APPLICATIONS OF FECAL METABOLOMICS FOR THE STUDY OF STRESS AND PREGNANCY

HANNAH SCOTT
Bachelor of Science, **University of Lethbridge, 2017**

A thesis submitted
in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in

NEUROSCIENCE

Department of Neuroscience
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

APPLICATIONS OF FECAL METABOLOMICS FOR THE STUDY OF STRESS AND PREGNANCY

HANNAH SCOTT

Date of Defence: December 3, 2020

Dr. Gerlinde Metz	Professor	Ph.D.
Thesis Supervisor		
Tony Montina	Instructor	M.Sc.
Thesis Examination Committee Member		
Dr. Douglas Inglis	Professor	Ph.D.
Thesis Examination Committee Member		
Dr. David Euston	Professor	Ph.D.
Chair, Thesis Examination Committee		

DEDICATION

To coffee, with love

ABSTRACT

The gut microbiota plays a substantial role in determining the health and behaviour of an individual. The study of this interaction requires tools with a broad, unbiased scope that can analyse both microbial and host activity; fecal metabolomics, the analysis of small-molecule metabolites, is one such method. This thesis will explore several considerations and applications of fecal metabolomics. First, a study comparing methods of sample preparation for fecal metabolomics displays the efficacy of *in silico* baseline correction combined with a Bligh-Dyer extraction technique. Next, the use of fecal metabolomics and 16S rRNA gene sequencing reveal no effect of prenatal stress on the postpartum maternal gut microbiota. Finally, the clinical applications of metabolomics is presented in a study displaying distinctions between the prenatal urinary metabolomes of obesity and gestational diabetes mellitus, identifying biomarkers of risk. These studies highlight the potential of fecal metabolomics in both research and clinical settings.

PREFACE

Hannah Scott is the primary author of all five chapters within this thesis. Edits and revisions were provided by Dr. Gerlinde Metz, Tony Montina and Dr. Douglas Inglis. Contributions of other individuals are listed below; individual designations are reflective of their position during their involvement with these studies. All shared content in this thesis is included with the permission of each co-author.

CHAPTER 2: Fecal Metabolomics: A Comparison of Sample Preparation Methods for NMR: Catherine Brown (MSc. Student, Department of Biology) contributed to the introduction, methods, and discussion of the main manuscript, and assisted with data analysis. Tony Montina contributed to the study design.

CHAPTER 3: No Effect of Prenatal Stress on the Postpartum Maternal Fecal Metabolome and Microbiota: Gerlinde Metz contributed to the study design. Jake Szamosi, Kathleen Kennedy, and Deborah Sloboda (Farncombe Metagenomics Facility, McMaster University) performed 16S rRNA gene sequencing and analysis.

CHAPTER 4: Metabolic Dysfunction in Pregnancy: Fingerprinting the Maternal Metabolome Using Proton Nuclear Magnetic Resonance Spectroscopy: Tony Montina, Gerlinde Metz, and Brenda Leung contributed to the manuscript and study design; Brenda Leung, Catherine Field and Nicole Letourneau led and recruited the cohort. Caylin Chadwick (BSc, Dept of Neuroscience) analysed samples, performed data analysis, and contributed to the manuscript introduction and methods. Marrissa Buchan (BSc, Dept of Neuroscience) assisted with data analysis and figure creation.

ACKNOWLEDGEMENTS

I would first like to thank my supervisory committee for their mentorship and guidance over the past three years; Douglas Inglis for sharing his expertise in the field of microbiology, Tony Montina for training me in all things metabolomics, and finally Gerlinde Metz for her continued support throughout my time at the University of Lethbridge. I am also incredibly grateful to Dr. Colleen Riphagen for her help proofreading my writing.

I would not have been able to complete much of this work without the help of some fantastic individuals. In particular I would like to acknowledge Erin Falkenberg, Jamshid Faraji, Nasrin Soultanpour, Mirela Ambeskovic, Marrissa Buchan, Caylin Chadwick, and Crystal Mulik from the Metz lab, Mike Opyr from the Nuclear Magnetic Resonance Facility, Kate Brown from the Department of Biology, Jake Szamosi, Kathleen Kennedy, and Deborah Sloboda of the Farncombe Metagenomics Facility at McMaster University, and Brenda Leung, Catherine Field and Nicole Letourneau for their expertise and hard work. A huge “thank you!” to all of the other students in the Metz and Montina labs for many shared laughs, tears, advice, and coffee breaks. Finally, thank you Mark for your constant warmth, support, and patience over the years; I would not have been able to do this without you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVATIONS

^{1}H-NMR	Proton Nuclear Magnetic Resonance
5-HT	Serotonin (5-hydroxytryptamine)
ANS	Autonomic Nervous System
APrON	Alberta Pregnancy Outcomes and Nutrition
AQ	Acquisition Time
AUC	Area Under the Curve
BCAA	Branched Chain Amino Acids
CNS	Central Nervous System
ddH_2O	Millipore Distilled Water
ENS	Enteric Nervous System
dNTPs	Deoxynucleoside Triphosphates
GABA	gamma-Aminobutyric Acid
GDM	Gestational Diabetes Mellitus
HDL	High Density Lipids
HPA Axis	Hypothalamic Pituitary Adrenal Axis
IBS	Inflammatory Bowel Syndrome
kDa	Kilodalton
LPS	Lipopolysaccharide
MS	Mass Spectrometry
MW	Mann-Whitney U test
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear Overhauser Effect Spectroscopy
o1p	Transmitter Offset
OPLS-DA	Orthogonal Projection to Latent Structures Discriminant Analysis
PCA	Principal Components Analysis
PCoA	Principal Coordinate Analysis
PCR	Polymerase Chain Reaction
ppm	Parts Per Million
RF	Radio Frequency
ROC	Receiver Operator Characteristic
rRNA	Ribosomal Ribonucleic Acid
SIDS	Sudden Infant Death Syndrome
SSRI	Selective Serotonin Reuptake Inhibitor
SW	Sweep Width
TD	Data Points
TSP	Trimethylsilylpropanoic Acid
VIAVC	Variable Importance Analysis based on random Variable Combination

CHAPTER 1: Introduction

Background

The microbiota within the gut play a significant role in many aspects of host health and behaviour. In recent years, understanding of the scope of its influence has increased, allowing for the identification of new factors by which these microorganisms interact with their host. As disciplines related to health transition away from the reductionist biological perspective of the past towards the broader field of omics, new techniques are constantly emerging that allow for the study of aspects of health as they relate to the whole body, not just within a single biofluid or organ. Fecal metabolomics, the study of all small molecule metabolites present within a fecal sample, is one such technique. The fecal metabolome provides valuable insights into how the biochemical pathways of both the host and microbiota affect each other, and how each are affected by additional factors such as stress.

The development of the gut microbiota

The gut is a sterile environment at birth, and immediately begins colonization at parturition (Foster & McVey Neufeld, 2013). It begins with direct exposure to the maternal microbiota, and becomes enriched over time by exposure to different environmental factors such as diet, hygiene, and the level of activity. Despite the fact that this colonization occurs after parturition, many aspects of the gut microbial community structure are affected by host genetics. Analysis of the bacterial components of fecal samples from the TwinsUK population, a cohort consisting of monozygotic and dizygotic twins based out of the United Kingdom, revealed that host genetics played a key role in the development of many microbial taxa (Goodrich et al., 2014). Mechanisms by which the microbiota and host interact are still being discovered, but are currently thought to occur primarily due to interactions with gastrointestinal function, the central nervous system (CNS), the autonomic nervous system (ANS), and the immune system; these in

turn impact fat storage, metabolic rate, cell barrier integrity, inflammation, the stress response, and other aspects of behaviour (Foster & McVey Neufeld, 2013). While there are several known mechanisms by which changes to gut microbial community structure may impact host behaviour, one of the most popular theories involves the absorption and metabolism of tryptophan into serotonin (5-HT). Tryptophan cannot be produced by the human body and must be obtained via diet, and as such is subject to the same microbiota dependent factors as any other nutrient (Agus et al., 2018). The majority of 5-HT production occurs in enterochromaffin cells and the enteric nervous system, and has been found to be impacted by various bacteria (Yano et al., 2015). The CNS can impact gastrointestinal microbiota by affecting aspects of their environment such as the rate of peristalsis, gap-junction size, and the secretion of signalling molecules, while gastrointestinal microbiota can in turn send signals to epithelial receptors and other host cells when gut permeability is increased (Rhee et al., 2009).

In human beings, the structure of the gut microbiota stabilizes at approximately two years of age; at this point, the gut microbiome is estimated to contain genetic material from over one thousand species of bacteria (Sears, 2005; Sommer & Backhed, 2013). Recent estimates count that the colon, which contains the majority of gut microbiota, houses around 38 trillion bacteria with a mass of just over 0.2 kg (Sender et al., 2016). The commensal and mutualistic relationships formed between gut microbiota and the host as both develop play surprisingly large roles in the eventual physiology of the gut. Germ free mice have been found to have enlarged ceca (Wostmann, 1981), decreased overall gut surface volume (Gordon & Bruckner-Kardoss, 1961), epithelial permeability (Lutgendorff et al., 2008), and a higher proliferation of enterochromaffin cells (O'Hara & Shanahan, 2006) when compared to their normally colonized counterparts. While some of these shortcomings, such as epithelial integrity resulting from the

effects of *Lactobacillus* on tight junctions, can be overcome with the eventual introduction of gut microbiota, insufficient exposure to diverse microbiota during development can have permanent effects on the health and behaviour of an organism.

The gut-brain axis

The gut-brain axis refers to the bidirectional linkage between the gut and the CNS. While is had been known for some time that the health of the brain and gut were interrelated, primarily due to correlation between disease states and perturbed gut function, the original definition included only the CNS, the ANS, the enteric nervous system (ENS) and the hypothalamic pituitary adrenal (HPA) axis (Carabotti et al., 2015). The gut microbiota was considered to be an active component of the gut-brain axis after a study by Sudo et al. in 2004 found that germ free mice had an abnormal stress response that was mitigated by colonization of the gastrointestinal tract with conventional microbes (Sudo et al., 2004). Today, the definition of the gut-brain axis has been expanded to include the gut microbiota, CNS, ENS, sympathetic and parasympathetic nervous system, the endocrine system, and the immune system (O'Mahony et al., 2011).

Methods of studying the gut-brain axis

Current understanding of the gut-brain axis necessitates the simultaneous study of multiple aspects of host health and microbial community structure. While the study of feces, known as coprology or scatology, has existed nearly as long as medicine itself, our current techniques allow for the great degree of sensitivity and specificity needed to make direct connections to the gut microbiota and host health. Current methods tend to focus on a single aspect of the gut-brain axis, such as gut microbial community structure, host physiology or behavioural outcomes. Correlation analysis performed on these factors can provide insights into the linkages between different constituents of the gut-brain axis, but often miss the mechanisms

by which these factors are related. New, broader techniques are currently in development which can more clearly display the multi-step processes by which the gut-brain axis functions.

16S ribosomal RNA analysis

One of the most popular techniques for studying gut microbiota is 16S rRNA (ribosomal ribonucleic acid) sequencing. This technique exploits the existence of a variable sequence within the highly conserved 16S rRNA gene in bacteria, and can be used to distinguish between different groups of bacteria at the species level with high specificity. Use of this technique can provide a valuable snapshot of the different proportions of gut microbiota present in a fecal sample which aids in identifying microbiota associated with specific disease states.

It is also of interest to examine several aspects of host health when exploring the gut-brain axis. While the complete reach of the impacts of the gut microbiota on host physiology is incompletely understood, several analytical measures have been identified that are particularly useful in its study. As previously mentioned, many aspects of gut physiology such as cellular makeup and gut surface volume are impacted by the presence of microbiota. These effects can either be directly measured, or indirectly determined using downstream effects. Intestinal permeability, for example, can be examined by serum cytokine expression; lipopolysaccharide (LPS), produced in the cell wall of pathogenic gram-negative bacteria, triggers an inflammatory response when it comes in contact with membranal receptors in epithelial tissue, triggering an inflammatory response resulting epithelial barrier damage (Bein et al., 2017). The production of several neurotransmitters including 5-HT, neuroactive amino acids glycine and gamma-aminobutyric acid (GABA), histamine, and catecholamines such as dopamine, norepinephrine, and epinephrine can also be affected by gut microbiota (Oleskin et al., 2017). The interplay between GABA and 5-HT in particular has been identified in multiple disorders including

multiple sclerosis and sudden infant death syndrome (SIDS) (Malinova et al., 2018; Praveen & Praveen, 2016), marking the microbial-modulated neurotransmitter as a mediator for the gut-brain axis and providing a potential biomarker for the diseases it is associated with.

While there is much to be learned about the gut-brain axis from microbial changes and host biomarkers, the nature of the system necessitates the use of techniques which can analyze both aspects simultaneously. Metabolomics, the study of the small-molecule metabolites composing and resulting from the body's metabolism, provides a strong candidate for such analysis due to its thoroughness and sensitivity. 70% of the variance of the fecal metabolome results from gut microbial composition, allowing fecal metabolomics to represent a snapshot of metabolites resulting from both the host and microbiota present at a specific time (Zierer et al., 2018) . The levels of expression of these metabolites provide insight into which biochemical pathways are involved in both disturbances to the gut-brain axis, and the composition of the gut-brain axis itself. The two primary methods of analysis used when performing metabolomics are mass spectrometry (MS), and nuclear magnetic resonance (NMR). Each method has its advantages; MS is more sensitive, and examines volatile metabolites, while NMR examines non-volatile metabolites, presents higher compound identification reproducibility, relatively simple sample preparation methods, and does not destroy the sample in question (Emwas, 2015).

Nuclear magnetic resonance metabolomics

NMR has been one of the primary methods used to determine chemical structure since the 1950s (Markley & Westler, 2017), exploiting the magnetic field generated by the characteristic known as the spin state of several nuclei. Proton NMR (^{1}H-NMR) utilizes hydrogen-1 nuclei and is common in metabolomics due to the high natural abundance of this isotope. In brief, the sample of interest is placed in an external magnetic field and the spins of the

hydrogen nuclei align with the magnetic field. An electromagnetic pulse in the radio frequency (RF) range is then applied at a frequency that resonates with the hydrogen spins. As a result, the proton spins enter a state of resonance in which their magnetic field is aligned with the applied RF pulse (instead of the external magnetic field). Once the RF pulse is turned off, the spins will re-align themselves with the external magnetic field and this re-alignment can be detected and produces the NMR signal. The frequency of the resonance signal is subject to the influence of the magnetic fields of other nearby atoms. The deviation of an influenced proton's energy signal from that of an uninfluenced proton is referred to as the chemical shift, which is presented as a parts per million (ppm) measurement. When interpreting NMR data, the chemical shift of a signal is plotted against its intensity, appearing as a spike, or peak. The area under this peak is indicative of the total number of protons influenced by the same local factors causing them to have the same chemical shift. As such, the peak area and chemical shift are interpretable as abundance and conformation, respectively. In the case of a small-molecule metabolite, often containing numerous hydrogen protons, multiple intensity peaks are plotted along the chemical shift axis. Instances of overlap can occur, during which peaks combine to form various shapes, such as doublets and shoulders. The resulting collection of peaks is referred to as a spectrum. It is from the interpretation of the shapes and abundances of these peaks that it is possible to obtain the information needed to identify the compound present in the sample by associating the peaks with previously determined conformational data. When examining a sample containing multiple metabolites, as it the goal in metabolomics, each compound produces a unique spectra comprised of the conformational information of the molecule. These spectra themselves overlap to form a more complex combinations of peaks that is representative of the sum of all conformations present in the sample. Spectra are interpreted and compared by parsing the data into units

containing information regarding a single peak, called a bin. The area under the curve of each bin, being representative of the abundance of the molecule creating it, can serve as an area of comparison for that molecule between different samples. When all bins are compared between different samples, bins that are significantly different can then have their associated molecule identified; ideally, a change in one part of a molecule's spectrum will be reflected in each other peak.

Fecal ^{1}H-NMR metabolomics

Despite the potential of ^{1}H-NMR metabolomics as a method to study the gut-host interactions of the gut-brain axis and other aspects of host health and microbiota, there is a large amount of discrepancy regarding how fecal metabolomics is performed, particularly in regard to sample preparation (Deda et al., 2018; Karu et al., 2018). A preferable method of preparation would be quick, affordable, effective, and perhaps most essentially, consistent. The least intensive method of preparation simply involves homogenizing a fecal sample with an appropriate buffer, centrifuging to remove debris, then performing ^{1}H-NMR on the sample. Unfortunately, fecal matter contains a large amount of material sourced from both diet and host shedding that can interfere with the analysis of water-soluble metabolites such as fibrous material, lipids, and macromolecules, which can distort the obtained spectra to an unusable extent. Most solid materials can be removed using centrifugation, but this still presents the issue of soluble contamination. Filtration using a cellulose filter of a known pore size is a popular method used to remove these materials, but can greatly increase the cost of sample processing. A Bligh-Dyer methanol-chloroform extraction, which removes lipids and macromolecules using shifts between the monophasic and biphasic states of a methanol-chloroform-water solution, is much more affordable, but can increase processing time, and introduces potential variation in

results. Optimization of methods of preparation are currently ongoing, and it may be some time before there is a standardized preparation method.

Despite being a relatively new technique, fecal metabolomics also has an advantage over several other traditional methods of analysis; it provides a broad and unbiased overview of all metabolites present within a sample, and is not limited to just those present in the host or the microbiota. The cellular membrane of microbiota is typically ruptured during sample preparation, allowing the intracellular contents to be released into the fecal water comprising the sample (O'Sullivan et al., 2018). While this may initially seem like an undesired side effect, it actually allows for correlations between intracellular content and host health outcomes, and can also potentially be used as an alternative to 16S rRNA gene sequencing when determining microbial structure (Zierer et al., 2018). Feces is also a fairly non-invasive sample to collect, presenting an alternative to painful blood draws. Pairing fecal metabolomic analysis with the examination of other biofluids allows correlations to be made between metabolite use and expression in various systems. Using biofluids that are more temporally sensitive, such as saliva, in conjunction with fecal metabolomics can also provide insights into the interactions between different systems in response to both external and internal factors.

During the transition into the age of personalized, precision medicine, techniques that can provide large amounts of information at a relatively low cost will become necessary; ^{1}H-NMR metabolomics, particularly that performed on biofluids which are comparatively easy to obtain, provides a promising candidate for the identification of diagnostic clinical biomarkers. Numerous studies have already been completed which identify relevant biomarkers of disease, many of which can be detected prior to the development of symptoms (Bahado-Singh et al., 2015; Wang et al., 2011); these biomarkers in particular, when paired with lifestyle and medical

interventions, suggest the transition of the field of medicine from one focused on treatment to one with the goal of risk identification and disease prevention (Collino et al., 2013). The next stages involve both the refinement of imaging technologies and the continued expansion of existing metabolite databases (Karu et al., 2018; Wishart et al., 2018).

Overview of thesis content

In this thesis I will test the hypothesis that the health outcomes of stress and pregnancy can be examined using fecal metabolomics. This will be performed over the course of three separate studies, each of which presents and examines a specific hypothesis related to the applications of metabolomics in the context of the overarching hypothesis. First, a study comparing different methods of preparation will be presented to illustrate that fecal metabolomics can and must be optimized and standardized prior to widespread use. Next, a study in which fecal metabolomics will be paired with 16S rRNA gene sequencing to study the effects of prenatal stress on the fecal metabolome and microbiota to partially demonstrate the value of fecal metabolomics in the study of the gut-brain axis. Finally, ^{1}H-NMR analysis will be performed on urine to determine biomarkers of obesity and gestational diabetes mellitus in the maternal prenatal urinary metabolome. While this final study does not involve fecal metabolomics, it serves to demonstrate how metabolomics can be used in the discovery of underlying metabolic mechanisms and biomarkers of risk. Together, these studies provide an overview of the required development, current application, and future potential of fecal metabolomics in both research and clinical settings.

REFERENCES

Agus, A., Planchais, J., & Sokol, H. (2018, Jun 13). Gut microbiota regulation of tryptophan metabolism in health and disease. *Cell Host Microbe, 23*(6), 716-724. https://doi.org/10.1016/j.chom.2018.05.003

Bahado-Singh, R. O., Syngelaki, A., Akolekar, R., Mandal, R., Bjondahl, T. C., Han, B., Dong, E., Bauer, S., Alpay-Savasan, Z., Graham, S., Turkoglu, O., Wishart, D. S., & Nicolaides, K. H. (2015, Oct). Validation of metabolomic models for prediction of early-onset preeclampsia. *Am J Obstet Gynecol, 213*(4), 530 e531-530 e510. https://doi.org/10.1016/j.ajog.2015.06.044

Bein, A., Zilbershtein, A., Golosovsky, M., Davidov, D., & Schwartz, B. (2017, Feb). LPS induces hyper-permeability of intestinal epithelial cells. *J Cell Physiol, 232*(2), 381-390. https://doi.org/10.1002/jcp.25435

Carabotti, M., Scirocco, A., Maselli, M. A., & Severi, C. (2015, Apr-Jun). The gut-brain axis: interactions between enteric microbiota, central and enteric nervous systems. *Ann Gastroenterol, 28*(2), 203-209. https://www.ncbi.nlm.nih.gov/pubmed/25830558

Collino, S., Martin, F. P., & Rezzi, S. (2013, Mar). Clinical metabolomics paves the way towards future healthcare strategies. *Br J Clin Pharmacol, 75*(3), 619-629. https://doi.org/10.1111/j.1365-2125.2012.04216.x

Deda, O., Gika, H. G., & Theodoridis, G. A. (2018). Rat fecal metabolomics-based analysis. *Methods Mol Biol, 1738*, 149-157. https://doi.org/10.1007/978-1-4939-7643-0_10

Emwas, A. H. M. (2015). The strengths and weaknesses of NMR spectroscopy and mass spectrometry with particular focus on metabolomics research. *Metabonomics* (pp. 161-193). Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-2377-9_13

Foster, J. A., & McVey Neufeld, K. A. (2013, May). Gut-brain axis: how the microbiome influences anxiety and depression. *Trends Neurosci, 36*(5), 305-312. https://doi.org/10.1016/j.tins.2013.01.005

Goodrich, J. K., Waters, J. L., Poole, A. C., Sutter, J. L., Koren, O., Blekhman, R., Beaumont, M., Van Treuren, W., Knight, R., Bell, J. T., Spector, T. D., Clark, A. G., & Ley, R. E. (2014, Nov 6). Human genetics shape the gut microbiome. *Cell, 159*(4), 789-799. https://doi.org/10.1016/j.cell.2014.09.053

Gordon, H. A., & Bruckner-Kardoss, E. (1961, Jul). Effect of normal microbial flora on intestinal surface area. *Am J Physiol, 201*, 175-178. https://doi.org/10.1152/ajplegacy.1961.201.1.175

Karu, N., Deng, L., Slae, M., Guo, A. C., Sajed, T., Huynh, H., Wine, E., & Wishart, D. S. (2018, Nov 7). A review on human fecal metabolomics: Methods, applications and the human

fecal metabolome database. *Anal Chim Acta, 1030*, 1-24. https://doi.org/10.1016/j.aca.2018.05.031

Lutgendorff, F., Akkermans, L. M., & Soderholm, J. D. (2008, Jun). The role of microbiota and probiotics in stress-induced gastro-intestinal damage. *Curr Mol Med, 8*(4), 282-298. https://www.ncbi.nlm.nih.gov/pubmed/18537636

Malinova, T. S., Dijkstra, C. D., & de Vries, H. E. (2018, Aug). Serotonin: A mediator of the gut-brain axis in multiple sclerosis. *Mult Scler, 24*(9), 1144-1150. https://doi.org/10.1177/1352458517739975

