UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

LIQUID CHROMATOGRAPHY-TANDEM MASS SPECTROMETRY-BASED METABOLITE PROFILING IN THE HEARTS OF GERM-FREE AND CONVENTIONALLY-RAISED MICE

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

Chaoyi Wu
Norman,
Oklahoma 2020

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/326825

LIQUID CHROMATOGRAPHY-TANDEM MASS SPECTROMETRY-BASED METABOLITE PROFILING IN THE HEARTS OF GERM-FREE AND CONVENTIONALLY-RAISED MICE

A THESIS APPROVED FOR THE
DEPARTMENT OF CHEMISTRY AND
BIOCHEMISTRY

BY THE COMMITTEE CONSISTING OF

Dr. Laura-Isobel McCall, Chair

Dr. Si Wu

Dr. Yihan Shao

Dr. Krithi Sankaranarayanan

© Copyright by Chaoyi Wu 2020
All Rights Reserved.

Table of Contents

1. Introduction....**1**
2. Results....6
   2.1. Overall metabolite profiling between heart sections of CONV-R and germ-free mice....6
   2.2. Impact of colonization status on short-, mid-, and long-chain acylcarnitines....8
   2.3. Impact of colonization status on the ratio of oxidized to reduced glutathione....10
   2.4. Random Forest analysis....12
   2.5. Parallel Reaction Monitoring (PRM)....12
   2.6. Glutamine....12
   2.7. *m/z* = 177.13, neighbor molecule of dodecanoic acid....13
   2.8. *m/z* = 230.1, ergothioneine....15
   2.9. *m/z* = 588.11, neighbor molecule of adenosine 5' diphosphoribose....17
3. Discussion....18
4. Conclusion....19
5. Materials and Methods....20
   5.1. *In vivo* experiment....20
   5.2. Sample preparation for LC-MS/MS....21
   5.3. Liquid chromatography-tandem mass spectrometry....21
   5.4. Data analysis....23
   5.5. Amino acids standards MS data collection....25
   5.6. Parallel monitoring reaction....25
6. References....27
7. Supplementary Information....30

Abstract

The microbiota is the community of microorganisms living on and in a biological system. It has shown to play a role in a broad range of medical conditions, including cardiovascular diseases. Germ-free (GF) mice grew under the condition lacking all microorganisms, in contrast to conventionally raised (CONV-R) mice colonized with a diverse microbiota. GF mice play a key role as tools to reveal the the causal relationship between microbiome and disease. In particular, GF mice present significant cardiac functional defects compared to CONV-R mice. The project's goal was to determine the spatial impact of GF vs. CONV-R status on cardiac metabolism across cardiac regions. 4 metabolite molecules showed significant differences by Random Forest analysis based on untargeted liquid chromatography-tandem mass spectrometry data. The four molecules are neighborhood of ADP-ribose (*m/z* = 588.11), ergothioneine (*m/z* = 230.1), neighborhood of dodecanoic acid (*m/z* = 177.13), and L-glutamine (*m/z* = 130.05). (Neighborhood results were generated from molecular networking.) Parallel reaction monitoring (PRM) was applied to provide more precise quantification results. In most locations, peak abundance significant difference was mapping in the locations: right ventricle free wall and left atrium part between germ-free and CONV-R conditions. Ergothioneine showed a difference in the left ventricle free wall bottom section. The metabolites may relate to the presence of microbiome and contribute to reducing the risk of cardiovascular disease.

1. Introduction

The microbiome is the community of all microbes living inside or on an environment, including bacteria, fungi, protozoa, and viruses[1]. The human microbiome is the community of the human body. The sponge microbiome is the community living on a sponge, and so on. It has an influence on various aspects such as metabolism, nutrition, and immunity of the host[2,3], so it has aroused great interest in biomedical research. The microbiome evolves from birth to death in a host. The continuous evolution of the human microbiome after birth is influenced by host factors (such as innate and adaptive immune systems), external factors (such as diet, drug and toxin exposure), and diseases[4]. The gut microbiome attracts particular interest because of its evolution across the entire lifespan, its connection to health, and its potential role in the development of cardiovascular disease (CVD)[5].

With the increasing importance of the gut microbiome as a necessary condition to maintain the balance of health vs. disease, interest in the manipulation of the gut microbiome in the mouse model has increased accordingly. Germ-free mice are bred in isolators, which prevent entirely microbial contact so that they are not affected by any detectable microbe community. Louis Pasteur and his colleague first brought out this concept in 1885, and germ-free mice colonies were initially generated in the 1940s. Germ-free mice models enable the study of animals that

are completely free of microorganisms or colonized with known microorganisms[6–8].

Prior work by collaborators at the University of Wisconsin-Madison (Dr. Federico Rey group) showed that microbiota colonization of previously germ-free mice for two weeks had minor influences on the left ventricle's thickness (LV) walls while significantly reducing the LV end volume. The mass of the LV of colonized mice was significantly increased compared to germ-free animals. Colonization also significantly increased the LV ejection fraction and had minor influences on stroke volume and cardiac output.

Cardiovascular disease (CVD) is caused by genetic, environmental, or a combination of genetic and environmental sources. Although extensive investigations have been carried out to find causal genetic variations, such as large-scale GWAS, genetic determinants cause less than one-fifth of cardiovascular risk. Now the diverse microbes residing in human gut have overwhelmed the number of 100 trillion, which is far more than the human host cells. In the past several decades, researchers have observed that intestinal microbiota, as an active participant, plays an active role in the growth of atherosclerosis and the adverse complications of cardiovascular disease via endocrine activity[9].

The mammalian microbiota represents a critical function in the pathogenesis of many illnesses. Due to the latest genomics, proteomics, and metabolomics, it is now possible to study the microbial components and metabolic activities in detail. The microbiota plays an essential role in the formation of healthy individuals, which changes the composition or function of intestinal microbiota (maladjustment), has specific harmful microorganisms (pathogenic bacteria), or lack of protective microorganisms may lead to immune and metabolism mediated diseases and cancers[10]. Due to the mechanism involved, these diseases may not be limited to digestive tract diseases[11].

Metabolomics is the scientific study of small molecules, commonly known as metabolites, such as ketone body β-hydroxybutyrate (BHB)[12], which is synthesized through fatty acids in the liver. When the glucose supply is too low to meet the body's energy needs, such as long-term exercise, starvation, or lack of dietary carbohydrates, BHB is an important energy carrier to the peripheral tissues from the liver. In 1987, Horrobin and his research group found metabolites of linoleic acid could decrease plasma cholesterol and then prevent cardiovascular disease[13].

The comprehensive collection of these metabolites within cells, biofluids, tissues, or organisms is known as metabolomics. Metabolite profiling can provide an instant snapshot of the physiology of the cell. Therefore, metabolomics can directly provide a functional plot of the organism's physiological state [14,15].

Liquid chromatography with tandem mass spectrometry (LC-MS/MS) is one of the most powerful commercial analytical techniques. Liquid chromatography provides separating power, and mass analysis is finished with high sensitivity and selectivity of detection using a mass spectrometer. An LC-MS/MS's general composition includes a sample injection unit, an ionization interface, a mass analyzer, an ion detector, and a data analysis module for output. According to different polarities, analytes are separated in liquid chromatography, causing different interaction times with the stationary phase and the mobile phase. Analytes are ionized at the interface between LC and MS systems, electrospray ionization (ESI) in this experiment[16]. Under high voltage, the mobile phase is transferred into an aerosol. The same charge on the liquid droplets expels each other. The solvent is evaporated in this process, and charged analytes are left and transmitted into the mass analyzer for subsequent analysis[16].

