Prediction, Identification, and Quantification of Microbiota Derived Metabolites in Murine Gut

An honors thesis for the

Department of Chemical and Biological Engineering

Long Bin Pan

TUFTS UNIVERSITY

May 2013

Acknowledgements

First, I would like to thank the thesis committee members Dr. Kyongbum Lee and Dr. Matthew Panzer of Tufts University Department of Chemical and Biological Engineering for their continuous support and guidance. Dr. Lee, through his knowledge and advice, has allowed me to successfully coordinate and complete my thesis. His revisions of my presentations and methods have greatly developed my communication, organization, and experimental development skills. Dr. Panzer, through his teachings and guidance, has honed my research skills and provided me with the fundamentals to achieve in a laboratory setting.

I would also like to thank Gautham V. Sridharan and Sara Manteiga for their assistance in the use the LCMS. Gautham V. Sridharan has aided me in the troubleshooting of various laboratory machinery and the development of methods that are optimized for the detection of specific analytes. Sara Manteiga has provided me with support for problems involving the analytical methods described in this report.

Table of Contents

Abstract

Recent studies have shown that the gastrointestinal (GI) tract microbiota plays an important role in the modulation of human health and disease. Disruptions in microbiota composition, dysbiosis, directly correlate with inflammatory bowel diseases, chronic diseases, and cardiovascular diseases (Nyangale et al., 2012). Beneficial processes of the microbiota include providing immune functions, maintaining epithelial barrier integrity, extracting nutrient and vitamin from host diets, and carrying out metabolic biotransformation reactions that are not readily available to the host organism (Bien et al., 2013). Substrates that escape digestion in the upper GI tract are utilized by the GI microbiota to produce a complex variety of metabolites. Current processes of identifying the spectrum of metabolites in the GI tract include isolation and characterization of individual bacterial species and untargeted metabolomics approaches. However, these methodologies do not account for the community-wide interactions of bacterial species and the host organism. These approaches also do not differentiate between metabolites produced by the GI microbiota and those by the host. In this study, a targeted metabolomics approach was utilized to predict, identify and quantify a set of metabolites that are nonnative to the host organism. A novel *in silico* approach was used to predict a set of potential anti-inflammatory metabolites derived from amino acid sources. Analyses with tandem mass spectrometry and high performance liquid chromatography were performed to quantify the levels of predicted metabolites in order to validate the accuracy of the *in silico* probabilistic algorithm. AhR activation experiments were used to verify the anti-inflammatory properties of metabolites because studies have shown that AhR, a ligand activated transcription factor, plays an important role in intestinal immune functions (Bjeldanes et al., 1991). Once anti-inflammatory metabolites have been determined and analyzed, they can be used for drug discovery and therapeutics.

Introduction

Gut Microbiota Diversity and Functions

Bacteria communities exist on the surfaces of all parts of the human body. In an adult human gastrointestinal tract, there are approximately 10^{14} bacteria from ~1000 identified species. Collectively, this diverse community of bacteria in the human gut is known as microbiota or commensal flora (Bien *et al*. 2013). Nucleic acid based studies have shown that the intestinal microbiota is comprised of several major bacterial phylums, *Firmicutes, Bacteriodetes, Actinobacteria, Proteobacteria, Fusobacteria, Verrucomicrobia, Cyanobacteria* and *Spirochaetes. Firmicutes* and *Bacteriodetes* phylums were identified as the dominant bacterial groups (Rajilic-Stojanovic et al. 2007).

Significant changes in the diversity of microbial species in the gut microbiota can occur based on the age, diet, and health status of the host organism. The proportions and composition of *Bacteriodetes* appeared to be relatively consistent from individual to individual while those of *Firmicutes* exhibited significant variations. A combination of genetic factors and environmental factors determine the unique set of bacterial species in each individual. There has been increasing evidence that shifts in microbiota compositions (dysbiosis) are correlated to inflammatory bowel diseases (IBD) such as Crohn's disease and colitis. One of the major functions of the microbiota is to provide a mucosal defense system in which commensal microbial species suppress the growth and colonization of pathogenic and pro-inflammatory bacteria. This mucosal defense mechanism against pathogenic bacteria is known as the barrier effect or colonization resistance. *Clostridium difficile* is an infectious Gram-positive spore-forming *Bacillus* microbe in the phylum *Firmicutes* that is known to cause IBD in large concentrations. With dysbiosis, the decrease in quantity of specific anaerobes results in the

inability to resist the growth and colonization of *C. difficile.* Toxin expression from *C. difficile* causes inflammation and damage to the gut musoca. Prolonged colonization and growth of *C. difficile* result in colitis, toxic megacolon, and in severe cases death (Bien et al. 2013). Similarly, an analysis performed by Chassaing and coworkers discovered that patients with Crohn's disease and colitis exhibit a higher concentration of pro-inflammatory bacteria while having fewer bacteria with anti-inflammatory properties. For example, fecal samples from mice with ulcerative colitis were reported to have reduced numbers of *Bacteriodes fragilis* that prevents colitis induction by *Helicobacter hepaticus.* In a healthy individual, anti-inflammatory bacteria interact with the host musoca to achieve homeostasis by mediating the pro-inflammatory effects of harmful bacteria (Figure 1). However, in dysbiosis, the decrease in anti-inflammatory bacteria and increase in harmful bacteria leads to a damaged epithelial layer, bacterial translocation, and chronic inflation. A fluorescence in situ hybridization analysis discovered that 30% of mucosal biopsy specimens from individuals with IBD showed the presence of bacteria that can penetrate the mucosal layer compared to 3% from those of healthy individuals (Chassaing et al, 2011).

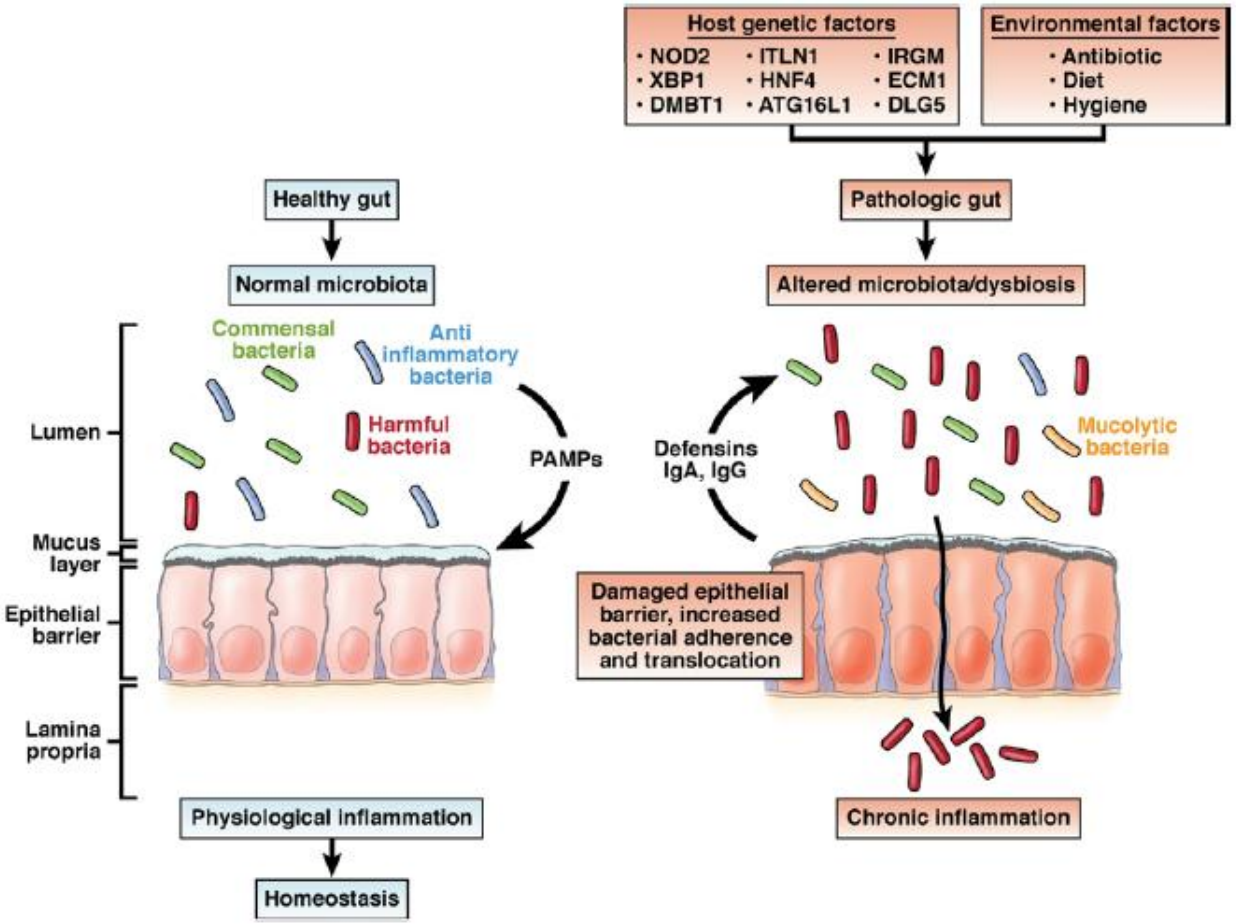


Figure 1: The diagram above depicts the interactions between the microbiota and the epithelial layer of the host organism. In the left panel, normal microbiota in healthy individuals interacts with the human mucosa to achieve homeostasis. In the right panel, quantitative and qualitative disturbances in the human microbiota results in chronic inflammation (Chassaing et al. 2011).

In addition to gut diseases, dysbiosis of the microbiota is also associated with a variety of chronic diseases and cardiovascular disorders such as colorectal cancer (CRC), obesity, type 2 diabetes, and heart disease (Chassaing et al., 2011; Mujico et al., 2012; Tang et al., 2013). Studies have shown that high levels of trimethylamine N-oxide (TMAO) in the blood are associated with an increased risk of heart attack and stroke. Trimethylamine N-oxide is a metabolite produced in the microbiota metabolism of choline. Mammals digest phosphatidylcholine (lecithin) from their diet to produce choline. Tang and coworkers

administered oral antibiotics to healthy individuals to promote dysbiosis in the human GI tract. It was discovered that individuals with induced dysbiosis exhibited a lower level of TMAO in their plasma samples (Tang et al., 2013). This is evidence that intestinal microbial species directly modulate the production of TMAO in the human GI tract. In a study conducted by Mujico and coworkers, it was determined that mice with high-fat diet induced obesity exhibited an increase in the phylum *Bacteriodetes* and a decrease in the phylum *Firmicutes* (Mujico et al., 2012).

The interest in gut microbiota stems from the symbiotic relationship between host cells and bacterial cells in the human GI tract. Beneficial characteristics of the gut microbiota include mucosal immune function, epithelial barrier integrity, nutrient and vitamin extraction from host diets, and metabolic activity that are not readily available to the host organism. Studies have shown that the human microbiome has 100 times more genes than those present in the human genetic sequence (Bien et al. 2013). Many of these genes encode for enzymes capable of converting substrates into anti-inflammatory metabolites and beneficial compounds. For example, microbes in the *C. coccoides* group are responsible for the production of short-chain fatty acids such as butyrate. In addition to butyrate's use as fuel for the host's colonocytes, it also has anti-inflammatory properties that inhibit the transcription nuclear factor κB (NFκB) involved in the secretion of pro-inflammatory cytokines (Pryde et al. 2002). Other metabolic functions encoded in the microbiota DNA allow the conversion of substrates that escape digestion in the upper GI tract. For example, in healthy individuals, 6-35% of ingested egg proteins escapes digestion in the upper GI tract and is later metabolized by the gut microbiota. It is estimated that approximately 6-18 g/day of proteins and amino acids reach the human colon and become ready for bacterial fermentation (Nyangale et al., 2012). Although metabolic activity occurs along the entire colon, it increases at the distal colon, an area of the large intestine that includes the

descending colon, the sigmoid colon, and the rectum (Figure 2). In these sections of the colon, the metabolites formed alter the pH condition from 5.5 in the cecum to approximately 6.6-6.9 (Cummings et al., 1991).

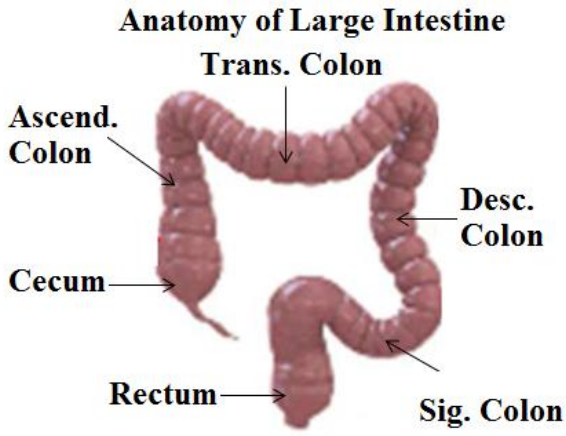


Figure 2: The diagram above illustrates the different sections of the human large intestine (Nyangale et al., 2012). For the metabolomics analysis of tryptophan derivatives, metabolites were extracted from cecum luminal contents. The cecum is the pouch located at the beginning of the large intestine.