Markley, J. L., & Westler, W. M. (2017, Aug 15). Biomolecular NMR: Past and future. *Arch Biochem Biophys, 628*, 3-16. https://doi.org/10.1016/j.abb.2017.05.003

O'Hara, A. M., & Shanahan, F. (2006, Jul). The gut flora as a forgotten organ. *EMBO Rep, 7*(7), 688-693. https://doi.org/10.1038/sj.embor.7400731

O'Mahony, S. M., Hyland, N. P., Dinan, T. G., & Cryan, J. F. (2011, Mar). Maternal separation as a model of brain-gut axis dysfunction. *Psychopharmacology (Berl), 214*(1), 71-88. https://doi.org/10.1007/s00213-010-2010-9

O'Sullivan, V., Madrid-Gambin, F., Alegra, T., Gibbons, H., & Brennan, L. (2018, Dec 31). Impact of sample storage on the nmr fecal water metabolome. *ACS Omega, 3*(12), 16585-16590. https://doi.org/10.1021/acsomega.8b01761

Oleskin, A. V., Shenderov, B. A., & Rogovsky, V. S. (2017, Sep). Role of Neurochemicals in the interaction between the microbiota and the immune and the nervous system of the host organism. *Probiotics Antimicrob Proteins, 9*(3), 215-234. https://doi.org/10.1007/s12602-017-9262-1

Praveen, V., & Praveen, S. (2016). Microbiome-gut-brain axis: A pathway for improving brainstem serotonin homeostasis and successful autoresuscitation in SIDS-A novel hypothesis. *Front Pediatr, 4*, 136. https://doi.org/10.3389/fped.2016.00136

Rhee, S. H., Pothoulakis, C., & Mayer, E. A. (2009, May). Principles and clinical implications of the brain-gut-enteric microbiota axis. *Nat Rev Gastroenterol Hepatol, 6*(5), 306-314. https://doi.org/10.1038/nrgastro.2009.35

Sears, C. L. (2005, Oct). A dynamic partnership: celebrating our gut flora. *Anaerobe, 11*(5), 247-251. https://doi.org/10.1016/j.anaerobe.2005.05.001

Sender, R., Fuchs, S., & Milo, R. (2016, Aug). Revised estimates for the number of human and bacteria cells in the body. *PLoS Biol, 14*(8), e1002533. https://doi.org/10.1371/journal.pbio.1002533

Sommer, F., & Backhed, F. (2013, Apr). The gut microbiota--masters of host development and physiology. *Nat Rev Microbiol, 11*(4), 227-238. https://doi.org/10.1038/nrmicro2974

Sudo, N., Chida, Y., Aiba, Y., Sonoda, J., Oyama, N., Yu, X. N., Kubo, C., & Koga, Y. (2004, Jul 1). Postnatal microbial colonization programs the hypothalamic-pituitary-adrenal system for stress response in mice. *J Physiol, 558*(Pt 1), 263-275. https://doi.org/10.1113/jphysiol.2004.063388

Wang, T. J., Larson, M. G., Vasan, R. S., Cheng, S., Rhee, E. P., McCabe, E., Lewis, G. D., Fox, C. S., Jacques, P. F., Fernandez, C., O'Donnell, C. J., Carr, S. A., Mootha, V. K., Florez, J. C., Souza, A., Melander, O., Clish, C. B., & Gerszten, R. E. (2011, Apr). Metabolite profiles and the risk of developing diabetes. *Nat Med, 17*(4), 448-453. https://doi.org/10.1038/nm.2307

Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vazquez-Fresno, R., Sajed, T., Johnson, D., Li, C. R., Karu, N., Sayeeda, Z., Lo, E., Assempour, N., Berjanskii, M., Singhal, S., Arndt, D., Liang, Y. J., Badran, H., Grant, J., Serra-Cayuela, A., Liu, Y. F., Mandal, R., Neveu, V., Pon, A., Knox, C., Wilson, M., Manach, C., & Scalbert, A. (2018, Jan 4). HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Research, 46*(D1), D608-D617. https://doi.org/10.1093/nar/gkx1089

Wostmann, B. S. (1981). The germfree animal in nutritional studies. *Annu Rev Nutr, 1*, 257-279. https://doi.org/10.1146/annurev.nu.01.070181.001353

Yano, J. M., Yu, K., Donaldson, G. P., Shastri, G. G., Ann, P., Ma, L., Nagler, C. R., Ismagilov, R. F., Mazmanian, S. K., & Hsiao, E. Y. (2015, Apr 9). Indigenous bacteria from the gut microbiota regulate host serotonin biosynthesis. *Cell, 161*(2), 264-276. https://doi.org/10.1016/j.cell.2015.02.047

Zierer, J., Jackson, M. A., Kastenmuller, G., Mangino, M., Long, T., Telenti, A., Mohney, R. P., Small, K. S., Bell, J. T., Steves, C. J., Valdes, A. M., Spector, T. D., & Menni, C. (2018, Jun). The fecal metabolome as a functional readout of the gut microbiome. *Nature Genetics, 50*(6), 790-+. https://doi.org/10.1038/s41588-018-0135-7

CHAPTER 2: Fecal Metabolomics: A Comparison of Sample Preparation Methods for NMR

INTRODUCTION

The gut microbiota can serve as both a causal agent and as an indicator of underlying health issues. These microorganisms impact host health through a variety of factors such as digestion and fermentation, modulation of immune responses such as inflammation, and protection against invading pathogens (Boulangé et al., 2016; Goldsmith & Sartor, 2014). Altered microbial community structure has been associated with a number of health conditions. While gut dysbiosis is to be expected in gastrointestinal disorders such as inflammatory bowel syndrome (IBS) (Casen et al., 2015), and forms of irritable bowel disease including ulcerative colitis and Crohn's disease (Danilova et al., 2019), a number of other chronic conditions also share an associated dysbiosis. Autoimmune disorders such as multiple sclerosis (Ochoa-Reparaz et al., 2017), mood disorders including major depressive disorder (Jiang et al., 2015), autism spectrum disorder (Mulle et al., 2013) and neurodegenerative diseases such as human motor neuron disease (Rowin et al., 2017) have all been found to have associated forms of gut dysbiosis. However, the mechanisms by which gastrointestinal microbiota affects overall health are still not fully understood.

Gastrointestinal bacteria have been shown to have an influence on neuronal development, leading to behavioural consequences during development and onward (Luczynski et al., 2016). For example, serotonergic neurotransmission can be affected by gastrointestinal microbiota (O'Mahony et al., 2015), and while the majority of bodily serotonin is produced in the digestive tract by enterochromaffin cells, serotonin itself is unable to pass the blood-brain barrier and must be synthesized in the CNS (Best et al., 2010). The precursor amino acid tryptophan is transported to the CNS after being absorbed in the gut; the availability of this compound is affected by factors regulated by gastrointestinal microbiota such as gut permeability or metabolism and

transformation into compounds other than serotonin (Zhang & Davies, 2016). The microbiota and the host's metabolism are intertwined due to the large role these microbes play in digestion and therefore nutrient absorption.

As the relationship between gastrointestinal microbiota and the host's health becomes more evident, there is increasing interest in using fecal composition as a diagnostic indicator of host health. Traditionally, the community structure of gut microbiota has been studied using 16S rRNA gene sequencing which can be used to determine the relative populations and community structure of the gastrointestinal microbiota (Johnson et al., 2019); however, this technique does not provide any insight into the actual metabolic processes affected by variations to the microbial community structure (Langille et al., 2013). Another disadvantage of 16S rRNA gene sequencing is that it is time intensive and can be cost prohibitive (Dethlefsen et al., 2008; Langille et al., 2013), however continued optimization has led to both of these factors being greatly improved in recent years.

An alternative method of analyzing the impact of dysbiosis on host health would be to examine the metabolism of both gastrointestinal microbial community structure and the host. Metabolomics, the study of the small-molecules forming and resulting from the biochemical pathways of an organism, can be used to determine the metabolomes present across the various tissues and biofluids of an organism (Lindon, 2000; Nicholson, 2008, 1999). The fecal metabolome receives up to 68% of its variance arising from gastrointestinal microbiota, not host genetics (Zierer et al., 2018). Metabolomic analysis of stool samples has the potential to be a non-invasive research and diagnostic tool of clinical relevance.

Currently, mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy are the primary tools used for metabolomics studies (Emwas, 2015; Markley et al., 2017; Probert

et al., 2004). MS provides a higher degree of sensitivity, and is often used for targeted or selective analysis; this technique also allows for the examination of volatile, nonpolar organic compounds. However, MS requires a large amount of demanding sample preparation and offers only moderate reproducibility. NMR allows for the analysis of both polar and nonpolar metabolites, and offers quantitative measurements, high reproducibility and requires relatively minimal sample preparation, but has a lower level of sensitivity and specificity. However, as technology is being optimized the sensitivity of NMR is expected to increase (Moser et al., 2017), so the current gap in sensitivity between MS and NMR spectroscopy can be expected to lessen with time. The methods by which stool samples can be prepared for metabolomics vary between analytical techniques; this study focuses on fecal metabolomic sample preparation for ^{1}H-NMR.

Despite the presence of established methods and databases for NMR metabolomic analysis of serum (Psychogios et al., 2011), urine (Bouatra et al., 2013) and saliva (Dame et al., 2015), the study of fecal metabolomics using NMR is still an emerging field with a large amount of variation in sample preparation methods between studies (Deda et al., 2018; N. Karu et al., 2018; Matysik et al., 2016). There is currently no universal consensus on the methods by which to prepare and analyze fecal samples for metabolomic analysis via proton nuclear magnetic resonance (^{1}H NMR) (Cui et al., 2020). The analysis of stool samples using metabolomic methods presents a series of unique challenges, namely the solid nature of the sample, and the presence of confounding factors and molecules such as water, fiber, proteins and lipids in varying amounts. Inconsistent and incomplete removal of fats and macromolecules from fecal samples can result in an increase in the number of broad signals in the baseline of NMR spectra and these signals can often span several parts per million on the NMR chemical shift scale. As a

direct result, there is a great deal of inaccuracy in the quantification of any metabolite that overlaps with these broad signals. The fecal metabolome is also understandably sensitive to variations in diet (Pfalzer et al., 2015), often requiring the use a controlled diet for subjects prior to sampling.

The methods used to prepare fecal samples for analysis with NMR tend to center around the removal of lipids, proteins, other macromolecules and solid matter such as fiber. The sample must be in liquid form prior to NMR metabolomics; this can be accomplished by simply centrifuging the sample to retrieve the fecal water which composes 74.6% of stool weight (Rose et al., 2015), and contains most bioactive compounds found in the original sample (Nyangale et al., 2012). Another common technique is to homogenize the sample in a PBS based buffer containing an antimicrobial agent; this can be done with or without lyophilization of the sample, and can be dependent on the weight to volume ratio (Wu et al., 2008). An important consideration to make at this point of sample preparation is the release of the intracellular components of bacteria present in the fecal sample; this can increase the number of branched-chain and aromatic amino acids present in a sample compared to those where the microbes are left intact (Karu et al., 2018). Whether the presence of these compounds is desired or not, it is critical that the respective technique selected treats these samples consistently. Techniques such as repeated freeze-thaw cycles, sonication or chemically lysing the bacteria will provide consistent lysis, whereas insufficient or inconsistent homogenization may increase the variability between samples.

Methods of storage must also be considered; the analysis of fresh samples is not always feasible, particularly if there is a large number of samples to analyze. Due to the high amount of variation between individual metabolomes, large sample sizes are often required in NMR

metabolomics to analyze the effect of a single factor, especially when diet and environment cannot be controlled (Emwas, 2015). Fortunately, most metabolites are stable and safe from microbial damage below -25°C (Lauridsen et al., 2007), but are most often stored at -80°C (Karu et al., 2018). An antimicrobial agent may be added, and if unnecessary freeze-thaw cycles are avoided samples may be stored indefinitely with minimal damage to metabolome integrity. Previous studies have also examined the impact of analyzing fresh fecal samples as opposed to frozen ones (Bezabeh et al., 2009; Saric et al., 2008). Freeze drying is also a popular technique with many comparisons to frozen samples (Chai et al., 2015; Deda et al., 2017; Naama Karu et al., 2018; Lamichhane et al., 2017; Lamichhane et al., 2015) Depending on the study design, samples may also be pooled to decrease day-to-day variability between individuals, but this may not always be possible due to aspects of study design such as limited timepoints and group size (Lamichhane et al., 2017).

Once the sample is in liquid form, immediately performing NMR comes at the risk of allowing confounding factors to interfere with the spectral analysis of the water-soluble metabolites present. In order to remove most solid factors, such as dietary fiber or cellular debris, centrifugation is often adequate (Karu et al., 2018). The removal of fats and macromolecules, when performed, often involves either ultrafiltration filtration, or a liquid-liquid extraction using chloroform and methanol. Ultrafiltration involves the use of a cellulose filter with a set pore size, often 3 kDa, which would permit the passing of most small molecule metabolites through, and remove any larger macromolecules or proteins. This technique has been recommended by The Metabolomics Innovation Center (TMIC) for many types of samples (N. Psychogios, 2011; Tiziani et al., 2008), because it yields a good signal-to-noise ratio, robust reproducibility without unwanted evaporation of volatile metabolites, and avoids loss of metabolites by dissolution in

solvents or the hydrolysis of metabolites due to the need to neutralize the solution. While this method is quick and effective, the high cost of the filters may make it prohibitively expensive.

A liquid-liquid extraction relies on the concentration-dependent phase separation of an organic solvent, methanol and water (Bligh & Dyer, 1959). The organic solvent, often chloroform, removes any organic soluble material, while the methanol precipitates any proteins in solution. While this method can be used to extract any one of these components, applying it to a sample for metabolomics allows for the removal of everything aside from the water-soluble metabolites. This may present an issue for those hoping to study lipid soluble metabolites, and the presence of methanol may alter the metabolome. There is also the issue of a required drying stage, which adds several days of preparation time to this method. However, the low relative cost of this method makes it a popular choice for metabolomics (Deda et al., 2018; N. Karu et al., 2018; Matysik et al., 2016).

An alternative approach for the consistent removal of broad signals from NMR spectra is software-based baseline correction. With respect to metabolomics, however, these methods have only been applied to smoothing or flattening distortions in the baseline of a spectrum, such as rolling baselines, and not the removal of the large broad signals that may be present due to residual macromolecules (Izquierdo-Garcia et al., 2011; Smolinska et al., 2012).

The current study was focused on comparing various sample preparation methods for fecal metabolomics analysis by ^{1}H-NMR on rat fecal samples. The methods tested include the following: (1) no extraction in which the fecal sample was added to buffer, vortexed, and centrifuged (hereafter referred to as “no extraction”); (2) the traditional Bligh-Dyer fat and protein extraction method (Bligh & Dyer, 1959b; Wu et al., 2008); (3) a modified version of Bligh-Dyer method that was intended for samples which have high fatty acid content (Cordoni et

al., 2016); (4) removal via ultrafiltration (Karu et al., 2018); (5) various combinations of the above methods. It was hypothesized that each technique used would improve the quality of the spectrum produced compared to the no-extraction method, but variations in efficacy and quality of spectra were anticipated. Lastly, the most effective of these methods was compared to results achieved using the no-extraction method in which the spectra were processed using a software-based baseline correction. The latter study examined a second hypothesis stating that a spectrum comparable to those produced using ultrafiltration could be produced using this *in silico* approach. All of the methods were evaluated both qualitatively and quantitatively based on signal-to-noise ratio, least amount of baseline variation and the most consistent differences between samples.

MATERIALS AND METHODS

Ethics approval

Procedures involving the handling of rats were approved by the University of Lethbridge Animal Care Committee under guidelines put forth by the Canadian Council on Animal Care following Animal Use Protocol #1715. The study was carried out in strict accordance with the recommendations established in the Canadian Council on Animal Care Guidelines.

Animal protocol

Long-Evans rats were born and housed in the University of Lethbridge vivarium on 12:12 hour day/night cycle with lights on at 7:30 AM in standard polycarbonate cages on corn chip bedding. Room temperature in the housing area was set at 22°C, relative humidity was 30%, and animals had access to food and water *ad libitum*.

Sample collection

Four feces samples were collected from adult Long Evans rats and stored at -80º C within 20 minutes of excretion. Frozen samples consisting of multiple pellets were mixed with metabolomics buffer at a 2:1 volume to mass ratio and vortexed until the sample became homogenous, creating equivalent sets of samples that could be used for comparing the reproducibility and variability of the fecal metabolome when samples are processed using different methodologies. This homogenous sample of feces in buffer is hereafter referred to as fecal water. The metabolomics buffer consists of a 4:1 ratio of mono and dibasic potassium phosphate salts (K_2HPO_4/KH_2PO_4) in H_2O w/ 3 mM sodium azide (NaN_3). The resulting pH of this buffer is 7.4 and the sodium azide was added as an anti-microbial agent. Fecal water samples were then frozen and stored at -80º C until being further processed.

Sample preparation – comparing methods without variation across individuals

Water soluble metabolites were isolated from a single fecal sample using the following six techniques: (1) no extraction; (2) single ultrafiltration; (3) Bligh-Dyer extraction; (4) fatty acid favoring Bligh-Dyer extraction; (5) filtration followed by Bligh-Dyer extraction, and (6) double ultrafiltration.

1. *No extraction*

   An aliquot of 275 µL of unaltered fecal water was set aside for NMR sample preparation.

2. *Single ultrafiltration*

   Amicon 0.5 mL 3 kDa centrifuge filters were washed 10 times with deionized water immediately prior to use in order to remove glycerol preservative from the filters. 350 µL of fecal water was loaded into hydrated 3kDa filters and centrifuged at 14 000g for 30 minutes at 4ºC. 275 µL of the filtered fecal water was set aside for NMR sample preparation.

3. *Bligh-Dyer methanol-chloroform extraction*

   An aliquot of 275µL of the fecal water sample was combined with 687.5µL of methanol and 343.75µL of chloroform in a 2mL tube. The tube was then vortexed and stored at -20º C for 15 minutes to allow for protein precipitation, after which it was centrifuged at 15 000g for 15 minutes at 4ºC. Afterwards the supernatant was decanted into a new tube, and equal volumes of 343.75µL of chloroform and Millipore water were added. This mixture was then vortexed briefly and centrifuged at 6700g for 5 minutes. 1000µL of the top, aqueous layer containing the water-soluble metabolites was then pipetted into new tubes that were placed in

a nitrogen gas flow box for 5-6 days to dry. Once dry, samples were rehydrated in 275 µL of Millipore water. This rehydrated sample was then set aside for NMR sample preparation.

4. *Fatty acid removal favoring Bligh-Dyer extraction*

An aliquot of 275 µL of fecal water was combined with 1500 µL of methanol and 1500 µL of chloroform. The sample tube was then vortexed briefly before being stored at -20°C for approximately 15 minutes. The resulting precipitated protein was removed via centrifugation at 15 000g for 15 minutes at 4°C. The supernatant was then collected and moved to a microcentrifuge tube containing 375 µL of methanol, 375 µL of chloroform and 475 µL of Millipore water. The sample was then vortexed before being centrifuged at 6800g for 5 minutes at 4°C. After centrifugation, 1000 µL of the aqueous layer of the sample, which contains the water-soluble metabolites, was placed in a new microcentrifuge tube and left to dry in a nitrogen gas flow chamber for 5-6 days. Once dry, samples were rehydrated in 275 µL of Millipore water. This rehydrated sample was then set aside for NMR sample preparation.

5. *Filtration followed by Bligh-Dyer Extraction*

The 275 µL of filtrate obtained at the end of the single ultrafiltration process outlined above is processed a second time using the protocol outlined for the Bligh-Dyer extraction, where the extraction begins with the filtrate rather than fecal water sample.

6. *Double ultrafiltration*

The 275µL of filtrate obtained at the end of the single ultrafiltration process outlined above was filtered again by loading the filtrate into a second filter and following the same filtration steps as outlined above.

Sample preparation – Replication of preparation methods across multiple samples

The following three methods were tested for variation across individuals: (1) single ultrafiltration; (2) Bligh-Dyer methanol-chloroform extraction; (3) no extraction processed. Each of these extraction methods were performed on three separate fecal samples obtained using the steps outlined in the sections above.

In silico baseline correction

In silico baseline correction, which is applied after NMR data acquisition and initial processing, was then implemented on the spectra obtained from the Bligh-Dyer and no-extraction methods. This step allowed for a comparison of the impact of a software-based baseline correction filter to the spectra obtained using the filtration method. The details of the baseline correction method are provided below following the NMR data acquisition and processing section.

Final preparation for ^{1}H-NMR spectroscopy

An aliquot of 275 µL of fecal water from each sample was mixed with 205 µL of KF buffer and 120 µL of D_2O containing 0.02709% w/v trimethylsilylpropanoic acid (TSP) and vortexed. Samples were then centrifuged at 12000 x g for 5 minutes. After centrifugation, 550

NMR data acquisition and processing

Spectra were obtained using the Bruker 1-D NOESY gradient pulse pre-saturation water suppression pulse sequence ‘noesygppr1d’ with 10 ms mixing time. Each sample was run for 512

scans to a total acquisition size of 128 k, a spectral window of 20.5 ppm, a transmitter offset of 4.7 ppm and a recycle delay of 4 seconds. All measurements were recorded using a Bruker triple resonance TBO-Z probe. The Bruker automation program "pulsecal" was used on each sample before data acquisition to guarantee that the 90-degree pulse was calibrated correctly, ensuring quantitative and comparable data across samples (Paxman et al., 2018). The spectra were zero filled to 256 k, automatically phased, baseline corrected and line-broadened by 0.3 Hz (Kiss et al., 2016).

In silico baseline correction

A software-based baseline correction method developed by Bruker was applied to spectra to remove broad baseline interference. The methods uses features within the Bruker Assure RMS (Raw Material Screening) software typically used for quality assurance and the screening of impurities in the manufacturing of pharmaceuticals. This process identified and fit a component to the baseline distortion present in a sample, then subtracted this component from each spectra it was applied to. This process was capable of removing variable levels of baseline distortion, making the baseline consistent between samples containing different amounts of interfering contaminants.

RESULTS

Variations in baseline interference and signal-noise ratio were observed between each of the 6 examined extraction methods (Figure 1).

No extraction

Unaltered fecal water preparation presented the least effective baseline distortion of all tested methods. This distortion was particularly visible around the 2ppm area. However, the signal to noise ratio appeared to be higher in the Bligh-Dyer extraction and single filtration spectra, particularly the peak at 7.5 ppm.

Bligh-Dyer extraction

Samples using variations of a Bligh-Dyer extraction resulting from different relative concentrations of methanol, chloroform, distilled water and fecal water displayed varying levels of both spectra distortion and metabolite signal to noise ratio. The standard Bligh-Dyer extraction method performed best in this study. The fatty acid favoring variant which initially appeared promising demonstrated a significant decrease in baseline distortion, but also reduced the signal to noise ration in several areas. Replication of this method also proved to be inconsistent, removing more baseline distortion from some samples more than others.

Filtration

All methods of sample filtration flattened the baseline in a consistent manner. Repeated sample filtration did not visibly improve spectra distortion or signal to noise ratio, indicating that this method removed all possible interfering material on the first pass. Filtering followed by Bligh-Dyer extraction provided a baseline comparable to single filtration but there were fewer peaks, and the peaks present were less intense than with filtration alone.

