sacs and appearance of capillaries), saccular (production of alveolar ducts and surfactant protein), and alveolar (maturation of the alveoli). [62] The timeline of the mouse stages is represented as embryonic (E) or postnatal (P) days, and the timeline of the human stages is shown by post-conception weeks (pcw). [62] Reprinted with permission.

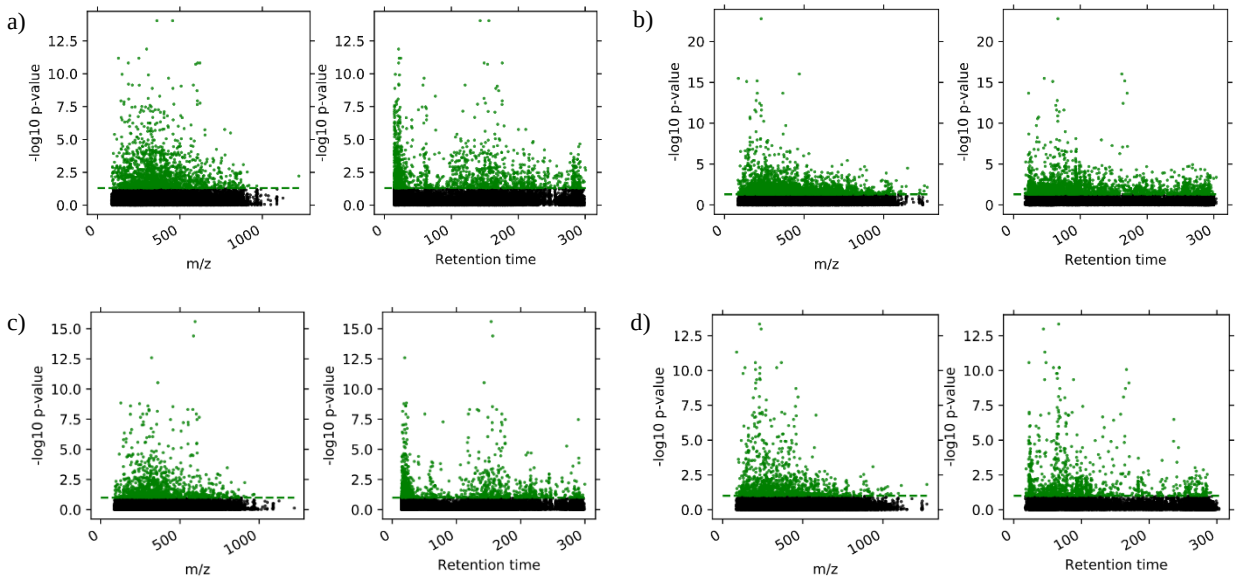

Figure 2. Manhattan plots of number of significant features for male pups. These plots display the significant metabolites that surpass the set threshold (shown in green dots, FDR corrected $q < 0.05$). Y-axis [-log(p-value)] represents the fold change (level of significance for each feature). Plots were generated by *mummichog 2.0*. a) Results from hillic positive column for male pups, comparing high dose, low dose, and control groups; 22650 original features, 2670 significant (FDR corrected $q < 0.05$) features. b) Results from c18 negative column for male pups, comparing high dose, low dose, and control groups; 13151 original features, 1550 significant features (FDR corrected $q < 0.05$). c) Results from hillic positive column for male pups, comparing high dose and control groups; 19966 original features, 1073 significant features (FDR corrected $q < 0.1$). d) Results from c18 negative column for male pups, comparing high dose and control groups; 12853 original features, 1076 significant features (FDR corrected $q < 0.1$). The different q value threshold shown here was to ensure there were comparable number of significant features for pathway analysis.

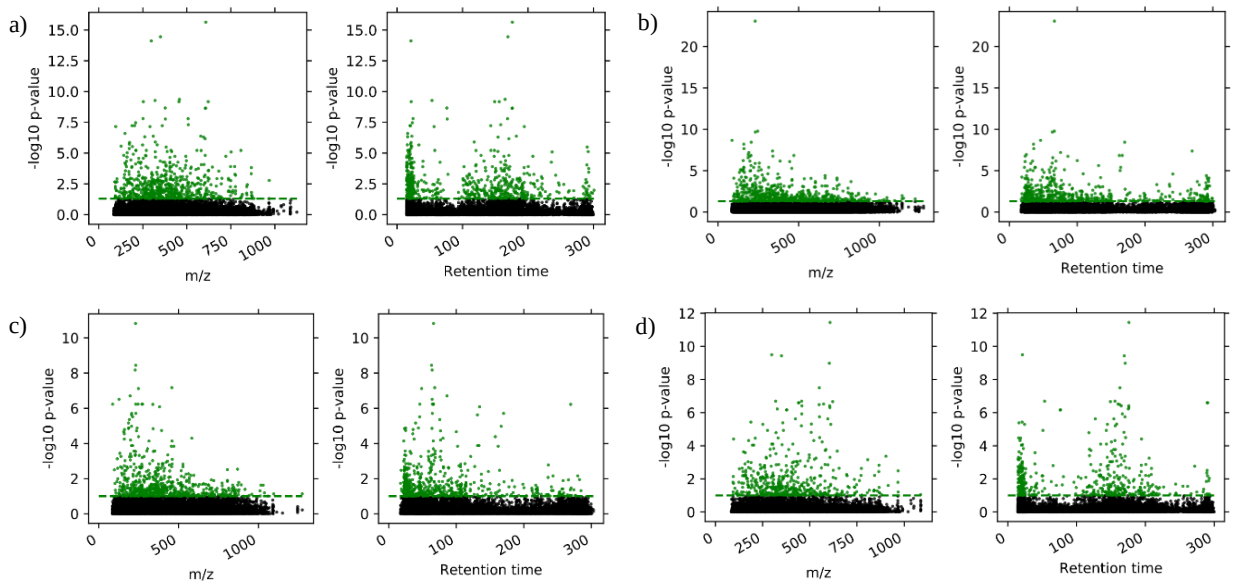

Figure 3. Manhattan plots of number of significant features for female pups. These plots display the significant metabolites that surpass the set threshold (shown in green dots, FDR corrected $q < 0.05$). Y-axis [-log(p-value)] represents the fold change (level of significance for each feature). Plots were generated by *mummichog 2.0*. a) Results from hillic positive column for female pups, comparing high dose, low dose, and control groups; 22160 original features, 633 significant features (FDR corrected $q < 0.05$). b) Results from c18 negative column for female pups, comparing high dose, low dose, and control groups; 12832 original features, 605 significant features (FDR corrected $q < 0.05$). c) Results from hillic positive column for female pups, comparing high dose and control groups; 17951 original features, 616 significant features (FDR corrected $q < 0.1$). d) Results from c18 negative column for female pups, comparing high dose and control groups; 12291 original features, 454 significant features (FDR corrected $q < 0.1$). The different q value threshold shown here was to ensure there were comparable number of significant features for pathway analysis.