The orbitrap combines the functions of analyzer and detector. With two outer electrodes and a central electrode, analyte ions are first trapped and then do a waving movement along the central electrode's axis. Ions move forward and backward following angular frequency, which has the following relation to their mass-to-charge ratios m/z and $\omega = \sqrt{k/(m/z)}$. A field inside the electrode reduction will cause ions ejection [17].

The best conditions for mass spectrometry acquisition depend on the device, including ion source and analyzer parameters, in addition to all the techniques

discussed earlier. The data is usually collected in full scan mode for an untargeted test, in which the instrument is set to scan the full mass spectrum from 50 to 1200 *m/z*.

Data pre-processing was finished on software MZmine2 version 2.40.1[18]. It has the feature of the ability to handle large file sizes of raw data. Standard processing procedures include the following steps: (1) peak detection for noise removal according to noise ratio (S/N), (2) chromatography builder and deconvolution for peak extraction, in which specific algorithms are selected under different circumstances, (3) isotope grouping for merging peaks with 1 *m/z* difference, (4) feature alignment to focus on the valid molecule peaks with *m/z* and retention time parameters. All pre-processing data procedures are aiming for convincing mass spectrometry analysis results for metabolites annotation[19].

Our project's most crucial dereplication method is molecular networking based on Global Natural Products Social Molecular Networking (GNPS)[20]. It contributes to annotate known metabolite results from complex biological extract analytes rapidly. The chemical similarity of MS/MS data is the main focus of helping organize molecular networking. The MS / MS spectra of mixtures of natural products and known standards, synthetic compounds, or organisms with good characteristics can be successfully reproduced using molecular networks, which should be organized into reliable databases. This method can be adapted to different ionization platforms and correlate MS / MS data from environmental

ionization, direct injection, and LC-based methods. Molecular networks can annotate known molecules from complex mixtures and capture related analogs, which is a challenge for many other annotation strategies[21].

We apply liquid chromatography-mass spectrometry-based molecular networking to germ-free and conventional mice's hearts section samples to illustrate significant profiling differences. Four metabolites, glutamine, ergothioneine, and neighborhood molecules of dodecanoic acid and adenosine 5' diphosphoribose, showed statistically significant difference by Random Forest analysis and PRM quantification validation. These molecules followed our hypothesis that metabolites would reduce cardiovascular disease risk, but this needs further validation.

2. Results

The principal coordinate analysis was applied to provide a snapshot for overall molecular diversity mapping across mice heart sections and conditions. Moreover, a machine learning method, the random forest, was used to pick up metabolic features from global natural product social (GNPS) molecular networking platform annotated results.

2.1. Overall metabolite profiling between heart sections of CONV-R and germ-free mice

Metabolites from six heart sections from 8 GF and 5 conventional 6mice were processed with aqueous and organic extraction and analyzed by LC-MS/MS. To

compare the overall metabolites profile between different heart sections and conditions, PCoA (principal coordinate analysis) was performed on MZmine2-processed LC-MS/MS results. Relative higher differences were shown between heart sections (Figure 1a, PERMANOVA p=0.001 R^2=48.89%), driven by the difference between the atrium and the other sections. No overall impact of germ-free vs. colonized condition was observed (Figure 1b. PERMANOVA p=0.786 R^2=5.152%). The limitation of PCoA is that an obvious mapping distinction appears only when there are large differences, without determining the specific molecules driving the impact of the heart section and colonization conditions.

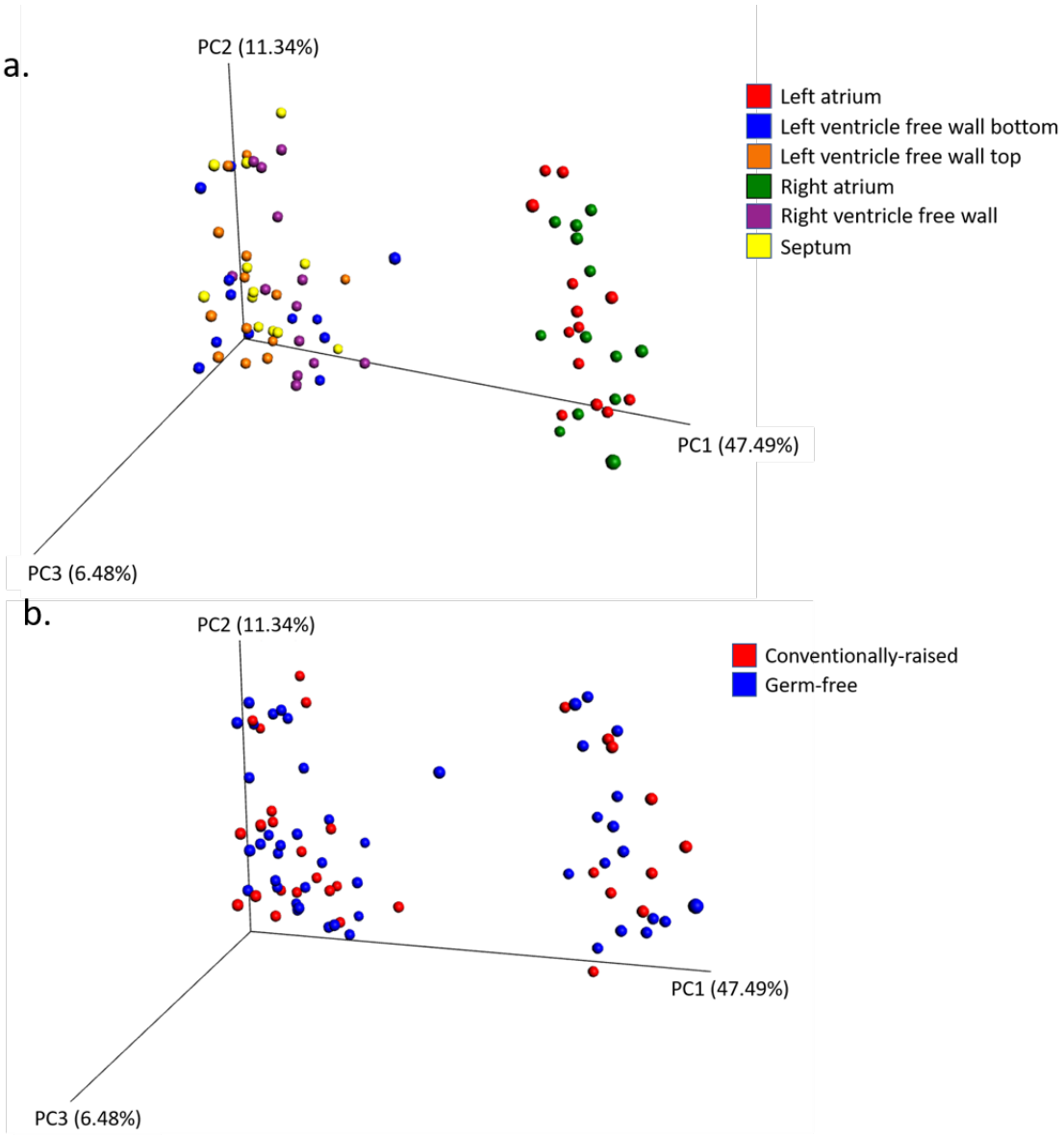

Figure 1. Principal coordinate analysis (PCoA) of comprehensive mapping to show differences between sections and conditions.

2.2. Impact of colonization status on short-, mid-, and long-chain acylcarnitines

Acylcarnitines, which are converted from carnitine by enzymes, play an important role as substrates in fat catabolism[22]. We studied acylcarnitines between CONV-R and germ-free conditions because of their role in fat

metabolism and other energy metabolism. Acylcarnitines with spectral annotation in GNPS (Global Natural Product Social Networking platform) were selected. The MS/MS spectra of all selected ones' neighboring molecules in molecular networking were visually examined to confirm the presence of a diagnostic peak at *m/z*=85.0288 with matching in MS/MS spectral fragment pattern. Different carbon chain lengths divided acylcarnitines into three groups: short-chain, medium length chain, and long-chain acylcarnitines. The normalized peak intensity of three acylcarnitines groups was summed up and analyzed by boxplot (Figure 2a-d.). As shown in figures, the left ventricle free wall had an apparent higher medium length acylcarnitines in CONV-R than germ-free mice. However, this was driven by an outlier in the CONV-R group. Acylcarnitines did not show any difference statistically by heart section and colonization condition.