Baseline correction

Figure 2 shows the results of the *in silico* baseline correction on the samples processed using the no-extraction and Bligh-Dyer methods (Figure 3A) and compared them to the single ultrafiltration method (Figure 3B). Baseline correction improved the distortion of both the no-extraction and Bligh-Dyer extraction methods, generating a baseline level comparable to filtration. However, some fanning was still evident in each technique used. The baseline correction resulted in a larger section of the spectrum around the water peak that was unusable compared with uncorrected samples. Whereas in Figure 3A, the water peak has been completely removed, in Figure 3B shows a section of the water peak that was wider than in the filtration samples as shown by the horizontal lines on either side of the break

DISCUSSION

The efficacy of an extraction method is determined by its ability to effectively remove unwanted macromolecules while preserving metabolites. Six different extraction methods were examined: single ultrafiltration, filtration followed by Bligh-Dyer extraction, double filtration, fatty acid favouring extraction, a standard Bligh-Dyer extraction, and unextracted fecal water. An *in silico* baseline correction was applied to samples undergoing both the Bligh-Dyer method and unextracted fecal water assessment.

Unaltered whole fecal water predictably resulted in the most severe baseline distortion. Comparatively, the Bligh-Dyer extraction, fatty acid extraction and the two filtration methods showed spectra with reduced baseline distortion. In the single sample comparison, the resultant spectrum after fatty acid extraction displayed the least baseline distortion of all extraction methods tried, signifying that the interfering fats and macromolecules have been nearly completely removed. However, this technique also removed some of the peaks almost entirely, particularly evident for the peaks at 5.25 and 2.4 ppm (Figure 1).

Single ultrafiltration produces a spectrum which is comparable to the Bligh-Dyer extraction spectrum, producing only slightly more distortion and much higher signal to noise ratio. Repetition of filtration did not produce a noticeable benefit; in fact, the double ultrafiltration method surprisingly produced slightly more distortion than a simple single filtration, possibly due to the extra handling and time spent at room temperature. The filtration followed by Bligh-Dyer extraction method displayed a baseline distortion similar to the single ultrafiltration method, suggesting that the filtration removed the majority of the interfering substances. The single ultrafiltration and the Bligh-Dyer extraction methods showed the most consistent levels of low baseline distortion without severely compromising signal to noise ratio.

The fanning caused by the variable presence of fats and macromolecules following a Bligh-Dyer extraction is demonstrated in Figure 3, which depicts an overlay of 80 rat fecal samples from a separate, unpublished study. A large, broad peak distorts the baseline of each spectrum, the size of which is dependant on the varying confounding factors. For comparison, filtration produces visibly less fanning due to the more consistent efficacy of the technique (Figure 4). The area under the peak inside a bin is ideally reflective of the abundance of its respective metabolite, so baseline distortion that varies between samples can compromise entire studies. The *in silico* baseline correction method applied removed this issue, allowing for the application of less expensive methods such as no extraction and Bligh-Dyer extraction.

The *in silico* baseline correction is an attractive choice for labs that cannot or do not want to pay for expensive filters. *In silico* correction applied to the Bligh-Dyer extracted samples produced results most comparable to the single ultrafiltration spectrum with regards to baseline distortion and fanning caused by the presence of fats in the samples. Single ultrafiltration is quickly becoming the standard in NMR metabolomics, though some groups have suggested that ultrafiltration can remove some metabolites from the sample (Snytnikova et al., 2019). Filtration is often praised for its excellent reproducibility, but the cost of the filters is quite substantial (Karu et al., 2018; Matysik et al., 2016). The Bligh-Dyer methanol-chloroform method is less expensive, though it does take longer to process the samples due to the 5-6 day wait while the samples dry, although the Bligh-Dyer method is not quite as reproducible as the filters, with the *in silico* correction it produces a comparable spectrum though there are fewer peaks and they tend to be a bit less intense than with filtration. With unaltered whole fecal water, a large amount of baseline distortion and fanning was observed because no attempt was made to remove any of the macromolecules that contribute to these. However, with the application of the *in silico*

correction a spectrum, comparable to filtration is produced. In fact, there are often more peaks that are more intense than in the filtered samples, though it should be mentioned that there were many more spectrum acquisition failures with the unaltered whole fecal water. It is possible that the fats and proteins still present interfered with the acquisition. The advantage of the unaltered whole fecal water is its speed and its low cost.

Whether unaltered whole fecal water or Bligh-Dyer followed by *in silico* baseline correction is best may depend on the host species. In the study of the rat fecal metabolome, the results of this study suggest that Bligh-Dyer followed by *in silico* correction provides the best results for the lowest cost. The downside to the *in silico* approach is that it can only be used for qualitative metabolomics; the software modifies the baseline, leaving no possibility of quantitative analyses. *In silico* corrections can also add a lot of time to analysis as it involves finding the correct filter size which is likely to be different for every species, but this additional time requirement becomes negligible once a size is selected for a specific sample type and extraction method.

The complex and heterogeneous macromolecule composition of feces introduces variation to water soluble metabolite expression between groups that may confound the results of studies relying exclusively on conventional protein and lipid extraction methods. The signal to noise variance among the results, especially differences observed in the fatty acid extraction and Bligh-Dyer extraction then filtration spectra, potentially indicate variability in metabolite recovery when a Bligh-Dyer extraction is used. Single ultrafiltration has little fanning and signal differences between samples while maintaining a strong signal to noise ratio. Based on the results presented, the single ultrafiltration method is the most effective method tested. This is consistent with recent review findings regarding human fecal metabolomics which also found

ultrafiltration to be an effective preparation method (Karu et al., 2018; Matysik et al., 2016). However, ultrafiltration is also the costliest method investigated in this comparison. The results of the current study suggest that comparable results can be obtained using the Bligh-Dyer extraction method and unaltered whole fecal water when combined with *in silico* baseline correction.

CONCLUSION

The current study demonstrates the relative effectiveness of several popular water-soluble metabolite extraction methods for ^{1}H-NMR analysis of the rat fecal metabolome. The hypothesis that each extraction technique would display improvements over the no-extraction method was supported in the results, with the most noticeable benefits observed in the ultrafiltration and Bligh-Dyer extraction methods. The results also indicate that ultrafiltration, and a combination of a Bligh-Dyer extraction and *in silico* baseline correction produce spectra with high reproducibility, good signal to noise ratio, and consistent baseline levels, supporting our second hypothesis that the *in silico* technique used can create results comparable to filtration. Selection of either technique will depend on whether a study values quick turnaround or low cost. The author recommends that further optimization be performed on additional species, and to draw comparisons to additional preparation and metabolite extraction methods as they become available.

Determining a consistent method for the preparation of samples for fecal metabolomic analysis will further the investigation into the fecal metabolome as a research and diagnostic tool. A method with which to standardize fecal metabolomic research, that is proven to consistently achieve accurate metabolite isolation, will allow for better metabolite profile cataloguing in the future. As the field evolves, it is crucial that a standardized methodology

becomes commonplace to allow the comparison of different studies as they emerge, rather than requiring the wasteful repetition of experiments prior to the translation of results. Reliable contributions to resources such as the human fecal metabolome database (HFMD) (Karu et al., 2018) offer researchers a better means of accessing and sharing information to further the field of fecal metabolomics.

FIGURES AND CAPTIONS

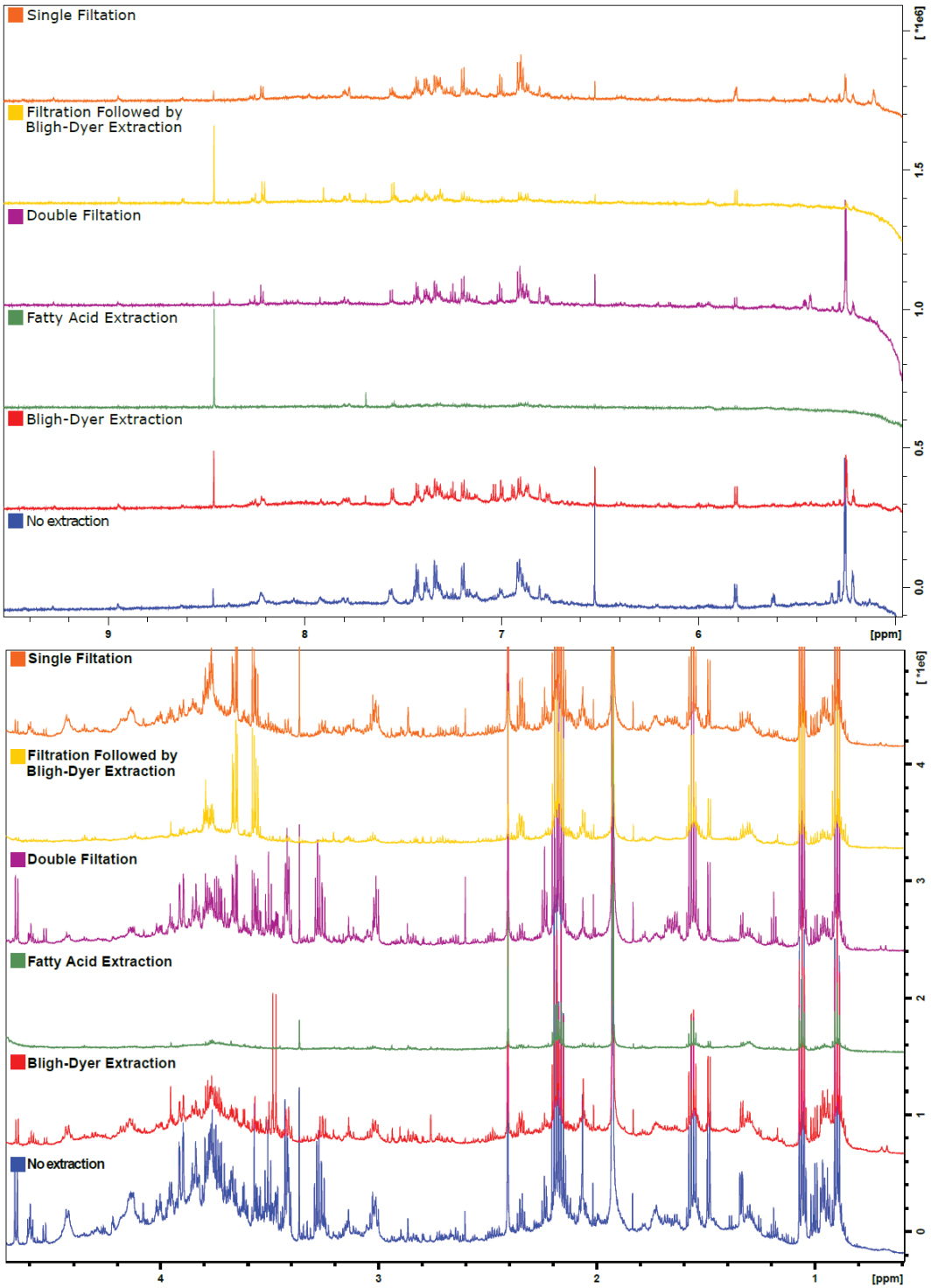

Figure 1. NMR spectra of a single rat fecal sample processed using six different techniques: (orange) single ultrafiltration; (yellow) filtration followed by Bligh-Dyer extraction; (purple) double ultrafiltration; (green) fatty acid favoring filtration; (red) Bligh-Dyer extraction, and (blue) no extraction. The water peak between 5.5 and 5.0 ppm has been removed.

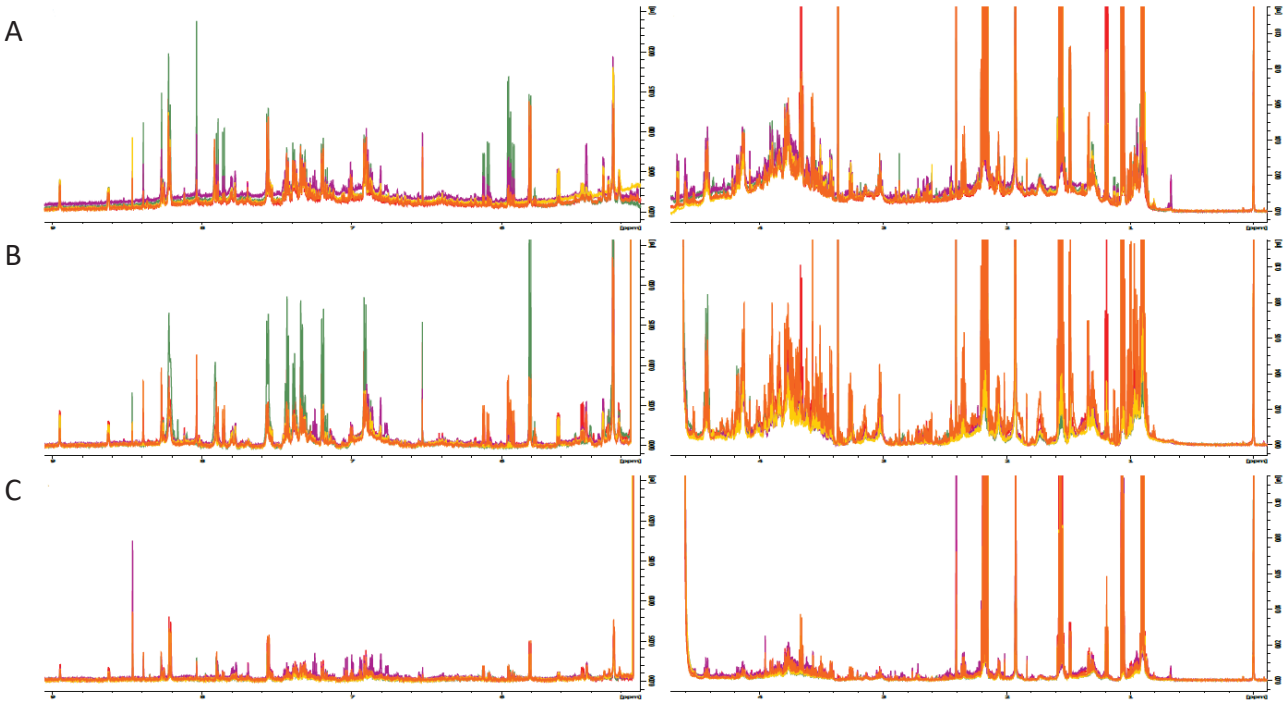

Figure 2. Comparison of filtration (A) to *in silico* baseline correction applied to the no-extraction and Bligh-Dyer method (B) in rat fecal spectra.

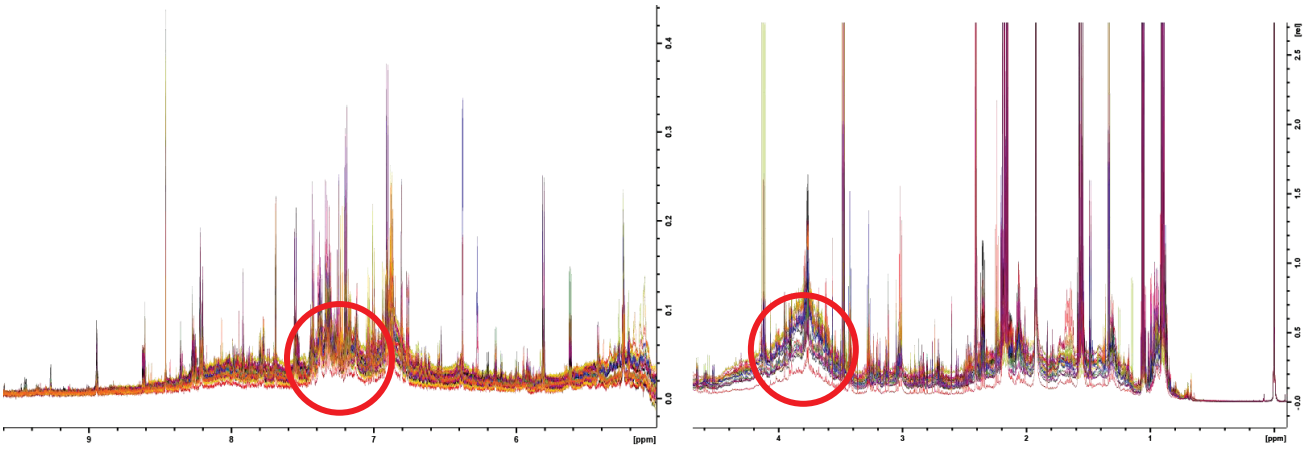

Figure 3. NMR spectra of rat fecal samples that were processed according to the Bligh-Dyer method showing fanning due to incomplete removal of fats most prominent between 5 – 3 ppm 10 ppm to the water peak, and from after the water peak to the TSP peak at 0 ppm. Areas displaying high levels of fanning are circled in red.

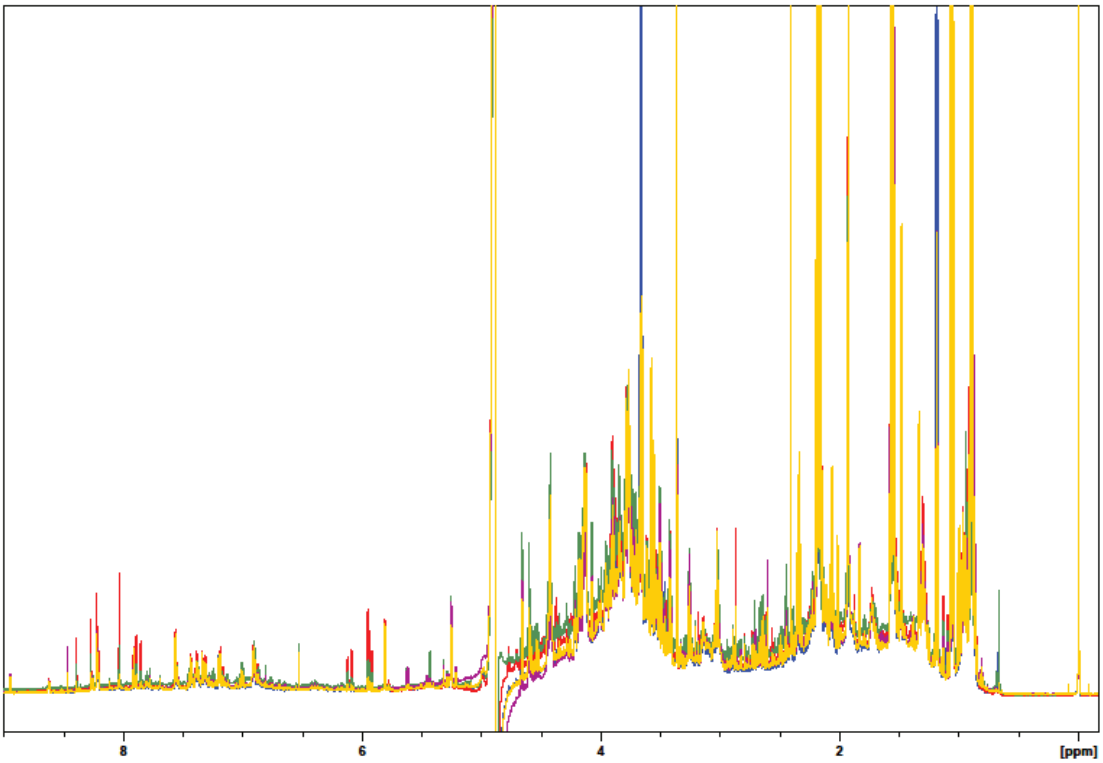

Figure 4. NMR spectra of 80 rat fecal samples that were processed by filtration, showing a more consistent baseline without significant fanning.

REFERENCES

Best, J., Nijhout, H. F., & Reed, M. (2010, Aug 19). Serotonin synthesis, release and reuptake in terminals: a mathematical model. *Theor Biol Med Model, 7*, 34. https://doi.org/10.1186/1742-4682-7-34

Bezabeh, T., Somorjai, R. L., & Smith, I. C. P. (2009, 2009-12-01). MR metabolomics of fecal extracts: applications in the study of bowel diseases. *Magnetic Resonance in Chemistry, 47*(S1), S54-S61. https://doi.org/10.1002/mrc.2530

Bligh, E. G., & Dyer, W. J. (1959a, 1959-08-01). A rapid method of total lipid extraction and purification. *Canadian Journal of Biochemistry and Physiology, 37*(8), 911-917. https://doi.org/10.1139/o59-099

Bligh, E. G., & Dyer, W. J. (1959b, Aug). A rapid method of total lipid extraction and purification. *Can J Biochem Physiol, 37*(8), 911-917. https://doi.org/10.1139/o59-099

Bouatra, S., Aziat, F., Mandal, R., Guo, A. C., Wilson, M. R., Knox, C., Bjorndahl, T. C., Krishnamurthy, R., Saleem, F., Liu, P., Dame, Z. T., Poelzer, J., Huynh, J., Yallou, F. S., Psychogios, N., Dong, E., Bogumil, R., Roehring, C., & Wishart, D. S. (2013). The human urine metabolome. *PLoS One, 8*(9), e73076. https://doi.org/10.1371/journal.pone.0073076

Boulangé, C. L., Neves, A. L., Chilloux, J., Nicholson, J. K., & Dumas, M.-E. (2016, 2016-12-01). Impact of the gut microbiota on inflammation, obesity, and metabolic disease. *Genome Medicine, 8*(1). https://doi.org/10.1186/s13073-016-0303-2

Casen, C., Vebo, H. C., Sekelja, M., Hegge, F. T., Karlsson, M. K., Ciemniejewska, E., Dzankovic, S., Froyland, C., Nestestog, R., Engstrand, L., Munkholm, P., Nielsen, O. H., Rogler, G., Simren, M., Ohman, L., Vatn, M. H., & Rudi, K. (2015, Jul). Deviations in human gut microbiota: a novel diagnostic test for determining dysbiosis in patients with IBS or IBD. *Aliment Pharmacol Ther, 42*(1), 71-83. https://doi.org/10.1111/apt.13236

Chai, Y., Wang, J., Wang, T., Yang, Y., Su, J., Shi, F., Wang, J., Zhou, X., He, B., Ma, H., & Liu, Z. (2015, 2015-11-01). Application of 1H NMR spectroscopy-based metabonomics to feces of cervical cancer patients with radiation-induced acute intestinal symptoms. *Radiotherapy and Oncology, 117*(2), 294-301. https://doi.org/10.1016/j.radonc.2015.07.037

Cordoni, G., Woodward, M. J., Wu, H., Alanazi, M., Wallis, T., & La Ragione, R. M. (2016, Nov 22). Comparative genomics of European avian pathogenic E. Coli (APEC). *BMC Genomics, 17*(1), 960. https://doi.org/10.1186/s12864-016-3289-7

Cui, M., Trimigno, A., Aru, V., Khakimov, B., & Engelsen, S. B. (2020, 2020-07-21). Human faecal 1H NMR metabolomics: evaluation of solvent and sample processing on coverage

and reproducibility of signature metabolites. *Analytical Chemistry, 92*(14), 9546-9555. https://doi.org/10.1021/acs.analchem.0c00606

Dame, Z. T., Aziat, F., Mandal, R., Krishnamurthy, R., Bouatra, S., Borzouie, S., Guo, A. C., Sajed, T., Deng, L., Lin, H., Liu, P., Dong, E., & Wishart, D. S. (2015, Dec). The human saliva metabolome. *Metabolomics, 11*(6), 1864-1883. https://doi.org/10.1007/s11306-015-0840-5

Danilova, N. A., Abdulkhakov, S. R., Grigoryeva, T. V., Markelova, M. I., Vasilyev, I. Y., Boulygina, E. A., Ardatskaya, M. D., Pavlenko, A. V., Tyakht, A. V., Odintsova, A. K., & Abdulkhakov, R. A. (2019, May 15). Markers of dysbiosis in patients with ulcerative colitis and Crohn's disease. *Ter Arkh, 91*(4), 17-24. https://doi.org/10.26442/00403660.2019.04.000211