Microbiota bioactivity can be categorized in two major areas: saccharolytic activity and proteolytic activity (Nyangale et al., 2012). The major products of saccharolytic activity include short chain fatty acids (predominantly butyrate, propionate, and acetate) and lactate. Bacterial species in the genera *Bacteriodes, Lactobacillus, and Bifidobacterium* convert glucose and carbohydrates to pyruvate which is in turn metabolized to short chain fatty acids and other organic acids through the Embden-Meyerof-Parnas pathway (Walker et al., 2005). Although short chain fatty acids have anti-inflammatory properties and nutritional benefits, it is often difficult to quantify significant amounts of these metabolites in fecal matter. A large proportion of short chain fatty acids are absorbed in the lumen and less than 5% escape into fecal excretion (Topping et al., 2001). Proteins that escape digestion in the small intestine are hydrolyzed in the colon to produce amino acids. Sulfur amino acids such as cysteine are converted by sulfate-reducing bacteria like *Desulfomanas* and *Desulfovibrio* to produce sulfur compounds such as

H_2S. Studies have shown that H_2S are linked to increased inflammation of the colonic lumen in patients with ulcerative colitis. At concentrations as low as 250 μmol/L, sulfide compounds caused significant genomic damage to CHO cells and human HT-29 colonic epithelial cells (Hughes et al., 2000). Nitrogen containing compounds also contribute to a large quantity of microbiota metabolic products. N-Nitroso compounds (NOC), heterocyclic amines (HCA), and ammonia are all products of bacterial metabolism. Both NOCs and HCAs have carcinogenic properties due to its ability to alkylate DNA. Patients with IBD and colorectal cancer have higher concentrations of NOCs and HCAs when compared to healthy individuals (Joosen et al., 2009). Ammonia has also been shown to affect DNA synthesis and promote the multiplication of damaged cells at concentrations as low as 5-10 mM (Visek et al., 1978). Phenolic and Indolic compounds are another class of metabolites produced from the microbiota conversion of aromatic amino acids such as tryptophan, phenylalanine, and tyrosine. Indole is one of the major products formed in the bacterial metabolism of tryptophan via the enzymatic activity of tryptophanase, which is expressed in the genes of multiple microbial species in the murine gut (Ku et al., 2006; Sridharan et al., 2013). A recent study by Banal and coworkers demonstrated that indole modulates the expression of pro-inflammatory genes through inhibition of NFκB and the increase in activity of anti-inflammatory genes. In addition, indole has also demonstrated the ability to increase epithelial barrier integrity through the expression of genes involved in tight junctions and actin cytoskeleton formation of intestinal epithelial cells (Banal et al., 2010).

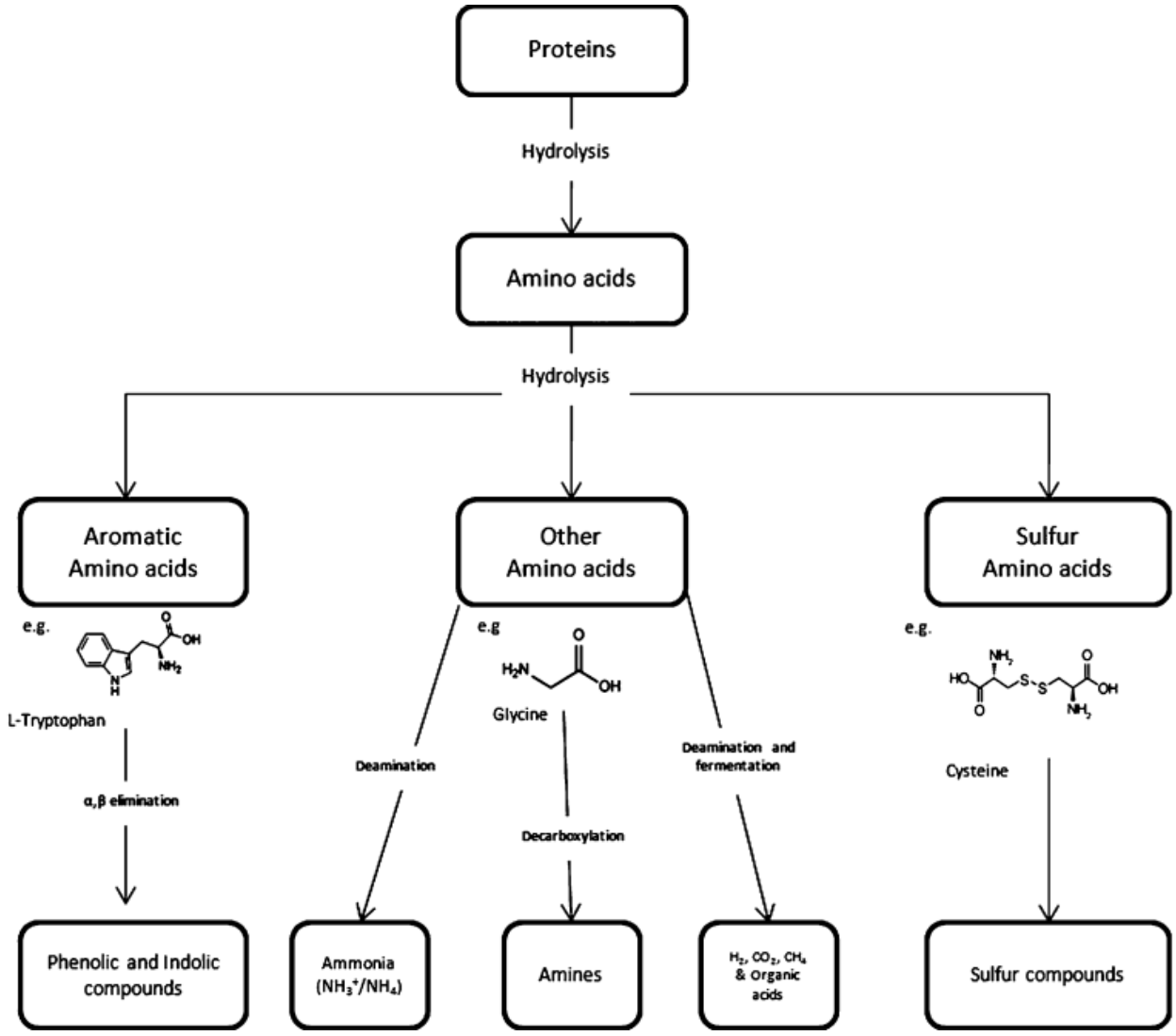


Figure 3: Proteins that escape the small intestine undergo a series of metabolic reactions in the colon. The host microbiota converts the protein-derived amino acids into a variety of metabolites with beneficial and harmful properties (Nyangale et al., 2012).

Analytical Strategies and Challenges for Targeted and Non-targeted Metabolomics

Currently, only a few bioactive microbiota metabolites in the GI tract have been identified and quantified. There are several challenges that attenuate the ease in the discovery of bioactive metabolites. First, the vast quantity of bacteria and the diversity of the commensal microbiota render it difficult to isolate and culture individual microbial species and analyze the metabolites they produce. To address this issue, novel culturing techniques have utilized

microbeads and multiplexed solid surfaces for high-throughput culturing analysis. These methods allow simultaneous single-cell cultivation of thousands of microbes. However, a major limitation of this approach is that it does not recognize the metabolic interactions between bacteria species themselves and those with the host cells. Many gastrointestinal tract microbes are dependent on the metabolic activity of other species and of the host organism. Therefore, an isolated culturing approach does not accurately model the dynamics of the complex gut microbiota (Zoetendal et al., 2008). Another major challenge is the difficulty in differentiating the metabolites produced by the host organism from the metabolites produced by the microbiota. There are a variety of metabolite intermediates that are conserved in the reaction pathways of both mammals and bacteria (Sridharan et al., 2013).

Metabolomics can be divided into two distinct categories, targeted approach and untargeted approach. In targeted metabolomics, a set of analytes are determined *a priori* to allow high selectivity and detection. However, a major drawback of targeted metabolomics is that previously unexplored analytes will not be discovered or analyzed. Several analytical approaches exist in the characterization of the metabolic profiles of mammalian GI tracts. Nuclear magnetic resonance spectroscopy (NMR), gas chromatography-mass spectrometry (GC-MS), and liquid chromatography-mass spectrometry (LC-MS) are three popular methods in metabolomics (Kloos et al., 2013; Wei et al., 2010). In a targeted metabolomics study conducted by Humblot and coworkers, ^{1}H NMR spectroscopy and high performance liquid chromatography (HPLC) were used to quantify and characterize levels of 2-Amino-3-methylimidazo[4,5-f]quinoline (IQ) and IQ derived metabolites in human and murine fecal samples to determine the metabolic activity of the gut microbiota (Humblot et al., 2005). One critical drawback of NMR is the lack of sensitivity to metabolites in low concentrations. However, it compensates for this disadvantage

by providing structural information during the detection of unknown compounds. For volatile organic metabolites, GC-MS is the preferred method of analysis. Garner and coworkers were able to identify 297 volatile organic compounds from human fecal samples through the use of GC-MS (Garner et al., 2007). Recently, LC-MS has dominated the field of metabolic profiling due to its robustness, speed, and minimal sample preparation (Kloos et al., 2013). In a targeted metabolomics study of murine plasma, Wei and coworkers were able to analyze 205 metabolites in 10 minutes. There are several other advantages with the use of MS. Compared to other methods of detection, MS offers high sensitivity (levels of quantification), high selectivity (MW detection), and coupling capabilities with separation technologies such as liquid chromatography, gas chromatography and capillary electrophoresis. Amongst the different types of MS, some of which are time-of-flight, ion-trap, orbitrap, and quadrupole mass spectrometers, a triple quadrupole mass spectrometer is optimal for targeted metabolomics. The first quadrupole (Q1) serves as a mass filter to select for a precursor ion with a known mass. The second quadrupole (Q2) is a collision cell where the precursor ion is fragmented. In Q3, the entire mass to charge ratio (m/z) is scanned to select for the specific daughter fragments of the precursor ion. Each metabolite of interest can be precisely identified and quantified due to their unique fragmentation pattern. Liquid chromatography is a popular selection amongst the MS-coupling separation techniques because of its wide selection of commercially available columns for different separations techniques such as reverse phase separation and hydrophilic interaction separations (Wei et al., 2010).

Experimental Design and Thesis Goals

In this study, a targeted metabolomics approach was utilized to identify potentially bioactive metabolites derived from tryptophan, phenylalanine, and tyrosine through microbiota

metabolism. An *in silico* prediction step is utilized to determine a set of high-frequency metabolites that exists within the human GI tract. The use of computational algorithms can efficiently conduct metabolite discovery experiments while bypassing challenges such as individual isolating and culturing of bacterial species. In addition, an *in silico* discovery approach can be potentially expanded into different host organisms through the substitution of new genomic models into the algorithm. To validate the accuracy of the probabilistic pathway construction algorithm, murine fecal samples were analyzed and quantified through a triple quadrupole mass spectrometer coupled with HPLC. Once the presence of the predicated metabolites has been validated, AhR analysis must be performed to indicate the anti-inflammatory properties of the predicted metabolites.

The primary goal of this study is to discover anti-inflammatory metabolites that regulate mucosal immune functions. Metabolites with beneficial properties can be used for drug discovery and therapeutics. Probiotics utilize consumed bacteria and microbes to produce a variety of metabolites, some of which are beneficial to humans. In this postbiotic approach, only beneficial products from microbiota metabolism are considered to serve as food supplements that can maintain healthy mucosal homeostasis.

Materials and Methods

Solvents and High Purity Metabolites

All high purity metabolite standards and HPLC-grade solvents were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise stated.

In Silico Prediction of Microbiota Metabolites

A probabilistic pathway algorithm was used to predict a set of microbiota-derived metabolites. The algorithm was previously utilized to analyze novel synthesis pathways for host and non-host metabolites for microbial species such as *Escherichia coli*. In this experiment, the algorithm was altered to trace metabolite biotransformation pathways that are nonnative to the mammalian host organism, in this case mice. This ensures that the predicted set of intermediates can only be produced by microbiota metabolism. Mouse metabolic reactions were derived from published mice genome models (Selvarasu et al., 2010; Sigurdsson et al., 2010) and assigned reaction and compound identification numbers based on the KEGG enzyme database, which encompasses all bacterial enzymes specified by gene entry.

The algorithm constructs a tree starting with the user-specified source metabolite as the initial root (Figure 4). A random reaction is chosen from a list of bacterial metabolic reactions cataloged in the KEGG database with the source metabolite as the initial reactant. The randomly selected reaction is represented by an edge addition to the tree and ends with a node that corresponds to the product metabolite or intermediate of the selected pathway. Each node becomes a new starting point for the recursion. An iteration of the algorithm discontinues when it reaches a node where the metabolite can also be produced by the host metabolism. The search space for each iteration was bound to an upper limit of 20 reactions. Variations in reaction number limits between 20 and 50 did not exhibit any observable difference in the set of predicted metabolites. When a reaction pathway violates the upper limit, the algorithm backtracks and proceeds in another previously unexplored pathway. Since the algorithm is based on probability, multiple iterations must be performed to produce a representative sample of the possible source-derived metabolites. The number of iterations specified for tryptophan metabolite prediction was

varied between 0 and 2000 iterations to determine the minimum amount of iterations required to produce a stable set of unique metabolites. For phenylalanine and tyrosine, iterations were varied between 0 and 2500.

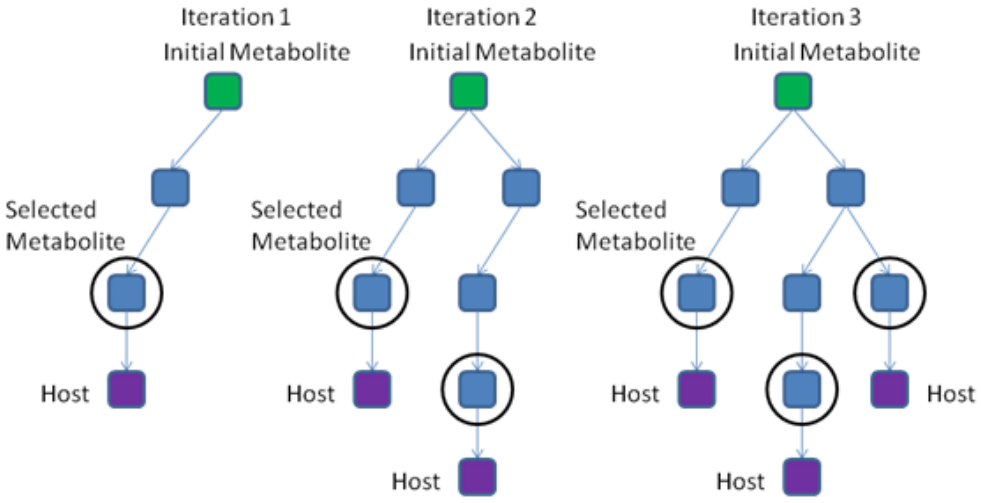


Figure 4: Tree branching formation from probabilistic pathway construction algorithm. The green squares represent the user-specified target metabolite. The arrow edge represents the random reaction pathway that is chosen with the initial metabolite as the reactant. Once the iteration reaches a metabolite that can be produced by the host (purple), the iteration halts and chooses the metabolite prior to the host metabolite as indicated by a circle.