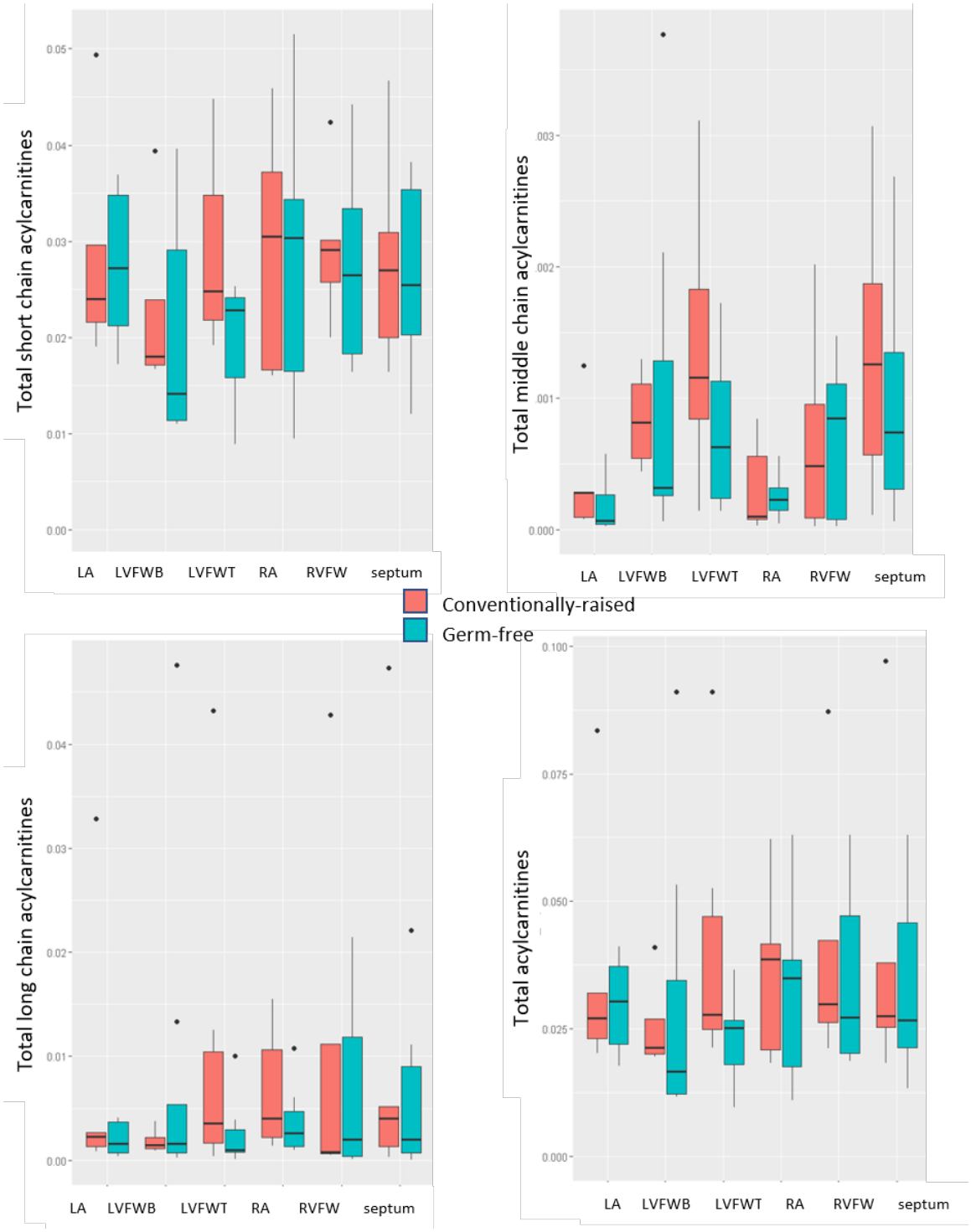

Figure 2. boxplots of acylcarnitine relative peak intensity between different sections and conditions.

2.3. Impact of colonization status on the ratio of oxidized to reduced glutathione

Glutathione has an ability on antioxidation, which can protect cells from being damaged by reactive oxygen species[23]. The glutathione groups were found to follow the same processing as described above for acylcarnitine. P values were applied to statistically analyze the existence of significant differences between different regional sections and conditions. However, there was no statistically significant difference shown in data.

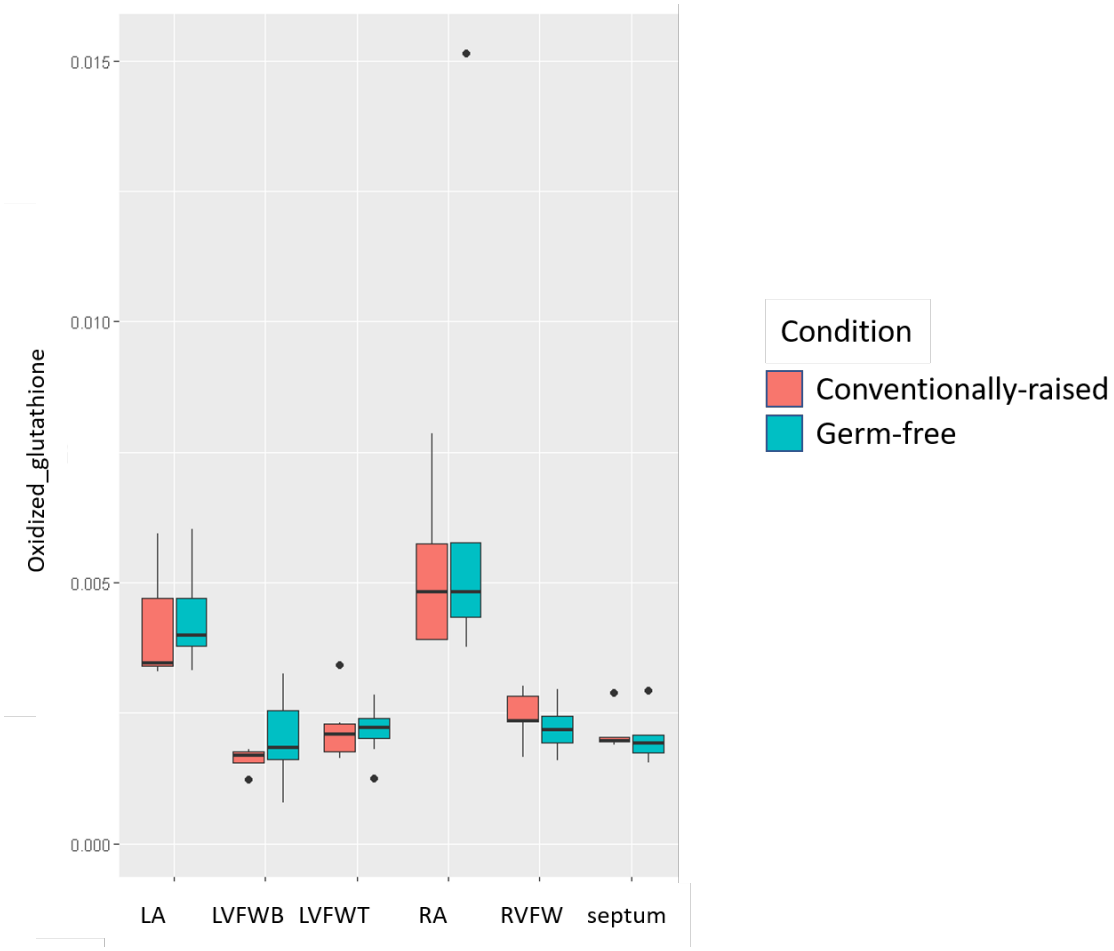

Figure 3. Boxplot of glutathione relative peak intensity between different sections and conditions in untargeted LC-MS analysis.