Deda, O., Chatziioannou, A. C., Fasoula, S., Palachanis, D., Raikos, N., Theodoridis, G. A., & Gika, H. G. (2017, 2017-03-01). Sample preparation optimization in fecal metabolic profiling. *Journal of Chromatography B, 1047*, 115-123. https://doi.org/10.1016/j.jchromb.2016.06.047

Deda, O., Gika, H. G., & Theodoridis, G. A. (2018). Rat fecal metabolomics-based analysis. *Methods Mol Biol, 1738*, 149-157. https://doi.org/10.1007/978-1-4939-7643-0_10

Dethlefsen, L., Huse, S., Sogin, M. L., & Relman, D. A. (2008, 2008-11-18). The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biology, 6*(11), e280. https://doi.org/10.1371/journal.pbio.0060280

Emwas, A. H. (2015). The strengths and weaknesses of NMR spectroscopy and mass spectrometry with particular focus on metabolomics research. *Methods Mol Biol, 1277*, 161-193. https://doi.org/10.1007/978-1-4939-2377-9_13

Goldsmith, J. R., & Sartor, R. B. (2014, 2014-05-01). The role of diet on intestinal microbiota metabolism: downstream impacts on host immune function and health, and therapeutic implications. *Journal of Gastroenterology, 49*(5), 785-798. https://doi.org/10.1007/s00535-014-0953-z

Izquierdo-Garcia, J. L., Villa, P., Kyriazis, A., del Puerto-Nevado, L., Perez-Rial, S., Rodriguez, I., Hernandez, N., & Ruiz-Cabello, J. (2011, Oct). Descriptive review of current NMR-based metabolomic data analysis packages. *Prog Nucl Magn Reson Spectrosc, 59*(3), 263-270. https://doi.org/10.1016/j.pnmrs.2011.02.001

Jiang, H., Ling, Z., Zhang, Y., Mao, H., Ma, Z., Yin, Y., Wang, W., Tang, W., Tan, Z., Shi, J., Li, L., & Ruan, B. (2015, Aug). Altered fecal microbiota composition in patients with major depressive disorder. *Brain Behav Immun, 48*, 186-194. https://doi.org/10.1016/j.bbi.2015.03.016

Johnson, J. S., Spakowicz, D. J., Hong, B.-Y., Petersen, L. M., Demkowicz, P., Chen, L., Leopold, S. R., Hanson, B. M., Agresta, H. O., Gerstein, M., Sodergren, E., & Weinstock, G. M. (2019, 2019-12-01). Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nature Communications, 10*(1). https://doi.org/10.1038/s41467-019-13036-1

Karu, N., Deng, L., Slae, M., Guo, A. C., Sajed, T., Huynh, H., Wine, E., & Wishart, D. S. (2018, Nov 7). A review on human fecal metabolomics: Methods, applications and the human fecal metabolome database. *Anal Chim Acta, 1030*, 1-24. https://doi.org/10.1016/j.aca.2018.05.031

Karu, N., Deng, L., Slae, M., Guo, A. C., Sajed, T., Huynh, H., Wine, E., & Wishart, D. S. (2018, 2018-11-01). A review on human fecal metabolomics: Methods, applications and the human fecal metabolome database. *Analytica Chimica Acta, 1030*, 1-24. https://doi.org/10.1016/j.aca.2018.05.031

Lamichhane, S., Sundekilde, U. K., Blædel, T., Dalsgaard, T. K., Larsen, L. H., Dragsted, L. O., Astrup, A., & Bertram, H. C. (2017, 2017-01-01). Optimizing sampling strategies for NMR-based metabolomics of human feces: pooled vs. unpooled analyses. *Analytical Methods, 9*(30), 4476-4480. https://doi.org/10.1039/c7ay01465a

Lamichhane, S., Yde, C. C., Schmedes, M. S., Jensen, H. M., Meier, S., & Bertram, H. C. (2015, 2015-06-16). Strategy for Nuclear-Magnetic-Resonance-Based Metabolomics of Human Feces. *Analytical Chemistry, 87*(12), 5930-5937. https://doi.org/10.1021/acs.analchem.5b00977

Langille, M. G. I., Zaneveld, J., Caporaso, J. G., Mcdonald, D., Knights, D., Reyes, J. A., Clemente, J. C., Burkepile, D. E., Vega Thurber, R. L., Knight, R., Beiko, R. G., & Huttenhower, C. (2013, 2013-09-01). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology, 31*(9), 814-821. https://doi.org/10.1038/nbt.2676

Lauridsen, M., Hansen, S. H., Jaroszewski, J. W., & Cornett, C. (2007, Feb 1). Human urine as test material in H-1 NMR-based metabonomics: Recommendations for sample preparation and storage. *Analytical Chemistry, 79*(3), 1181-1186. https://doi.org/10.1021/ac061354x

Lindon J. C., J. K. N., E. Holmes, J. R. Everett. (2000). Metabonomics: metabolic processes studied by nmr spectroscopy of biofluids. *Concepts in Magnetic Resonance, 12*(5), 289-320.

Luczynski, P., Whelan, S. O., O'Sullivan, C., Clarke, G., Shanahan, F., Dinan, T. G., & Cryan, J. F. (2016, Nov). Adult microbiota-deficient mice have distinct dendritic morphological changes: differential effects in the amygdala and hippocampus. *European Journal of Neuroscience, 44*(9), 2654-2666. https://doi.org/10.1111/ejn.13291

Markley, J. L., Bruschweiler, R., Edison, A. S., Eghbalnia, H. R., Powers, R., Raftery, D., & Wishart, D. S. (2017, Feb). The future of NMR-based metabolomics. *Curr Opin Biotechnol, 43*, 34-40. https://doi.org/10.1016/j.copbio.2016.08.001

Matysik, S., Le Roy, C. I., Liebisch, G., & Claus, S. P. (2016, Nov). Metabolomics of fecal samples: A practical consideration. *Trends in Food Science & Technology, 57*, 244-255. https://doi.org/10.1016/j.tifs.2016.05.011

Moser, E., Laistler, E., Schmitt, F., & Kontaxis, G. (2017, Sep 12). Ultra-high field NMR and MRI-the role of magnet technology to increase sensitivity and specificity (vol 5, pg 33, 2017). *Frontiers in Physics, 5*. https://doi.org/ARTN 4110.3389/fphy.2017.00041

Mulle, J. G., Sharp, W. G., & Cubells, J. F. (2013). The gut microbiome: a new frontier in autism research. *Current Psychiatry Reports, 15*(2). https://doi.org/10.1007/s11920-012-0337-0

N. Psychogios, e. a. (2011). The human serum metabolome. *PLoS ONE, 6*(2), 1- 23. https://doi.org/10.1371/journal

Nicholson J. K. , J. C. L. (2008). Metabonomics. *Nature, 455*, 1054-1056.

Nicholson J. K. , J. C. L., and E. Holmes. (1999). 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica, 29*(11), 1181-1189. https://www.tandfonline.com/doi/abs/10.1080/004982599238047

Nyangale, E. P., Mottram, D. S., & Gibson, G. R. (2012, Dec). Gut microbial activity, implications for health and disease: the potential role of metabolite analysis. *Journal of Proteome Research, 11*(12), 5573-5585. https://doi.org/10.1021/pr300637d

O'Mahony, S. M., Clarke, G., Borre, Y. E., Dinan, T. G., & Cryan, J. F. (2015, Jan 15). Serotonin, tryptophan metabolism and the brain-gut-microbiome axis. *Behav Brain Res, 277*, 32-48. https://doi.org/10.1016/j.bbr.2014.07.027

Ochoa-Reparaz, J., Magori, K., & Kasper, L. H. (2017, Mar). The chicken or the egg dilemma: intestinal dysbiosis in multiple sclerosis. *Ann Transl Med, 5*(6), 145. https://doi.org/10.21037/atm.2017.01.18

Pfalzer, A. C., Nesbeth, P. D., Parnell, L. D., Iyer, L. K., Liu, Z., Kane, A. V., Chen, C. Y., Tai, A. K., Bowman, T. A., Obin, M. S., Mason, J. B., Greenberg, A. S., Choi, S. W., Selhub, J., Paul, L., & Crott, J. W. (2015). Diet- and genetically-induced obesity differentially affect the fecal microbiome and metabolome in Apc1638N mice. *PLoS One, 10*(8), e0135758. https://doi.org/10.1371/journal.pone.0135758

Probert, C. S., Jones, P. R., & Ratcliffe, N. M. (2004, Jan). A novel method for rapidly diagnosing the causes of diarrhoea. *Gut, 53*(1), 58-61. https://doi.org/10.1136/gut.53.1.58

Psychogios, N., Hau, D. D., Peng, J., Guo, A. C., Mandal, R., Bouatra, S., Sinelnikov, I., Krishnamurthy, R., Eisner, R., Gautam, B., Young, N., Xia, J., Knox, C., Dong, E., Huang, P., Hollander, Z., Pedersen, T. L., Smith, S. R., Bamforth, F., Greiner, R., McManus, B., Newman, J. W., Goodfriend, T., & Wishart, D. S. (2011, Feb 16). The human serum metabolome. *PLoS One, 6*(2), e16957.https://doi.org/10.1371/journal.pone.0016957

Rose, C., Parker, A., Jefferson, B., & Cartmell, E. (2015, Sep 2). The characterization of feces and urine: a review of the literature to inform advanced treatment technology. *Crit Rev Environ Sci Technol, 45*(17), 1827-1879. https://doi.org/10.1080/10643389.2014.1000761

Rowin, J., Xia, Y., Jung, B., & Sun, J. (2017, Sep). Gut inflammation and dysbiosis in human motor neuron disease. *Physiol Rep, 5*(18). https://doi.org/10.14814/phy2.13443

Saric, J., Wang, Y., Li, J., Coen, M., Utzinger, J., Marchesi, J. R., Keiser, J., Veselkov, K., Lindon, J. C., Nicholson, J. K., & Holmes, E. (2008, 2008-01-01). Species variation in the fecal metabolome gives insight into differential gastrointestinal function. *Journal of Proteome Research, 7*(1), 352-360. https://doi.org/10.1021/pr070340k

Smolinska, A., Blanchet, L., Buydens, L. M. C., & Wijmenga, S. S. (2012). NMR and pattern recognition methods in metabolomics: From data acquisition to biomarker discovery: A review. *Analytica Chimica Acta, 750*, 82-97. https://doi.org/10.1016/j.aca.2012.05.049

Snytnikova, O. A., Khlichkina, A. A., Sagdeev, R. Z., & Tsentalovich, Y. P. (2019, May 24). Evaluation of sample preparation protocols for quantitative NMR-based metabolomics. *Metabolomics, 15*(6), 84. https://doi.org/10.1007/s11306-019-1545-y

Tiziani, S., Emwas, A.-H., Lodi, A., Ludwig, C., Bunce, C. M., Viant, M. R., & Günther, U. L. (2008, 2008-06-01). Optimized metabolite extraction from blood serum for 1H nuclear magnetic resonance spectroscopy. *Analytical Biochemistry, 377*(1), 16-23. https://doi.org/10.1016/j.ab.2008.01.037

Wu, H., Southam, A. D., Hines, A., & Viant, M. R. (2008, 2008-01-01). High-throughput tissue extraction protocol for NMR- and MS-based metabolomics. *Analytical Biochemistry, 372*(2), 204-212. https://doi.org/10.1016/j.ab.2007.10.002

Zhang, L. S., & Davies, S. S. (2016, Apr 21). Microbial metabolism of dietary components to bioactive metabolites: opportunities for new therapeutic interventions. *Genome Med, 8*(1), 46. https://doi.org/10.1186/s13073-016-0296-x

Zierer, J., Jackson, M. A., Kastenmuller, G., Mangino, M., Long, T., Telenti, A., Mohney, R. P., Small, K. S., Bell, J. T., Steves, C. J., Valdes, A. M., Spector, T. D., & Menni, C. (2018, Jun). The fecal metabolome as a functional readout of the gut microbiome. *Nature Genetics, 50*(6), 790-+. https://doi.org/10.1038/s41588-018-0135-7

CHAPTER 3: Stability of the Postpartum Maternal Fecal Metabolome and Microbiota Following Repeated Acute Prenatal Stress

INTRODUCTION

Prenatal stress refers to stress that is experienced by a mother, and in turn, the fetus, during pregnancy. Maternal stress during pregnancy has lasting effects on the health and behaviour of both mother and offspring, some of which can resonate down the maternal germline for generations (Yao et al., 2014). Prenatal stress comes in many forms, and can result from a single, acute exposure such as a single traumatic event, or from a chronic, long-term exposure such as that resulting from a disease or socioeconomic instability. The impacts of prenatal stress are also not limited to the mother and child; transgenerational stress inheritance occurs when the effects of a single stress, be it acute or chronic, lasts for multiple generations down a single geneline (Yao et al., 2014).

Despite its common use, the term stress is often ambiguously defined, and tends to vary considerably between disciplines. One of the first empirical designations was coined by Cannon in 1915 where he defined stress as the physiological response of the sympathetic nervous system to a perceived threat (Cannon, 1915). As time passed, the colloquial definition expanded to include more social and psychological traits, but during this period the term lost its ability to apply to a specific biological phenomenon. In 1998, McEwen proposed a change to the definition of stress that linked it to the maintenance of homeostasis in an organism (McEwen, 1998b). When homeostasis is disrupted, a subject enters a state referred to as allostasis during which behavioural and physiological changes occur to counterbalance the change and return the organism to its' original homeostatic state (McEwen, 1998b; McEwen & Wingfield, 2003). The extent to which an organism is undergoing allostatic processes is referred to as allostatic load. Allostasis can involve the hypothalamic–pituitary–adrenal (HPA) axis, the sympathetic and parasympathetic nervous systems, catecholamines, cytokines, blood pressure, and potentially any

biological system that adapts for the return to homeostasis following a disturbance. For the purpose of this thesis, allostasis will be defined using McEwen's current definition, 'the process for actively maintaining homeostasis' (McEwen & Wingfield, 2003), and stress will be used to describe a state that triggers allostasis in subsequent physiological response.

The HPA axis is the primary mechanism by which a stress response occurs. When a subject is exposed to a stressful stimulus, be it physiological or psychological, the hypothalamus releases corticotropin-releasing hormone, which in turn stimulates the anterior pituitary gland. Adrenocorticotropic hormone is then released into the circulatory system, and binds to the adrenal glands to stimulate the release of the glucocorticoid cortisol. The primary functions of cortisol are to increase the body's metabolic rate and support gluconeogenesis, as well as reduce suppress aspects of the immune response such as inflammation (Dickerson & Kemeny, 2004). This does not, however, lead to a negative impact on immunity, as acute stress can actually increase the presence of lymphocytes and macrophages in a potentially compromised area (Dhabhar & McEwen, 1996). While essential for the short term stress response, chronically high levels of corticosterone, the murine equivalent of cortisol, can damage immune function, memory, cognitive performance, and exacerbate disease states such as hypertension or diabetes (McEwen, 1998a). These states are thought to be due to a developed insensitivity to cortisol, as well as a growing difficulty ending a stress response. While both of these issues arise with age, chronic stress exposure, smoking, and poor diet are found to increase their severity.

One of the most novel discoveries regarding the effects and underlying mechanisms of stress lies not within host physiology, but in the activity of completely separate organisms, the commensal microbiota. In recent years, correlations between changes to gut microbiota, physiology, and behaviour led to the discovery of the gut-brain axis. This bidirectional pathway

includes gut microbiota, CNS, ENS, sympathetic and parasympathetic nervous systems, the endocrine system, and the immune system (O'Mahony et al., 2015). Microbial community structure responds to changes in host physiology (Rhee et al., 2009) and genetics (Goodrich et al., 2014), and in turn affects host health by producing molecules which interact with gastrointestinal function, the CNS, the ANS, and the immune system (Foster & McVey Neufeld, 2013).

The relationship between gut microbiota and stress was initially established through animal models. For example, germ free mice raised without a exposure to or colonization by microbiota display an abnormal stress response and high HPA activation following a stressor which can be mitigated and even reversed with inoculation and colonization with *Bifidobacteria* (Sudo et al., 2004). This bidirectional nature of the gut-brain axis as it relates to stress has been reflected in multiple other studies. Early-life stress (O'Mahony et al., 2009) and chronic stress exposure (Bharwani et al., 2016) can cause changes in microbial communities, while transplantation of bacteria from stress-exposed mice to germ free animals causes an exacerbation of the inflammatory response to facilitate a *Citrobacter rodentium* infection (Willing et al., 2011).

The gut was previously thought to be a sterile environment at birth, however several studies suggest that colonization of the fetus may begin prior to birth via the microbiota present in the placenta (Aagaard et al., 2014) due to the presence of microbiota in the meconium (Hu et al., 2013). After parturition and throughout early life, individuals are exposed to increasingly large numbers of unique communities; in humans, the gut microbial community structure reaches stability at 2 years of age (Sears, 2005). Its development typically begins with exposure to the vaginal maternal microbiota at birth, and exposure to the mother’s other microbiota during the

postpartum period. Exposure to stress during the prenatal period can cause perturbations to the maternal gastrointestinal microbiota which are reflected in adult female offspring in mice; changes in exposure of offspring to this microbiota can result not only from alterations to the maternal microbial community structure, but from alterations to maternal behaviour such as feeding patterns and hygiene, as well as other aspects of maternal health such as inflammation (Gur et al., 2017). In order to explore the gastrointestinal microbiota as a mechanism by which the effects of prenatal stress can be transmitted, we must first understand the extent to which prenatal stress affects the maternal gastrointestinal microbiota. Pregnancy itself can cause changes in the maternal microbiota (Gohir et al., 2015; Smid et al., 2018), and the extent to which these changes interact with other external factors such as stress or diet may lead to unexpected microbial and health outcomes.

Analyses of questions relating to the gut-brain axis often require the correlation of changes to microbial communities with other aspects of host health. Analysis of gut microbiota community structure is normally performed using 16S rRNA (ribosomal ribonucleic acid) gene sequencing, but the related aspects of host health and the techniques used to study them tend to vary more widely between studies. Metabolomics, the study of the small-molecules forming and resulting from the biochemical pathways of an organism, has been emerging as a popular choice for the study of the gut-brain axis due to the broad range of biochemical processes and pathways that the studied metabolites reflect (Karu et al., 2018). Fecal metabolomics examines metabolites resulting from both microbial activity and host activity, presenting a unique view into the bidirectional interactions that occur in the gut; however, it is worth noting that the majority of the fecal metabolome's variance stems from the gastrointestinal microbiota, not host genetics (Zierer et al., 2018). While trailing behind the analysis of other biofluids such as serum and blood, recent

advancements in sample preparation and database creation has made fecal metabolomics a viable and non-invasive technique for studying the microbiota-host interactions occurring in the gut (Karu et al., 2018).

This study utilizes both proton nuclear magnetic resonance (^{1}H-NMR) metabolomics and 16S rRNA gene sequencing to study how prenatal stress effects the maternal gut microbiota during the postpartum period using a cohort of Long-Evans rats. A difference would support the theory that gastrointestinal microbiota act as a method by which prenatal stress can influence the resulting geneline. It was hypothesized that repeated acute stress exposure during pregnancy would significantly alter fecal metabolomes and microbial community structure. Understanding these effects can assist in discovering the underlying connections between stress, pregnancy, and gastrointestinal microbiota, which in turn may help develop interventions for negative health and pregnancy outcomes. The current study also serves to demonstrate the use of fecal metabolomics in conjunction with 16S rRNA gene sequencing in analysis involving the gut-brain axis.

MATERIALS AND METHODS

Animals

This study involved a cohort of Long-Evans rats bred and raised at the University of Lethbridge vivarium. Animals were housed in groups of at least two in ventilated cages. Animals belonging to different groups were housed in separate rooms. Conditions were maintained at 20°C and roughly 30% humidity, with a 12-hour dark/light cycle beginning at 7:30 AM. All procedures were performed according to standards set by the Canadian Council on Animal Care and approved by the University of Lethbridge Animal Welfare Committee following Animal Use Protocol #1715.

Experimental design

Sixteen female rats were split into a stress group (n = 8) and a control group (n=8) and bred at 95 days of age. After successful implantation, individuals in the stress group were exposed to a seven-day stress program consisting of daily restraint stress from gestational days 12-18 and additional and overnight isolation stress on gestational days 14 and 17, while control individuals experienced an undisturbed pregnancy (Figure 1). After parturition occurring around the 22nd day of gestation, fecal samples were collected form mothers.

Prenatal stress administration

Animals selected for the stress group were exposed to varying amounts of restraint stress and overnight isolation stress between the 12th and 18th days of their pregnancies (Table 1). Restraint stress was administered as previously described using custom-made transparent Plexiglas tubes with a 6 cm inner diameter and an adjustable length which was determined by body size (Metz et al., 2005). These tubes render the rat unable to turn or move without

compressing the body or restricting breathing. Social isolation stress consisted of housing the rat in a standard housing unit away from its regular housing partner in a separate room.

Sample preparation

Fecal pellets were obtained by one of two methods; by allowing the rat to defecate onto a sterile surface, or by gently squeezing the perineum to force defecation of a pellet if one was present. Once the pellet was obtained, it was frozen at -80°C within 30 minutes of excretion.

1H- NMR Analysis

Preparation of feces samples for 1H – NMR analysis

300mg-500mg of frozen feces was placed in a 2:1 ratio of KF buffer to sample mass ratio with a buffer containing a 4:1 ratio of dibasic potassium phosphate (K_2HPO_4) to monobasic potassium phosphate (KH_2PO_4) with a combined concentration of 0.5M in H_2O, containing 3mM NaN_3 anti-microbial agent. The sample was then briefly mixed using a Bullet Blender (Next Advance, #3ISBBX24) without beads to ensure all parts of the fecal pellet were exposed to the buffer. Samples were then centrifuged at 20000g for 30 minutes to pellet the debris before decanting the supernatant.

Extraction of fats and macromolecules via chloroform-methanol extraction

275 µL of the newly obtained sample suspension had 687.5 µL of methanol and 343.75 µL of chloroform methanol and chloroform added in a 1:2.5:1.25 sample/water to methanol to chloroform ratio to induce protein precipitation. Samples were stored at -20°C for 15 minutes to encourage precipitation, then centrifuged for 15 minutes at 20000g. The supernatant was decanted into a new tube, and equal volumes of 343.75µL of chloroform and Millipore distilled water (ddH_2O) were added for a final ratio of 2.25:2.5:2.5 of sample/water:methanol:chloroform.

This caused a phase separation between the bottom, organic chloroform layer and top, aqueous water/sample/methanol layer, which contained the metabolites. The sample was then centrifuged at 20000 x g for 5 minutes to guarantee complete separation before 1 mL of the aqueous layer containing that water soluble metabolites was decanted off. Samples were left open in a container with $N_{2(g)}$ airflow for four days before being re-suspended in a volume of ddH_2O equivalent to their original volume before precipitation.

^{1}H- NMR analysis

275 µL of resuspended samples were mixed with 205 of KF buffer and 120 µL of D_2O containing 0.02709% w/v trimethylsilylpropanoic acid (TSP) as a chemical shift reference for a final volume of 600 µL and a respective ratio of 2.3:1.7:1. Samples were vortexed, then centrifuged at 12000 x g for 5 minutes. After centrifugation, 550 µL from each sample was transferred to an 8" 5mm glass NMR tube. Sample spectra were then collected on a 700 MHz Bruker Avance III HD spectrometer using a Bruker 1-D NOESY gradient water suppression pulse sequence 'noesygppr1d'. Each sample was run for a total of 512 scans. Spectra underwent automatic phase and baseline correction, and were referenced to TSP.