Cecum and Fecal Sample Collection

Female C57BL/6 mice at 5 weeks and 14 weeks of age were purchased from Jackson Laboratories (Bar Harbor, Maine) and raised in a pathogen free animal facility located in Texas A&M Health Science Center. All mice were allowed to orient to colony for 1 week. These animals were handled under the specifications of the Institutional Animal Care and Use Committee guidelines with appropriate animal use protocol. For tryptophan quantification, young and old mice (n=7) were sacrificed at 6 and 15 weeks of age, respectively. The cecum tissue and luminal content and fecal pellets were collected from each mouse. Samples were

weighed, flash frozen, and stored at -80 ℃ prior to metabolite extraction. For phenylalanine and tyrosine quantification, mice at 6 weeks of age (n=5) were treated with 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) for 1 day, 3 days, and 5 days. Fecal samples from control and TCDD treated mice were weighed, flash frozen, and stored at -80 ℃ prior to metabolite extraction.

Metabolite Extraction from Cecum and Fecal Samples

A solvent-based method (Sellick et al. 2010) was used to extract metabolites from the cecum luminal contents and fecal pellets. First, 1.5 ml of ice-cold methanol/chloroform (2:1 v/v) was added to the pre-weighed murine samples. This step was followed by homogenization on ice. The test tubes were centrifuged at 15,000 x g at 4 ℃ for 10 minutes. The supernatants for each sample were transferred to a new test tube through a 70-µm cell strainer. Next, 0.6 ml of ice-cold water was added to the supernatants. Each sample was vortexed and centrifuged at 15,000 x g at 4 ℃ for 5 minutes. With a syringe, the upper (polar phase - methanol and water) and the bottom (organic phase - chloroform) were separated into different test tubes. In order to improve detection in MS, 500 µL of the polar phase was concentrated in a Savant speedvac concentrator (Thermo Scientific, Asheville, NC). The samples were stored in 50 µL of 50% methanol/water (v/v) at -80 ℃ until analysis.

Quantification and Identification through LC-MS/MS

A triple quadrupole linear ion trap mass spectrometer (3200 QTRAP, AB SCIEX, Foster City, CA) coupled with a binary pump HPLC (1200 Series, Agilent, Santa Clara, CA) was used to detect and quantify the set of predicted metabolites. The predicted set of tryptophan,

phenylalanine and tyrosine-derived metabolites were optimized for the detection in mass spectrometry prior to sample analysis. Utilizing a direct-syringe injection method at 10 µL/ min, MS parameters were optimized for each metabolite to obtain the highest peak intensity in multiple reaction monitoring (MRM) analysis. Depending on peak intensity, metabolites were analyzed either through positive mode or negative mode. Declustering potential (DP), entrance potential (EP), collision cell entrance potential (CEP), collision energy (CE), and collision cell exit potential (CXP) were the five parameters optimized for each metabolite. To determine peak area and retention time, Analyst software (version 5, Agilent, Foster City, CA) was used. HPLC samples were stored in the autosampler at 4 ℃ prior to injection.

Two solvent gradient methods were utilized to achieve chromatographic separation. For tryptophan (indole, indole-3-acetate, indole-3-acetamide, 7-hydroxyindole, and tryptamine) and phenylalanine derivatives (phenylpyruvate and 2-hydroxyphenylacetate), chromatographic separation was performed on a hydrophilic interaction column (Luna 5 µm NH_2 100 Å 250 mm x 2 mm, Phenomenex, Torrance, CA) with a solvent gradient method (Bajad et. al, 2006). Solvent A was a solution of water and 5% acetonitrile (v/v) with 20 mM of ammonium acetate. The pH of solvent A was adjusted to 9.45 through addition of ammonium hydroxide. Solvent B was pure acetonitrile. For tyrosine derivatives (tyramine and 4-phenylbenzoate) chromatographic separation was achieved on a reverse phase hydrophobic interactions column (Synergi 4 µm Fusion-RP 80 Å 150 mm x 2 mm, Phenomenex, Torrance, CA) with a solvent gradient method (Lu et al., 2005). Solvent A was a solution of water with 0.1% formic acid. Solvent B was a solution of methanol with 0.1% formic acid. Detailed solvent composition changes are outlined in Appendix A.

A standard curve for each metabolite was produced to quantify the concentrations in the murine samples. Metabolites samples of 10^9 ng/L were produced by adding 10 mg of high purity metabolite solids in 10 ml of 50% methanol in water (v/v). Two serial dilutions at 10x were performed to achieve a concentration of 10^7 ng/L. Standard curve samples were prepared in HPLC vials at concentrations of 10^7 ng/L, 0.75 x 10^7 ng/L, 0.50 x 10^7 ng/L, and 0.25 x 10^7 ng/L.

Results and Discussion

The LCMS analysis method was specified to detect only one metabolite at a time for each sample instead of multiple metabolites at once. This specification avoids two major complications that may arise: ion suppression and metabolites that elute at similar times. For phenylalanine metabolites, both 2-hydroxyphenylacetate and phenylpyruvate had retention times between 9 and 10 minutes. Detection of both analytes could compromise peak sensitivity and shape. Ion suppression is often another major concern in LCMS experiments. When specifying MS parameters for more than one analyte in a single run, ionization competition may lead to decreases in peak area and inaccurate measurements. Two mechanisms are proposed to explain ion suppression. First, the presence of multiple analytes can decrease evaporation efficiency and reduce the capability of analytes converting to its gas phase. Second, competing analytes may reach the maximum ionization limit of the MS (Antignac et al., 2005). The metabolite extraction method utilized in this experiment was implemented to reduce potential sources of ion suppression. Lipids and fatty acids are concentrated in the organic chloroform layer of the extraction process. By concentrating samples only in the polar phase, significant amounts of undesired compounds were removed.

Tryptophan Metabolite Prediction and Quantification

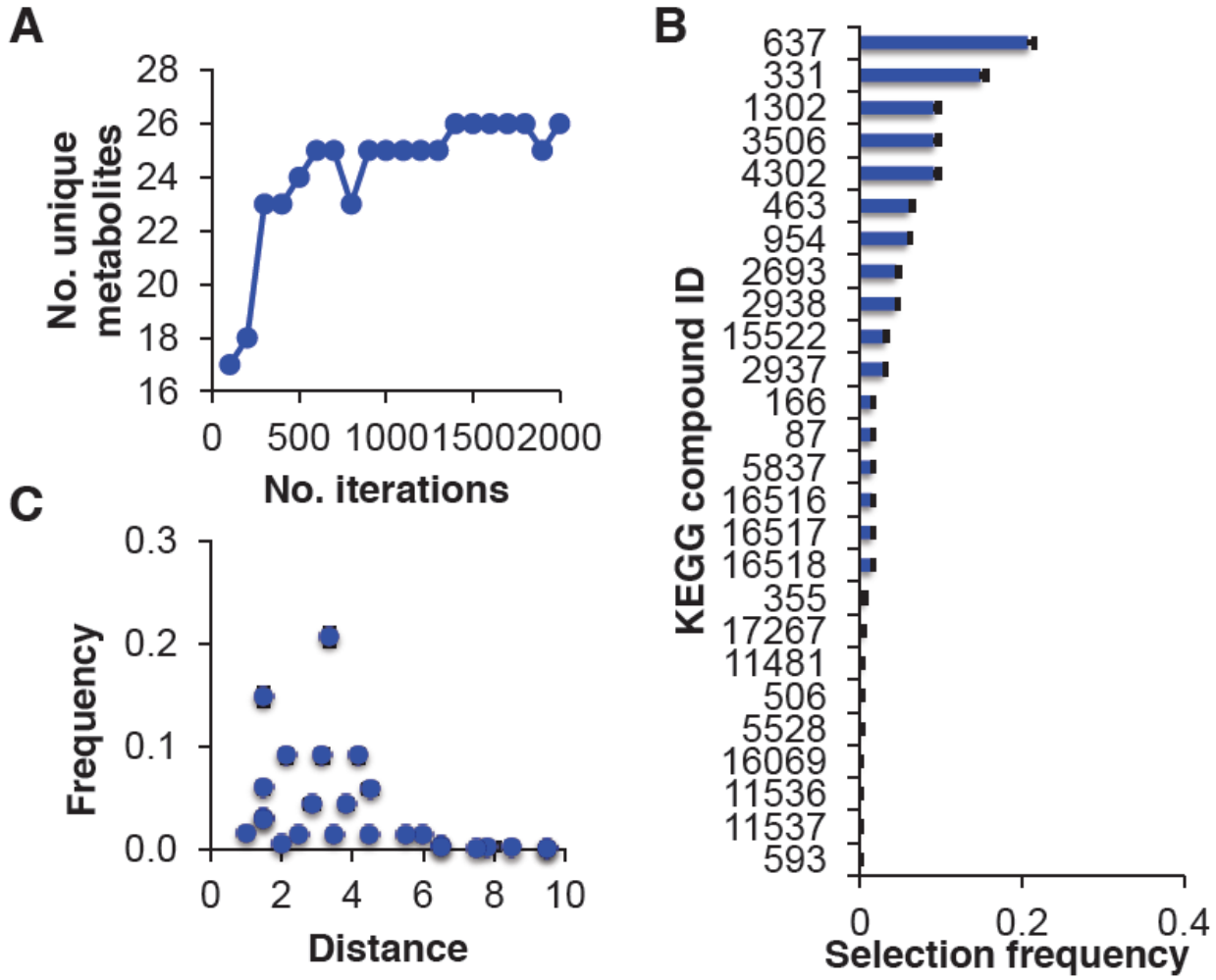


Figure 5: (A) Plots the number of unique metabolites produced versus the number of iterations on the probabilistic pathway algorithm. (B) Indicates the average selection frequency of specific compounds identified through its KEGG ID. (C) Illustrates the average distance of metabolites based on their selection frequency (Sridharan et al., 2013).

A subset of five high frequency tryptophan-derived metabolites predicted in the probabilistic algorithm was selected for analysis. High frequency compounds such as indole-3-acetaldehyde were excluded because a high purity standard was not commercially available. Other compounds including indole-3-pyruvate were excluded due to poor ionization which resulted in a low limit of detection (LOD).

Due to the probabilistic nature of the pathway construction algorithm, many iterations were required to acquire a representative set of high frequency metabolites. The increase in iterations correlated to the increase in the set of unique metabolites. For tryptophan, it was determined that an iteration number of 1500 and above provided a stable set of unique metabolites. With the iteration number set at 1500, 50 simulations were performed to calculate for average frequencies of the predicted set of metabolites. It was determined that the top 12 high-frequency metabolites accounted for more than 90% of the total metabolite selections. Highest frequency metabolites were also on average within a shorter distance to the source metabolite (Figure 5).

The following five tryptophan derivatives were analyzed: indole, indole-3-acetate, indole-3-acetamide, 7-hydroxyindole, and tryptamine. Each compound was optimized individually to produce a set of MS parameters that were specific to the detection of its parent/daughter masses (see Appendix B.1). In cecum samples, it was determined that the five metabolites existed in levels of concentrations around 30 μM (Figure 6.A). However, these reported concentrations are only conservative estimates of the actual concentrations within the cecum due to efficiencies in the methanol/chloroform/water metabolite extraction method. With the exception of indole-3-acetate and tryptamine, young mice exhibited statistically higher concentrations of the tryptophan-derived metabolites in their cecum samples. In fecal samples, the five metabolites were present in concentrations ranging from 10 to 10^2 nmol/g-feces. Indole-3-acetate exhibited a higher concentration in older mice in fecal matter. However, the other metabolites were present in relatively equal amounts for the two age sets. It was expected that young mice and old mice would have different concentrations of certain metabolites. A study by Vaahtovuo and coworkers indicated that there were significant differences in bacteria-derived

cellular metabolites in the fecal samples of 5-7 and 15-19 week old mice (Vaahtovuo et al., 2001). Many environmental factors, specifically age and diet, affect the microbiota composition of a mammalian gut. After birth, the gut of an unborn mouse is considered sterile and bacteria colonization begins with the inoculation from the maternal microbiota. Subsequently, the gut becomes more developed through the diet of the host organism. In humans, the commensal microbiota resembles that of an adult by the age of two (Nyangale et al., 2012). It was hypothesized that young mice would exhibit less of the predicted set of mice due to the fact that their microbiota is less developed and populated than that of an older mice. However, the analysis of the cecum samples illustrated the opposite. Several possible explanations can deduce the cause of such deviations. The microbiota and the host intestinal cells have developed a symbiotic relationship where the metabolites produced by one species of bacteria may be utilized by the host cells or other bacterial species in their metabolism. A more developed microbiota has higher levels of metabolic interactions between different species of bacteria, resulting in the consumption of tryptophan-derived metabolites. Another source of deviation can result from the differences in the ability of young and old mice to digest tryptophan in the upper GI tract. Young mice and old mice can have considerable differenced in the uptake of tryptophan through their diet. In cecum samples, it was shown that young mice exhibited a higher concentration of tryptophan. If the concentrations of the other tryptophan-derived metabolites were normalized based on the quantity of tryptophan detected, the differences in concentrations between these metabolites would most likely also vary to a lesser degree.