2.4. Random Forest analysis

Based on the limitation of PCoA, a machine learning approach called Random Forest (RF) was performed to identify specific metabolite molecules that differ between CONV-R and germ-free mice hearts. The top 20 molecular features in 6 different sections were listed according to MeanDecreaseAccuracy values. We also manually made boxplots (Figure S1.) for these molecules to exclude ones with outliers.

2.5. Parallel Reaction Monitoring (PRM)

The results from the untargeted LC-MS experiment had a limited precision for quantification. Furthermore, the Random Forest result also needed a more convincing validation. The Parallel Reaction Monitoring (PRM) experiment was designed to make up for the limitations. The following boxplots (Figure 4-7.) were generated from PRM results and normalized by the internal standard sulfadimethoxine added to PRM run with the same amount in LC-MS experiment[24].

2.6. Glutamine

L-glutamine is the human body's most concentrated amino acid. In different applications, it is used in human body as protein's biosynthesis and substance transportation in blood and defense against viruses and bacteria[25]. In random forest results, it showed differences in several sections between CONV-R and germ-free conditions. Its abundance was higher in CONV-R mice right ventricle free wall and lowered in the left

ventricle free wall bottom. Glutamine can drive the essential processes of vascular cells, including proliferation, replication, apoptosis, ageing, and extracellular matrix deposition, as a substrate for the synthesis of DNA, ATP, proteins and lipids. In addition, through stimulating the expression of heme oxygenase-1, heat shock protein and glutathione, glutamine plays a potent anti-inflammatory and anti-inflammatory function in circulation. Significantly, the risk factors for certain cardiovascular diseases, such as hypertension, hyperlipidemia, glucose intolerance, obesity and diabetes, may be minimized by glutamine[26].

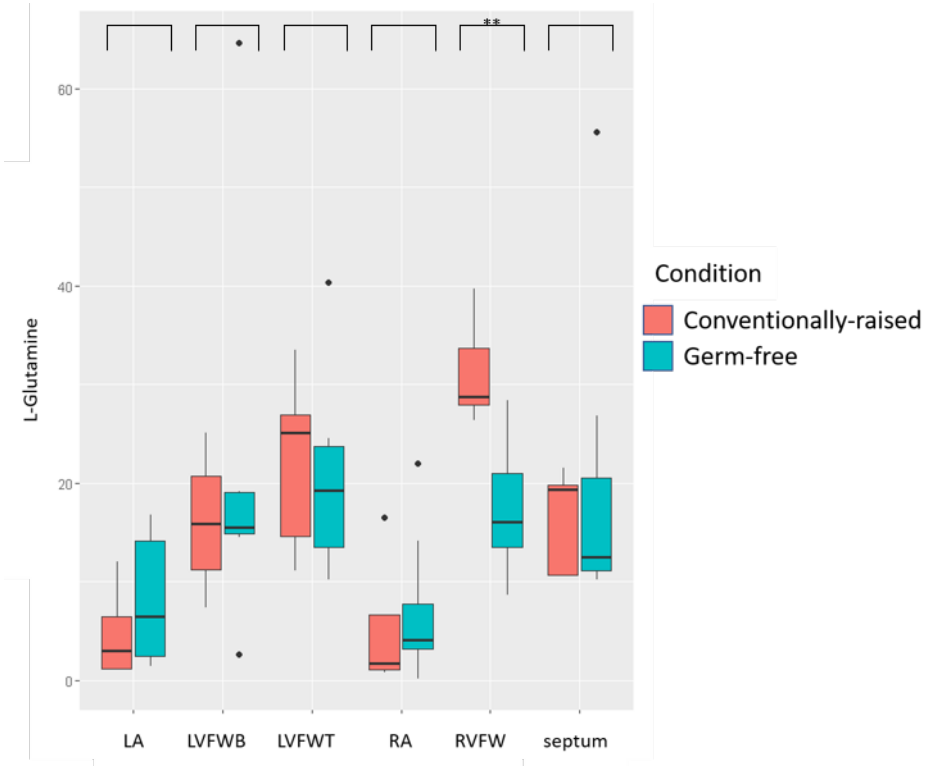

Figure 4. L-glutamine (*m/z* = 130.05) boxplot.

2.7. *m/z* = 177.13, neighbor molecule of dodecanoic acid

This molecule is a neighborhood of dodecanoic acid in molecular networking, which means a similar structure and main function groups. Dodecanoic acid, also known as lauric acid, is a saturated fatty acid usually delivered to the liver and transformed into energy instead of fat. That is because lauric acid is a medium-chain fatty acid covering C6-C12, different from long-chain fatty acids covering C14 and longer. In terms of lipoprotein metabolism, lauric acid higher concentration tends to increase low-density lipoprotein (LDL) levels. However, they have a more significant proportion of high-density lipoprotein (HDL) levels, resulting in decreased total cholesterol/HDL cholesterol levels. Lauric acid increases the ratio of total to HDL cholesterol, which may reduce cardiovascular disease[27,28].

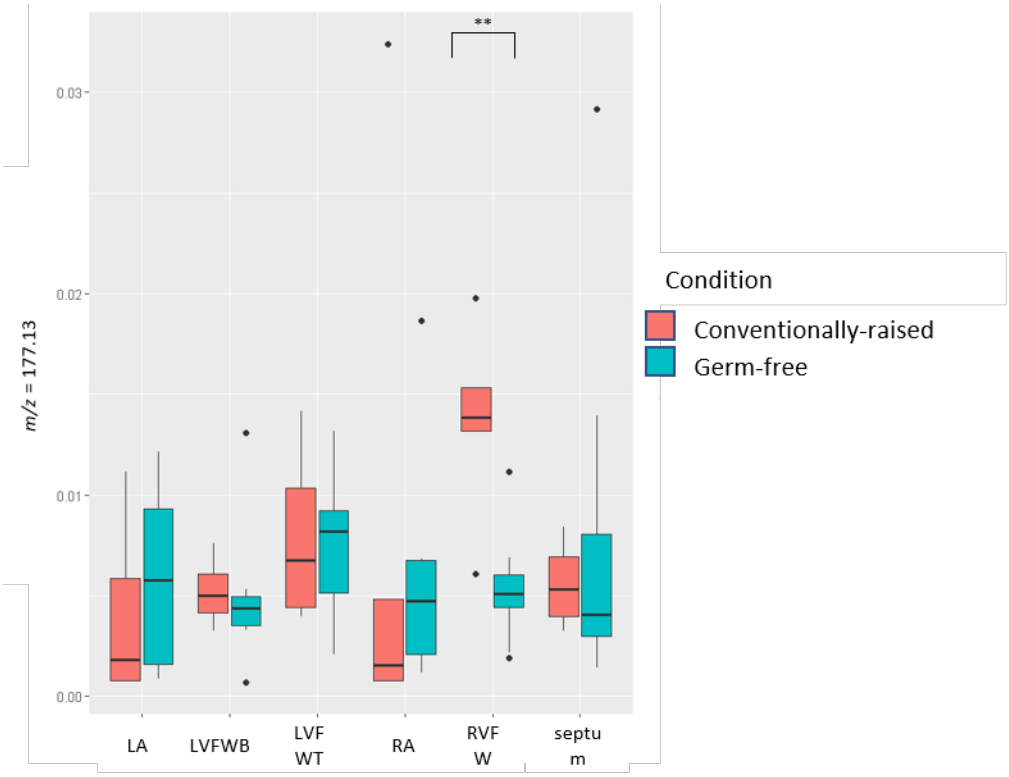

Figure 5. neighbor molecule of dodecanoic acid (*m/z* = 177.13) boxplot of peak abundances from PRM results between different conditions and heart sections.