Baseline correction

A software-based baseline correction method developed by Bruker was applied to spectra to remove broad baseline interference. The methods uses features within the Bruker Assure RMS (Raw Material Screening) software typically used for quality assurance and the screening of impurities in the manufacturing of pharmaceuticals. This process identifies and fits a component to the baseline distortion present in a sample, then subtracts this component from each spectra it is applied to. This process is capable of removing variable levels of baseline distortion, making

the baseline consistent between samples containing different amounts of interfering contaminants.

NMR data acquisition and processing

NMR spectra of the fecal samples were acquired at room temperature using a 700 MHz Bruker Avance III HD NMR Spectrometer (Bruker, Ontario, Canada) equipped with a 5 mm triple resonance TBO-Z probe. The one-dimensional Nuclear Overhauser Effect Spectroscopy (NOESY) gradient water suppression pulse sequence (noesygppr1d) was utilized with the following parameters: mixing time of 10 ms; 128 k data points (TD); sweep width (SW) of 20.52 ppm, acquisition time (AQ) of 4.56 sec, transmitter offset (o1p) of 4.6 ppm; recycle delay (D1) of 1 sec; 512 scans (NS). Spectra were then processed using zero-filling to 256 k points, line broadening with a 0.3 Hz exponential multiplication, automatic phased and baseline-correction, and chemical shift referenced with respect to the TSP peak at 0 ppm. All spectra were converted to ASCII files and exported to MATLAB (MathWorks, MA, USA) for further analysis. The spectra underwent dynamic adaptive binning (Anderson et al., 2011), followed by manual adjustment to correct for any errors in the algorithm. The bins containing the water peaks were removed resulting in a total of 539 bins for all spectra. The spectra were normalized to the total area of all bins (the total metabolome), log transformed, and Pareto scaled(Craig et al., 2006).

NMR statistical analysis

Bins underwent a decision tree algorithm to determine whether bins were normally distributed using a Shapiro-Wilk test (Goodpaster et al., 2010). If bins were normally distributed, they underwent a t-test; if they were not, they underwent a Mann-Whitney U test (MW). Significant bins then underwent Variable Importance Analysis based on random Variable

Combination analysis (VIAVC) using MATLAB to perform a Receiver Operator Characteristic (ROC) test and determine the area under the curve (AUC) (Yun et al., 2015). This method uses binary matrix resampling to create 1000 subsets of variables containing 50% original bins and 50% random variables. Each variable is sampled individually, and removed if its exclusion has no effect on the AUC for the ROC curve. The final list of remaining bins is referred to as the best subset. Principal Components Analysis (PCA) was used to show the degree of separation between groups without the presence of a supervising algorithm. In addition, orthogonal projection to latent structures discriminant analysis (OPLS-DA) was used to visualize between-group separation as a function of within-group separation.

16S rRNA gene sequencing

DNA extraction

Fecal pellets were sent to the Farncombe Metagenomics Facility at McMaster University for 16S rRNA gene sequencing. One sample from the stress group was unable to be processed due to insufficient volume, reducing this group's size to n=7. Genomic DNA was extracted as previously described (Whelan et al., 2014) with the addition of a few modifications. 0.1-0.2g of sample was lysed in 800 µL of 200 mM $NaPO_4$ and 100 µL of guanidine thiocyanate-ethylenediaminetetraacetic acid-Sarkosyl using 0.2g of 2.8 mm ceramic beads (Mo Bio Laboratories, #13114-50) using a bead based homogenizer (PowerLyzer, Medicorp Inc., #13155). An additional 0.2g of 0.1mm glass beads (Mo Bio Laboratories, #13118-50) were added, and samples were homogenized in the same Powerlyzer. Samples then underwent enzymatic lysis to effectively lyse gram positive bacteria. 50 ml of lysozyme (100 mg/ml, [Sigma-Aldrich, #L6876-10G]) and 10 ml RNase A (10 mg/ml, [Qiagen, #19101]) were added to each microcentrifuge tube, which were then vortexed and incubated in a 37°C water bath for 1.5

hours. 25 µL 25% sodium dodecyl sulfate, 25 µL proteinase K (Sigma-Aldrich, #P2308-1G), and 62.5 ml 5 M NaCl were then added, followed by incubation at 65°C for 1 hour. Samples were then centrifuged at 13 000 g for 10 minutes; the supernatant was transferred to a new tube, and the pellet was discarded. 200 µL of DNA binding buffer (DNA Clean and Concentrator-25, Cedarlane Laboratories, #D4034) was added to each sample, which were then transferred to a DNA column (DNA Clean and Concentrator-25, Cedarlane Laboratories, #D4034) and spun at 12 000 x g until passed through. 200 µL of wash buffer (DNA Clean and Concentrator-25, Cedarlane Laboratories, #D4034) was then run through the column to extract any remaining impurities. The wash step was repeated once more before samples were transferred to new tubes and resuspended in 50 µL of sterile DNase/RNase free ddH20 preheated to 65°C by reversing the columns and spinning at 12,000g for 1min.

DNA amplification

Polymerase chain reaction (PCR) amplification of the variable 3 and 4 (V3-V4) regions of the 16S rRNA gene was subsequently performed on the extracted DNA from each sample using methods previously described (Bartram et al., 2011). Each reaction contained 5 pmol of primer (341F – CCTACGGGNGGCWGCAG, 806R – GGACTACNVGGGTWTC-TAAT), 200 mM of deoxynucleoside triphosphate (dNTPs), 1.5 µL 50 mM $MgCl_2$, 2 µL of 10 mg/ml bovine serum albumin (irradiated with a transilluminator to eliminate contaminating DNA) and 0.25 µL Taq polymerase (Life Technologies, #10342020) for a total reaction volume of 50 µL. Samples were prepared in a sterile PCR workstation (CanadaWide Scientific, #95-0438-01). Amplification occurred in a Thermal cycler (Veriti 96 well thermal cycler, Applied Biosystems, #437586) using the following program specifications:

- 94°C – 2min

- 30 Cycles
    - 94°C – 30s
    - 50°C – 30s
    - 72°C – 30s
- Final Cycle
    - 72°C – 10 min
    - 4°C – Hold

Five µL of the resulting product was visualized on a 2% agarose gel to ensure successful replication. 341F and 806R rRNA gene primers were modified to include adapter sequences specific to the Illumina technology and six-base pair barcodes were used to allow multiplexing of samples. 16S DNA products of the PCR amplification were subsequently sequenced twice on two separate runs of the Illumina MiSeq platform (2x300bp) at the Farncombe Genomics Facility (McMaster University, Hamilton ON, Canada).

16S rRNA gene sequence and analysis

Primers were trimmed from FASTQ files using Cutadapt (Martin, 2011) and DADA2 (Callahan et al., 2016) was used to derive amplicon sequence variants (ASVs). The resulting ASV table is a high resolution representation of how many times each unique amplicon sequence variant occurs within each sample. Taxonomy was assigned using the Silva 132 reference database (Quast et al., 2013). Non-bacterial ASVs were culled (kingdom Eukaryota, family Mitochondria, or no assigned phylum), as was any taxa with a mean count less than 10.

Alpha and beta diversity analyses was performed in R using phyloseq (McMurdie & Holmes, 2013). Species diversity reflects both the number of different species present, or the richness, and their relative abundance, or evenness. Alpha diversity is a measure of within sample species diversity, while beta diversity instead measures the variations in species diversity between samples. Both alpha and beta diversity was then tested for whole community differences across groups using vegan's (Oksanen et al., 2013) implementation of permutational multivariate

analysis of variance (PERMANOVA) in the adonis command. The resulting Bray-Curtis dissimilarities serve as a quantitative representation of beta diversity, and were visualized via Principal Coordinate Analysis (PCoA) ordination using R's ggplot2 package (Wickham, 2016). Bray-Curtis dissimilarity values fall between 0 and 1, with higher values being indicative of greater differences in species abundance between samples. Significance of alpha diversity was analyzed by a linear model with group as a fixed effect, and represented using a Shannon diversity index. The Shannon diversity index values reflect both the species richness variations in abundance; larger values (those approaching n where n = the total number of species) are indicative of higher community diversity.

.

RESULTS

Metabolomics

^{1}H-NMR analysis of fecal samples found no effect of prenatal stress of the maternal fecal metabolome during the postpartum period. A total of the 539 unique bins were identified in the spectra during comparison. The VIAVC test found two bins that contributed to separation between the groups while Mann-Whitney test identified 72 for a combined total of 74 unique bins. While some separation can be observed between the groups in both the unsupervised PCA (Figure 2A) and supervised OPLS-DA (Figure 2B) comparisons, this separation was not found to be significant, a finding which which was confirmed by cross-validation and permutation tests ($q^2 = 0.454$ ($p = 0.012$), $r^2 = 0.753$ ($p = 0.054$)). It is important to note that these analyses were performed only on bins labelled as significant by at least one of Mann-Whitney or VIAVC. ROC curves generated using variables included in the VIAVC best subset were used to determine the specificity, sensitivity, and accuracy of these comparisons, resulting in an AUC of 0.609 and a predictive accuracy of 62.3% (Figure 3). This lack of separation can also be visualized using heat maps displaying the change in metabolite expression in both all bins and in bins identified as significant by either paired t-test/Mann-Whitney or VIAVC analysis (Figure 4).

16S rRNA gene sequencing

There was no observable effect of prenatal stress on the maternal fecal microbiota during the postpartum period as determined by 16S rRNA gene sequencing. In both the control and stress groups, bacteria the *Firmicutes* phylum, particularly the *Lachnospiracae* and *Ruminococcus* families, was dominant, followed by bacteria in the *Bacteroidetes* phylum (Figure 5). No distinction between the two treatments was observed in both the Principal Coordinate Analysis (PCoA) of Bray-Curtis dissimilarities (Figure 6), and measured differences in alpha

diversity using the Shannon diversity index (Figure 7). Differences between the groups only accounted for 10.5% of the variation observed in the sample's Bray-Curtis dissimilarities (r^2 = 0.105, p = 0.078).

DISCUSSION

This study utilized ^{1}H-NMR fecal metabolomics alongside 16S rRNA gene sequencing to determine if there was an observable effect of repeated prenatal stress on the maternal fecal metabolome and microbial community structure. While bins contributing towards separation were identified by ^{1}H-NMR analysis, their effect was not large enough to produce significant separation between the stress and control treatments. This observation was confirmed by 16S rRNA sequencing, and no difference was observed between the stress and control groups using either technique. Bray-Curtis dissimilarities resulting from 16S rRNA gene sequencing accounted for only 10.5% of the variation between samples, less than could be accounted for by chance, suggesting that individual variation was greater than that explained by variation between groups.

Our results suggest that previously observed stress-induced differences in the maternal fecal metabolome and microbiota during pregnancy (Gur et al., 2017) are not reflected during the postpartum period. There are four possible explanations for this finding that will be considered in this discussion. First, that there was no effect of stress on the fecal microbiota during pregnancy in rats; second, the applied stressor was insufficient to induce changes to the fecal microbiota. Third was that there is an effect of stress that reverts after parturition, and finally that there is an effect that is overshadowed by other factors present during the postpartum period.

The first two explanations are possibly the most straightforward to address. The impact of stress on gut microbiota is a well-documented phenomenon, having been observed in a variety of species including mice (Bharwani et al., 2016; Sudo et al., 2004), rats (O'Mahony et al., 2009), and human beings (Hantsoo et al., 2019), and the work of Gur et al. (2017)on a mouse model of prenatal stress demonstrates that pregnancy does not preclude these alterations from affecting the

fecal metabolome (Gur et al., 2017). It may be that the stressor applied in this study was not of a sufficient magnitude to case stress-induced changes to gastrointestinal microbiota. A similar 6-day administration of multifactor prenatal stress including restraint stress and swimming stress was observed to affect gestational weight gain in rats, but not maternal tail-chasing behaviour or microRNA (miRNA) profiles (Yao et al., 2014). However, this and similar stressors was observed to have effects on the health and behaviour of subsequent generations, impacting preterm birth risk, gestational length, blood glucose, and maternal behaviour. Acute restraint stress was also the technique applied by Sudo et al. in their foundational study on the impact of stress on the gut microbiota (Sudo et al., 2004). Increasing the severity of the stress program or using new forms of stressors such as an acoustic startle response or social crowding may cause an observable effect in the gastrointestinal microbiota of pregnant rats, but it is important to note that exposure to extreme stress can compromise birth outcomes (Baibazarova et al., 2013; Coussons-Read et al., 2005) which may not be desirable in an animal cohort.

Determining whether the fecal microbial community structure and metabolome revert after parturition, or are overshadowed by other postpartum factors require the consideration of additional time points both before and during pregnancy. It may also be of use to sample other biofluids with a higher level of temporal sensitivity, such as serum or urine. While multiple studies have already been conducted on the effects of pregnancy on gastrointestinal microbiota at various time points (Gohir et al., 2015), comparisons of the maternal fecal metabolome and microbiota of the prenatal and postnatal period are much rarer. One recent study by Ramsteijn et al. (2020) utilized both 16S rRNA gene sequencing and metabolomic analysis to examine the effects of early life stress and selective serotonin reuptake inhibitor (SSRI) administration on the maternal fecal metabolome and microbiota during both pregnancy and lactation (Ramsteijn et al.,

2020). Their results interestingly indicate that early life stress increases microbial alpha diversity prior to pregnancy, but that this effect does not persist over pregnancy and parturition. Animals in both groups in fact display a decrease in alpha diversity during the transition from the prenatal period to the lactation period. Another study by Jost et al. (2014) found maternal gut microbial community structure remains stable over the perinatal period spanning from late pregnancy to early lactation, despite the presence of altered metabolic activity and low-grade inflammation (Jost et al., 2014). This stability may be sufficient to overshadow the ability of prenatal stress to impact the fecal microbiota in a significant manner, providing support for our third possible explanation for the lack of sufficient differences in this study.

The sources of gut microbiota stability are as of yet unknown. Many states of health and disease have an associated gut microbiota fingerprint, and it is not an improbable notion that pregnancy, parturition and lactation all have such a state resulting from underlying physiological changes in the mother. It is also likely that this stability can result from behavioural changes in the mother during and after pregnancy. In the case of rats, alterations to feeding patterns, sleep cycles, grooming, and baseline activity all occur during pregnancy. After birth, proximity to and caring for pups may cause the ingestion of microbes that overshadow other factors of gut microbial community structure such as grooming, nursing, and coprophagia. Regardless of its cause, the stability of the gut microbiota during the perinatal period are sufficient to mask the effects of prenatal stress.

One of the strengths of the current study was the use of 16S rRNA gene sequencing alongside ^{1}H-NMR fecal metabolomics. Examining only the changes to the community structure of the microbiota in response to stress overlooks potential physiological and metabolic changes occurring in the host which could be reflected in the metabolome. In the case of the current

study, it is unsurprising that both techniques yielded similar results given that 67.7% of the fecal metabolome is determined by gastrointestinal microbiota (Zierer et al., 2018). The cellular membrane of microbiota is typically ruptured during sample preparation, allowing the intracellular contents to be released into the sample alongside the metabolites produced by the host (O'Sullivan et al., 2018). As such, the metabolites present in the sample can reflect both host and microbial activity.

In this study, we hypothesized that prenatal stress would induce changes in the maternal fecal metabolome and microbiota during the postpartum period. Our results indicate that prenatal stress did not significantly impact either the microbiota or metabolome on the 21st day following parturition. While it is possible that the stress program used was not of sufficient magnitude to induce a change, the data suggests that observed stability in the perinatal gut microbial environment overshadowed any observable effect. This effect needs to be taken into account when examining the gut microbiota and metabolome in and around the perinatal period.

CONCLUSION

The results of this study suggest negligible metabolomic and genomic effects of prenatal stress on the maternal fecal microbiota during the postnatal period. 16S rRNA gene sequencing was used to examine changes to the fecal microbial community structure and diversity while ^{1}H-NMR metabolomics examined the metabolites resulting from both gastrointestinal microbiota and host metabolism. This finding disproves the initial hypothesis that prenatal stress would influence the postpartum maternal gut microbiota which were based on previous observations of stress exposure induced disruptions to gastrointestinal microbiota. Several explanations for this finding were explored, including insufficiency of the stressor and a reversion of any effect of prenatal stress during the perinatal period. However, literature describing how gut microbiota experiences stability during the perinatal period suggest that changes due to stress are overshadowed by other factors present prior to and during the postpartum period are the most likely explanation for our findings.

The identification of these confounding factors necessitates a greater understanding of both the full physiological and psychological scope of the stress response and the underlying physiological mechanisms of the gut-brain axis. The broad scope of metabolomics, particularly when applied to the study of the fecal metabolome, may assist with the discovery of these factors when used alongside traditional methods of microbial population and community structure analysis.

FIGURES AND CAPTIONS

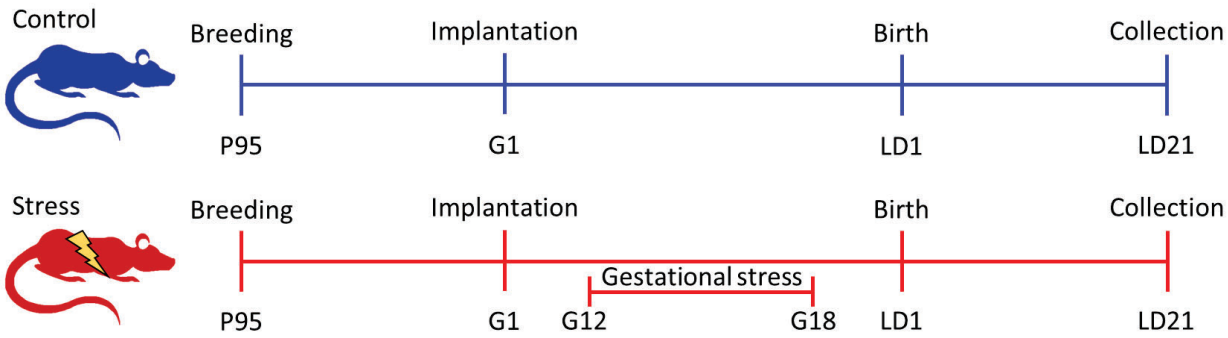

Figure 1. Two almost identical groups of female Long-Evans rats were bred. One group was exposed to social isolation and restraint stressors during pregnancy, while the other underwent no stress. The three categories representing animal are P: postnatal age, G: day of gestation (after implantation), and LD: lactation day (after birth).

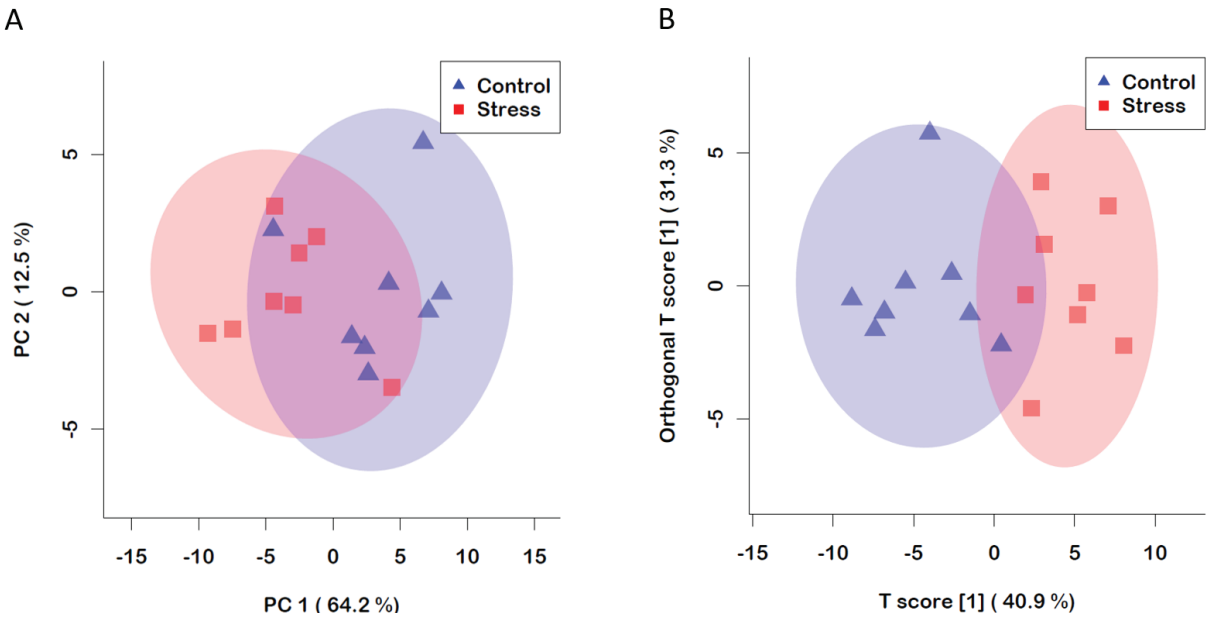

Figure 2. Principal component analysis (PCA) (A) and orthogonal partial least squares discriminant analysis (B) comparing the fecal metabolomes of stressed and control pregnant female rats. Each ellipsis represents the 95% confidence interval of their respective group. These comparisons were performed on bins identified as significant by either paired t-test/Mann-Whitney or inclusion in the VIAVC best subset list. The OPLS-DA cross-validation and permutation measures are: $q^2 = 0.454$ ($p = 0.012$), $r^2 = 0.753$ ($p = 0.054$).

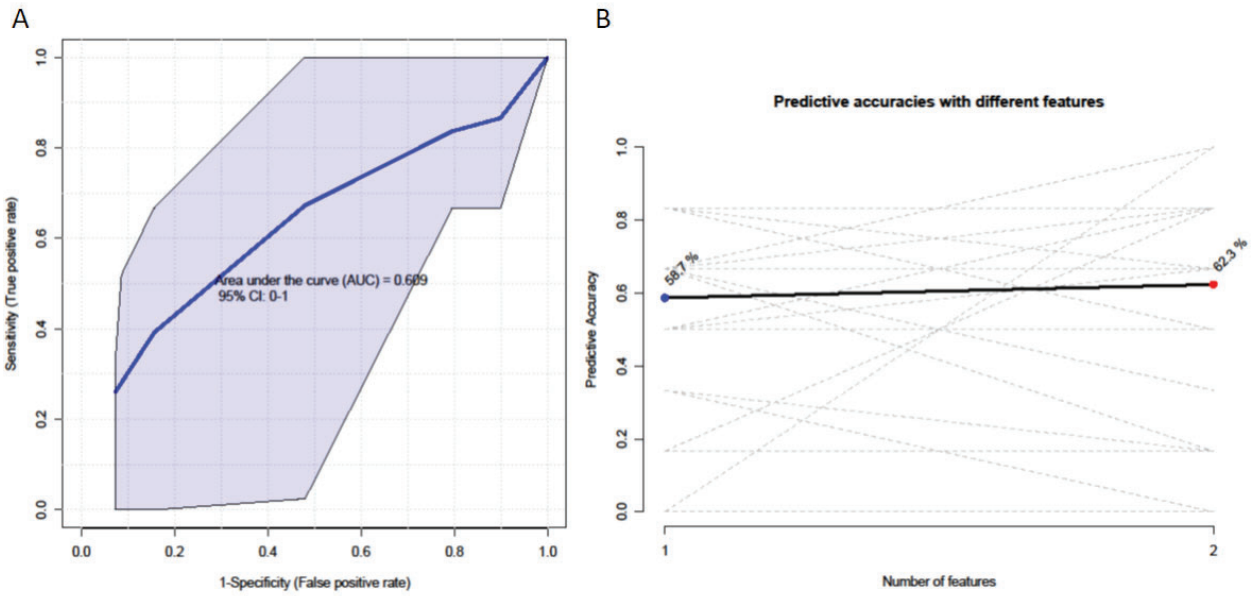

Figure 3. The Receiver Operator Characteristic (ROC) curve (A) and predictive accuracy (B) for the comparison of the fecal metabolomes of stressed and control pregnant female rats. The area under the curve (AUC) and confidence interval are displayed on the figure. The ROC and predictive accuracy was determined using the 2 bins belonging to the VIAVC best subset group.