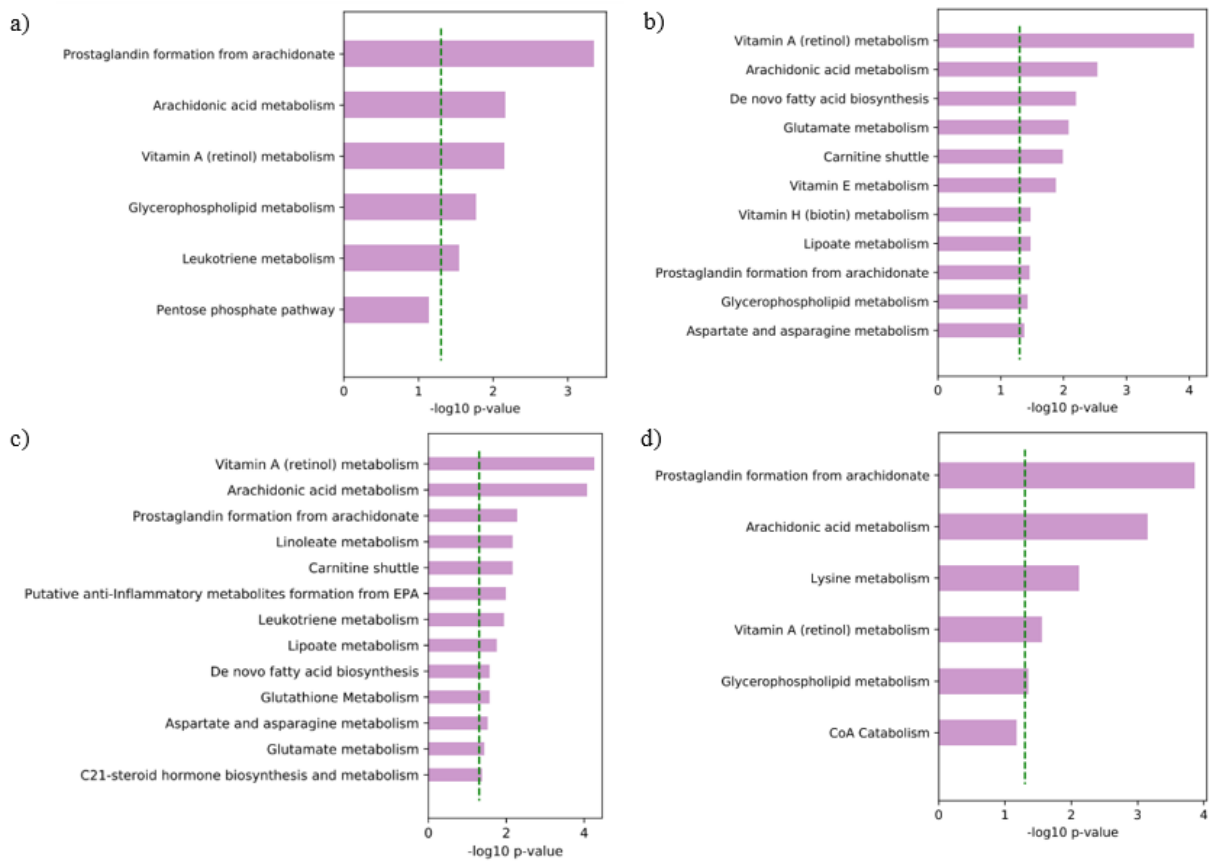

Figure 4. Top pathways with significant changes in metabolite features for male pups. These plots illustrate the top pathways with the most significant metabolites for each column and sex. a) hilic positive male pups with comparison between high dose, low dose, and control groups; full feature list was matched to 800 empirical compounds with 174 of them being significant. b) c18 negative male pups with comparison between high dose, low dose, and control groups; full feature list was matched to 852 empirical compounds with 156 of them being significant. c) hilic positive male pups with comparison between low dose and control groups; full feature list was matched to 769 empirical compounds with 128 of them being significant. d) c18 negative male pups with comparison between low dose and control groups; full feature list was matched to 832 empirical compounds with 131 of them being significant.

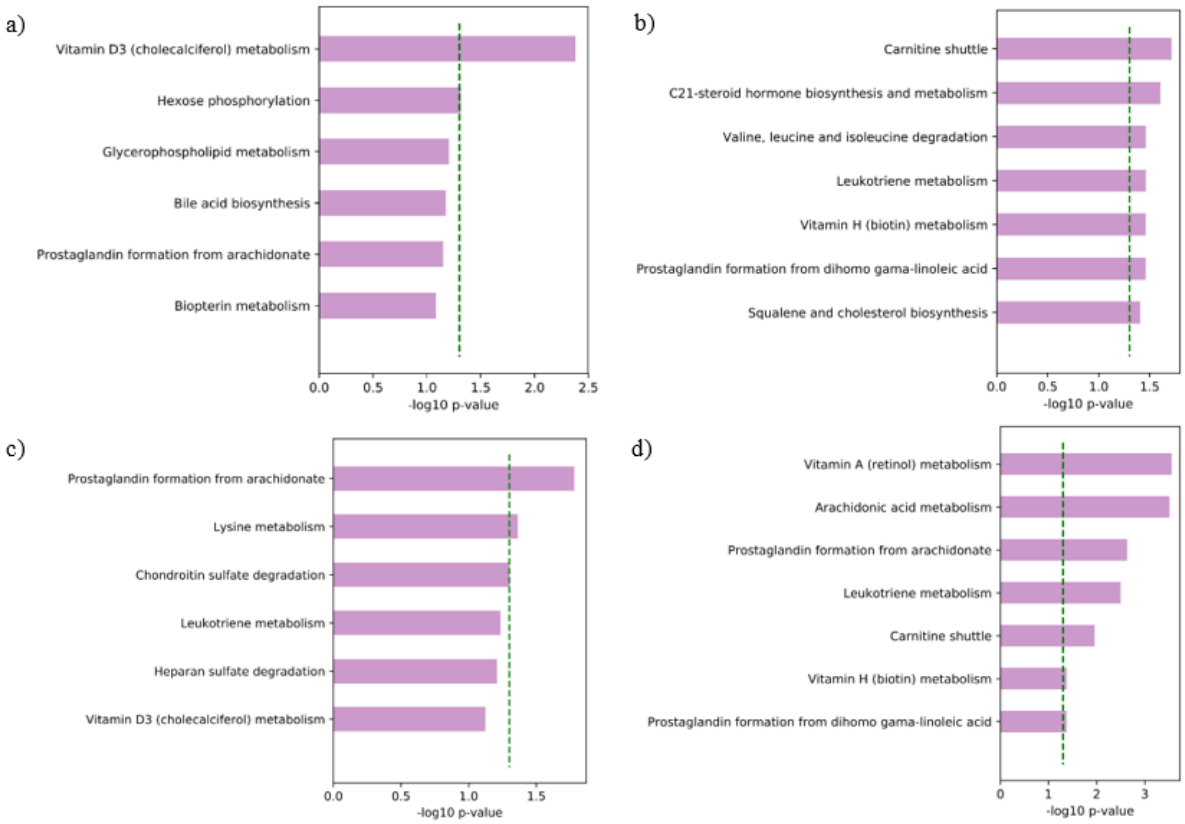

Figure 5. Top pathways with significant changes in metabolite features for female pups. These plots illustrate the top pathways with the most significant metabolites for each column and sex. a) hilic positive male pups with comparison between high dose, low dose, and control groups; full feature list was matched to 788 empirical compounds with 56 of them being significant. b) c18 negative male pups with comparison between high dose, low dose, and control groups; full feature list was matched to 841 empirical compounds with 75 of them being significant. c) hilic positive male pups with comparison between low dose and control groups; full feature list was matched to 737 empirical compounds with 66 of them being significant. d) c18 negative male pups with comparison between low dose and control groups; full feature list was matched to 807 empirical compounds with 54 of them being significant