2.8. *m/z* = 230.1. ergothioneine

Ergothioneine is an amino acid present mainly in mushrooms, and there is no way to synthesize it in the human body; it can also be found in oat bran. Ergothioneine accumulates to nanomoles in human and animal tissues through its high-affinity transporters. The daily diet is the primary source to accumulate it by a human. The chemical redox reaction of ergothioneine and alkylthiols, such as cysteine and glutathione, provides a new perspective for ergothioneine's role in oxidative damage.

Furthermore, it has just proved in recent work that higher ergothioneine could be a biomarker at a lower risk of cardiovascular disease[29,30]. In this last project, ergothioneine was statistically proved to indicate a significant association with a reduced risk of cardiovascular and CAD mortality. Multivariable Cox regression analyses were applied to verify the relation. Moreover, the previous research found that ergothioneine could be suggested to be a more controlled antioxidant because of its specific transporter with a function upregulated in inflammation[31].

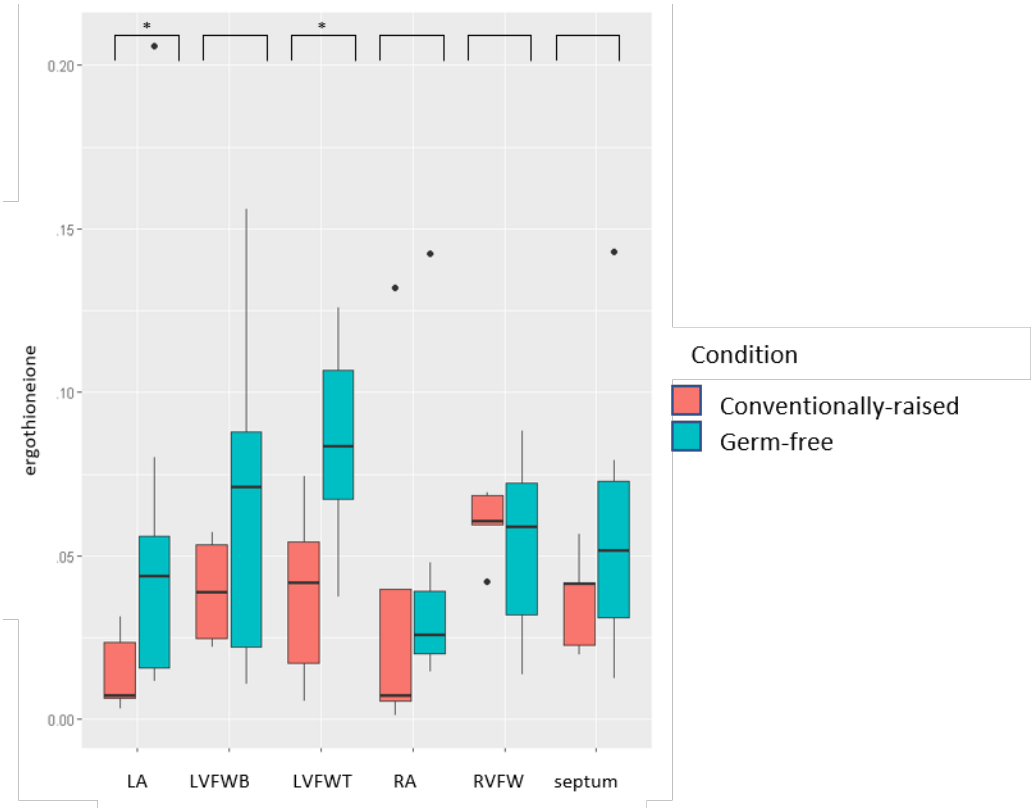

Figure 6. ergothioneine boxplot of peak abundance results from PRM quantification between different conditions and heart sections

2.9. *m/z* = 588.11, neighbor molecule of adenosine 5’ diphosphoribose

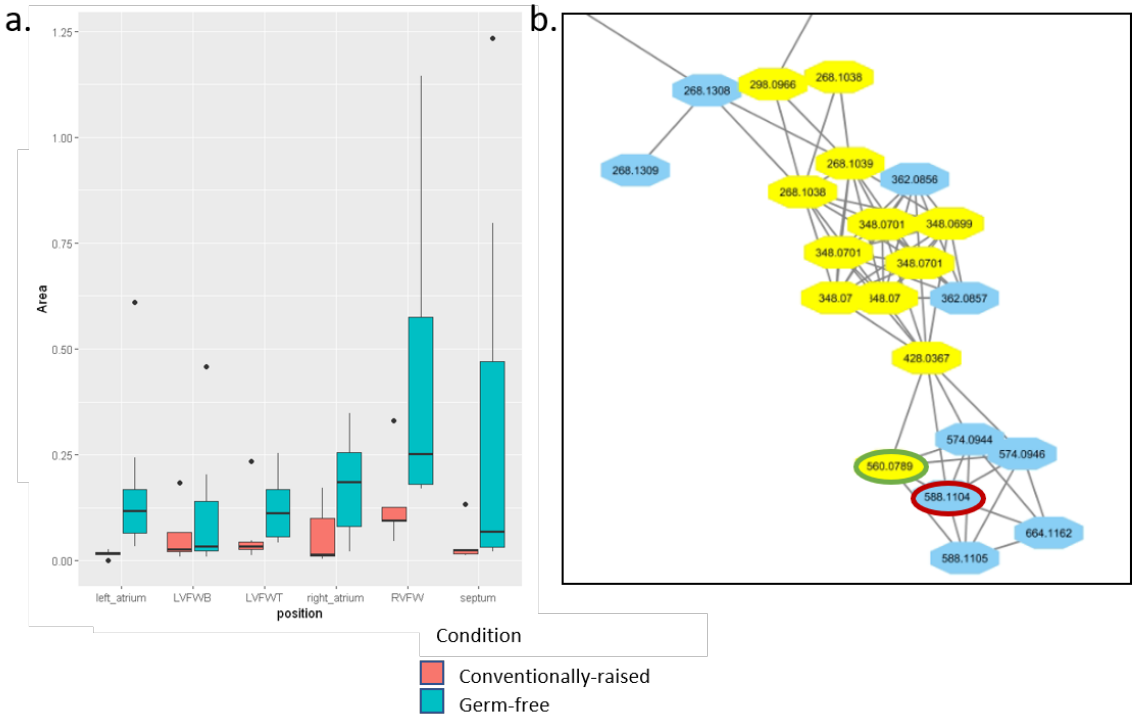

Figure 7a. ADP-ribose neighbor molecule PRM results in normalized boxplot between different conditions and sections. 7b. red circle labeled molecule (*m/z* = 588.11) and the green circle labeled annotated adenosine 5’-diphosphoribose in molecular networking. Each node on molecular networking represents one molecule. Edge between nodes represents similarity of the MS2 fragment patterns.

The *m/z* = 588.11 molecule has an annotated neighborhood, which is Adenosine 5’-diphosphoribose, also known as ADP-ribose. The short edge between them means they have a similar molecule structure and function groups. The most abundant isoform in the PARP molecule family is the polymerase 1 of ADP-ribose (PARP-1). PARP-1 was shown to accumulate evidence that certain reactive oxygen and nitrogen species that are formed in endothelial and cardiomyocyte cells during heart failure,

cardiovascular ageing, and cardiomyopathy may be activated. Recently, PARP was developed to be inhibitors and applied in clinical usage for various cardiovascular indications. It has proved to have a significant advantage in animal models of cardiovascular disorders[32]. In our project, the neighborhood molecule of ADP-ribose showed relatively higher abundance in the right ventricle free wall and left atrium sections in germ-free condition than the CONV-R condition.