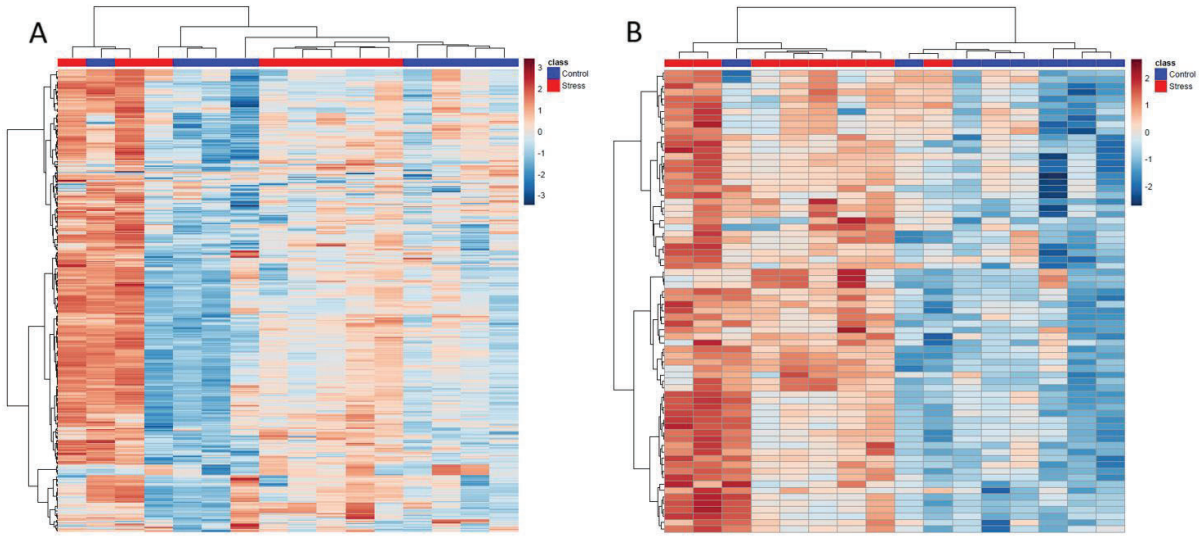

Figure 4. A heat map displaying the relative up-regulation or down-regulation of all metabolites (A), and those belonging to the best subset of bins identified as significant by either paired t-test/Mann-Whitney or VIAVC analysis (B). Metabolites are organized on the y-axis, while subjects are organized on the x-axis.

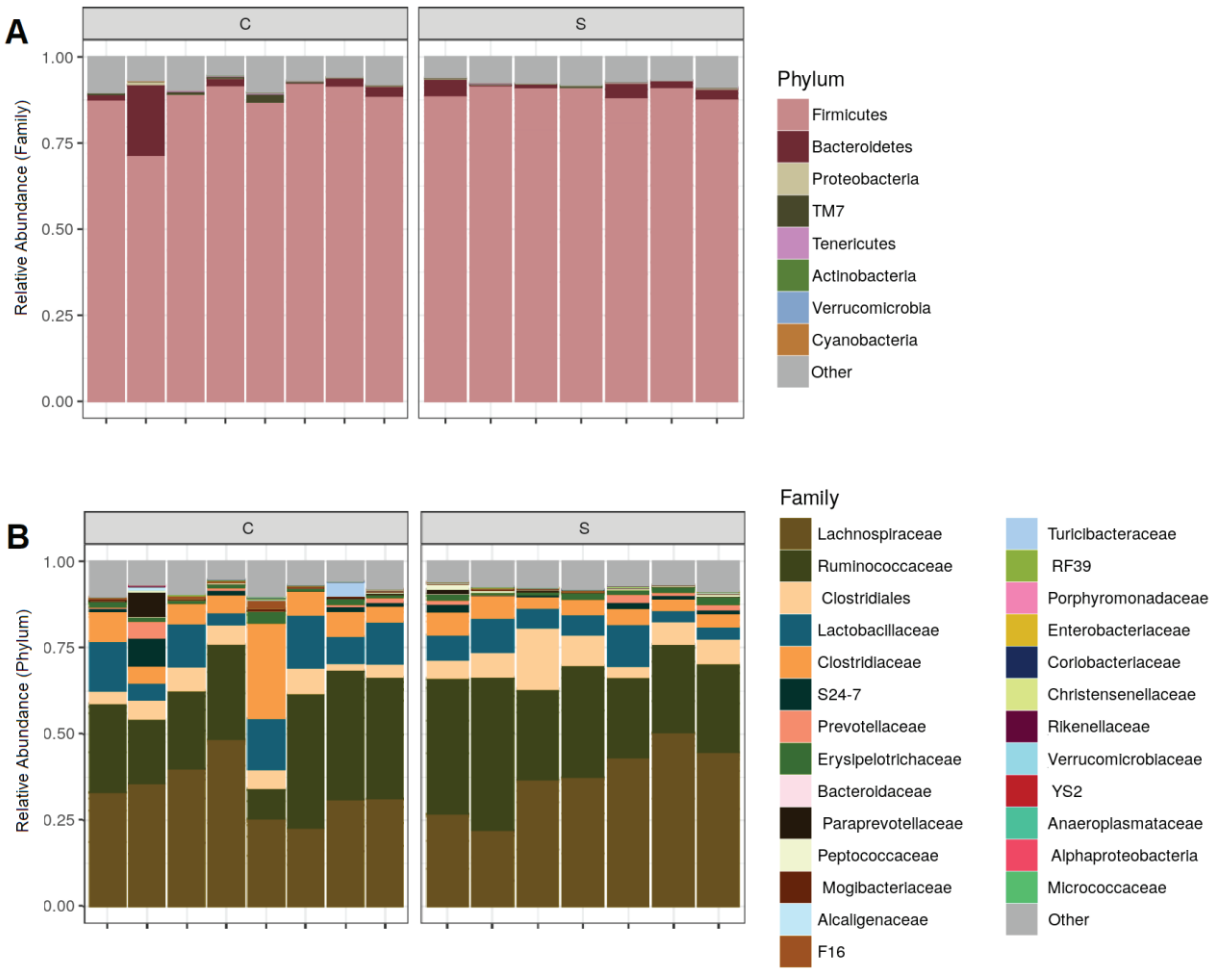

Figure 5. Taxonomic classification of bacterial phyla (A) and families (B) present in female rats in the control (C) and stress (S) groups.

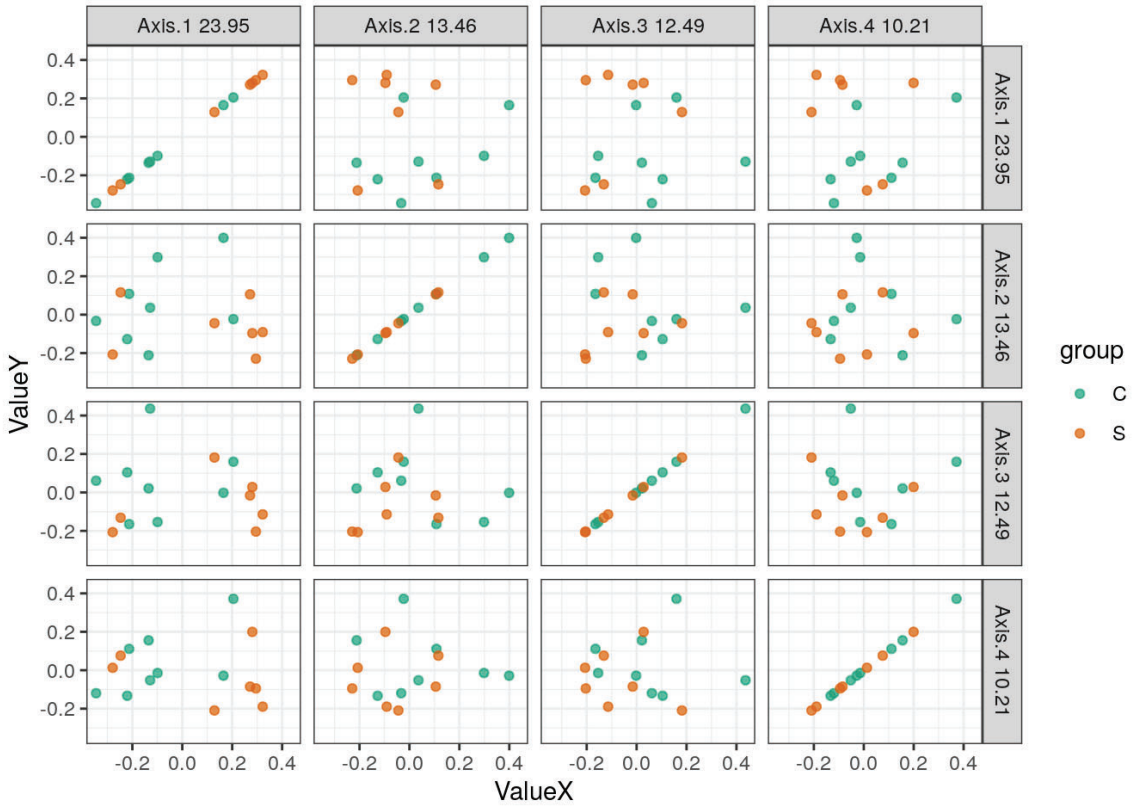

Figure 6. Principal Coordinate Analysis (PCoA) representing the Bray- Curtis dissimilarities of bacterial communities present in the fecal samples of prenatal stress and control rat groups. These values represent the beta diversity, or between sample diversity differences. Group differences account for 10.5% of the among-sample variation, less than would be expected by chance (r2 = 0.105, p = 0.078).

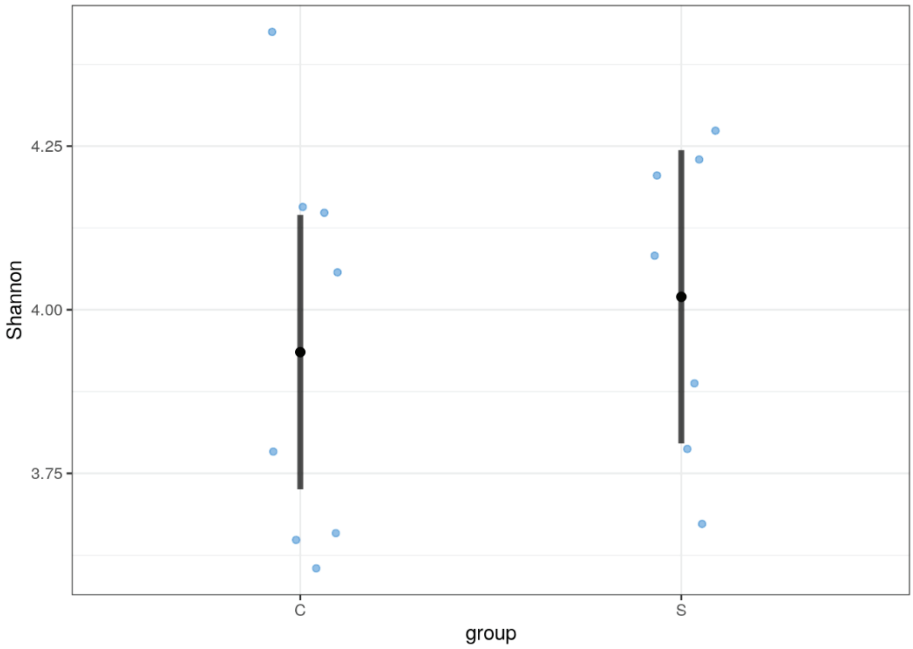

Figure 7. Shannon diversity indexes displaying the alpha diversity of bacterial communities present in the fecal samples of prenatal stress and control rat groups. Shannon diversity index values are displayed by blue dots, while the group's mean value is represented by a black dot. The 95% confidence interval is represented by the blue bar.

Table 1. Schedule for the administration of restrain stress and overnight social isolation stress to animals in the prenatal stress group between the 12th and 16th day of gestation.

Day of Gestation	Treatment (am)	Treatment (pm)
12	Restraint 60 minutes	N/A
13	N/A	Restraint 30 minutes
14	Restraint 45 minutes	Isolation Overnight
15	N/A	Restraint 60 minutes
16	Restraint 15 minutes	N/A
17	N/A	Restraint 30 minutes
		Isolation Overnight
18	Restraint 45 minutes	N/A

REFERENCES

Aagaard, K., Ma, J., Antony, K. M., Ganu, R., Petrosino, J., & Versalovic, J. (2014, May 21). The placenta harbors a unique microbiome. *Sci Transl Med, 6*(237), 237ra265. https://doi.org/10.1126/scitranslmed.3008599

Anderson, P. E., Mahle, D. A., Doom, T. E., Reo, N. V., DelRaso, N. J., & Raymer, M. L. (2011, Jun). Dynamic adaptive binning: an improved quantification technique for NMR spectroscopic data. *Metabolomics, 7*(2), 179-190. https://doi.org/10.1007/s11306-010-0242-7

Baibazarova, E., van de Beek, C., Cohen-Kettenis, P. T., Buitelaar, J., Shelton, K. H., & van Goozen, S. H. (2013, Jun). Influence of prenatal maternal stress, maternal plasma cortisol and cortisol in the amniotic fluid on birth outcomes and child temperament at 3 months. *Psychoneuroendocrinology, 38*(6), 907-915. https://doi.org/10.1016/j.psyneuen.2012.09.015

Bartram, A. K., Lynch, M. D., Stearns, J. C., Moreno-Hagelsieb, G., & Neufeld, J. D. (2011, Jun). Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end illumina reads. *Appl Environ Microbiol, 77*(11), 3846-3852. https://doi.org/10.1128/AEM.02772-10

Bharwani, A., Mian, M. F., Foster, J. A., Surette, M. G., Bienenstock, J., & Forsythe, P. (2016, Jan). Structural & functional consequences of chronic psychosocial stress on the microbiome & host. *Psychoneuroendocrinology, 63*, 217-227. https://doi.org/10.1016/j.psyneuen.2015.10.001

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J., & Holmes, S. P. (2016, Jul). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods, 13*(7), 581-583. https://doi.org/10.1038/nmeth.3869

Cannon, W. (1915). Bodily Changes in Pain, Hunger, Fear and Rage. An Account of Recent Researches into the Function of Emotional Excitement. *Journal of the American Medical Association, LXIV*(23), 1933-1933. https://doi.org/10.1001/jama.1915.02570490049033

Coussons-Read, M. E., Okun, M. L., Schmitt, M. P., & Giese, S. (2005, Jul-Aug). Prenatal stress alters cytokine levels in a manner that may endanger human pregnancy. *Psychosom Med, 67*(4), 625-631. https://doi.org/10.1097/01.psy.0000170331.74960.ad

Craig, A., Cloareo, O., Holmes, E., Nicholson, J. K., & Lindon, J. C. (2006, Apr 1). Scaling and normalization effects in NMR spectroscopic metabonomic data sets. *Analytical Chemistry, 78*(7), 2262-2267. https://doi.org/10.1021/ac0519312

Dhabhar, F. S., & McEwen, B. S. (1996, Apr 1). Stress-induced enhancement of antigen-specific cell-mediated immunity. *J Immunol, 156*(7), 2608-2615. https://www.ncbi.nlm.nih.gov/pubmed/8786326

Dickerson, S. S., & Kemeny, M. E. (2004, May). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychol Bull, 130*(3), 355-391. https://doi.org/10.1037/0033-2909.130.3.355

Foster, J. A., & McVey Neufeld, K. A. (2013, May). Gut-brain axis: how the microbiome influences anxiety and depression. *Trends Neurosci, 36*(5), 305-312. https://doi.org/10.1016/j.tins.2013.01.005

Gohir, W., Whelan, F. J., Surette, M. G., Moore, C., Schertzer, J. D., & Sloboda, D. M. (2015). Pregnancy-related changes in the maternal gut microbiota are dependent upon the mother's periconceptional diet. *Gut Microbes, 6*(5), 310-320. https://doi.org/10.1080/19490976.2015.1086056

Goodpaster, A. M., Romick-Rosendale, L. E., & Kennedy, M. A. (2010, Jun 1). Statistical significance analysis of nuclear magnetic resonance-based metabonomics data. *Analytical Biochemistry, 401*(1), 134-143. https://doi.org/10.1016/j.ab.2010.02.005

Goodrich, J. K., Waters, J. L., Poole, A. C., Sutter, J. L., Koren, O., Blekhman, R., Beaumont, M., Van Treuren, W., Knight, R., Bell, J. T., Spector, T. D., Clark, A. G., & Ley, R. E. (2014, Nov 6). Human genetics shape the gut microbiome. *Cell, 159*(4), 789-799. https://doi.org/10.1016/j.cell.2014.09.053

Gur, T. L., Shay, L., Palkar, A. V., Fisher, S., Varaljay, V. A., Dowd, S., & Bailey, M. T. (2017, Aug). Prenatal stress affects placental cytokines and neurotrophins, commensal microbes, and anxiety-like behavior in adult female offspring. *Brain Behav Immun, 64*, 50-58. https://doi.org/10.1016/j.bbi.2016.12.021

Hantsoo, L., Jasarevic, E., Criniti, S., McGeehan, B., Tanes, C., Sammel, M. D., Elovitz, M. A., Compher, C., Wu, G., & Epperson, C. N. (2019, Jan). Childhood adversity impact on gut microbiota and inflammatory response to stress during pregnancy. *Brain Behav Immun, 75*, 240-250. https://doi.org/10.1016/j.bbi.2018.11.005

Hu, J., Nomura, Y., Bashir, A., Fernandez-Hernandez, H., Itzkowitz, S., Pei, Z., Stone, J., Loudon, H., & Peter, I. (2013). Diversified microbiota of meconium is affected by maternal diabetes status. *PLoS One, 8*(11), e78257. https://doi.org/10.1371/journal.pone.0078257

Jost, T., Lacroix, C., Braegger, C., & Chassard, C. (2014, Apr). Stability of the maternal gut microbiota during late pregnancy and early lactation. *Curr Microbiol, 68*(4), 419-427. https://doi.org/10.1007/s00284-013-0491-6

Karu, N., Deng, L., Slae, M., Guo, A. C., Sajed, T., Huynh, H., Wine, E., & Wishart, D. S. (2018, Nov 7). A review on human fecal metabolomics: Methods, applications and the human fecal metabolome database. *Anal Chim Acta, 1030*, 1-24. https://doi.org/10.1016/j.aca.2018.05.031

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal, 17*(1), 10-12.

McEwen, B. S. (1998a, Jan 15). Protective and damaging effects of stress mediators. *N Engl J Med, 338*(3), 171-179. https://doi.org/10.1056/NEJM199801153380307

McEwen, B. S. (1998b, May 1). Stress, adaptation, and disease. Allostasis and allostatic load. *Ann N Y Acad Sci, 840*, 33-44. https://www.ncbi.nlm.nih.gov/pubmed/9629234

McEwen, B. S., & Wingfield, J. C. (2003, Jan). The concept of allostasis in biology and biomedicine. *Horm Behav, 43*(1), 2-15. https://www.ncbi.nlm.nih.gov/pubmed/12614627

McMurdie, P. J., & Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One, 8*(4), e61217. https://doi.org/10.1371/journal.pone.0061217

Metz, G. A., Jadavji, N. M., & Smith, L. K. (2005, Sep). Modulation of motor function by stress: a novel concept of the effects of stress and corticosterone on behavior. *European Journal of Neuroscience, 22*(5), 1190-1200. https://doi.org/10.1111/j.1460-9568.2005.04285.x

O'Mahony, S. M., Clarke, G., Borre, Y. E., Dinan, T. G., & Cryan, J. F. (2015, Jan 15). Serotonin, tryptophan metabolism and the brain-gut-microbiome axis. *Behav Brain Res, 277*, 32-48. https://doi.org/10.1016/j.bbr.2014.07.027

O'Mahony, S. M., Marchesi, J. R., Scully, P., Codling, C., Ceolho, A. M., Quigley, E. M., Cryan, J. F., & Dinan, T. G. (2009, Feb 1). Early life stress alters behavior, immunity, and microbiota in rats: implications for irritable bowel syndrome and psychiatric illnesses. *Biol Psychiatry, 65*(3), 263-267. https://doi.org/10.1016/j.biopsych.2008.06.026

O'Sullivan, V., Madrid-Gambin, F., Alegra, T., Gibbons, H., & Brennan, L. (2018, Dec 31). Impact of Sample Storage on the NMR Fecal Water Metabolome. *ACS Omega, 3*(12), 16585-16590. https://doi.org/10.1021/acsomega.8b01761

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P., O'hara, R., Simpson, G., Solymos, P., Stevens, M., & Wagner, H. (2013). Community ecology package. *R package version*, 2.0-2.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glockner, F. O. (2013, Jan). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research, 41*(Database issue), D590-596. https://doi.org/10.1093/nar/gks1219

Ramsteijn, A. S., Jasarevic, E., Houwing, D. J., Bale, T. L., & Olivier, J. D. (2020, Jul 3). Antidepressant treatment with fluoxetine during pregnancy and lactation modulates the gut microbiome and metabolome in a rat model relevant to depression. *Gut Microbes, 11*(4), 735-753. https://doi.org/10.1080/19490976.2019.1705728

Rhee, S. H., Pothoulakis, C., & Mayer, E. A. (2009, May). Principles and clinical implications of the brain-gut-enteric microbiota axis. *Nat Rev Gastroenterol Hepatol, 6*(5), 306-314. https://doi.org/10.1038/nrgastro.2009.35

Sears, C. L. (2005, Oct). A dynamic partnership: celebrating our gut flora. *Anaerobe, 11*(5), 247-251. https://doi.org/10.1016/j.anaerobe.2005.05.001

Smid, M. C., Ricks, N. M., Panzer, A., McCoy, A. N., Azcarate-Peril, M. A., Keku, T. O., & Boggess, K. A. (2018, Jan). Maternal gut microbiome biodiversity in pregnancy. *Am J Perinatol, 35*(1), 24-30. https://doi.org/10.1055/s-0037-1604412

Sudo, N., Chida, Y., Aiba, Y., Sonoda, J., Oyama, N., Yu, X. N., Kubo, C., & Koga, Y. (2004, Jul 1). Postnatal microbial colonization programs the hypothalamic-pituitary-adrenal system for stress response in mice. *J Physiol, 558*(Pt 1), 263-275. https://doi.org/10.1113/jphysiol.2004.063388

Whelan, F. J., Verschoor, C. P., Stearns, J. C., Rossi, L., Luinstra, K., Loeb, M., Smieja, M., Johnstone, J., Surette, M. G., & Bowdish, D. M. (2014, May). The loss of topography in the microbial communities of the upper respiratory tract in the elderly. *Ann Am Thorac Soc, 11*(4), 513-521. https://doi.org/10.1513/AnnalsATS.201310-351OC

Wickham, H. (2016). *ggplot2: elegant graphics for data analysis*. springer.