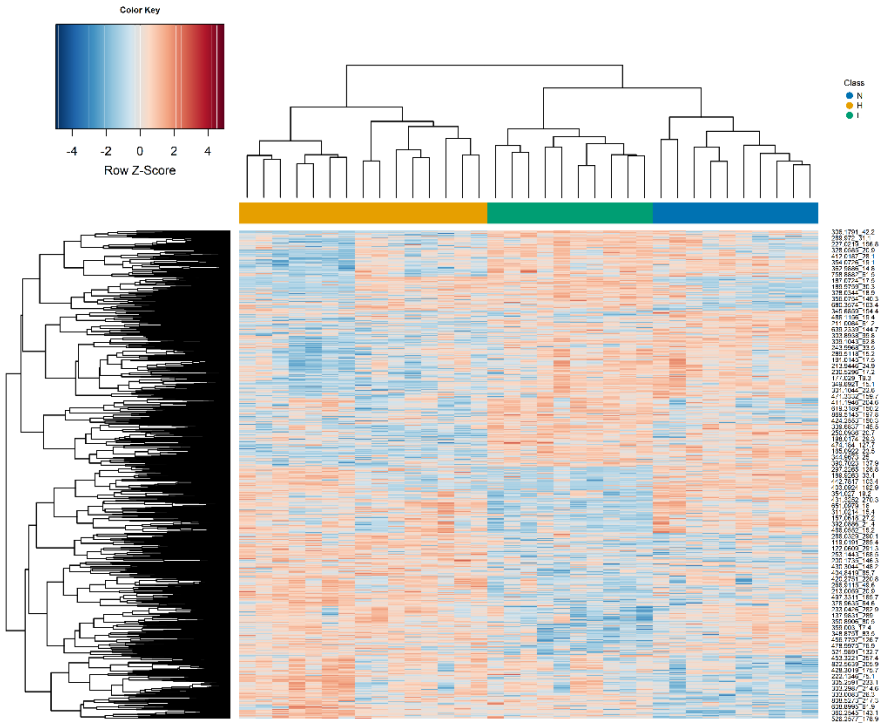

Figure 6. Dendrogram unsupervised two-way hierarchical cluster analysis (HCA) for male pups using c18 negative data. Metabolic features, shown on the y-axis, that differed among three groups ($q<0.05$ by limma test) were hierarchically grouped into clusters based on similarities without any prior classification of the samples. The top bar, on the x-axis, displays the group that the significant features belong to, labelled after the analysis. The yellow bar (far left) represents the high dosage group, the green bar (middle) represents the low dosage group, and the blue bar (far right) represents the control group. Row Z score represents normalized peak intensity of metabolic features (red: high; blue: low). Red box indicates an increase in metabolites and a blue box indicates a decrease in metabolites.

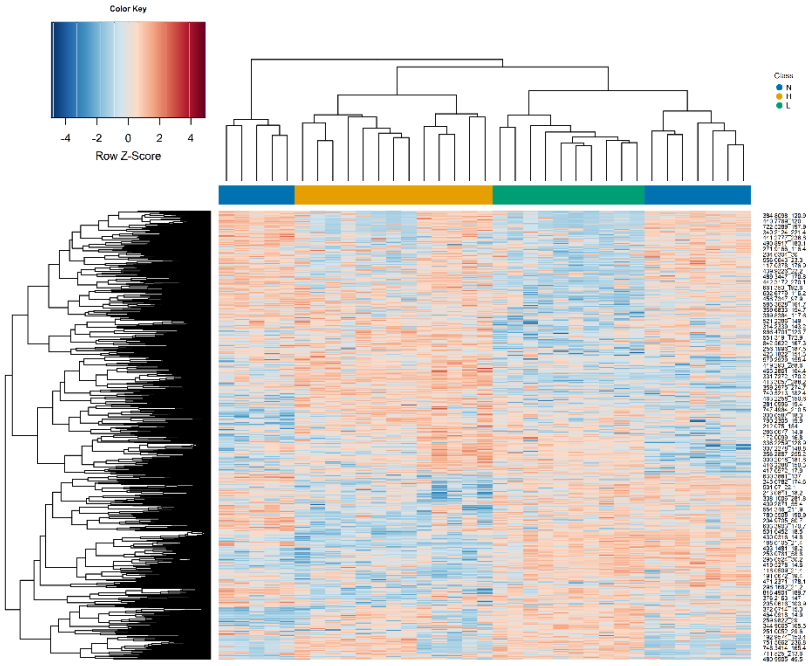

Figure 7. Dendrogram of unsupervised two-way hierarchical cluster analysis (HCA) for female pups using c18 negative data. Metabolic features, shown on the y-axis, that differed among three groups ($q<0.05$ by limma test) were hierarchically grouped into clusters based on similarities without any prior classification of the samples. The top bar, on the x-axis, displays the group that the significant features belong to, labelled after the analysis. The yellow bar (far left) represents the high dosage group, the green bar (middle) represents the low dosage group, and the blue bar (far right) represents the control group. Row Z score represents normalized peak intensity of metabolic features (red: high; blue: low). Red box indicates an increase in metabolites and a blue box indicates a decrease in metabolites.