3. Discussion

In this project, we analyzed metabolites in germ-free and CONV-R mice hearts regionally. Because there are some specific kinds of metabolites interested in researchers, we also made boxplots and statistical analysis on acylcarnitine with different chain lengths and glutathione with the oxidized glutathione[33,34]. Nevertheless, the two kinds of metabolites did not show a statistically significant difference in mice heart samples. With Random Forest analysis, we found 26 metabolite molecules with different amounts between germ-free and CONV-R mice hearts. We did PRM test and found the following results. Four molecules included L-glutamine, ergothioneine, and neighbor molecules of dodecanoic acid and ADP-ribose in molecular networking. These four metabolites all showed a statistically significant difference in the right ventricle free wall, left atrium, or two sections. Ergothioneine had a statistical difference in the left ventricle free wall top section. All four metabolites had a relationship

with a low risk of cardiovascular disease. L-glutamine can reduce the risk factors directly with the anti-inflammatory function. Lauric acid could increase the ratio of total cholesterol to HDL to affect the risk of CVD. The ergothioneine was also a biomarker that represents a lower risk of CVD[26].

Our major strength was the mice model design and metabolomic study combined with regional analysis. Other studies have performed mapping of metabolites. For example, for metabolite mapping analysis of brain, interleaved water reference was applied to combined with a volumetric MR spectroscopic imaging (MRSI) acquisition[35]. In contrast, our project specifically focused on the heart. Compared with other similar works before, our project combined the 3D structure of mice hearts to analyze the metabolite difference.

The main limitation is the sample amount. Some of the heart sections were too-small sample to reduce the sample loss in the preparation step. Furthermore, that is why we ran PRM to decrease the adverse effects. The results turned out to be optimized to a convincing quantification now.

4. Conclusion

We controlled the microbiome by utilizing of germ-free mice model and re-colonized them to generate our CONV-R mice as a comparison. Twenty-six metabolites were picked up and related to microbiome presence. The microbiome correlated with metabolite changes in different sections of mice's hearts. Parallel reaction monitoring (PRM) was a critical

quantification step to assess our relative peak abundance differences between conditions and heart sections. Four metabolites selected from PRM results interested us because they or their neighbors are known as molecules that could prevent cardiovascular disease.

Moreover, we could also learn the germ-free mice model's limitation because the lack of microbiome might induce an increased risk of certain diseases. To dig deeper based on our current research, microbiomes will be the primary study direction because the mechanism to generate our interesting metabolites is essential to learn. However, the project represents a step forward to more understanding the microbiome.

5. Materials and Methods

5.1. *In vivo* experiment

Half of 12-week-old GF C57/BL6 mice were colonized with microbiota from conventionally raised (CONV-R) mice, and the other half of mice were kept GF. 8 GF and 5 CONV-R mice were alive until 14 weeks. All animals were fed a chow diet and were echoed at 14 weeks of age to assess structural and functional differences. Mice were euthanized by CO_2 asphyxiation at 14 weeks. Hearts were cut into six sections: left atrium (LA), right atrium (RA), right ventricle free wall (RVFW), septum, left ventricle free wall top (LVFWT) and left ventricle free wall bottom (LVFWB). Heart samples were transported to McCall lab at the University of Oklahoma and stored at -80°C.

5.2. Sample preparation for LC-MS/MS

All solutions in this section were LC-MS grade. Heart samples were weighed and mixed with water with volume normalization by 50mg sample to 250 μL 50% methanol. For some specific samples with limited weights, 50% of methanol volumes were adjusted to normalize within 10 μL error. Samples were then homogenized in a TissueLyzer II (Qiagen) with a 5 mm stainless steel bead (Qiagen) at 25Hz for 5 min. Aqueous fractions were extracted from supernatants after centrifugation (14,800 rpm, 10min) at 4℃ and plated out up to 160-180 μL (according to different sample amounts) 96 well plates. The plate was then dried in a vacuum concentrator (SpeedVac, Savant SPD 111V, ThermoFisher Scientific). Pellets were resuspended with 3:1 dichloromethane: methanol (spiked with 0.5 μg/mL sulfachloropyridazine) and homogenized at 25 Hz 5 min, followed with centrifugation at 14800 rpm for 10 min for organic extraction. The supernatants were dried into and stored at -80℃ for later LC-MS/MS analysis.

5.3. Liquid chromatography-tandem mass spectrometry

The pellets were resuspended to 150 μL with 1:1 methanol: water, spiked with 0.5 μg/mL sulfadimethoxine (internal control). Liquid chromatography was performed on a ThermoScientific Vanquish UHPLC system equipped with 1.7 μm 100 Å Kintex C8 column (50 X 2.1 mm) (Phenomenex). The column effluent was analyzed using a Q Exactive Plus

(ThermoScientific) high-resolution mass spectrometer fitted with heated electrospray ionization (HESI) interface in positive ion mode. Calibration was done with commercial Calmix standard (ThermoFisher) consisting of MRFA, caffeine, Ultramark 1621, and n-butylamine in acetonitrile/MeOH/acetic acid solution.

Table 1. Instrument Parameters for Full MS/dd-MS^2

Properties of Full MS/dd-MS^2	
General	
Runtime	0 to 7.5 min
Polarity	Positive
Default Charge	1
Inclusion	-
Full MS	
Resolution	70,000
AGC target	1e6
Scan Range	70 to 1050 *m/z*
Maximum IT	246 ms
dd-MS^2	
Resolution	17,500
AGC target	2e5
Maximum IT	54 ms
Loop count	5
TopN	5
Isolation window	1.0 *m/z*
Fixed mass	-
(N)CE/stepped	NCE: 20, 40, 60
dd Settings	
Minimum AGC	8.00e3
Peptide Match	Preferred

Exclude isotopes	on
Dynamic exclusion	10.0 s

Table 2. Instrument Parameters for HESI source

Properties of HESI source	
General	
Sheath gas flow rate	35
Aux gas flow rate	10
Sweep gas flow rate	0
Spray voltage (|kV|)	3.80
Spray current (μA)	
Capillary temp. (°C)	320
S-lens RF level	50.0
Aux gas heater temp (°C)	350

5.4. Data analysis

Raw MS data were converted to mzXML format using MSconvert software[36,37]. The output mzXML files were processed in MZmine version 2.40.1[18] (see Table 3 for parameters). Principal coordinate analysis (PCoA, bray-curtis-faith distance metric) and Random Forest were performed in Jupyter notebook with R (http://jupyter.org). Metabolites feature annotation was finished on the Global Natural Products Social Networking (GNPS) platform[20] and visualized as molecular networking in Cytoscape[38,39].

Table 3. MZmine processing parameters.