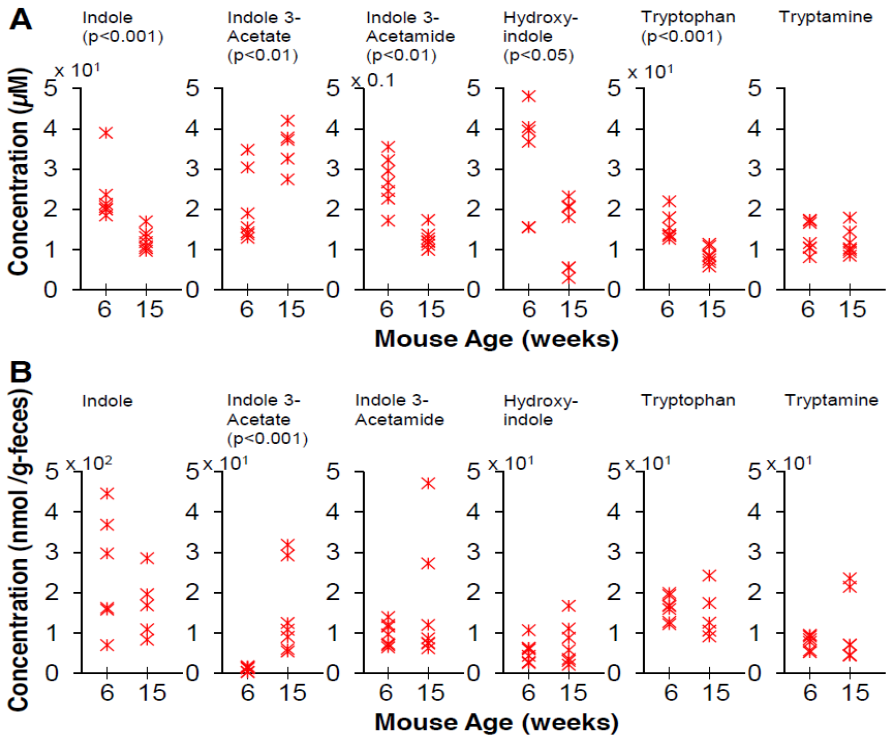


Figure 6: (A) Comparison of tryptophan-derived metabolites in young and old mice cecum samples. (B) Comparison of tryptophan-derived metabolites in young and old mice fecal samples. Data points were normalized to the mass of the fecal sample.

AhR activation studies were conducted to determine the bioactivity of tryptophan-derived metabolites. The AhR is a ligand activated transcription factor that modulates the mucosal immune system. Past studies have shown that several tryptophan-derived metabolites were determined as AhR ligands (Bjeldanes et al., 1991). AhR induction was measured through the rate of AhR-driven luciferase activity. It was determined that indole-3-acetate, tryptamine, and indole-3-pyruvate were able to act as agonists for the AhR in a dose dependent manner. However, indole and indole-3-acetamide did not exhibit AhR activation at a dose of 100 μM (Figure 7).

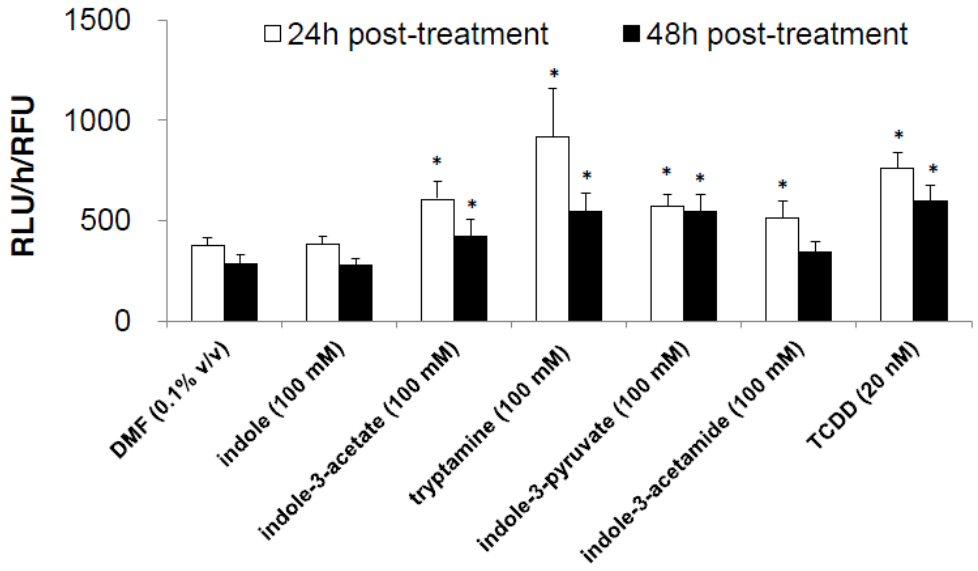


Figure 7: The induction of AhR activity through 100 µM dose of tryptophan-derived metabolites. DMF (0.1% v/v) and TCDD (20 nM) were performed to act as negative and positive control respectively. An asterisk denotes AhR activation (Sridharan et al., 2013).

Phenylalanine Metabolite Prediction and Detection

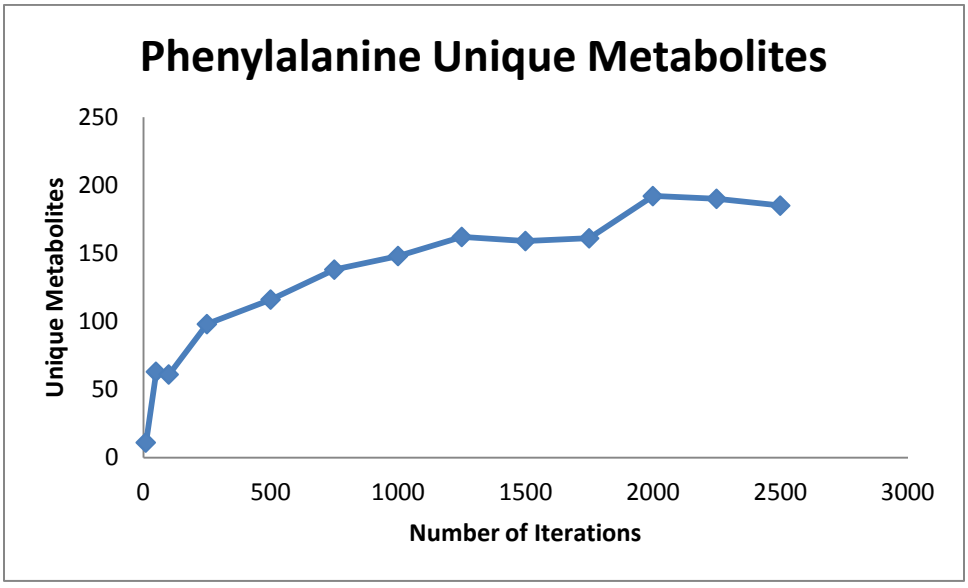


Figure 8: Quantity of unique metabolites produced by the probabilistic pathway construction algorithm with phenylalanine (algorithm ID: 79) as the source metabolite.

Similar to tryptophan-derivatives, an experiment was performed to determine the amount of iterations required to obtain a constant set of unique metabolites. For phenylalanine prediction, iterations were varied from 0 to 2500 (Figure 8). At 2000 iterations and above, a relatively stable quantity of unique metabolites were predicted in the probabilistic algorithm. However, the differences in the list of highest frequency metabolites did not vary from 1000 iterations to 2000 iterations. Due to a lengthy run-time (1 hour or more per run with iterations above 1000), an iteration number of 1000 was chosen for phenylalanine. Subsequently, 3 runs of 1000 iterations were performed to determine the average selection frequency of PA-metabolites. It was determined that the top 7 metabolites produced in the probabilistic algorithm accounted for 57% of the total selection frequency (Table 1).

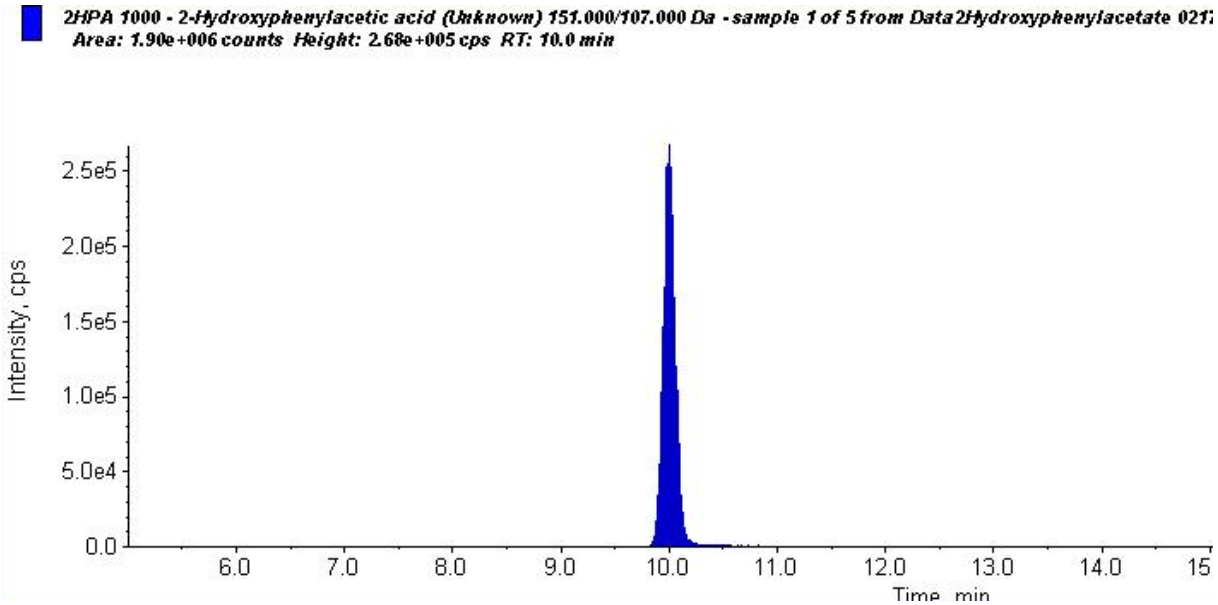
Table 1: Predicted set of Phenylalanine derived Metabolites

Name	Distance	Frequency	% Frequency
'Phenylpyruvate'	1.56	1100	18.73
'Prephenate'	2.51	740	12.60
'L-Arogenate'	2.41	556	9.47
'2-Hydroxyphenylacetate'	2.59	392	6.67
'trans-Cinnamate'	1.50	224	3.81
'2-Hydroxy-2,4-pentadienoate'	6.42	196	3.34
'Chorismate'	3.57	136	2.32

In the set of 7 high frequency metabolites only phenylpyruvate (PP) and 2-hydroxyphenylacetate (2HPA) were chosen for murine sample analysis. MS optimization of prephenate and chorismate indicated that these compounds had poor ionization. High purity products of L-arogenate and 2-hydroxy-2,4-pentadienoate were not commercially available. A HILIC column was chosen to elute the phenylalanine derived metabolites. Due to the polar nature of phenylpyruvate and 2-hydroxyphenylacetate, the use of hydrophilic interaction chromatography can achieve adequate separation and detection within the MS. In a LC method

optimization study conducted by Bajad and coworkers, it was determined that an amino HILIC column with solvent gradient method (Appendix A.1), could accurately identify and quantify levels of tryptophan, phenylalanine, phenylalanine derivatives (specifically phenylpyruvate), tyrosine, and tyrosine derivatives (specifically *p*-hydroxybenzoate) (Bajad et al., 2006).