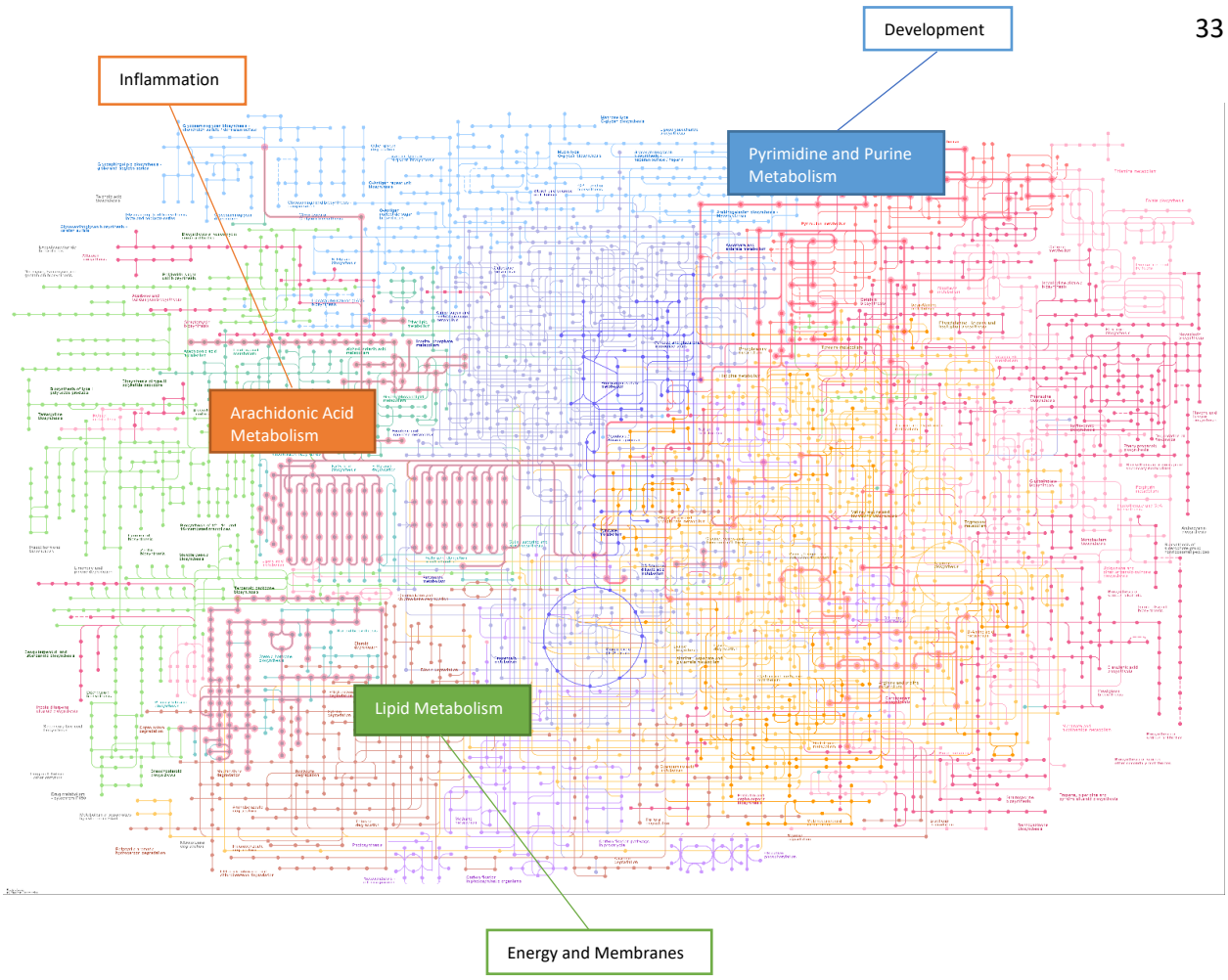

Figure 8. Annotated KEGG (Kyoto Encyclopedia of Genes and Genomes) metabolism pathways figure. (https://www.kegg.jp/pathway/map01100) The highlights pathways represent the pathways most changed from the early life heavy metal exposure done in the experiment. Arachidonic acid, lipid, and pyrimidine and purine metabolism pathways are all labeled. Reprinted with permission. [63]

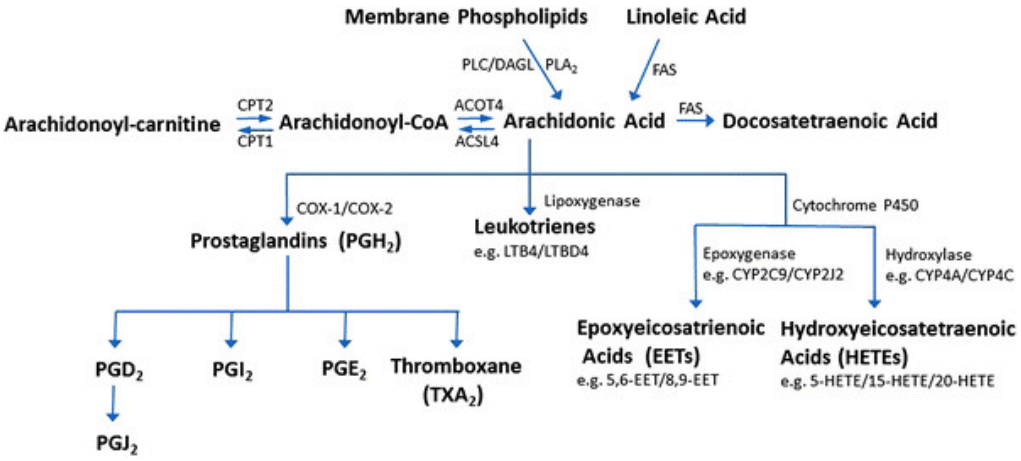

Figure 9. Full arachidonic acid metabolic and biosynthetic pathway. 3 main downstream pathways of arachidonic acid: cyclooxygenases (COXs), lipoxygenases (LOXs), and cytochrome P450 (CYPs). Also displays the links to lipid metabolism and the carnitine shuttle. Reprinted with permission.[64]

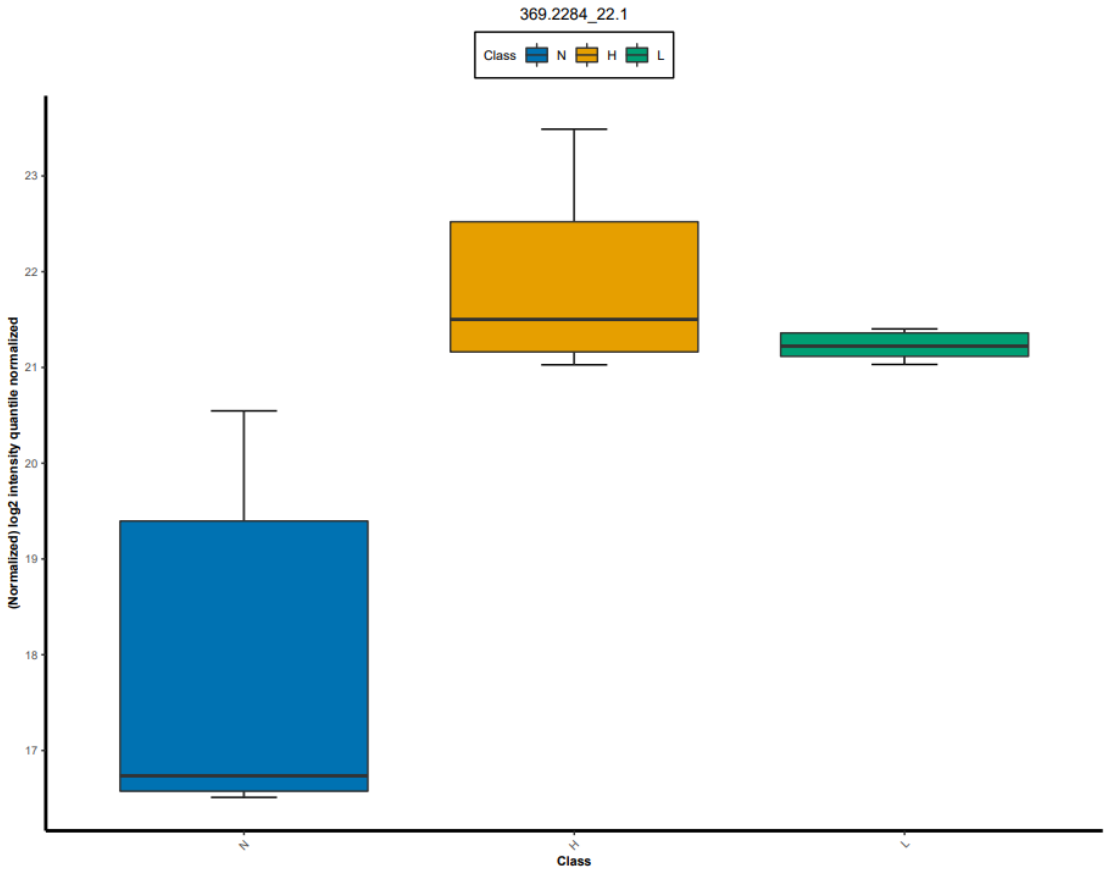

Figure 10. Box plot of Thromboxane B2 significant increase. In the high (H) and low (L) dosage treatment groups compared to the control group (N), Thromboxane B2 has significantly increased.