Willing, B. P., Vacharaksa, A., Croxen, M., Thanachayanont, T., & Finlay, B. B. (2011). Altering host resistance to infections through microbial transplantation. *PLoS One, 6*(10), e26988. https://doi.org/10.1371/journal.pone.0026988

Yao, Y. L., Robinson, A. M., Zucchi, F. C. R., Robbins, J. C., Babenko, O., Kovalchuk, O., Kovalchuk, I., Olson, D. M., & Metz, G. A. S. (2014, Aug 7). Ancestral exposure to stress epigenetically programs preterm birth risk and adverse maternal and newborn outcomes. *Bmc Medicine, 12*. https://doi.org/ARTN 121 10.1186/s12916-014-0121-6

Yun, Y. H., Liang, F., Deng, B. C., Lai, G. B., Goncalves, C. M. V., Lu, H. M., Yan, J., Huang, X., Yi, L. Z., & Liang, Y. Z. (2015, Dec). Informative metabolites identification by variable importance analysis based on random variable combination. *Metabolomics, 11*(6), 1539-1551. https://doi.org/10.1007/s11306-015-0803-x

Zierer, J., Jackson, M. A., Kastenmuller, G., Mangino, M., Long, T., Telenti, A., Mohney, R. P., Small, K. S., Bell, J. T., Steves, C. J., Valdes, A. M., Spector, T. D., & Menni, C. (2018, Jun). The fecal metabolome as a functional readout of the gut microbiome. *Nature Genetics, 50*(6), 790-+. https://doi.org/10.1038/s41588-018-0135-7

CHAPTER 4: Metabolic Dysfunction in Pregnancy: Fingerprinting the Maternal Metabolome Using Proton Nuclear Magnetic Resonance Spectroscopy

At the time of thesis submission, this manuscript has been submitted for publication at Endocrinology, Diabetes & Metabolism, ID EDM2-2020-0093 under a Creative Commons License (Scott et al., 2020). Author contributions can be found in the preface of this thesis.

INTRODUCTION

An estimated 20% of Canadians are currently diagnosed with metabolic syndrome, a cluster of features that puts an individual at higher risk of developing diabetes and cardiovascular disease (Riediger & Clara, 2011). During pregnancy, metabolic disorders such as obesity and gestational diabetes mellitus (GDM) have potentially adverse long-term consequences for both mother and child, such as increasing the risk of preeclampsia, preterm birth, caesarean section, and neonatal intensive care unit admissions (Chatzi et al., 2009; Horvath et al., 2013). Obesity can be a consequence of metabolic dysfunction, or a precursor to metabolic syndrome as it increases an individual's risk of developing other metabolic conditions, such as GDM or hypertension (Horvath et al., 2013). During pregnancy, the recommended weight gain for a woman of normal weight is between 25 to 35 pounds (11 to 16 kg); this amount decreases with higher pre-pregnancy Body Mass Index (BMI) levels (Rasmussen et al., 2009). Weight gain beyond these recommendations, or preexisting obesity, may lead to adverse pregnancy outcomes such as preeclampsia, miscarriage, congenital anomalies, preterm birth, and/or fetal complications (e.g. macrosomia) (Leddy et al., 2008). Furthermore, poor post-partum weight loss serves as a predictor of future obesity (Rooney & Schauberger, 2002).

GDM is characterized by glucose intolerance with onset or first recognition during pregnancy and it affects an estimated 3.7% of all pregnancies in Alberta (Nerenberg et al., 2012). Despite its prevalence, the initial pathogenic mechanisms of GDM is not fully understood. While the disease is strongly associated with the development of type 2 diabetes later in life, many individuals diagnosed with GDM have no prior known metabolic dysfunction (Damm et al., 2016). Inter- and transgenerational inheritance of GDM risk, however, has previously been suggested (Litwin et al., 2020; Yao et al., 2014). While both GDM and obesity are individually associated

with adverse pregnancy outcomes, in conjunction they have synergistic effects. Hence, the 2012 Hyperglycemia and Adverse Pregnancy Outcome (HAPO) study reported that mothers with both GDM and obesity had significantly increased birth weight, newborn body fat, cesarean delivery rates, and prevalence of preeclampsia when compared to individuals with only one, or neither, risk factor (Catalano et al., 2012).

Despite the increased risk of the aforementioned negative pregnancy and health outcomes associated with metabolic syndrome, obesity and GDM, not all individuals with either condition go on to develop further disorders. The underlying mechanisms that cause conditions to worsen in some individuals but not others are not fully understood, however, cross-sectional analysis of at-risk individuals may address this gap by identifying factors of risk that have not yet been considered (Butte et al., 2015; Cobb et al., 2013; McCormack et al., 2013). In many cases, such as prediabetes, mitigation of these factors can improve health outcomes, and even prevent the disease state (Cobb et al., 2013; Perreault et al., 2012). Metabolomics, the study of the metabolism and associated pathways, presents an effective tool for the discovery of relevant and quantifiable biomarkers that can be utilized to identify high risk individuals for targeted interventions and disease prevention (Ho et al., 2016; Wang et al., 2019). A recent study investigating the early pregnancy serum metabolomic profile of overweight and obese women identified several biomarkers that indicated an increased risk of developing GDM; these included small high density lipid (HDL) particles, branched chain amino acids (BCAAs), and inflammatory markers (Mokkala et al., 2020).

The present study aims to identify robust biomarkers associated with obesity and GDM status in the urinary metabolome. Urinary metabolomics enables the non-invasive detection of metabolic diseases during pregnancy without the need for a blood test. This study uses ^{1}H nuclear

magnetic resonance (NMR) spectroscopy to investigate metabolomic signatures linked to obesity and GDM in the urine of pregnant women collected in the second trimester from the Alberta Pregnancy Outcomes and Nutrition (APrON) study (Leung et al., 2016). We hypothesized that women with obesity or GDM will show characteristic urinary metabolomic signatures that differ from controls.

MATERIALS AND METHODS

Study design

Urine samples were collected from pregnant women in the Alberta Pregnancy Outcomes and Nutrition (APrON) study, a Canadian pregnancy cohort study (Kaplan et al., 2014). The APrON study was created to examine the links between perinatal nutrition intake and birth outcomes, child development, and maternal mental health(Leung et al., 2016). The full cohort consists of 2,140 women, 2,172 infants, and 1,417 biological fathers recruited in Calgary (population 1.1 million) and Edmonton (population 0.9 million), Alberta, Canada. The methodology details for the APrON study have been published elsewhere (Kaplan et al., 2014; Leung et al., 2016).

The present study used urine samples collected between 14 to 27 weeks of pregnancy from 29 women diagnosed with obesity, 37 who later developed GDM, and 36 healthy controls. The three groups were identified as followed: 1) pregnant women with GDM were extracted from hospital records whose diagnosis of GDM comes from routine prenatal screening for GDM according to the Canadian Diabetes Association criteria; 2) pregnant women (non-diabetic) with obesity were classified according to BMI status >30, or waist circumference >30); 3) controls were pregnant women without diabetes or obesity, matched on age, family income, education. Pregnant women were screened for GDM according to the Canadian Diabetes Association criteria of a 1-hour plasma glucose (1hPG) measurement following a 50-gram glucose load given at any time of the day. If the 1hPG was ≥10.3 mmol/L, or if it was >7.9 mmol/L and a subsequent 2 hour 75gm glucose load was positive, GDM was confirmed. Screening for GDM is part of routine prenatal care at ~24 to 28 weeks gestation. The GDM cases were identified from hospital records. Each case of obesity or GDM was counted only once. Women with obesity did

not go on to develop GDM in the cases we selected. Controls were matched by age-, income-, and education level. Table 1 provides the characteristics of the participants in this study.

Sample collection and preparation

Urine samples were obtained midstream during the first passage of the day following an overnight fasting period and were stored at -80°C. Samples were thawed at room temperature and 400 μL of urine were added to 200 μL of phosphate buffer containing a 4:1 ratio of dibasic potassium phosphate (K_2HPO_4) to monobasic potassium phosphate (KH_2PO_4) with a combined concentration of 0.5 M (pH 7.4) in 80% H_2O and 20% D_2O. The D_2O contained 0.03% (w/v) 3-(trimethylsilyl)propionic acid (TSP) to be used as an internal chemical shift reference and quantification standard. Three mM of sodium azide (NaN_3) were added as anti-microbial agent. The urine/buffer mixture was gently vortexed until homogenous, then centrifuged at 10,600 g for 5 min at 4°C. 550 μL of supernatant were pipetted into 5 mm NMR tubes before proceeding with ^{1}H-NMR spectroscopy.

NMR data acquisition and processing

NMR spectra of the urine samples were acquired at room temperature using a 700 MHz Bruker Avance III HD NMR Spectrometer (Bruker, Ontario, Canada) equipped with a 5 mm triple resonance TBO-Z probe. The one-dimensional NOESY gradient water suppression pulse sequence (noesygppr1d) was utilized with the following parameters: mixing time of 10 ms; 128 k data points (TD); sweet width (SW) of 20.52 ppm, acquisition time (AQ) of 4.56 sec, transmitter offset (o1p) of 4.6 ppm; recycle delay (D1) of 1 sec; 128 scans (NS). Spectra were then processed using zero-filling to 256 k points, line broadening with a 0.3 Hz exponential multiplication, automatic phased and baseline-correction, and chemical shift referenced with respect to the TSP peak at 0 ppm. All spectra were converted to ascii files and exported to

MATLAB (MathWorks, MA, USA) for further analysis. The spectra underwent dynamic adaptive binning (Anderson et al., 2011), followed by manual adjustment to correct for any errors in the algorithm. The bins containing the water and urea peaks were removed resulting in a total of 277 bins for all urine spectra. The spectra were normalized to the total area of all bins (the total metabolome), log transformed, and Pareto scaled (Craig et al., 2006).

Statistical analyses

For each comparison (obese vs control; GDM vs control; obese vs GDM) bins underwent both univariate and multivariate testing. On one hand, univariate testing provides a statistical measure by which each of the spectral bins can be tested to determine if they have been significantly altered across the comparison groups on an individual basis. On the other hand, multivariate testing offers a method by which each of the bins can be statistically assessed with respect to their importance to class separation when considered as part of a complete set of variables. Thus, univariate and multivariate testing provide different but complimentary information about the importance of a bin, or metabolite, to observed group differences.

Univariate testing

In order to ensure that the appropriate univariate test was applied to the bins, the decision tree algorithm outlined by Goodpaster et al. (Goodpaster et al., 2010) was utilized. This decision tree algorithm first uses a Shapiro-Wilk test to determine if the bins are normally distributed. If the bins are determined to have a non-normal distribution, they then undergo a Mann-Whitney U test (MW) for significance and are classified as significant or not-significant based on having a p-value of less than or greater than 0.05, respectively. In the case of this study, the data for all comparisons was determined to have non-normal distributions and the Mann-Whitney U test was performed.

Multivariate testing

Variable Importance Analysis based on random Variable Combination (VIAVC) (Yun et al., 2015) was utilized to assess variable significance when considered as part of the total set of variables. The VIAVC algorithm utilizes binary matrix resampling to create several random subsets of variables containing 50% of the original bins. In the case of this study, 1000 subsets of random variables were used. Binary matrix resampling ensures that each variable has been selected with the same probability when creating the random subsets. The importance of each bin is then assessed by calculating the corresponding area under the curve (AUC) for the receiver operator characteristic (ROC) curve with and without a variable included in the model. If the inclusion of the variable increases or decreases the area under the curve it is either retained or removed from the overall dataset, respectively. This process is repeated until no more variables can be removed and the resulting list of bins is known as the best subset. By doing this, the VIAVC algorithm can determine if any synergistic effects exist between variables that appear non-significant based on traditional univariate and multivariate measures. In addition, it can remove bins that are significant based on a univariate test but reduce group separation when they are included in the whole set of variables. Ultimately, this machine learning approach is an ideal method for determining the subset of variables which lead to the best separation between two classes or groups.

Metaboanalyst (Chong et al., 2018; Chong et al., 2019) was utilized to carry out and visualize both the Orthogonal Partial Least Squares Discriminant Analysis (OPLS-DA) and Receiver Operating Characteristic (ROC) curves. The former was carried out using bins identified as significant by the MW or VIAVC tests. The latter was performed using only the

bins identified as significant by the VIAVC test. All modeling underwent permutation testing (2,000 permutations) and double ten-fold cross validation(Szymanska et al., 2012).

Metabolite identification and pathway analysis

Both Chenomx (Chenomx, Alberta, Canada) and the Human Metabolome Database (HMDB) (Wishart et al., 2018; Wishart et al., 2013; Wishart et al., 2009; Wishart et al., 2007) were used to identify the metabolites present in each of the significant bins. The complete list of metabolites identified as significantly altered by either the MW test or VIAVC algorithm were used for pathway topology analysis (Xia & Wishart, 2010). Pathway topology analysis was carried out in Metaboanalyst by selecting the hypergeometric test for the over-representation analysis, relative-betweenness centrality for the topology analysis, and using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database for Home sapiens (Kanehisa, 2019; Kanehisa & Goto, 2000; Kanehisa et al., 2019) as the pathway library.

RESULTS

Of the 277 total spectral bins, both VIAVC and MW tests identified which bins led to significant group separation. Analyses were applied to three comparison groups and resulted in the following number of significant bins: obesity vs. control (33 MW, 31 VIAVC, and 9 common bins); GDM vs. control (61 MW, 30 VIAVC, and 12 common bins); and obesity vs. GDM (30 MW, 19 VIAVC, and 8 common bin). The supervised OPLS-DA comparisons of the obese and control (Fig. 1A), GDM and control (Fig. 1B), and GDM and obesity (Fig. 1C) all showed significant group separation with the largest separation between the obese and control groups. Permutation and cross-validation tests confirmed this observed separation for the three comparisons ($p<0.05$). Receiver operator characteristics (ROC) curves were used to determine the specificity, sensitivity, and accuracy of each comparison model based on metabolites identified as significant by VIAVC. The comparisons between obesity and control, GDM and control, and GDM and obesity gave an area under the curve (AUC) and predictive accuracy (in brackets) of 0.906 (80.8%), 0.803 (74.5%), and 0.846 (75.9%), respectively (Fig. 2).

Supplementary Tables 1, 2 and 3 provide the metabolites that were found to be significantly altered by either Mann-Whitney U or VIAVC. The last column on each table indicates whether the metabolite was up- or downregulated in the first group of the comparison compared to the second. A total of 58 metabolite bins significantly contributed to the distinction between the obesity vs control group, 79 in the GDM vs control comparison, and 41 to the obesity vs GDM comparison. A total of 50 out of 58 and 65 out of 79 significantly altered bins were downregulated compared to the controls in both the obesity and GDM groups, respectively. No such trend was observed in the obesity vs GDM comparison.

The most impacted three metabolites for the comparison of obesity and control groups were pantothenate, formic acid, and glycine with all three metabolites downregulated in the obesity group. The top three metabolites for the comparison of GMD and control groups were formic acid, dimethylamine, and galactose with all three metabolites downregulated in the GDM group. The top three metabolites for the comparison of GDM and obesity groups were creatine/caffeine, sarcosine/dimethylamine, and maltose/sucrose.

For the obese and control comparison, the metabolomic pathway analysis identified 10 significantly altered pathways (Fig. 3A), with glycine, serine and threonine metabolism ($p<0.05$), phenylalanine metabolism ($p<0.05$), and methane metabolism ($p<0.05$) being the most significantly altered. In the GDM and control comparison, 13 significantly altered pathways were identified (Fig. 3B), including glycine, serine and threonine metabolism ($p<0.05$), galactose metabolism ($p<0.05$), and phenylalanine metabolism ($p<0.05$). In the obesity and GDM comparison, 16 metabolic pathways were significantly altered (Fig. 3C), with the greatest impact on glycine, serine and threonine metabolism ($p<0.05$), starch and sucrose metabolism ($p<0.05$), and the citrate cycle ($p<0.05$).

DISCUSSION

This study demonstrated that ^{1}H-NMR metabolomics can be used to distinguish and compare the urinary metabolomes of healthy pregnant woman to those diagnosed with obesity and may be predictive of women who develop GDM. Metabolites or impacted pathways found to be distinct between groups may serve as indicators of risk, allowing for the use of early clinical or lifestyle intervention prior to manifestation of a disease state (Kohler et al., 2017). Characterization of the metabolic processes altered by obesity and GDM makes an important contribution to understanding their pathogenic mechanisms (Damm et al., 2016). The present results provide insight into the comorbidity existing between obesity and GDM; the similarity of the metabolomes between these disease states indicate shared biochemical pathway disruptions which may be detectable in individuals prior to the development of related negative health outcomes.

The high predictive accuracy of the three ROC curves resulting from this study (80.8% for obesity vs control, 74.5% for GDM vs control, and 75.9% for obesity vs GDM) demonstrate that the best subset of metabolites as determined by the VIAVC analysis can be used to distinguish between the groups. While the R^2X values for the GDM vs control comparisons and obesity vs control indicated a large degree of variability (>0.7), the obesity vs GDM comparison was 0.465, demonstrating only a moderate difference between the groups. This indicates the presence of similar underlying mechanisms and metabolic disturbances shared between obesity and GDM. No single metabolite or pathway was responsible for the majority of this separation, which was not surprising considering the multifaceted pathogenic mechanisms underlying obesity and GDM. While the best subset of metabolites provides the most accurate diagnostic tool, examining the metabolites and pathways individually provides insight into the underlying mechanisms of obesity and GDM, and can help provide a better understanding of the risk factors behind each condition.

While multiple metabolic pathways were impacted by the presence of obesity and GDM, altered glycine, serine and threonine metabolism pathway function was among the most distinct signatures for both conditions. In this pathway, serine is derived from glycolysis, and in turn is converted into glycine. Threonine is an essential amino acid derived from diet, which is also converted into glycine (Razak et al., 2017). Glycine deficiency in particular has been found to be associated with increased abdominal adipose tissue (Lustgarten et al., 2013), potentially contributing to the downregulation of the metabolite observed in the obese group.

The phenylalanine metabolism was found in the top three significantly impacted pathways when comparing both obesity and GDM to control groups. Phenylalanine is an aromatic amino acid that acts as a precursor to tyrosine, along with multiple catecholamines including epinephrine, norepinephrine and dopamine. Metabolic disorders such as obesity and GDM have been found to lead to elevations in phenylalanine and several of its metabolic products (Libert et al., 2018). Elevated levels of aromatic amino acids have been associated with obesity (Butte et al., 2015), and the development of insulin resistance in non-diabetic individuals (Yamada et al., 2015). BCAAs, particularly valine, leucine and isoleucine, are often presented as indicators of risk for insulin resistance alongside aromatic amino acids (Tai et al., 2010; Wang et al., 2011), and have been implicated in the development of GDM in overweight and obese pregnant women (Mokkala et al., 2020). In this study's comparison between individuals with obesity and GDM, the BCAA leucine was significantly upregulated in the GDM group alongside phenylalanine (Table 4), and valine was found to be higher in individuals with GDM versus their control counterparts (Table 3), supporting the role of BCAAs in the development of insulin resistance. Thus, phenylalanine, leucine, and valine provide valuable urinary biomarkers of GDM risk in obese individuals.

It should be noted that caution must be taken when comparing studies using different biofluids, as metabolite expression can be dependent on a variety of cohort factors such as circadian rhythms, diet, and physical activity. In the present study, the downregulation of galactose in individuals with GDM seemingly conflicts with previous findings that serum d-galactose is upregulated in response to the disease (Enquobahrie et al., 2015). In addition, pantothenate, the most impacted individual metabolite in the obesity vs control comparison, represents another case of discrepancy between serum and urinary expression. While the present downregulation of urinary pantothenate may serve as a urinary indicator of obesity, serum pantothenate has been found to be upregulated in response to obesity (Patel et al., 2017). As mentioned above, these discrepancies may be due to differences between the study cohorts; however, it cannot be ruled out that excreted galactose and pantothenate in the urine may reflect serum levels. For example, the metabolite being upregulated in blood could serve as an indicator for an increased demand for the metabolite, which would result in the observed decrease of the metabolite in the urine. As mentioned earlier, urinary metabolomics enables the non-invasive detection of metabolites during pregnancy without the need for a blood test. In addition, NMR based urinary metabolomics provides that added benefit of containing much more metabolite information when compared to serum metabolomics, with the former and later providing information on 209 versus 49 metabolites, respectively (Bouatra et al., 2013; Psychogios et al., 2011)

The regulation of sucrose expression appears to also vary between the serum and urinary metabolomes. Increased intake of sucrose has been linked to insulin resistance in mice (Cao et al., 2007). However, both the obese and GDM groups of this study experienced reduced urinary sucrose levels when compared to control individuals. This may be partially explained by sex differences; sucrose-induced insulin resistant models have only been successfully created in male

animals, and females appear to be resistant to sucrose induced insulin resistance (Horton et al., 1997). When comparing individuals with obesity to those with GDM, sucrose was the third most impacted metabolite, being increased in GDM individuals. This finding indicates that starch and sucrose metabolism does indeed affect females and potentially contributes to the development of GDM, which highlights the need for a better understanding of sex differences in diabetes-related health outcomes, such as coronary heart disease (Kanaya et al., 2002).

The most impacted pathway distinction between obesity and GDM concerned starch and sucrose metabolism. High starch diets appear to have the opposite effect of sucrose on insulin resistance, reducing its severity and decreasing adipose tissue weight (Harazaki et al., 2014). Variations in starch and sucrose metabolism may reflect dietary variations in the subjects, which in turn may correlate with the presence of obesity and GDM states (Beckmann et al., 2016). It should also be noted that, while this study controlled for age, income, and education level, subjects underwent no dietary restrictions prior to sampling, so impacts of diet on the metabolome act as a potential confound.

The results of this study allow for further analysis of the comorbidity of obesity and GDM with their associated health risks. Preeclampsia is one of the most severe pregnancy complications associated with GDM. The high blood pressure of this complication results from the improper formation of blood vessels in the placenta, and can cause organ damage, and even maternal and fetal death if not addressed. The presence of GDM more than doubles the risk of a mother developing preeclampsia (Ostlund et al., 2004). Several of the most impacted urinary metabolites in the GDM vs control (formic acid, ethanol, propylene glycol) and obesity vs control (formic acid, glycine) comparisons of this study are known to be associated with alterations to the serum

metabolome found in early preeclampsia (Bahado-Singh et al., 2012), and may be indicative of shared pathway disruptions between the conditions.

Spontaneous preterm birth, when the infant is delivered prior to 37 weeks of gestation, is another pregnancy complication that is associated with metabolic syndrome, obesity, and GDM obesity (Chatzi et al., 2009; Cnattingius et al., 2013; Hedderson et al., 2003). Previous urinary metabolic analysis of preterm birth revealed an association between spontaneous preterm birth and decreased levels of formic acid (Maitre et al., 2014). This trend was also observed in the obese and GDM groups of this study when compared to control individuals. The mechanisms by which this trend is associated with preterm birth are not yet known, but it was suggested that diminished urinary formic acid raises the risk of hypertension (Holmes et al., 2008), which in turn is positively associated with preterm birth (Chatzi et al., 2009).

CONCLUSION

This exploratory study contributes to the understanding of the metabolic profiles of both obesity and GDM, and supports the role of ^{1}H NMR spectroscopy metabolomics in the developing field of precision medicine. The individual metabolites that contributed most to separation between disease groups have the highest potential to provide simple diagnostic tests. The creation of new tests for precision medicine approaches for intervention necessitates a deeper understanding of the underlying metabolic mechanisms behind obesity, GDM and pregnancy. The identified pathways provide insight into the underlying mechanisms of obesity and GDM and potential therapeutic targets.

While the present findings suggest a composite metabolic profile to be the most robust predictor, the findings also indicate that the glycine, serine and threonine metabolism pathway may provide the most feasible potential to provide a single component biomarker to distinguish obesity and GDM from healthy individuals. Moreover, the starch and sucrose metabolism pathway provides the most distinction between the urinary metabolome of individuals with GDM and those who are obese. The upregulation of BCAA and aromatic amino acids also provides an effective biomarker for the development of GDM in obese individuals and suggests possible causal mechanisms for insulin resistance. This study also supports further investigation of the urinary metabolome as a non-invasive diagnostic tool, as it delivers powerful results without the need for a serum sample. Using biomarkers to determine and potentially prevent metabolic disease ultimately may not only improve quality of life for mothers and their children, but also assist in the transition to preventative and precision medicine.