Procedure	**Parameter**

Mass Peak Detection	MS level 1: Noise level	2.0E5
	MS level 2: Noise level	1.0E3
	Mass detector	Centroid
Chromatogram Builder	Scans	MS1
	Mass list	Masses
	Min time span (min)	0.1
	Min height	5.0E5
	m/z tolerance	1.0E-6 *m/z* (or 10 ppm)
Chromatogram Deconvolution	Algorithm	Local minimum search
	Chromatographic threshold	20%
	Search minimum in RT range (min)	0.20
	Minimum relative height	30.0%
	Minimum absolute height	5.0E5
	Min ratio of peak top/edge	1.19
	Peak duration range (min)	0.01-2.40
	m/z center calculation	Median
	m/z range (Da)	0.01
	RT range (min)	0.2
Isotopic Grouping	*m/z* tolerance	1.0E-6 *m/z* (or 10 ppm)
	Absolute RT tolerance (min)	0.05
	Monotonic shape	Yes
	Maximum charge	3
Feature Alignment	*m/z* tolerance	1.0E-6 *m/z* (or 10 ppm)
	Weight for *m/z*	10
	Absolute RT tolerance	0.5
	Weight for RT	1
Peak List Row Filter	Minimum peaks in a row	3
	RT (min)	0.20-7.51
	Keep only peaks with MS2 scan (GNPS)	Yes
	Reset the peak number ID	Yes

Gap filling	*m/z* tolerance	0.001*m/z* (or 5 ppm)
	Intensity tolerance	25%
	Absolute RT tolerance (min)	0.1
	RT correction	Yes

5.5. Amino acids standards MS data collection

Ten amino acid standards (Sigma-Aldrich) were selected, consisting of arginine, methionine, tryptophan, histidine, lysine, threonine, valine, isoleucine, leucine, and phenylalanine. 100 mM solution was prepared with 3:1 dichloromethane: methanol (spiked with 0.5 μg/mL sulfadimethoxine) for each standard. Serial dilutions were performed to make lower concentration solutions as 1 μM and 10 mM. Diluted solutions were plate out 150 μL in 96 well plates and analyzed with untargeted LC-MS/MS.

5.6. Parallel monitoring reaction

Top 37 metabolites were selected from boxplots combined with Random Forest[40] results, classifying based on the germ-free and conventionally-raised status between different heart sections. The transition list was generated to include *m/z* and RT of the 25 metabolites and 2 internal standards. The parallel monitoring reaction was applied as an LC-MS/MS run under the same method with the transition list uploaded for targeting. The output PRM raw dataset was imported into Skyline (SL) software for quantification results processing (see Table 4 for parameters).

Table 4. Instrumental parameters for PRM

Properties of PRM	
General	
Runtime	0 to 7.5 min
Polarity	Positive
Default Charge	1
Inclusion	-
PRM	
Resolution	17,500
AGC target	2e5
Maximum IT	54 ms
Isolation window	1.0 m/z
Fixed first mass	-
(N)CE / stepped nce	20, 40, 60

Table 5. PRM Skyline processing parameters.

Procedure		**Parameter**
Prediction	Precursor mass	Monoisotopic
	Product ion mass	Monoisotopic
	Collision energy	None
	Declustering potential	None
	Optimization library	None
	Compensation voltage	None
Filter	Precursor adducts	[M+H]
	Fragment adducts	[M+]
	Ion types	f, p
Library	Ion match tolerance	0.5 *m/z*
	If a library spectrum is available, pick its most intense ions.	Yes
	Pick	3 product ions
	From filtered ion adducts and types	Yes

Instrument	Min *m/z*	50
	Max *m/z*	1500
	Method match tolerance *m/z*	0.055
Full-scan	Isotope peaks included	Count
	Precursor mass analyzer	Orbitrap
	Peaks	3
	Resolving power	70 At: 200 *m/z*
	Isotope labeling enrichment	Default
	Acquisition method	Targeted
	Product mass analyzer	Orbitrap
	Resolving power	17,500 At: 200 *m/z*
	Use only scans within 5 minutes of MS/MS IDs.	Yes

References

1. Turnbaugh, P.J.; Ley, R.E.; Hamady, M.; Fraser-Liggett, C.M.; Knight, R.; Gordon, J.I. The Human Microbiome Project. *Nature* **2007**, *449*, 804–810, doi:10.1038/nature06244.
2. Donohoe, D.R.; Garge, N.; Zhang, X.; Sun, W.; O'Connell, T.M.; Bunger, M.K.; Bultman, S.J. The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. *Cell Metab.* **2011**, *13*, 517–526.
3. Kau, A.L.; Ahern, P.P.; Griffin, N.W.; Goodman, A.L.; Gordon, J.I. Human nutrition, the gut microbiome and the immune system. *Nature* **2011**, *474*, 327–336.
4. Davenport, E.R.; Sanders, J.G.; Song, S.J.; Amato, K.R.; Clark, A.G.; Knight, R. The human microbiome in evolution. *BMC Biol.* **2017**, *15*, 1–12.
5. Jie, Z.; Xia, H.; Zhong, S.-L.; Feng, Q.; Li, S.; Liang, S.; Zhong, H.; Liu, Z.; Gao, Y.; Zhao, H. The gut microbiome in atherosclerotic cardiovascular disease. *Nat. Commun.* **2017**, *8*, 1–12.
6. Wostmann, B.S. The germfree animal in nutritional studies. *Annu. Rev. Nutr.* **1981**, *1*, 257–279.

7. Yi, P.; Li, L. The germfree murine animal: an important animal model for research on the relationship between gut microbiota and the host. *Vet. Microbiol.* **2012**, *157*, 1–7.

8. Wostmann, B.S. *Germfree and gnotobiotic animal models: background and applications*; CRC Press, 1996; ISBN 084934008X.

9. Wang, Z.; Klipfell, E.; Bennett, B.J.; Koeth, R.; Levison, B.S.; DuGar, B.; Feldstein, A.E.; Britt, E.B.; Fu, X.; Chung, Y.-M. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **2011**, *472*, 57–63.

10. Tlaskalova-Hogenova, H.; Vannucci, L.; Klimesova, K.; Stepankova, R.; Krizan, J.; Kverka, M. Microbiome and colorectal carcinoma: Insights from germ-free and conventional animal models. *Cancer J. (United States)* **2014**, *20*, 217–224, doi:10.1097/PPO.0000000000000052.

11. Blaut, M.; Clavel, T. Metabolic diversity of the intestinal microbiota: implications for health and disease. *J. Nutr.* **2007**, *137*, 751S-755S.

12. Newman, J.C.; Verdin, E. β-Hydroxybutyrate: A Signaling Metabolite. *Annu. Rev. Nutr.* **2017**, *37*, 51–76, doi:10.1146/annurev-nutr-071816-064916.

13. Horrobin, D.F.; Huang, Y.S. The role of linoleic acid and its metabolites in the lowering of plasma cholesterol and the prevention of cardiovascular disease. *Int. J. Cardiol.* **1987**, *17*, 241–255, doi:10.1016/0167-5273(87)90073-8.

14. Merchant, A.; Richter, A.; Popp, M.; Adams, M. Targeted metabolite profiling provides a functional link among eucalypt taxonomy, physiology and evolution. *Phytochemistry* **2006**, *67*, 402–408.

15. Villas-Bôas, S.G.; Mas, S.; Åkesson, M.; Smedsgaard, J.; Nielsen, J. Mass spectrometry in metabolome analysis. *Mass Spectrom. Rev.* **2005**, *24*, 613–646.

16. Takáts, Z.; Wiseman, J.M.; Gologan, B.; Cooks, R.G. Mass spectrometry sampling under ambient conditions with desorption electrospray ionization. *Science (80-.).* **2004**, *306*, 471–473.

17. Hu, Q.; Noll, R.J.; Li, H.; Makarov, A.; Hardman, M.; Graham Cooks, R. The Orbitrap: a new mass spectrometer. *J. mass Spectrom.* **2005**, *40*, 430–443.

18. Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **2010**, *11*, doi:10.1186/1471-2105-11-395.

19. Ivanisevic, J.; Want, E.J. From samples to insights into metabolism: Uncovering biologically relevant information in LC- HRMS metabolomics data. *Metabolites* **2019**, *9*, 1–30, doi:10.3390/metabo9120308.

20. Wang, M.; Carver, J.J.; Phelan, V. V.; Sanchez, L.M..; Garg, N..; Peng, Y..; Nguyen, D.D..; Watrous, J..; Kapono, C.A..; Luzzatto-Knaan, T..; et al. Sharing and community curation of mass spectrometry data with GNPS. *Nat. Biotechnol.* **2017**, *34*, 828–837, doi:10.1038/nbt.3597.Sharing.