Thromboxanes are produced by platelets and have been linked to many cardiovascular conditions, such as asthma.[51] This feature was taken from the c18 negative male feature table.

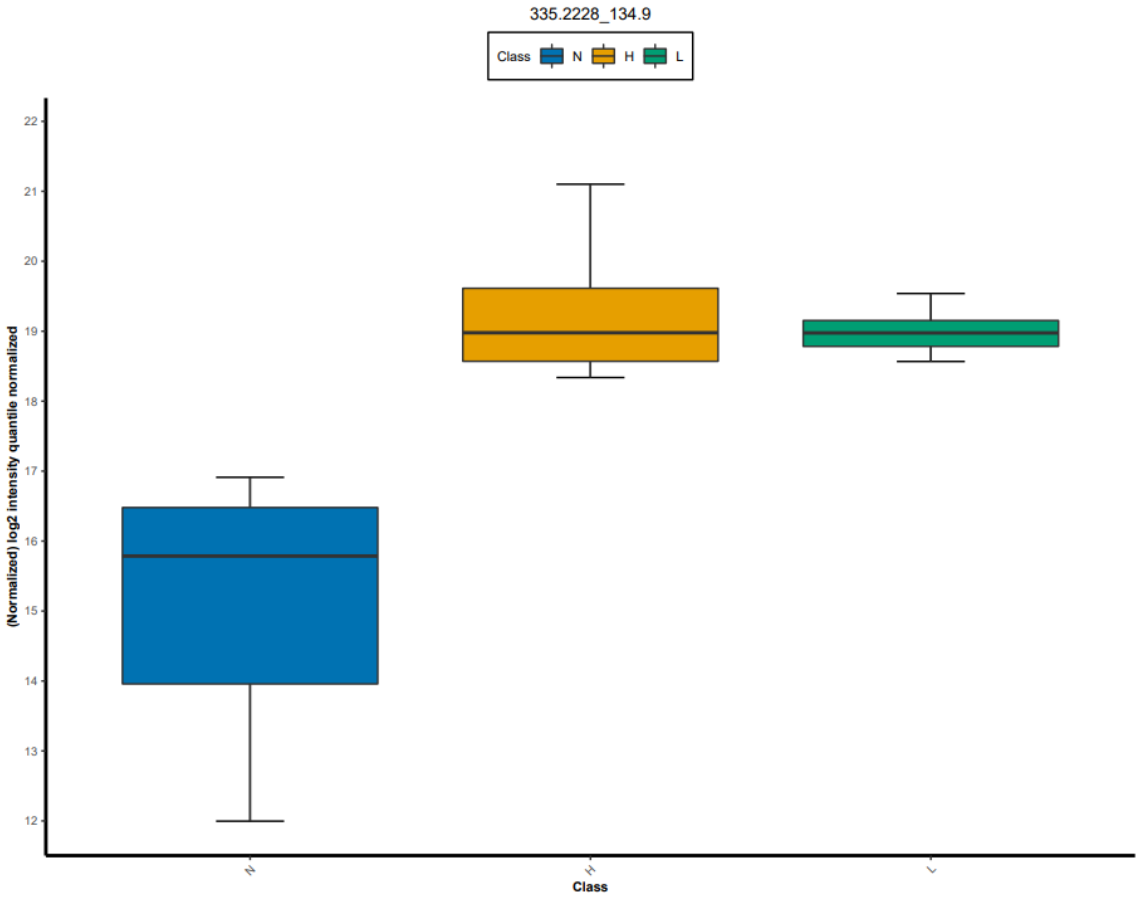

Figure 11. Box plot of Prostaglandin B1 significant increase. In the high (H) and low (L) dosage treatment groups compared to the control group (N), Prostaglandin B2 has significantly increased. Prostaglandins are lipid metabolites derived from arachidonic acid through the COX pathway, which sustain homeostasis and mediate pathogenic mechanisms, including the inflammatory response.[52] This feature was taken from the c18 negative male feature table.

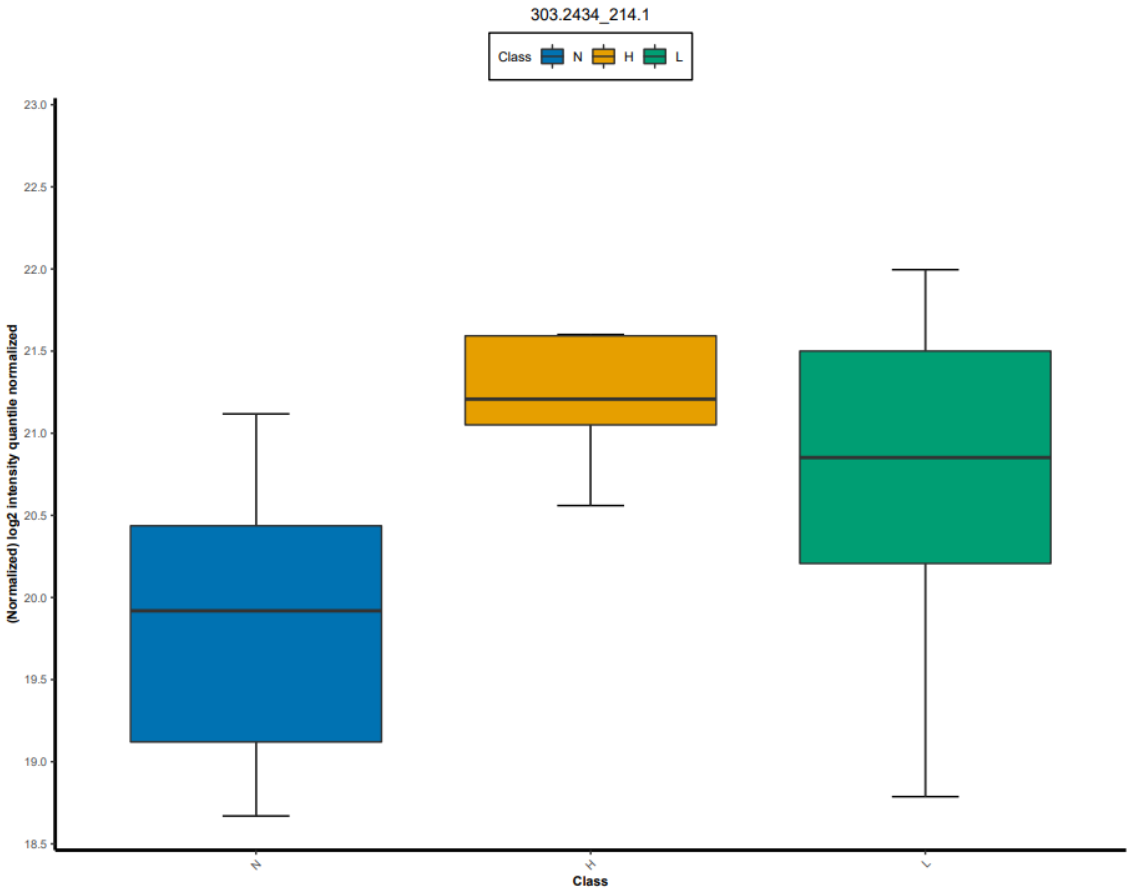

Figure 12. Box plot of Arachidonic Acid significant increase. In the high (H) and low (L) dosage treatment groups compared to the control group (N), Arachidonic Acid has significantly increased. Arachidonic acid is the core metabolite that is the foundation of the inflammation pathways. There is slight overlap in the ranges. This feature was taken from the c18 negative male feature table.