A



B

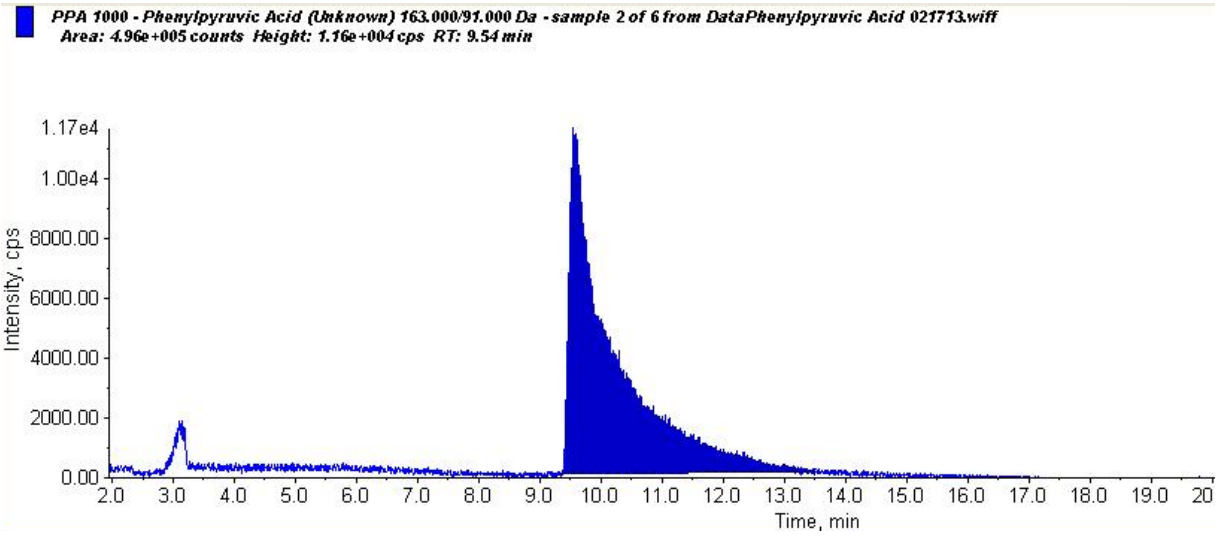


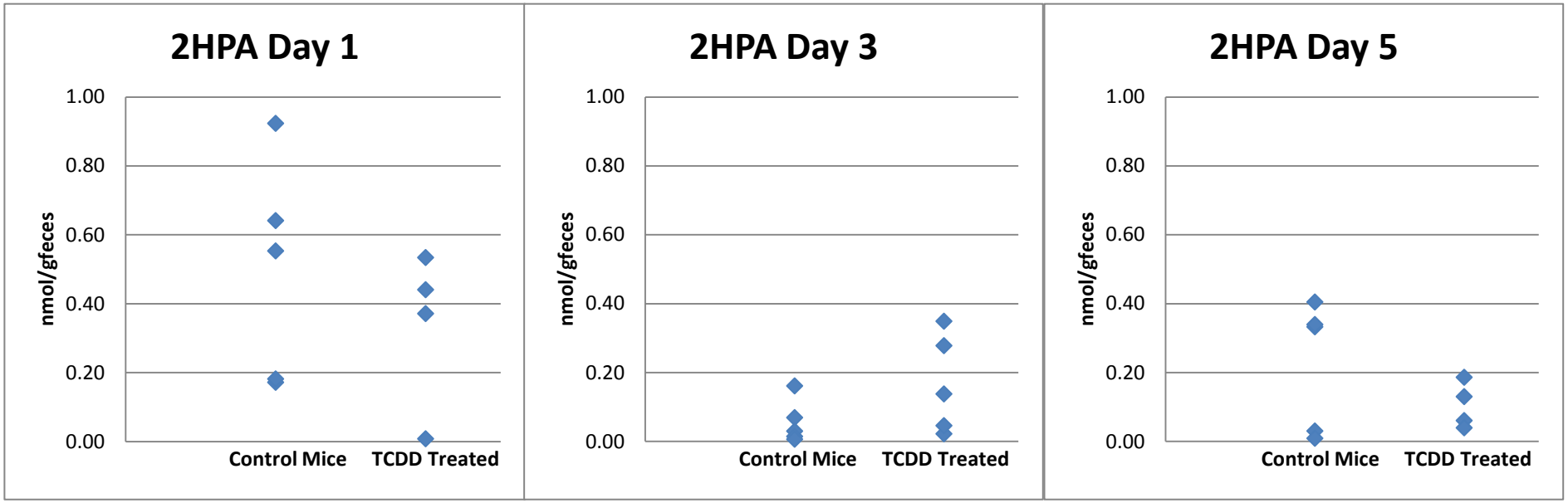
Figure 9: (A) Chromatogram of standard solution of 2HPA at 10^7 ng/L. (B) Chromatogram of standard solution of PP at 10^7 ng/L.

The retention time of 2-hydroxyphenylacetate and phenylpyruvate were determined to be from 9.7 min to 10.2 min and from 9.5 min to 13 min respectively. As opposed to 2-hydroxyphenylacetate (2HPA), phenylpyruvate (PP) exhibits noticeable peak broadening. Several factors concerning the LC-method result in peak broadening during analyte elution. Excessive tubing is one of the common errors that cause substantial peak broadening. However, since analysis of 2HPA and PP were run on the same day, this was unlikely to be the cause. Another factor that causes peak broadening is large injection volumes that cause column overloading. For long chromatographic columns (i.e. 250mm), an injection volume of 10 μL is adequate. The use of gradient solvent methods as opposed to isocratic (constant composition) methods usually reduces peak broadening issues. However, improper design of gradient methods can result in peak broadening during the initial stages of the run, due to the fact that there is column volume delay. Therefore, it is vital to equilibrate the column prior to any sample run. The most likely cause of the peak broadening of PP is the interaction of the analyte with the amino stationary phase of the HILIC column. A stronger polar solvent may be required to reduce the peak broadening of PP. Despite the peak broadening issue, accurate standard curves were created to quantify murine fecal sample (Appendix C.1, C.2).

Fecal murine samples were extracted from mice of 6 weeks of age. C56BL/6 mice (n=5) were treated with TCDD and fecal samples were collected on days 1, 3, and 5. These fecal samples were tested for the presence of phenylalanine-derived metabolites, phenylpyruvate and 2-hydroxyphenylacetate. Concentrations of these two PA-derived metabolites were compared to those of the control group. TCDD is a high affinity AhR agonist that can increase the AhR-driven luciferase activity by 2-fold (Sridharan et al., 2013). High concentrations of TCDD have been shown to increase and decrease the expression of hundreds of rat genes. Some of these

genes encode for enzymes that digests carcinogenic polycyclic hydrocarbons (Tijet et al., 2006). TCDD is a highly potent environmental carcinogen that affects the immune response of humans and mammals. Studies have shown that a single dose (30μg/kg) of TCDD to mice can alter T- and B- cell dependent immune responses and impair the ability of mice and mammalian hosts to resist infectious microbes and antigens (Singh et al., 2009). Recently, Jayaraman and coworkers demonstrated that a 3-day exposure to TCDD decreases the levels of tryptophan, indole, and indole-3-acetate. However, a 1-day exposure and a 5-day exposure to TCDD resulted in no significant differences between TCDD treated mice and control mice.

No statistically difference in concentrations of 2-hydroxyphenylacetate and phenylpyruvate were observed between fecal samples of TCDD treated mice and control mice. A Mann-Whiteney U-Test (See Table 3), was performed to compare the medians of the two independent mice sample sets. According to the calculated p values, which where all greater than p=0.05, the differences in the concentrations of phenylalanine derived metabolites were not significant. The average concentration of 2HPA found in murine fecal samples 6 weeks of age was 0.231 nmol/gfeces. The average concentration of PP found in murine fecal samples 6 weeks of age was 0.425 nmol/gfeces. Both TCDD treated and control mice exhibited decreased concentrations of 2HPA and PP when plotted over time. However, the concentration between the two groups did not vary significantly over time (Figure 10 and 11, bottom left graph).



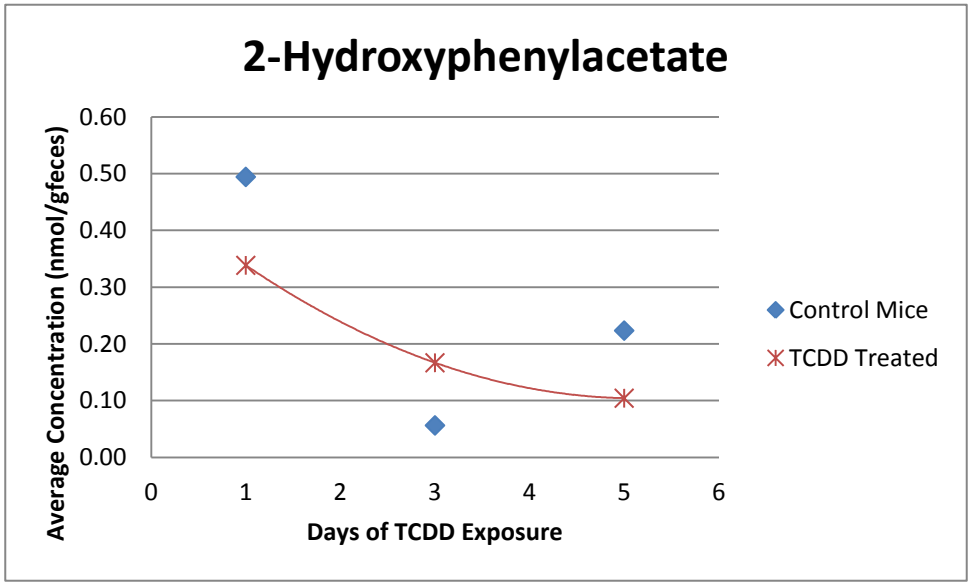
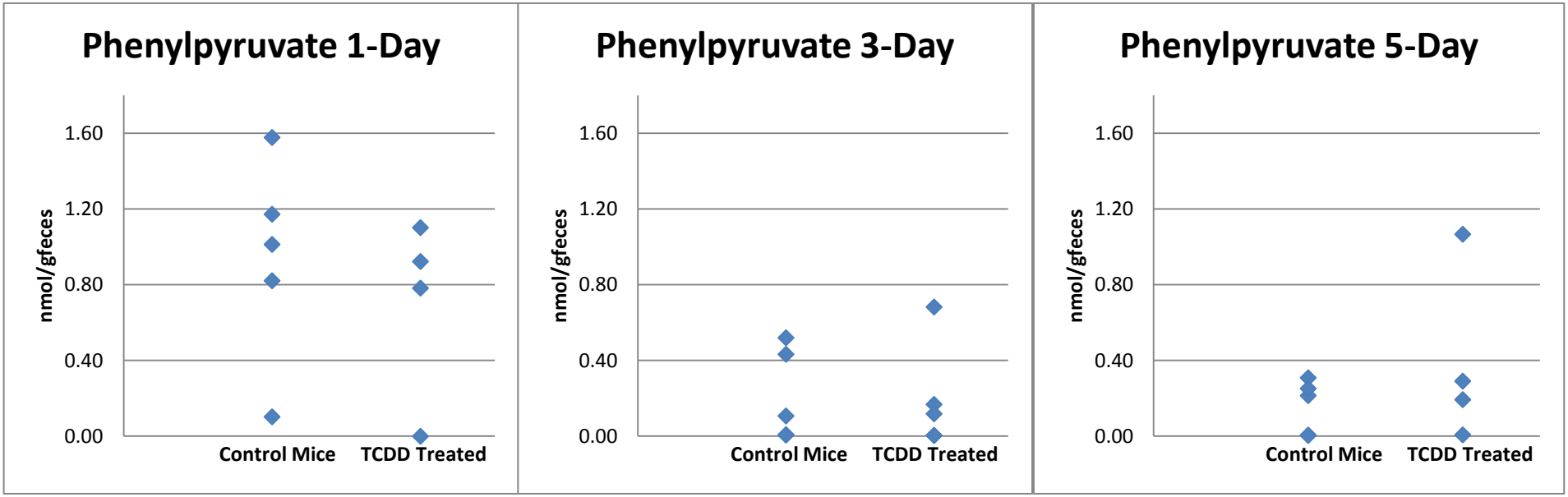


Figure 10: The three graphs above illustrate the concentrations of 2-hydroxyphenylacetate in control mice (n=5) and TCDD treated mice (n=5) at three different exposure durations. The graph on the left plots the average concentrations of 2HPA over the days of TCDD exposure.



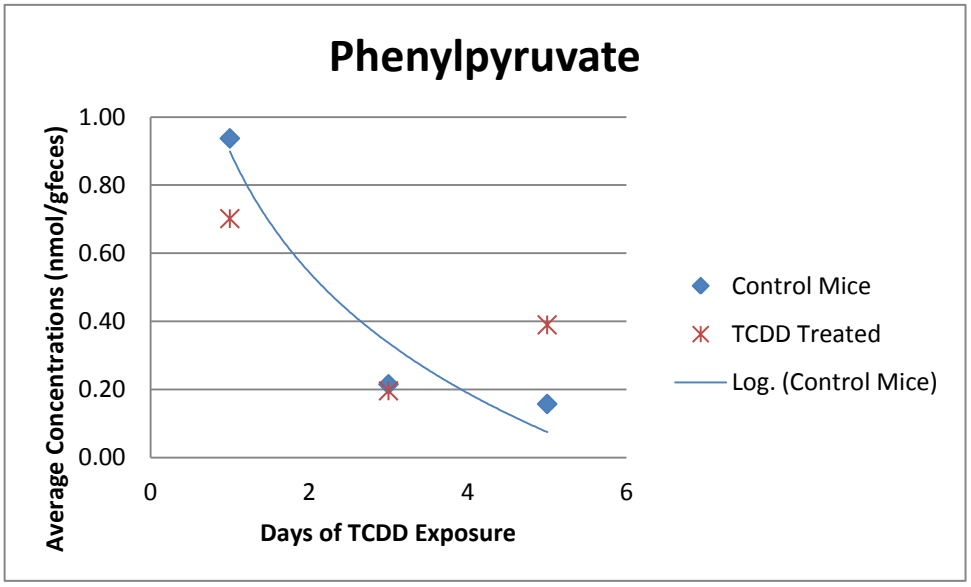


Figure 11: The three graphs above illustrate the concentrations of phenylpyruvate in control mice (n=5) and TCDD treated mice (n=5) at three different exposure durations. The graph on the left plots the average concentrations of PP over the days of TCDD exposure.

Tyrosine Metabolite Prediction and Quantification

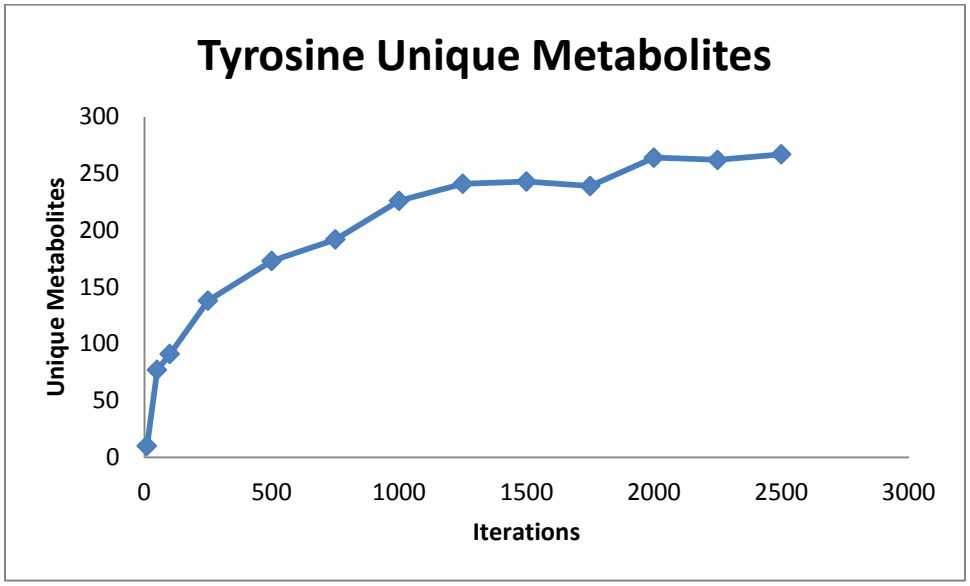


Figure 12: Quantity of unique metabolites produced by the probabilistic pathway construction algorithm with phenylalanine (algorithm ID: 82) as the source metabolite.