21. Wang, M.; Carver, J.J.; Phelan, V. V.; Sanchez, L.M.; Garg, N.; Peng, Y.; Nguyen, D.D.; Watrous, J.; Kapono, C.A.; Luzzatto-Knaan, T.; et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34*, 828–837, doi:10.1038/nbt.3597.

22. Reuter, S.E.; Evans, A.M. Carnitine and acylcarnitines. *Clin. Pharmacokinet.* **2012**, *51*, 553–572.

23. Meister, A.; Anderson, M.E. Glutathione. *Annu. Rev. Biochem.* **1983**, *52*, 711–760.

24. Peterson, A.C.; Russell, J.D.; Bailey, D.J.; Westphall, M.S.; Coon, J.J. Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Mol. Cell. Proteomics* **2012**, *11*, 1475–1488, doi:10.1074/mcp.O112.020131.

25. Ziegler, T.R.; Benfell, K.; Smith, R.J.; Young, L.S.; Brown, E.; Ferrari-Baliviera, E.; Lowe, D.K.; Wilmore, D.W. Safety and metabolic effects of L-glutamine administration in humans. *J. Parenter. Enter. Nutr.* **1990**, *14*, 137S-146S.

26. Durante, W. The emerging role of l-glutamine in cardiovascular health and disease. *Nutrients* **2019**, *11*, 2092.

27. Mensink, R.P.; Zock, P.L.; Kester, A.D.M.; Katan, M.B. Effects of dietary fatty acids and carbohydrates on the ratio of serum total to HDL cholesterol and on serum lipids and apolipoproteins: A meta-analysis of 60 controlled trials. *Am. J. Clin. Nutr.* **2003**, *77*, 1146–1155, doi:10.1093/ajcn/77.5.1146.

28. Vergroesen, A.J. Dietary fat and cardiovascular disease: possible modes of action of linoleic acid. *Proc. Nutr. Soc.* **1972**, *31*, 323–329.

29. Smith, E.; Ottosson, F.; Hellstrand, S.; Ericson, U.; Orho-Melander, M.; Fernandez, C.; Melander, O. Ergothioneine is associated with reduced mortality and decreased risk of cardiovascular disease. *Heart* **2020**, *106*, 691–697, doi:10.1136/heartjnl-2019-315485.

30. Servillo, L.; D'Onofrio, N.; Balestrieri, M.L. Ergothioneine antioxidant function: from chemistry to cardiovascular therapeutic potential. *J. Cardiovasc. Pharmacol.* **2017**, *69*, 183–191.

31. Halliwell, B.; Cheah, I.K.; Tang, R.M.Y. Ergothioneine–a diet-derived antioxidant with therapeutic potential. *Febs Lett.* **2018**, *592*, 3357–3366.

32. Pacher, P.; Szabó, C. Role of poly (ADP-ribose) polymerase 1 (PARP-1) in cardiovascular diseases: the therapeutic potential of PARP inhibitors. *Cardiovasc. Drug Rev.* **2007**, *25*, 235–260.

33. Guasch-Ferré, M.; Zheng, Y.; Ruiz-Canela, M.; Hruby, A.; Martínez-González, M.A.; Clish, C.B.; Corella, D.; Estruch, R.; Ros, E.; Fitó, M. Plasma acylcarnitines and risk of cardiovascular disease: effect of Mediterranean diet interventions. *Am. J. Clin. Nutr.* **2016**, *103*, 1408–1416.

34. Blankenberg, S.; Rupprecht, H.J.; Bickel, C.; Torzewski, M.; Hafner, G.; Tiret, L.; Smieja, M.; Cambien, F.; Meyer, J.; Lackner, K.J. Glutathione peroxidase 1 activity and cardiovascular events in patients with coronary artery disease. *N. Engl. J. Med.* **2003**, *349*, 1605–1613.

35. Maudsley, A.A.; Domenig, C.; Govind, V.; Darkazanli, A.; Studholme, C.; Arheart, K.; Bloomer, C. Mapping of brain metabolite distributions by volumetric proton MR spectroscopic imaging (MRSI). *Magn. Reson. Med. An Off. J. Int. Soc. Magn. Reson. Med.* **2009**, *61*, 548–559.

36. Chambers, M.C.; MacLean, B.; Burke, R.; Amodei, D.; Ruderman, D.L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; et al. A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* **2012**, *30*, 918–920, doi:10.1038/nbt.2377.

37. Kessner, D.; Chambers, M.; Burke, R.; Agus, D.; Mallick, P. ProteoWizard: Open source software for rapid proteomics tools development. *Bioinformatics* **2008**, *24*, 2534–2536, doi:10.1093/bioinformatics/btn323.

38. Smoot, M.E.; Ono, K.; Ruscheinski, J.; Wang, P.L.; Ideker, T. Cytoscape 2.8: New features for data integration and network visualization. *Bioinformatics* **2011**, *27*, 431–432, doi:10.1093/bioinformatics/btq675.

39. Paul Shannon, 1; Andrew Markiel, 1; Owen Ozier, 2 Nitin S. Baliga, 1 Jonathan T. Wang, 2 Daniel Ramage, 2; Nada Amin, 2; Benno Schwikowski, 1, 5 and Trey Ideker2, 3, 4, 5; 山本隆久; 豊田直平; 深瀬吉邦; 大森敏行 Cytoscape: A Software Environment for Integrated Models. *Genome Res.* **1971**, *13*, 426, doi:10.1101/gr.1239303.metabolite.

40. Pavlov, Y.L. Random forests. *Random For.* **2019**, 1–122, doi:10.1201/9780367816377-11.

Supplementary Information

Table S1. 25 metabolite molecules selected from Random Forest results.

m/z	**Retention time**	**Library hit (neighbor annotation)**
106.05	0.28	Serine
130.05	0.27	L-glutamine
134.04	1.73	Aspartic acid
135.08	2.27	Azelaic acid
136.05	0.26	Adenosine monophosphate
137.1	2.41	(2-tert-Butryl-p-quinone)
146.12	0.26	(Spermidine|N-(3-aminopropyl)butane-1,4-diamine)
147.08	0.27	Glutamine|2,5-diamino-5-oxopentanoic acid
148.06	0.27	alpha.-Guanidinoglutaric acid
150.06	0.27	Methionine
156.08	0.26	Histidine
175.12	0.25	Arginine
177.13	2.63	(Dodecanedioic acid)
182.08	0.29	Tyrosine
189.16	0.27	NEPSILON,NEPSILON,NEPSILON-TRIMETHYLLYSINE
205.1	0.7	L-tryptophan
230.1	2.17	Ergothioneine
251.04	0.53	(Guanidinoethyl sulfonate)
280.26	2.99	(13-Docosenamide, (Z)-)
482.42	3.03	(Linoleoyl carnitine)
588.11	0.5	(Adenosine 5'-diphosphoribose)
742.57	3.76	(1-hexadecyl-2-(9Z-octadecenoyl)-sn-glycero-3-phosphocholine)
748.53	3.67	(1-(1Z-Octadecenyl)-2-(4Z,7Z,10Z,13Z,16Z,19Z-

		docosahexaenoyl)-sn-glycero-3-phosphoethanolamine)
796.58	3.69	1-Heptadecanoyl-2-(5Z,8Z,11Z,14Z-eicosatetraenoyl)-sn-glycero-3-phosphocholine from NIST14
886.54	3.53	(1-Docosahexaenoyl-2-stearoyl-sn-glycero-3-phosphocholine)

Figure S1. Untargeted LCMS results boxplots of 25 molecules selected from Random Forest results.