m/z	Retention Time (sec)	Potential Metabolite Names	Pathways	m/z diff	Test Statistic	Confirmed
424.3419	23.3	Linoelaidyl carnitine	Carnitine shuttle	-2.00E-04	1.7	
1048.306	129.2	timnodonyl coenzyme A	De novo fatty acid biosynthesis, Carnitine shuttle	6.00E-04	-2.07	
665.5133	23.7	Diacylglycerol	Glycerophospholipid metabolism	0.003	2.59	
385.1292	76.5	S-Adenosyl-L-homocysteine	Glycerophospholipid metabolism, Methionine and cysteine metabolism	-0.0011	-1.61	Yes
398.3264	23.4	Hexadec-enoyl carnitine	Carnitine shuttle	-1.00E-04	1.19	
337.2373	25.7	Prostaglandin B1	Arachidonic acid metabolism, Prostaglandin formation from arachidonate, Leukotriene metabolism	-1.00E-04	2.57	
476.3732	23.7	Adrenyl carnitine	Carnitine shuttle	-2.00E-04	1.43	
376.2596	24.6	9'-carboxy-gama-chromanol	Vitamin E metabolism	-0.0018	-1.45	
297.2425	24.8	12(13)-EpOME	Linoleate metabolism	0.0023	2.88	
420.3103	24.3	stearidonyl carnitine	Carnitine shuttle	-5.00E-04	1.29	
302.3053	23.1	Sphinganine	Glycerophospholipid metabolism, Glycosphingolipid metabolism	0.0025	1.52	Yes
428.3735	23.9	stearoylcarnitine	Carnitine shuttle	-0.0013	1.32	

Table 1. Top annotations for potential metabolites using hilic positive male pup data. Listed are the top metabolites (with corresponding m/z and retention time) in important pathways for inflammation, lipid, and developmental pathways. “m/z diff” is the difference between the m/z from the data gathered and the exact mass of the actual compound. A positive test statistic represents an increase in the metabolite and a negative test statistic represents a decrease in the metabolite, after early life heavy metal exposure. Annotations were made by *mummichog* and manually. The mass-to-charge ratios and their corresponding retention times were compared between this list of potential metabolites and a reference list (from Dean Jones Laboratory). A strong match between the two lists allowed for a strong confirmation (marked in the confirmed column) of a changed abundance of the metabolite in the metabolome.

m/z	Retention Time (sec)	Potential Metabolite Names	Pathways	m/z diff	Test Statistic	Confirmed
179.0562	25.3	Galactose	Glycerophospholipid metabolism, Glycolysis and Gluconeogenesis	1.00E-04	-0.62	Yes
369.2284	22.1	thomboxane b2	Prostaglandin formation from arachidonate	2.00E-04	4.07	
353.2334	25	11,12,15-THETA	Prostaglandin formation from arachidonate, Arachidonic acid metabolism	1.00E-04	2.88	
335.2228	32.8	Prostaglandin B1	Prostaglandin formation from arachidonate, Arachidonic acid metabolism, Leukotriene metabolism	0	4.25	
377.27	76.1	2-Arachidonoylglycerol	Prostaglandin formation from arachidonate	3.00E-04	5.69	
315.1983	16.3	all-trans-18-Hydroxyretinoic acid	Vitamin A (retinol) metabolism	0.0013	3.18	
394.2601	211.2	Prostaglandin E2 ethanolamide	Prostaglandin formation from arachidonate	2.00E-04	7.02	
171.0065	19.8	sn-Glycerol 3-phosphate	Glycerophospholipid metabolism	1.00E-04	0.86	
242.0799	23.1	Cytidine	Pyrimidine metabolism	0.0017	-0.84	
333.2071	214.6	Prostaglandin A2	Prostaglandin formation from arachidonate	0	1.17	
227.0674	22.4	Deoxyuridine	Pyrimidine metabolism	1.00E-04	1.24	Yes
367.2149	149.5	Prostaglandin G2	Prostaglandin formation from arachidonate	0.0025	1.22	
267.0733	24.4	Inosine	Purine metabolism	-2.00E-04	-0.88	Yes
317.2131	19.4	Leukotriene A4	Arachidonic acid metabolism, Leukotriene metabolism	9.00E-04	1.32	
303.2434	214.1	eicosatetranoic acid (Arachidonic Acid)	Arachidonic acid metabolism, Fatty acid activation	0.01117	1.36	Yes
637.5397	245.3	Triacylglycerol	Glycerophospholipid metabolism	-0.0016	2.39	
699.502	221.7	Phosphatidate	Vitamin A (retinol) metabolism, Glycerophospholipid metabolism	0.005	2.2	

Table 2. Top annotations for potential metabolites using c18 negative male pup data. Listed are the top metabolites (with corresponding m/z and retention time) in important pathways for inflammation, lipid, and developmental pathways. “m/z diff” is the difference between the m/z from the data gathered and the exact mass of the actual compound. A positive test statistic represents an increase in the metabolite and a negative test statistic represents a decrease in the metabolite, after early life heavy metal exposure. Annotations were made by *mummichog* and manually. The mass-to-charge ratios and their corresponding retention times were compared between this list of potential metabolites and a reference list (from Dean Jones Laboratory). A strong match between the two lists allowed for a strong confirmation (marked in the confirmed column) of a changed abundance of the metabolite in the metabolome.

m/z	Retention Time (sec)	Potential Metabolite Names	Pathways	m/z_diff	Test Statistic	Confirmed
283.2056	24.2	Leukotriene A4	Leukotriene metabolism, Arachidonic acid metabolism	-1.00E-04	1.92	
331.2268	26.8	17alpha-Hydroxyprogesterone	C21-steroid hormone biosynthesis and metabolism	0	-2.17	Yes
351.2528	24.9	Tetrahydrocorticosterone	C21-steroid hormone biosynthesis and metabolism	-2.00E-04	5.82	
218.1386	34.5	propionyl-carnitine	Carnitine shuttle	-1.00E-04	-1.16	Yes
229.0473	47.3	(R)-5-Phosphomevalonate	Squalene and cholesterol biosynthesis	1.00E-04	-3.81	
291.2318	24.4	5beta-Dihydrotestosterone	C21-steroid hormone biosynthesis and metabolism, Androgen and estrogen biosynthesis and metabolism	-1.00E-04	2.36	

Table 3. Top annotations for potential metabolites using hilic positive female pup data. Listed are the top metabolites (with corresponding m/z and retention time) in important pathways for inflammation, lipid, and developmental pathways. "m/z diff" is the difference between the m/z from the data gathered and the exact mass of the actual compound. A positive test statistic represents an increase in the metabolite and a negative test statistic represents a decrease in the metabolite, after early life heavy metal exposure. Annotations were made by *mummichog* and manually. The mass-to-charge ratios and their corresponding retention times were compared between this list of potential metabolites and a reference list (from Dean Jones Laboratory). A strong match between the two lists allowed for a strong confirmation (marked in the confirmed column) of a changed abundance of the metabolite in the metabolome.