An experiment was performed to determine the amount of iterations required to obtain a constant set of unique metabolites. For tyrosine prediction, iterations were varied from 0 to 2500 (Figure 12). At 2000 iterations and above, a relatively stable quantity of unique metabolites were predicted in the probabilistic algorithm. However, the differences in the list of highest frequency metabolites did not vary from 1000 iterations to 2000 iterations. Due to a lengthy run-time, an iteration number of 1000 was also chosen for tyrosine. Subsequently, 3 runs of 1000 iterations were performed to determine the average selection frequency of TYR-metabolites. It was determined that the top 7 metabolites produced in the probabilistic algorithm accounted for 25% of the total selection frequency (Table 2). On the contrary to tryptophan and phenylalanine derived metabolites, high frequency tyrosine-derived metabolites did not exhibit a frequency above 7% in any run. Tyrosine has more reaction pathways available and therefore a greater

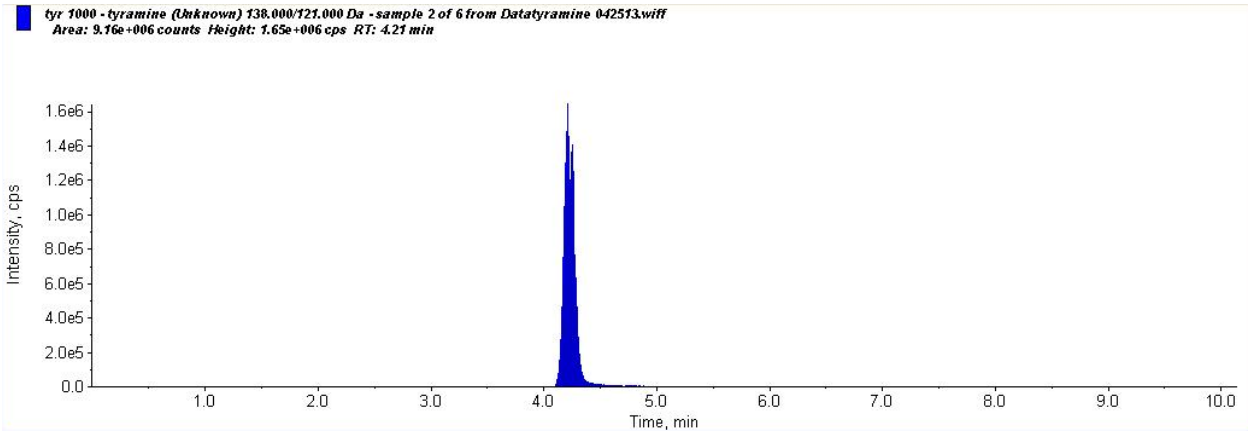
quantity of unique metabolites can be obtained through the probabilistic pathway construction algorithm.

Table 2: High Frequency Metabolites from Tyrosine

Name	Distance	Frequency	% Frequency
'Prephenate'	2.6429	238	6.62
'L-Arogenate'	1.6102	236	6.56
'4-Hydroxybenzoate'	5.2042	142	3.95
'4-Coumarate'	1.6404	114	3.17
'Phenol'	2.3302	106	2.95
'Tyramine'	1.5	86	2.21

Of the 6 top high frequency metabolites, only 4-hydroxybenzoate (4HB) and tyramine were selected for MS analysis. Prephenate and phenol were known to have poor ionization in the MS. A commercial product of L-arogenate was not available. Although 4-coumarate was able to ionize properly in the MS, the LC method used was not appropriate in the elution of 4-coumarate. The LC method chosen for the analysis of tyrosine-derived metabolites was a reverse-phase separation method. The Synergy Fusion RP column used for this analysis has the capability to separate both polar and non-polar compounds. Previous experiments had shown that the solvent-based method used in this separation process (Lu et al., 2005) was suitable for the separation of amino acids and amino acid-derivatives. However, there is not one single method that suites the separation of all metabolites due to the difference in polarity and interactions with the column solid phase.

A



B

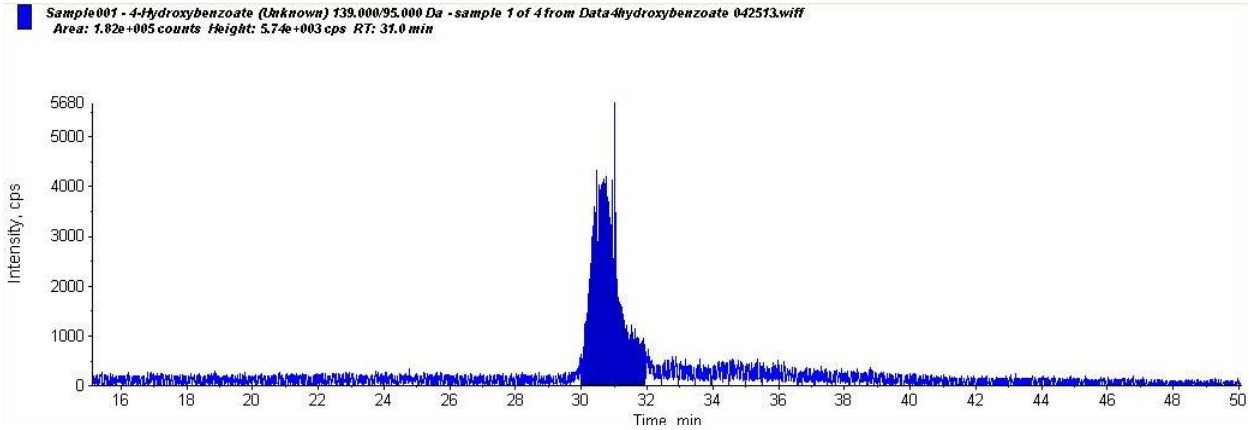
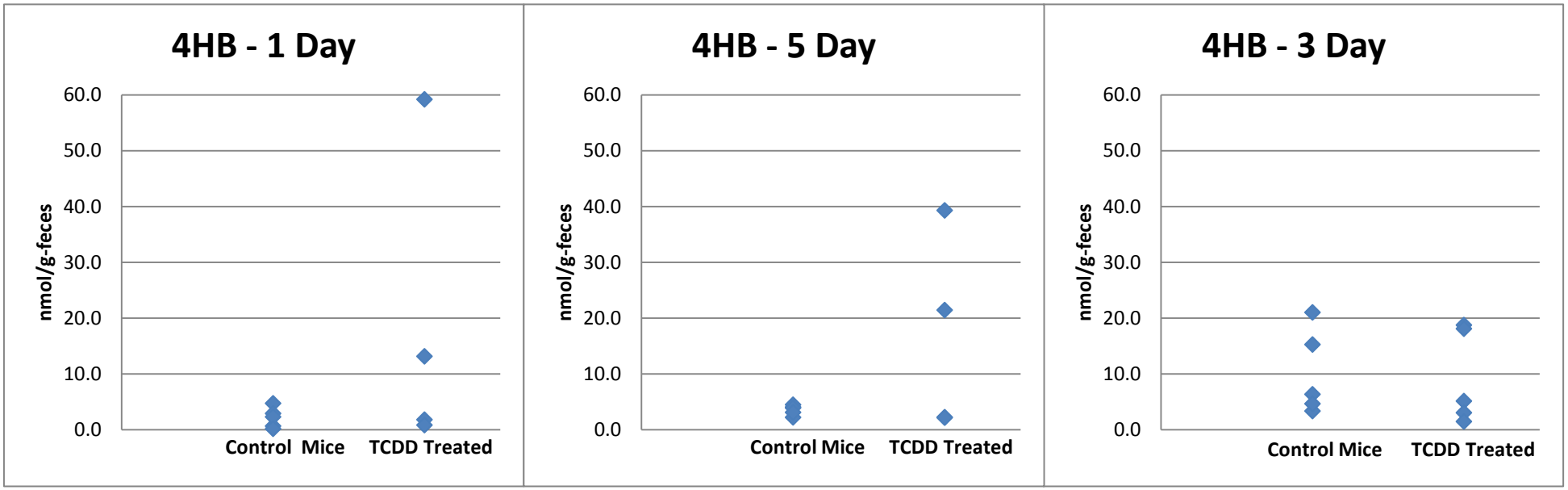


Figure 13: (A) Chromatogram of standard solution of tyramine at 10^7 ng/L. (B) Chromatogram of standard solution of 4HBA at 10^7 ng/L.

Utilizing reverse phase interactions chromatography, tyramine and 4-hydroxybenzoate retention times were determined to be between 4.1 and 4.4 minutes and between 30 and 32 minutes respectively. Tyramine was able to elute at 4 minutes due to its polar amine group. Between t=0 and t=38, the mobile phase is primarily the polar solvent (water). In addition, the hydrophobic solid phase allows for a rapid separation of tyramine. On the other hand, 4-hydroxybenzoate was more hydrophobic than tyramine. A late retention time indicated that the compound had stronger interactions with the solid phase.



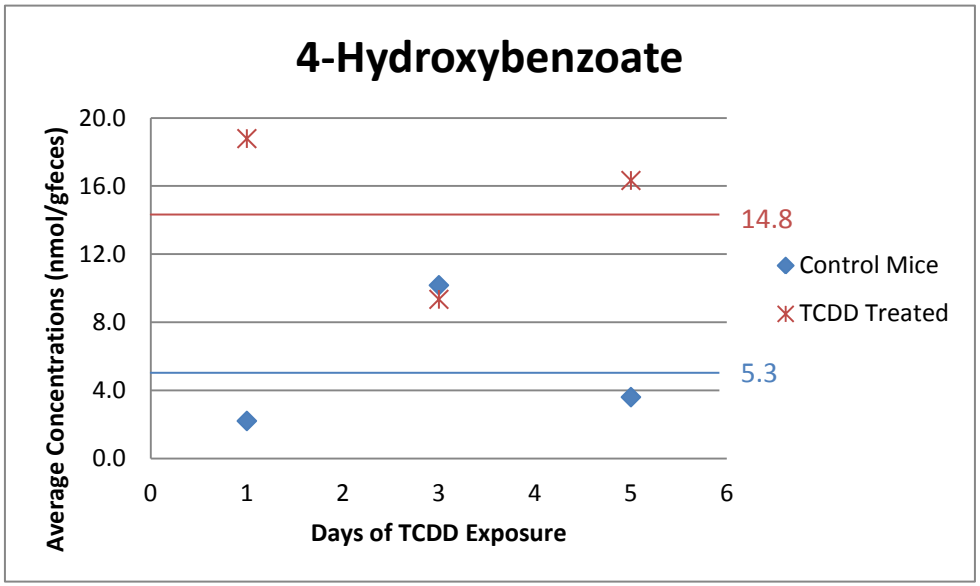
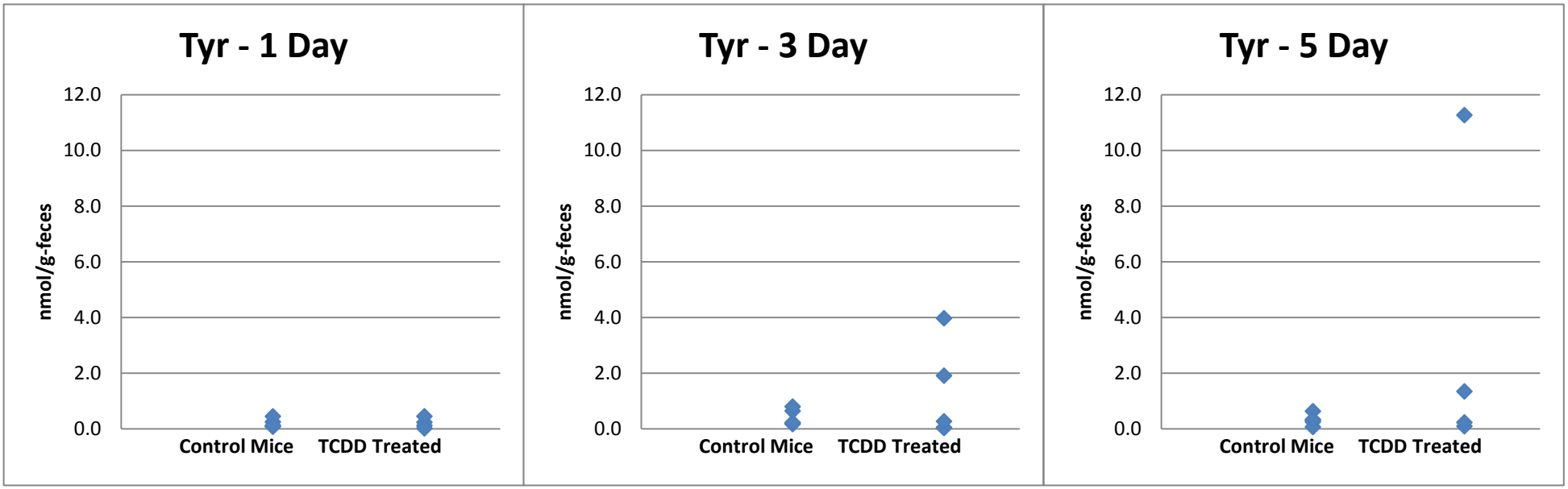


Figure 14: The three graphs illustrate the concentrations 4-hydroxybenzoate in control mice (n=5) and TCDD treated mice (n=5) at three different exposure durations. The graph on the left plots the average concentrations of 4HB over the days of TCDD exposure. The red line illustrates the average 4HB concentration in TCDD fecal samples. The blue line illustrates the average 4HB concentration in control mice fecal samples.