m/z	Retention Time (sec)	Potential Metabolite Names	Pathways	m/z_diff	Test Statistic	Confirmed
421.0751	17.1	Lactose 6-phosphate	Galactose metabolism, Sialic acid metabolism	-1.00E-04	1.88	
699.502	221.7	Phosphatidate	Glycerophospholipid metabolism, Vitamin A (retinol) metabolism	0.005	1.37	
171.0065	19.8	sn-Glycerol 3-phosphate	Glycerophospholipid metabolism	1.00E-04	0.94	
298.2751	166.9	Sphingosine	Glycerophospholipid metabolism, Glycosphingolipid metabolism	0	5.3	
258.0384	17.8	D-Glucosamine 6-phosphate	Hexose phosphorylation, Pentose phosphate pathway, Aminosugars metabolism	0	1.55	
214.0487	21.4	sn-glycero-3-Phosphoethanolamine	Glycerophospholipid metabolism	1.00E-04	-0.83	

253.0931	22.7	3-beta-D-Galactosyl-sn-glycerol	Glycerophospholipid metabolism, Galactose metabolism	2.00E-04	-1.75	
179.0562	25.3	Galactose	Glycerophospholipid metabolism, Glycolysis and Gluconeogenesis	1.00E-04	-0.73	Yes

Table 4. Top annotations for potential metabolites using c18 negative female pup data. Listed are the top metabolites (with corresponding m/z and retention time) in important pathways for inflammation, lipid, and developmental pathways. "m/z diff" is the difference between the m/z from the data gathered and the exact mass of the actual compound. A positive test statistic represents an increase in the metabolite and a negative test statistic represents a decrease in the metabolite, after early life heavy metal exposure. Annotations were made by *mummichog* and manually. The mass-to-charge ratios and their corresponding retention times were compared between this list of potential metabolites and a reference list (from Dean Jones Laboratory). A strong match between the two lists allowed for a strong confirmation (marked in the confirmed column) of a changed abundance of the metabolite in the metabolome.

References

1. Xie, M., et al., *Trends in prevalence and incidence of chronic respiratory diseases from 1990 to 2017.* Respiratory Research, 2020. **21**(1): p. 49.
2. Çakır Edis, E., *Chronic Pulmonary Diseases and COVID-19.* Turk Thorac J, 2020. **21**(5): p. 345-349.
3. Ferraro, V.A., S. Zanconato, and S. Carraro, *Impact of COVID-19 in Children with Chronic Lung Diseases.* Int J Environ Res Public Health, 2022. **19**(18).
4. Agency for Toxic Substances and Disease Registry, *2007 CERCLA Priority List of Hazardous Substances*, A.f.T.S.a.D. Registry, Editor. 2009, CDC.
5. Woodruff, T.J., A.R. Zota, and J.M. Schwartz, *Environmental Chemicals in Pregnant Women in the United States: NHANES 2003–2004.* Environmental Health Perspectives, 2011. **119**(6): p. 878-885.
6. Gilbert-Diamond, D., et al., *Rice consumption contributes to arsenic exposure in US women.* Proceedings of the National Academy of Sciences, 2011. **108**(51): p. 20656-20660.
7. Gundacker, C. and M. Hengstschläger, *The role of the placenta in fetal exposure to heavy metals.* Wiener Medizinische Wochenschrift, 2012. **162**(9): p. 201-206.
8. Hendryx, M. and J. Luo, *Children's environmental chemical exposures in the USA, NHANES 2003–2012.* Environmental Science and Pollution Research, 2018. **25**(6): p. 5336-5343.
9. Hu X, L.K., Ma J, Orr M, Tran V, Go Y-M, Jones D, Marsit C, Wongtrakool C., , *Low Dose Heavy Metals Disrupt Metabolome and Inhibits Branching Morphogenesis in Fetal Lung Development.* P110 LUNG DEVELOPMENT: American Thoracic Society, 2021: p. A4308-A.
10. Metallo, C.M. and M.G. Vander Heiden, *Understanding metabolic regulation and its influence on cell physiology.* Mol Cell, 2013. **49**(3): p. 388-98.
11. Lee, H.A., et al., *The association between metabolic components and markers of inflammatory and endothelial dysfunction in adolescents, based on the Ewha Birth and Growth Cohort Study.* PLOS ONE, 2020. **15**(5): p. e0233469.
12. Fu, Z. and S. Xi, *The effects of heavy metals on human metabolism.* Toxicol Mech Methods, 2020. **30**(3): p. 167-176.
13. Mazumder, D.N.G., *Arsenic and non-malignant lung disease.* Journal of Environmental Science and Health, Part A, 2007. **42**(12): p. 1859-1867.
14. U.S. Environmental Protection Agency, *National Primary Drinking Water Regulations*, U.S.E.P. Agency, Editor. 2023.
15. Guha Mazumder, D.N., *Chronic arsenic toxicity & human health.* Indian Journal of Medical Research, 2008. **128**(4).
16. Bernard, A., *Cadmium & its adverse effects on human health.* Indian Journal of Medical Research, 2008. **128**(4): p. 557-564.
17. Bernhoft, R.A., *Cadmium Toxicity and Treatment.* The Scientific World Journal, 2013. **2013**: p. 394652.
18. Hu, X., et al., *Environmental Cadmium Enhances Lung Injury by Respiratory Syncytial Virus Infection.* Am J Pathol, 2019. **189**(8): p. 1513-1525.
19. Hudson, K.M., S.M. Belcher, and M. Cowley, *Maternal cadmium exposure in the mouse leads to increased heart weight at birth and programs susceptibility to hypertension in adulthood.* Scientific Reports, 2019. **9**(1): p. 13553.
20. Ramsey, K.A., et al., *In utero exposure to arsenic alters lung development and genes related to immune and mucociliary function in mice.* Environ Health Perspect, 2013. **121**(2): p. 244-50.
21. Rydell-Törmänen, K. and J.R. Johnson, *The Applicability of Mouse Models to the Study of Human Disease.* Methods Mol Biol, 2019. **1940**: p. 3-22.
22. Krystel-Whittemore, M., K.N. Dileepan, and J.G. Wood, *Mast Cell: A Multi-Functional Master Cell.* Front Immunol, 2015. **6**: p. 620.