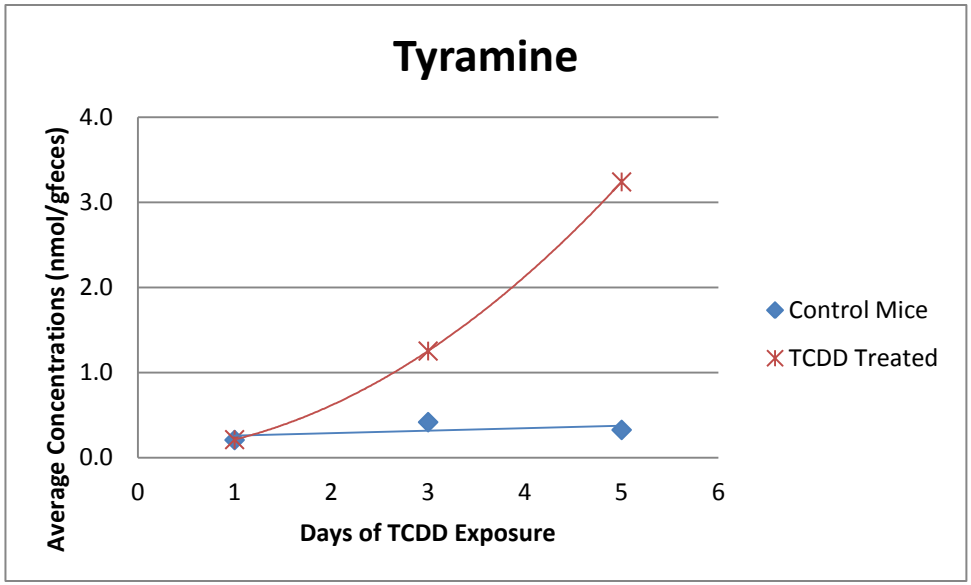


Figure 15: The three graphs illustrate the concentrations tyramine in control mice (n=5) and TCDD treated mice (n=5) at three different exposure durations. The graph on the left plots the average concentrations of tyramine over the days of TCDD exposure.

For 4-hydroxybenzoate, metabolite concentration levels ranged from 0.2 ng/g-feces to 59 ng/g-feces. Concentrations of 4HB with TCDD treated mice on 1 and day 3 appeared to be greater than that of control mice. However, the difference was not statistically significant as shown in the Mann-Whitney calculations (see Table 3). Many factors could have contributed to the wide distribution of 4HB in the TCDD treated samples. Two subgroups(n=3 and n=2) of TCDD treated mice samples were run on the LCMS on separate occasions, due to a system malfunction, which closed the LC and exposed the samples to room temperature overnight. Exposure to room temperature may have degraded the metabolites. In the first subset (n=3), the TCDD fecal samples were stored at 4 ℃ in the autosampler prior to sample analysis. This refrigerated subset generated high concentrations of 4HB in TCDD fecal samples at day 1 (13.2 nmol/gfeces and 59.2 nmol/gfeces) and at day 5 (39.3 nmol/gfeces and 21.4 nmol/gfeces). These concentrations were significantly higher than TCDD fecal samples (n=2) exposed to room temperature, which had concentrations average concentrations of approximately 2 nmol/g-feces, and those of the control mice which had average concentrations of approximately 2-3 nmol/g-feces. The average concentration (Figure 14, bottom left graph) of TCDD fecal samples appeared to be higher than that of control mice over the 5 day period. However, unless more samples are analyzed, the differences are not statistically significant.

For tyramine, metabolite concentration levels ranged from 0.02 ng/g-feces to 11.2 ng/g-feces. A similar pattern to 4HB occurred for tyramine at day 3 and day 5. The refrigerated TCDD fecal subset (n=3), exhibited higher concentrations (3.98 nmol/g-feces and 1.92 nmol/g-feces) on day 3 and (11.26 nmol/g-feces and 1.35 nmol/g-feces) on day 5. Exposure to room temperature may have also degraded tyramine metabolites. When the average concentrations of tyramine were plotted over time, tyramine concentrations increased in TCDD fecal samples while

tyramine concentrations remained constant for control mice samples. However, the drastic increase in tyramine concentrations on day 5 for TCDD samples was due to one high concentration outlier of 11.26 nmol/gfeces measured for one of the samples on day 5.

Statistical Analysis:

To determine the statistical differences between control mice and TCDD treated mice, non-parametric two-sided Mann-Whitney tests were performed to compare the medians between the metabolite levels of control and TCDD treated mice. The Mann-Whitney test can only be used when four assumptions are satisfied. First, your dependent variable, in this case concentration of metabolites, should be ordinal, or in other words measured at interval levels. Second, your independent variables should be in two independent, categorical groups. In this experiment the two categorical groups are control mice and TCDD treated mice. Third, there should be an independence of observations. Mice in each group were entirely independent of each other and measured concentration levels were independent of concentration levels of other mice. Lastly, there should be a non-normal distribution in the data (Hart, 2001). The following table summarizes the p values with their respective analyzed groups. The null hypothesis that the two medians are the same was rejected for $p < 0.05$. In all subgroups, p was greater than 0.05. Therefore, the difference in concentration levels of 2HPA, PP, 4HB, and tyramine were not statistically significant.

Table 3: P Values calculated for Mann-Whitney U-test

2-Hydroxyphenylacetate	**P Values**
Day 1	0.4009
Day 3	0.2113
Day 5	0.6745
Phenylpyruvate	
Day 1	0.2983
Day 3	0.2089
Day 5	0.2983
4-Hydroxybenzoate	
Day 1	0.2113
Day 3	0.6745
Day 5	0.5287
Tyramine	
Day 1	1.0
Day 3	1.0
Day 5	0.4009

Conclusions:

TRP, PA, and TYR-derived metabolites were all successfully quantified in murine specimens 6 weeks of age to 15 weeks of age. For tryptophan derivatives, cecum and fecal samples of two subsets of mice (young = 6 weeks, old = 15 weeks) were analyzed with the LCMS. For cecum samples, the average concentration in the luminal contents was determined to

be approximately 30 μM. For fecal samples, the concentration of the pellets was determined to be in the range of $10 – 10^2$ nmoles/g-feces. Concentrations of indole-3-acetamide, hydroxyindole, indole, and tryptophan in cecum samples were significantly higher in young mice than those in old mice. Contrarily, concentrations of indole-3-acetate were higher in both cecum and fecal samples of old mice. Other metabolites did not show any statistical difference in fecal samples. AhR studies indicated that indole-3-acetate and tryptamine were AhR agonists. This is evidence that indole-3-acetate and tryptamine have beneficial anti-inflammatory properties. Therefore, these metabolites have the potential to be used for therapeutics, specifically the modulation of mucosal immune functions.

PA and TYR-derived metabolites fecal samples were gathered from mice samples treated with TCDD for 1, 3, and 5 days. Treatment to TCDD may alter gene expression in mice which in turn regulates the production of certain metabolites. The metabolite concentrations were compared to that of control mice. Although there were no significant differences between the control and the TCDD treated mice, further tests are required to determine the possible increases in concentrations of tyrosine-derived metabolites (4-hydroxybenzoate and tyramine). When eliminating the erroneous subset data, 4-hydroxybenzoate and tyramine exhibited higher concentrations in TCDD treated mice.

AhR is a diverse ligand-activated transcription factor that can accept a variety of different agonists. One of the major known classes of AhR agonists is aromatic hydrocarbons. In addition, it is reported that phenylalanine and tyrosine can also affect the binding affinity of different ligands to AhR (Zhao et al., 2010). For future work, the bioactivity of 2-hydroxyphenylacetate, phenylpyruvate, 4-hydroxybenzoate, and tyramine must be determined. If these metabolites act

as AhR agonists, they may be used for drug discovery and therapeutics. Compounds that serve as AhR ligands have anti-inflammatory properties that modulate commensal immune functions.

Another area of future work can be made through the implementation of new murine genomic models. One of the major limitations of the probabilistic pathway algorithm is the accuracy of the genomic model integrated into the MATLAB code. For example, Wikoff and coworkers identified indole-3-propionate as a microbiota-produced tryptophan derivative while the algorithm did not (Wikoff et al., 2009). In addition, changes in the algorithm stopping criteria may also produce a different set of unique high-frequency metabolites. One of the stopping criteria recently implemented in the MATLAB algorithm is the selection of metabolites prior to encountering a reaction pathway that is native to the host organism. The implementation of a new stopping criterion introduces new high frequency metabolites that were not previously predicted. For example, phenylethylamine was predicted as a high frequency phenylalanine-derived metabolite in the new algorithm.

Short chain fatty acids are another potential source of anti-inflammatory metabolites that can be investigated. It well known that SCFAs, such as butyrate exhibit anti-inflammatory properties. In addition, changes in sample collection can be made to increase metabolite detection in the LCMS. It is reported that microbiota activity is most prominent in the distal colon. Therefore, murine samples can be obtained near the descending colon as opposed to the cecum (Nyangale et al., 2012).

Appendices:

A: LC Solvent Methods

A.1 HILIC Solvent Method

Time (min)	Solvent A	Solvent B
0	15%	85%
15	100%	0%
28	100%	0%
30	15%	85%
40	15%	85%
50	15%	85%

A.2 Reverse Phase Solvent Method

Time (min)	Solvent A	Solvent B
0	97%	3%
8	97%	3%
38	5%	95%
45	5%	95%
47	97%	3%
55	97%	3%

B: Mass Spectrometer Optimized Parameters for Metabolites of Interest

B.1 Tryptophan Optimized Parameters

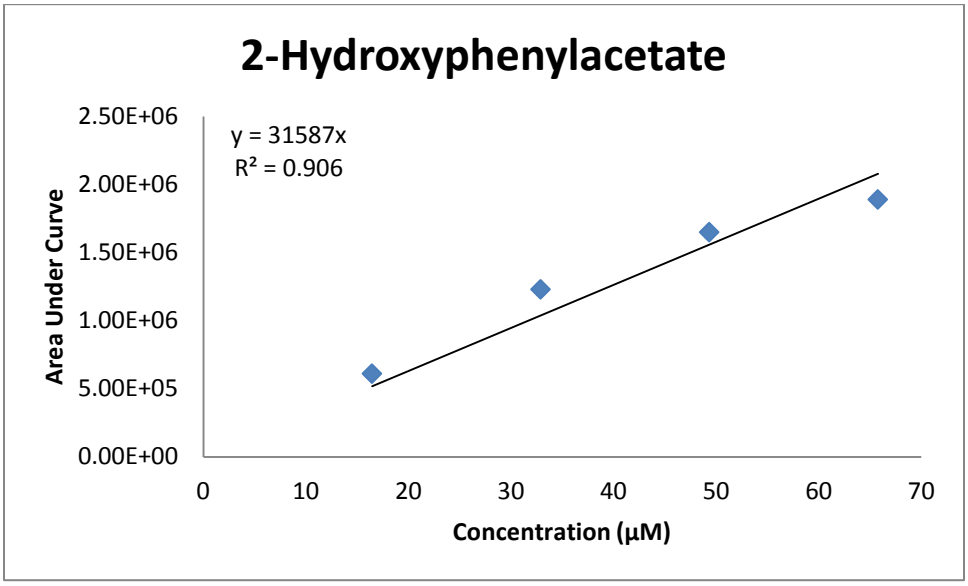
Compound	**Parent Mass (Da)**	**Daughter Mass (Da)**	**DP**	**EP**	**CE**	**CXP**
Indole	118	91	41	10	27	2.5
Hydroxyindole	134.1	77.1	26	10	37	2.5
Tryptamine	161	144	11	4	15	4
Indole 3-Acetate	176	130	31	9	19	4
Indole 3-Acetamide	175	130	26	10	19	4
Tryptophan	205	188	21	10	13	4

B.2 Phenylalanine and Tyrosine Optimized Parameters

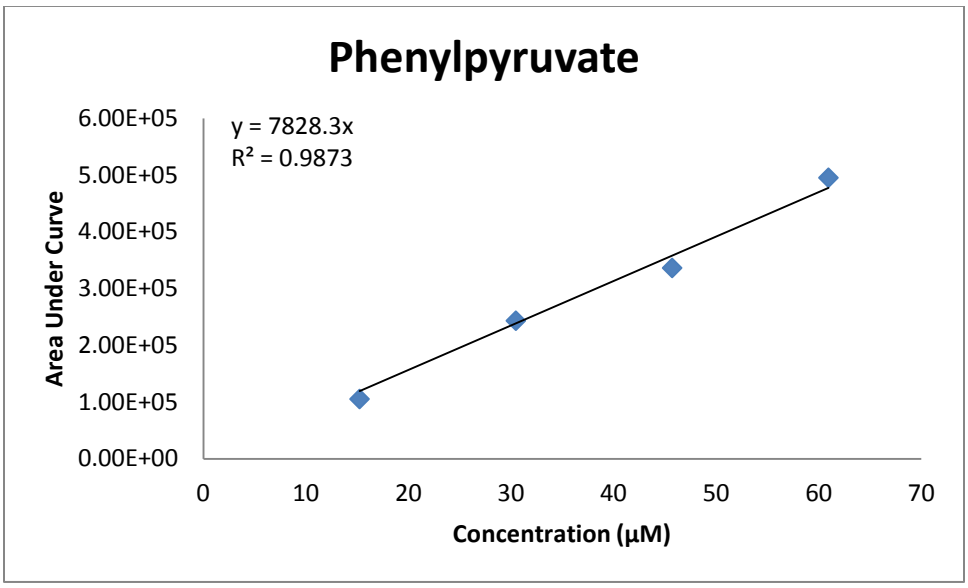
Compound	**Parent (Da) Mass (Da)**	**Daughter Mass (Da)**	**DP**	**EP**	**CEP**	**CE**	**CXP**
Phenyl pyruvate	163	91	-20	-4.5	-88	-14	0
2-Hydroxyphenylacetate	151	107	-20	-4	-12	-18	-4
Tryptamine	138	121	11	8	14	13	4
Indole 3-Acetate	153	107	16	11	14	21	8

C: Standard Curves of Metabolites

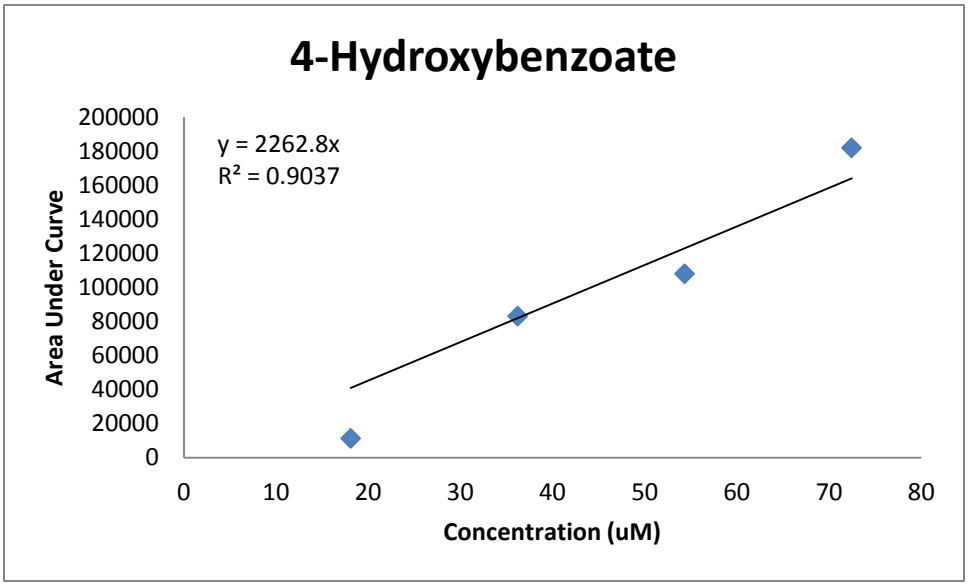
C.1 2-Hydroxyphenylacetate Standard Curve



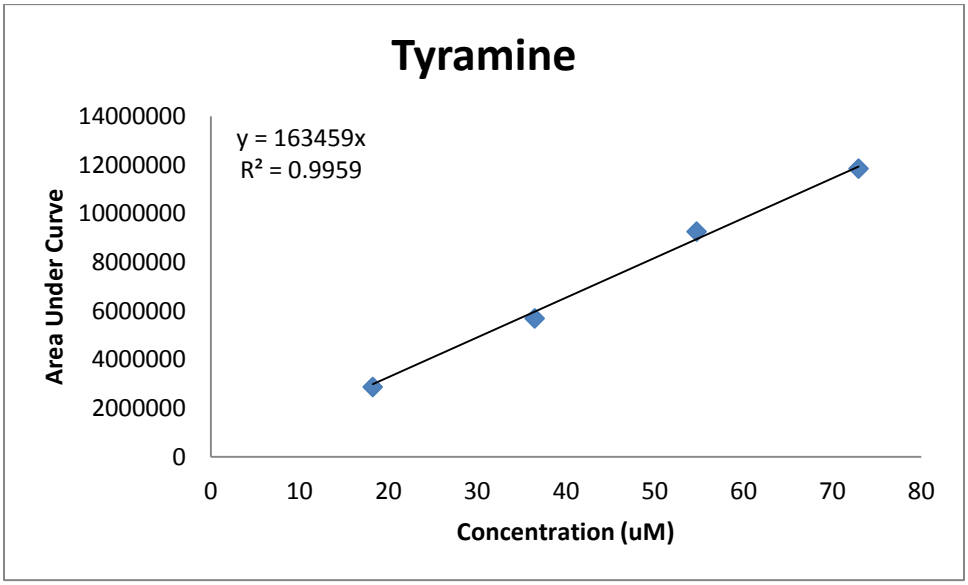
C.2 Phenylpyruvate Standard Curve



C. 3 4-Hydroxybenzoate Standard Curve



C.4 Tyramine Standard Curve



D: Raw Data for Metabolite Concentrations in Control and TCDD treated Mice

4hba	1day
1	0.6916
1	4.7625
1	2.9504
1	2.3407
1	0.2079
2	13.1852
2	59.2143
2	1.8411
2	0.8692
4hba	3day
1	15.2991
1	3.3845
1	6.382
1	4.6906
1	21.0482
2	18.782
2	5.1707
2	18.1318
2	3.0824
2	1.5099
4hba	5day
1	4.0912
1	2.254
1	4.5167
1	3.9415
1	3.1489
2	21.4652
2	39.321
2	2.2806
2	2.2026
tyr	1day
1	0.2554
1	0.1232

2hpa	day 1
1	0.5532
1	0.6413
1	0.1718
1	0.1817
1	0.9235
2	0.3713
2	0.5340
2	0.0079
2	0.4406
2hpa	day 3
1	0.0693
1	0.1613
1	0.0300
1	0.0146
1	0.0064
2	0.2781
2	0.1380
2	0.0223
2	0.0460
2	0.3491
2hpa	day 5
1	0.0304
1	0.0092
1	0.4049
1	0.3328
1	0.3396
2	0.1303
2	0.1863
2	0.0607
2	0.0397
PP	1day
1	1.5777
1	1.0132

1	0.0846
1	0.4545
1	0.1184
2	0.1224
2	0.2421
2	0.4566
2	0.0258
tyr	3 day
1	0.2239
1	0.1794
1	0.2253
1	0.6493
1	0.8065
2	3.9765
2	0.0385
2	1.9157
2	0.0545
2	0.2773
tyr	5 day
1	0.3142
1	0.2592
1	0.0741
1	0.6399
1	0.3415
2	11.2683
2	1.3513
2	0.1014
2	0.2369

1	1.1724
1	0.1029
1	0.8213
2	0.9229
2	0.7823
2	0.0000
2	1.1017
PP	3day
1	0.4331
1	0.5201
1	0.0085
1	0.0063
1	0.1073
2	0.0047
2	0.1685
2	0.0045
2	0.1187
2	0.6827
PP	5day
1	0.0064
1	0.0043
1	0.3089
1	0.2147
1	0.2520
2	0.0075
2	1.0668
2	0.2912
2	0.1934

References:

Antignac et al. (2005) The ion suppression phenomenon in liquid chromatography-mass spectrometry and its consequences in the field of residue analysis. *Analytica Chimica Acta*. 529: 129-136.

Bajad, S.U. et al. (2006) Separation and quantitation of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass spectrometry. *J Chromatogr A.*1125:76–88.

Bansal, T., Alaniz, R.C., Wood, T.K., and Jayaraman, A. (2010). The bacterial signal indole increases epithelial-cell tight-junction resistance and attenuates indicators of inflammation. *Proc Natl Acad Sci.* 107:228–33.

Bien, J., Palagani, V., and Bozko, P. (2013) The intestinal microbiota dysbiosis and *Clostridium difficile* infection: is there a relationship with inflammatory bowel disease? *Therapeutic Advances in Gastroenterology*. 6(11): 53-68.

Bjeldanes, L.F., Kim, J., Grose, K.R., Bartholomew, J.C., and Bradfield, C.A. (1991) Aromatic hydrocarbon responsiveness-receptor agonists generated from indole-3-carbinol in vitro and in vivo: comparisons with 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Proc Natl Acad Sci USA.* 88:9543–7.

Chassaing, B., and Farfeuille-Michaud, A. (2011) The Commensal Microbiota and Enterpathogens in the Pathogenesis of Inflammatory Bowel Diseases. *Gastroenterology*. 140: 1720-1728.

Cummings, J. H., and Macfarlane, G. T. The control and consequences of bacterial fermentation in the human colon. *Journal of Applied Bacteriology.* 1991, 70 (6): 443−459.

Garner, C. E., Smith, S., Costello, B., White, P., Spencer, R., Probert, C. S. J., and Ratcliffe, N. M. (2007) Volatile organic compounds from feces and their potential for diagnosis of gastrointestinal disease. *The FASEB Journal*. 21: 1675-1689.

Hart, Anna. (2001) Mann-Whitney test is not just a test of median: differences in spread can be important. *BMJ*. 323: 391-393.

Hughes, R., Magee, E. A. M., and Bingham, S. (2000) Protein degredation in the large intestine: Relevance to colorectal cancer. *Current Issues Intestinal Microbiology*. 1 (2): 51−58.

Humblot, C., Combourieu, B., Vaisanen, M., Furet, J., Delort, A., and Rabot S. (2005) ^{1}H Nuclear Magnetic Resonance Spectroscopy-Based Studies of the Metabolism of Food-Borne Carbinogen 2-Amino-3-Methylimidazo[4,5-f]Quinoline by Human Intestinal Microbiota. *Applied and Environmental Microbiology*. 71(9): 5116-5123.

Joosen, A., Kuhnle, G. G. C., Aspinall, S. M., Barrow, T. M., Lecommandeur, E., Azqueta, A., Collins, A. R., and Bingham, S. A. (2009) Effect of processed and red meat on endogenous nitrosation and DNA damage. *Carcinogenesis*. 30 (8): 1402−1407.

Kloos, D. P., Lingeman, H., Niessen, W. M. A., Deelder, A. M., Giera, M., and Mayboroda, O. A. (2013) Evaluation of different column chemistries for fast urinary metabolic profiling. *Journal of Chromatography B.*

Ku, S. Y., Yip, P., and Howell, P. L. (2006) Structure of Escherichia coli tryptophanase. *Acta Crystallogr.D*. 62: 814−823.

Lu, W., Kimball, E., and Rabinowitz, J. D. (2005) A High-Performance Liquid Chromatography-Tandem Mass Spectrometry method for Quantitation of Nitrogen-Containing Intracellular Metabolites. *American Society for Mass Spectrometry*. 17: 37-50.

Mujico, J., Baccan, G., Gheorghe, A., Diaz, L., and Marcos, A. (2012) Changes in gut microbioata due to supplemented fatty acids in diet-induced obese mice. *British Journal of Nutrition*. 10: 1-10.

Nyangale, E., Mottram, D., and Gibson, G. (2012) Gut Microbial Activity, Implications for Health and Disease: The Potential Role of Metabolite Analysis. *Journal of proteome research.*11: 5573-5585.

Pryde, S., Duncan, S., Hold, G., Stewart, C. and Flint, H. (2002) The microbiology of butyrate formation in the human colon. *FEMS Microbiology Letters. 217*: 133–139.

Rajilic-Stojanovic, M., Smidt, H., and Vos, W. (2007) Diversity of the human gastrointestinal tract microbiota revisited. *Environmental Microbiology*. 9(9): 2125–2136.

Sellick C. A. et al. (2010) Evaluation of extraction processes for intracellular metabolite profiling of mammalian cells: matching extraction approaches to cell type and metabolite targets. *Metabolomics.* 6:427–438.

Selvarasu, S., Karimi, I. A., Ghim, G-H., and Lee, D-Y. (2010) Genome-scale modeling and in silico analysis of mouse cell metabolic network. *Molecular Biosystem.* 6:152–61.

Sigurdsson, M.I., Jamshidi, N., Steingrimsson, E., Thiele, I., and Palsson, B. Ø. (2010) A detailed genomewide reconstruction of mouse metabolism based on human Recon 1. *BMC Syst Biol.* 4:140.

Singh, K.P., Wyman, A., Casado, F.L., Garrett, R.W., and Gasiewicz, T.A. (2009) Treatment of mice with the Ah receptor agonist and human carcinogen dioxin results in altered numbers and function of hemotopoietic stem cells. *Carcinogenesis*. 30(1):11-19.

Sridharan, G. V., Choi, K., Prabakaran, D., Pan, L. B., Alaniz, R., Lee, K., and Jayaraman A. (2013) Prediction, Identification, and Quantification of Microbiota Metabolites in Murine Gut using *In Silico* Target Metabolomics. *Proceedings of the National Acadamy of Science.*

Tang, W., Wang, Z., Levison, B., Koeth, R., Britt, E., Fu, X., Wu, Y., and Hanzen, S. (2013) Intestinal Microbial Metabolism of Phosphatidylcholine and Cardiovascular Risk. *The New England Journal of Medicine*. 368(17): 1575-1584.

Topping, D. L., and Clifton, P. M. (2001) Short-chain fatty acids and human colonic function: Roles of resistant starch and nonstarch polysaccharides. *Physiology Review*. 81 (3): 1031−1064.

Vaahtovuo, J., Toivanen, P., and Eerola, E. (2001) Study of murine faecal microflora by cellular fatty acid analysis; effect of age and mouse strain. *Antonie van Leeuwenhoek.* 80:35–42.

Visek, W. J. (1978) Diet and cell growth modulation by ammonia. *American Journal of Clinical. Nutrition*. 31: 216S−220S.

Walker, A. W., Duncan, S. H., Leitch, E. C. M., Child, M. W., and Flint, H. J. (2005) pH and peptide supply can radically alter bacterial populations and short-chain fatty acid ratios within microbial communities from the human colon. *Applied Environmental Microbiology*. 71(7): 3692−3700.

Wei, R., Li, G., and Seymour, A. B. (2010) High Throughput and Multiplexed LC/MS/MRM Method for Targeted Metabolomics. *Analytical Chemistry*. 82(13): 5527-5534.

Wikoff et al. (2009) Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc Natl Acd Sci USA.* 106: 3698-703.

Zhao et al. (2010) CH223191 is a Ligand-Selective Antagonist of the Ah (Dioxin) Receptor). *Toxicological Sciences*. 117(2): 393-403.

Zoetendal, E. G., Rajilic-Stojanovic, M., and de Vos, W. M. (2008) High-throughput diversity and functionality analysis of the gastrointestinal tract microbiota. *Gut*. 57: 1605-1615.