23. National Human Genome Research Institute. *Why Mouse Matters*. 2010; Available from: https://www.genome.gov/10001345/importance-of-mouse-genome.
24. Blais, E.M., et al., *Reconciled rat and human metabolic networks for comparative toxicogenomics and biomarker predictions.* Nature Communications, 2017. **8**(1): p. 14250.
25. Rodell, R., et al., *Use of High-Performance Liquid Chromatography-Mass Spectrometry (HPLC-MS) to Quantify Modified Nucleosides.* Methods Mol Biol, 2022. **2444**: p. 125-140.
26. Holcapek, M., L. Kolárová, and M. Nobilis, *High-performance liquid chromatography-tandem mass spectrometry in the identification and determination of phase I and phase II drug metabolites.* Anal Bioanal Chem, 2008. **391**(1): p. 59-78.
27. Thermo Fisher Scientific. *What is HPLC?* 2022; Available from: https://www.thermofisher.com/us/en/home/industrial/chromatography/chromatography-learning-center/liquid-chromatography-information/hplc-basics.html.
28. Buszewski, B. and S. Noga, *Hydrophilic interaction liquid chromatography (HILIC)--a powerful separation technique.* Anal Bioanal Chem, 2012. **402**(1): p. 231-47.
29. Ruiz-Angel, M.J., et al., *Performance of different C18 columns in reversed-phase liquid chromatography with hydro-organic and micellar-organic mobile phases.* Journal of Chromatography A, 2014. **1344**: p. 76-82.
30. Siuzdak, G., *The emergence of mass spectrometry in biochemical research.* Proceedings of the National Academy of Sciences, 1994. **91**(24): p. 11290-11297.
31. Liu, K.H., et al., *High-Resolution Metabolomics Assessment of Military Personnel: Evaluating Analytical Strategies for Chemical Detection.* J Occup Environ Med, 2016. **58**(8 Suppl 1): p. S53-61.
32. Taylor, M.J., J.K. Lukowski, and C.R. Anderton, *Spatially Resolved Mass Spectrometry at the Single Cell: Recent Innovations in Proteomics and Metabolomics.* J Am Soc Mass Spectrom, 2021. **32**(4): p. 872-894.
33. Zubarev, R.A. and A. Makarov, *Orbitrap Mass Spectrometry.* Analytical Chemistry, 2013. **85**(11): p. 5288-5296.
34. Raffaelli, A. and A. Saba, *Ion scanning or ion trapping: Why not both?* Mass Spectrometry Reviews. **n/a**(n/a): p. e21746.
35. Dunn, W.B., *Chapter two - Mass Spectrometry in Systems Biology: An Introduction*, in *Methods in Enzymology*, D. Jameson, M. Verma, and H.V. Westerhoff, Editors. 2011, Academic Press. p. 15-35.
36. Lu, X., L.-J. Ji, and J.-L. Chen, *Chapter Thirteen - Metabolomic Profiling of Neoplastic Lesions in Mice*, in *Methods in Enzymology*, L. Galluzzi and G. Kroemer, Editors. 2014, Academic Press. p. 261-273.
37. Anton Kaufmann, M.B., *Selecting the best Q Exactive Orbitrap mass spectrometer scan mode for your application.* Thermo Scientific, 2018.
38. Michalski, A., et al., *Mass Spectrometry-based Proteomics Using Q Exactive, a High-performance Benchtop Quadrupole Orbitrap Mass Spectrometer*.* Molecular & Cellular Proteomics, 2011. **10**(9): p. M111.011015.
39. Clish, C.B., *Metabolomics: an emerging but powerful tool for precision medicine.* Cold Spring Harb Mol Case Stud, 2015. **1**(1): p. a000588.
40. Uppal, K., et al., *Computational Metabolomics: A Framework for the Million Metabolome.* Chem Res Toxicol, 2016. **29**(12): p. 1956-1975.
41. Yu, T., et al., *apLCMS--adaptive processing of high-resolution LC/MS data.* Bioinformatics, 2009. **25**(15): p. 1930-6.
42. Uppal, K., et al., *xMSanalyzer: automated pipeline for improved feature detection and downstream analysis of large-scale, non-targeted metabolomics data.* BMC Bioinformatics, 2013. **14**(1): p. 15.
43. Moradzadeh, K., et al., *Analysis of time-course microarray data: Comparison of common tools.* Genomics, 2019. **111**(4): p. 636-641.

44. Li, S., et al., *Predicting Network Activity from High Throughput Metabolomics.* PLOS Computational Biology, 2013. **9**(7): p. e1003123.
45. Institute for Quality and Efficiency in Health Care (IQWiG) *What is an inflammation?* 2010.
46. Wang, B., et al., *Metabolism pathways of arachidonic acids: mechanisms and potential therapeutic targets.* Signal Transduction and Targeted Therapy, 2021. **6**(1): p. 94.
47. Emory University Division of Pulmonary Allergy and Critical Care Medicine, *Dual column/polarity liquid chromatography operation for 5 minute high-resolution metabolomics, Q-Exactive HF*. 2017, Clinical Biomarkers Laboratory: 615 Michael St. Ste. 225, Atlanta GA, 30322.
48. Chadeau-Hyam, M., et al., *Metabolic profiling and the metabolome-wide association study: significance level for biomarker identification.* J Proteome Res, 2010. **9**(9): p. 4620-7.
49. Committee on Nutritional Status During Pregnancy and Lactation of the Institute of Medicine in the United States, *Nutrition During Pregnancy: Part I Weight Gain: Part II Nutrient Supplements*. Vol. 17. 1990, Washington, DC: National Academies Press.
50. Herrera, E. and E. Amusquivar, *Lipid metabolism in the fetus and the newborn.* Diabetes Metab Res Rev, 2000. **16**(3): p. 202-10.
51. Otimenyin, S.O., *Chapter 15 - Antiinflammatory Medicinal Plants: A Remedy for Most Disease Conditions?*, in *Natural Products and Drug Discovery*, S.C. Mandal, V. Mandal, and T. Konishi, Editors. 2018, Elsevier. p. 411-431.
52. Ricciotti, E. and G.A. FitzGerald, *Prostaglandins and inflammation.* Arterioscler Thromb Vasc Biol, 2011. **31**(5): p. 986-1000.
53. Müller, L., *Consequences of cadmium toxicity in rat hepatocytes: Mitochondrial dysfunction and lipid peroxidation.* Toxicology, 1986. **40**(3): p. 285-295.
54. Thomas, D.J., *Arsenolysis and Thiol-Dependent Arsenate Reduction.* Toxicological Sciences, 2010. **117**(2): p. 249-252.
55. Tchounwou, P.B., et al., *Heavy metal toxicity and the environment.* Exp Suppl, 2012. **101**: p. 133-64.
56. Nigra, A.E., et al., *Inequalities in Public Water Arsenic Concentrations in Counties and Community Water Systems across the United States, 2006–2011.* Environmental Health Perspectives, 2020. **128**(12): p. 127001.
57. World Health Organization. Water, S. and T. Health, *Guidelines for drinking-water quality [electronic resource] : incorporating first addendum. Vol. 1, Recommendations*. 2006, World Health Organization: Geneva.
58. Nair, A.B. and S. Jacob, *A simple practice guide for dose conversion between animals and human.* J Basic Clin Pharm, 2016. **7**(2): p. 27-31.
59. Jara, E.A. and C.K. Winter, *Dietary exposure to total and inorganic arsenic in the United States, 2006–2008.* International Journal of Food Contamination, 2014. **1**(1): p. 3.
60. Xue, J., et al., *Probabilistic Modeling of Dietary Arsenic Exposure and Dose and Evaluation with 2003-2004 NHANES Data.* Environ Health Perspect, 2010. **118**(3): p. 345-50.
61. Krumsiek, J., et al., *Gender-specific pathway differences in the human serum metabolome.* Metabolomics, 2015. **11**(6): p. 1815-1833.
62. Bartman, C.M., A. Matveyenko, and Y.S. Prakash, *It's about time: clocks in the developing lung.* The Journal of Clinical Investigation, 2020. **130**(1): p. 39-50.
63. Kanehisa, M. and S. Goto, *KEGG: kyoto encyclopedia of genes and genomes.* Nucleic Acids Res, 2000. **28**(1): p. 27-30.
64. Weng, L., et al., *Presence of arachidonoyl-carnitine is associated with adverse cardiometabolic responses in hypertensive patients treated with atenolol.* Metabolomics, 2016. **12**: p. 160.