FIGURES AND CAPTIONS

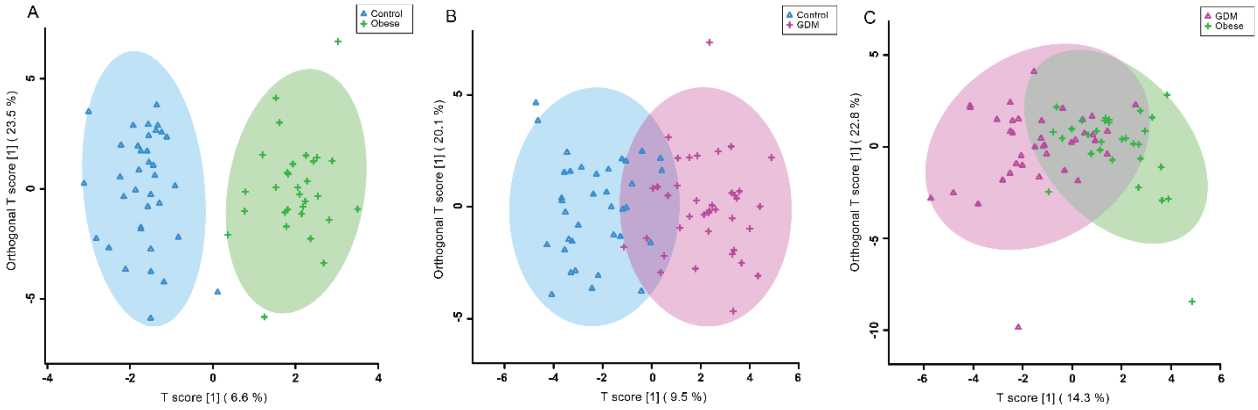

Figure 1. Orthogonal partial least squares discriminant analysis (OPLS-DA) score plots showing supervised group seperation between (A) obesity and controls ($R^2X = 0.89$, $Q^2 = 0.574$, R2Y p-value = 0.0005, Q2 p-value = 0.0005), (B) GDM and controls ($R^2X = 0.722$, $Q^2 = 0.277$, R2Y p-value = 0.0005, Q2 p-value = 0.0005), and (C) obesity and GDM ($R^2X = 0.465$, $Q^2 = 0.258$, R2Y p-value = 0.0005, Q2 p-value = 0.0005). Each triangle or cross represents one individual under study, plotted using a list of urinary metabolites found to be statistically significant by either MW or VIAVC testing. The x- and y-axis show the predictive (across group variation) and orthogonal (within group) components, respectively.

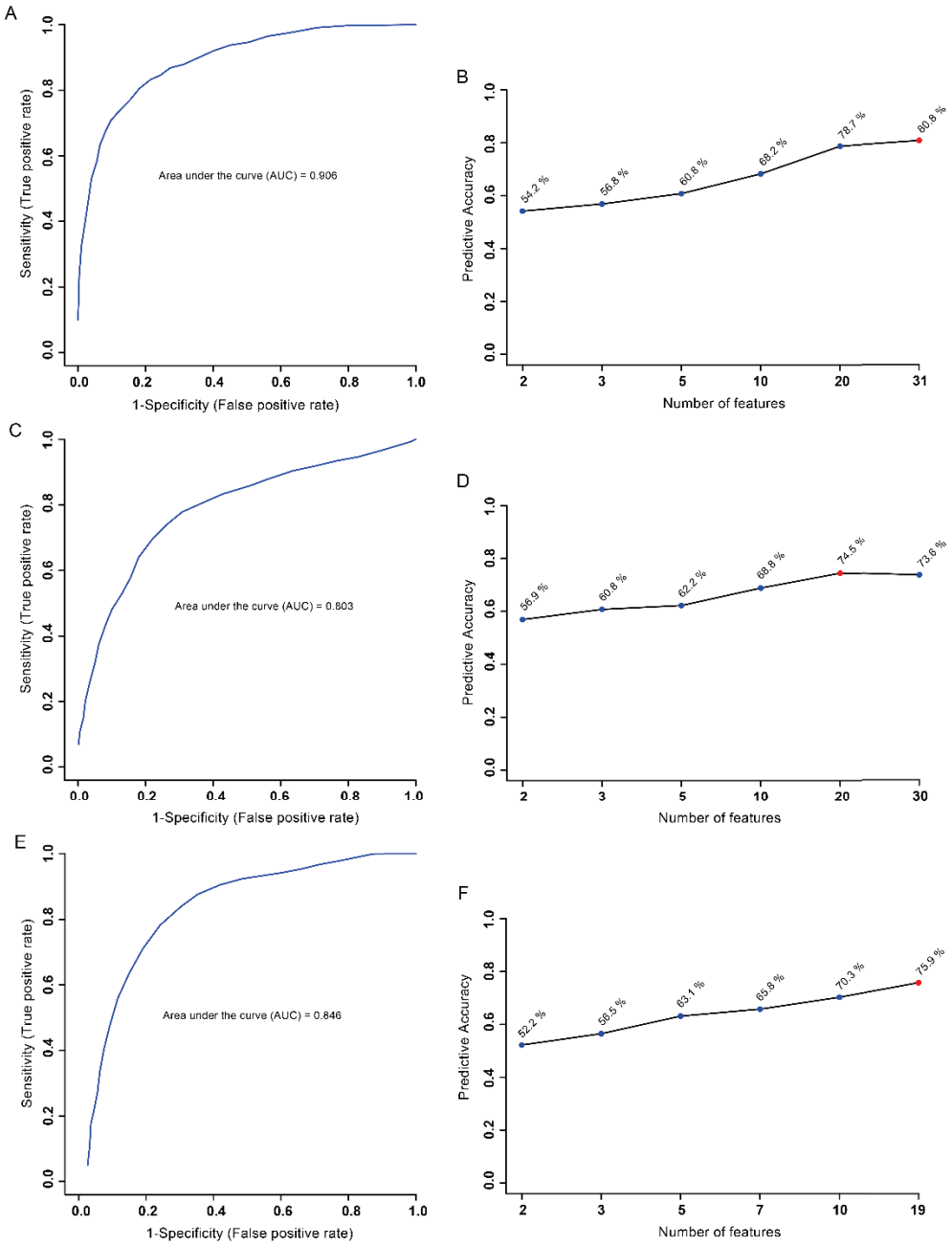

Figure 2. Receiver operating characteristic (ROC) curve (A, C, E) and predictive accuracy (B, D, F) for the comparison of (A, B) obesity and controls, (C, D) GDM and controls, and (E, F) obesity and GDM. These figures were created using the best subset of metabolites as determined by the VIAVC analysis. The 95% confidence intervals for A, C, and E are 0.772-0.986, 0.655-0.942, and 0.603-0.962, respectively.

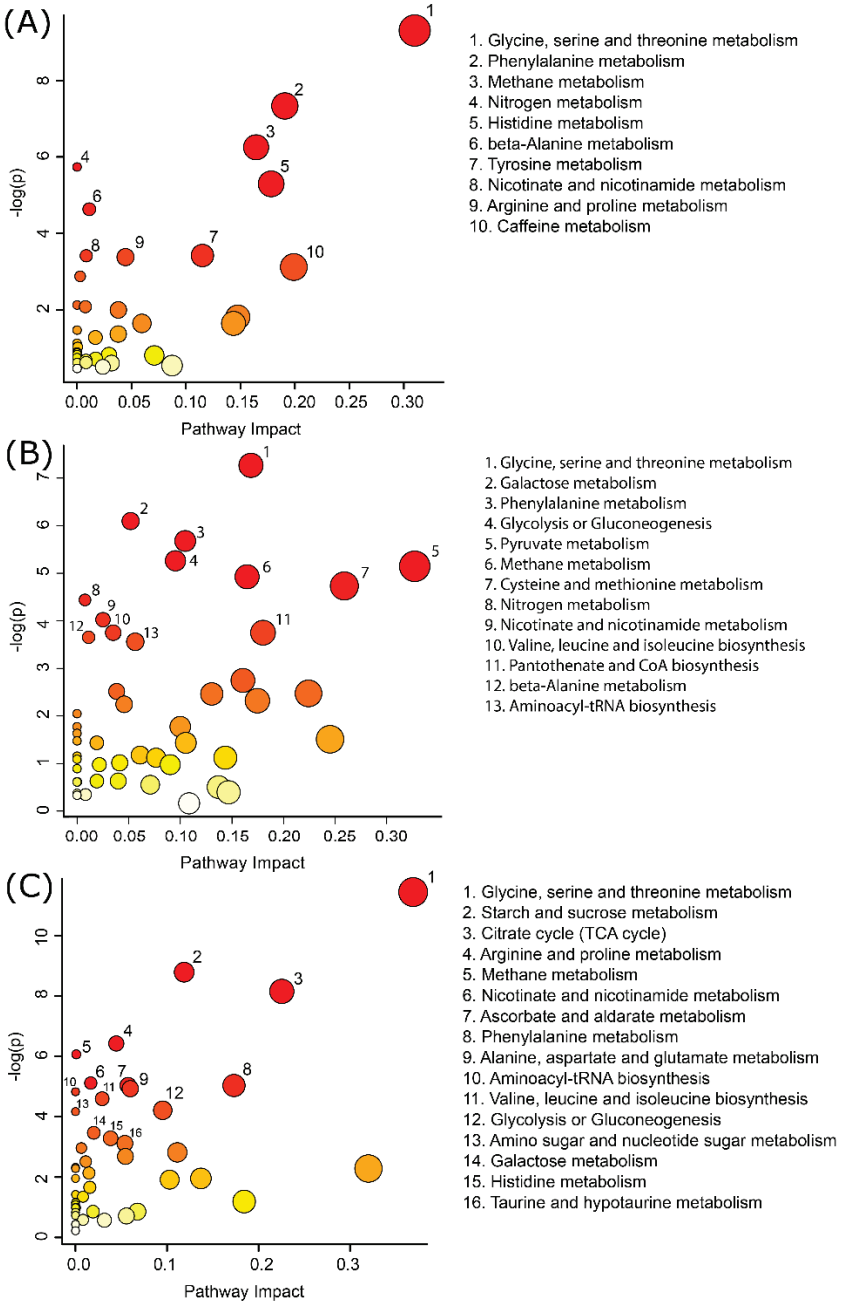

Figure 3. Metabolomic Pathway Analysis for the comparison of the obesity vs control groups (A), GDM vs control groups (B), and obesity vs GDM groups (C). A higher value on the y-axis indicates a lower p-value for the pathway. The x-axis gives the pathway impact, which indicates how affected the pathway is by the metabolites identified as significantly altered. Only metabolic pathways with $p < 0.05$ are labeled. This figure was created using the lists of metabolites identified as significantly altered between the obese and control groups by either MW or VIAVC testing.

TABLES AND CAPTIONS

Table 1

Characteristics of the study participants

Participant Information		Control (n=36)	GDM (n=37)	Obesity (n=29)
Age	Mean (Std Dev)	33.4 (4.5)	33.8 (3.9)	32.8 (4.2)
Pre-Pregnancy BMI	Mean (Std Dev)	23.1 (2.2)	27.7 (6.5)	33.2 (4.0)
Income	$20,000-$39,999	2	3	3
	$40,000-$69,999	9	7	4
	$70,000-$99,999	11	11	11
	$100,000 or More	14	16	11
Education	Completed High School Diploma	3	5	3
	Completed Trade, Technical	5	6	4
	Completed University	16	15	13
	Completed Post-Grad	12	11	9
Marital Status	Single	0	0	2
	Married	35	36	26
	Common-Law	1	0	1
	Unknown		1	
Ethnicity	Caucasian	28	28	26
	Other	8	9	3

Table 2

P-values of urinary metabolites found to be significant in separation of obese and control groups in either a Mann-Whitney U test, the Variable Importance Analysis based on random Variable Combination (VIAVC), or both. Regulation is shown for the metabolite levels in the obese group. Metabolites for which more than one NMR resonance peak was identified as significant are represented as metabolite.1, metabolite.2, … metabolite.n.

Metabolite	Mann Whitney U Test	VIAVC p-value	Regulation
Pantothenate.1	1.02E-02	3.51E-149	Down
Formic acid	2.64E-03	6.48E-116	Down
Glycine	-	1.64E-110	Down
Caffeine.2, Methanol	6.71E-03	1.34E-106	Down
3,4-Dihydroxybenzeneacetate	-	7.39E-103	Up
Sarcosine, Dimethylamine	-	1.49E-67	Down
3-Hydroxyphenylacetic acid	-	5.08E-64	Down
Anserine.2	1.52E-02	7.72E-63	Down
N,N-Dimethylformamide.2, 3-Phenylpropionate	-	3.63E-56	Up
Unidentified metabolite (Doublet at 6.488ppm)	3.12E-03	2.96E-46	Down
Glucuronic acid	-	8.00E-46	Down
m-Thymol	6.21E-03	1.78E-41	Down
2-Octenoate.2, Caprate, 2-Hydroxyisovalerate	-	1.92E-37	Down
Fucose.2	-	7.35E-33	Up
Fucose.1	-	5.05E-25	Up
N-Acetylaspartate.3	-	1.32E-23	Down
Homoserine	-	8.99E-21	Down
Methylsuccinic acid, Isobutyric acid.2	-	8.91E-17	Up
Fumaric acid	1.69E-02	4.23E-15	Down
5-Aminolevulinate	-	1.90E-14	Up
N,N-Dimethylformamide.1	-	6.39E-11	Down
Sucrose.2	2.46E-02	4.37E-09	Down
N-Acetylglutamine	-	2.97E-08	Down
3-Hydroxymandelic acid	-	1.53E-07	Down
2-Methylglutaric acid, 3-Hydroxyisobutyric acid, Isobutyric acid.1, 3-Methyl-2-oxovaleric acid	-	2.32E-07	Down
Caffeine.1, 3-Methylxanthine	-	6.63E-06	Down
Chlorogenate	-	9.35E-06	Up
Malonate	-	1.09E-04	Down
Cholate	-	2.87E-04	Down

Unidentified metabolite (Peak at 6.502ppm)	2.75E-03	6.51E-03	Down
Creatinine	-	1.34E-02	Up
1,3-Dihydroxyacetone	3.75E-02	-	Down
2-Hydroxybutyric acid, Histidine, Phenylalanine.1	1.81E-02	-	Down
2-Octenoate.1	3.10E-02	-	Down
ADP, Anserine.1	3.20E-02	-	Down
Ascorbic acid	1.14E-02	-	Down
Cinnamic acid.1	8.14E-03	-	Down
Cinnamic acid.2	4.00E-02	-	Down
Galactose	9.82E-03	-	Down
Hippuric acid, 2-Furoylglycine, Creatine	4.00E-02	-	Down
Histamine, 1-Methylhistidine, Caffeine.3	3.39E-03	-	Down
Homogentisic acid.1	6.98E-03	-	Down
Homogentisic acid.2	4.25E-02	-	Down
Hydroxyacetone	2.72E-02	-	Down
Maleic acid	4.25E-02	-	Down
N-Acetylaspartate.1	1.69E-02	-	Down
N-Acetylaspartate.2	3.10E-02	-	Down
Pantothenate.2	4.95E-02	-	Down
Phenylalanine.2	1.52E-02	-	Down
Sucrose.1	3.87E-02	-	Down
Tartrate	4.25E-02	-	Down
Trigonelline.1	2.91E-02	-	Down
Trigonelline.2	3.64E-02	-	Down
Trigonelline.3	3.64E-02	-	Down
Trigonelline.4	3.87E-02	-	Down

Table 3

P-values of urinary metabolites found to be significant in separation of GDM and control groups in either a Mann-Whitney U test, the Variable Importance Analysis based on random Variable Combination (VIAVC), or both. Regulation is shown for the metabolite levels in the GDM group. Metabolites for which more than one NMR resonance peak was identified as significant are represented as metabolite.1, metabolite.2, … metabolite.n.

Metabolite	Mann Whitney U Test	VIAVC p-value	Regulation
Formic acid	8.82E-04	1.21E-173	Down
Dimethylamine	2.21E-04	2.79E-111	Down
Galactose	9.01E-03	6.23E-84	Down
Anserine.3	2.62E-04	3.15E-75	Down
2,3,4-Trihydroxybenzoic acid.2	2.69E-03	3.37E-65	Down
N-Methylhydantoin, 5-Aminolevulinic acid.2		1.06E-63	Up
1,3-Dimethyluric acid	1.02E-02	4.72E-58	Down
3-Hydroxymandelic acid	5.01E-03	1.39E-52	Down
Succinylacetone	-	3.46E-51	Up
Propylene glycol	-	8.04E-48	Up
5-Aminolevulinic acid.1	-	6.71E-40	Up
Caffeine.1	-	1.50E-36	Down
3-Phenylpropionic acid.2		1.55E-34	Down
2-Phenylpropionic acid	-	1.07E-28	Down
Trigonelline.1	-	3.22E-27	Down
Lactose.1	-	6.25E-26	Up
Theophylline	-	7.06E-23	Down
Trigonelline.2	-	8.72E-23	Down
Desaminotyrosine, N-Acetyltyrosine, 4-Hydroxyphenylacetic acid	7.68E-03	2.14E-18	Down
Homogentisic acid.2	6.97E-03	6.03E-18	Down
Methanol	-	2.07E-17	Down
Valine	-	9.05E-14	Up
Pyruvic acid	-	1.00E-13	Down
Pantothenic acid	4.15E-02	3.87E-12	Down
Lysine	-	8.17E-12	Down
Asparagine	-	1.27E-11	Up
Mannose	5.72E-03	2.61E-09	Down
3,4-Dihydroxybenzeneacetic acid	-	1.86E-08	Up
S-Adenosylhomocysteine, 2'-Deoxyadenosine	5.01E-03	4.51E-08	Down
Ascorbic acid	-	9.17E-07	Up
ADP, Anserine.1	2.63E-02		Down

Anserine.2	4.48E-02		Down
Caffeine.2, 1-Methylhistidine, Histamine	1.02E-02		Down
Indoxyl sulfate.2, Tryptophan	4.53E-03		Down
Indoxyl sulfate.1	4.15E-02		Down
Protocatechuic acid	9.59E-03		Down
Thymol	5.07E-04		Down
Homogentisic acid.1	4.48E-02		Down
2-Octenoic acid.1	3.18E-02		Down
2-Octenoic acid.2	1.05E-02		Down
2-Octenoic acid.3	2.55E-02		Down
2,3,4-Trihydroxybenzoic acid.1	3.36E-02		Down
Maleic acid, Allantoin	5.01E-03		Down
Levoglucosan	4.04E-02		Down
Glucose	4.72E-02		Up
Unidentified metabolite	2.55E-02		Down
1-Methylnicotinamide.1	2.17E-03		Down
Lactose.2	3.18E-02		Down
Lactose.3	1.88E-03		Down
Unidentified metabolite	1.15E-03		Down
Hydroxyacetone	4.48E-02		Down
Unidentified metabolite	4.09E-03		Down
1-Methylnicotinamide.2	1.71E-02		Down
Tartrate	3.36E-02		Down
Adenosine	4.84E-02		Down
Sucrose.1	1.66E-02		Down
Sucrose.2	7.93E-03		Down
o-Hydroxyphenylacetic acid, Caffeine.3	3.18E-02		Up
Myoinositol, Phenylacetic acid	4.84E-02		Up
Malonic acid	4.48E-02		Down
3-Phenylpropionic acid.1	3.69E-03		Down
Methylguanidine	4.15E-02		Down
Levulinic acid	1.23E-02		Down
Homocysteine	9.30E-03		Down
3-Aminoisobutanoic acid.1	4.53E-03		Down
3-Aminoisobutanoic acid.2	3.66E-04		Down
3-Aminoisobutanoic acid.4, Methylamine	4.26E-02		Down
3-Aminoisobutanoic acid.3	2.97E-04		Down
Glutathione	1.39E-02		Down
Homoserine	4.48E-02		Down
Threonine, Lactic acid	1.71E-02		Up
Unidentified metabolite	1.81E-02		Down

Fucose	4.60E-02		Down
Ethanol.1	4.97E-02		Down
Ethanol.2	4.84E-02		Down
Methylsuccinic acid, 3-Hydroxyisobutyric acid	4.15E-02		Up
2-Octenoic acid.4, Capric acid, 2-Hydroxyisovaleric acid	1.39E-02		Down
Unidentified metabolite	2.99E-03		Down
Unidentified metabolite	2.55E-02		Down

Table 4

P-values of urinary metabolites found to be significant in separation of GDM and obese groups in either a Mann-Whitney U test, the Variable Importance Analysis based on random Variable Combination (VIAVC), or both. Regulation is shown for the metabolite levels in the GDM group with respect to the obese group. Metabolites for which more than one NMR resonance peak was identified as significant are represented as metabolite.1, metabolite.2, … metabolite.n.

Metabolite	Mann Whitney U Test	VIAVC p-value	Regulation
Creatine, Caffeine.1	1.80E-02	4.04E-241	Up
Sarcosine, Dimethylamine	-	2.81E-96	Down
Maltose, Sucrose	-	5.64E-64	Up
Methionine.2	2.86E-03	2.91E-55	Down
Methionine.1	2.43E-03	6.08E-52	Down
Leucine	-	1.27E-51	Up
Ethanolamine	-	6.79E-50	Up
Methylsuccinic acid.2, 2-Methylglutaric acid.2, 3-Hydroxyisobutyric acid.2, Isobutyric acid.2	2.06E-03	9.64E-48	Up
Alpha-Aminoadipic acid	1.08E-02	5.46E-44	Down
3-Hydroxyisovaleric acid, Pyruvic acid	-	2.96E-40	Up
Methionine.3, Methylamine	1.11E-02	1.15E-31	Down
Methylsuccinic acid.1, 2-Methylglutaric acid.1, 3-Hydroxyisobutyric acid.1, Isobutyric acid.1	3.35E-03	4.62E-31	Up
N-Acetylglutamine.1	-	2.39E-20	Up
1-Methylnicotinamide	-	1.34E-19	Up
Citric acid	-	2.27E-18	Up
Dimethylglycine.2, N-Methylhydantoin	1.53E-03	1.95E-14	Down
Trigonelline.2	-	6.61E-13	Up
Trigonelline.1	-	3.48E-12	Up
Trigonelline.3	-	2.30E-09	Up
2-Hydroxyisovalerate	2.66E-02	-	Down
Alanine	4.74E-02	-	Up
Anserine, 1-Methylhistidine, Dimethylglycine.1	1.33E-02	-	Up
cis-Aconitic acid, Malonic acid	2.20E-02	-	Down
Citraconic acid	3.86E-02	-	Down
Creatinine.1	2.15E-03	-	Down

Creatinine.2	1.42E-02	-	Down
Ethanol, Fucose.3	1.28E-02	-	Down
Fucose.1	3.52E-02	-	Down
Fucose.2	4.74E-02	-	Down
Fumaric acid	3.35E-03	-	Up
Glucuronic acid.1, Glucose.1	1.99E-02	-	Up
Glycine, Glucose.2, Glucuronic acid.2, Fructose.2	3.63E-03	-	Up
Histamine, Glucuronic acid.4, Xylose	2.93E-02	-	Up
Lactic acid, 5-Aminolevulinic acid, Fructose.1	4.09E-02	-	Up
Levulinic acid, Acetone	4.74E-02	-	Down
Methylguanidine	3.11E-02	-	Down
N-Acetylglutamine.2	6.46E-03	-	Down
Phenylacetic acid, myoinositol, Glucuronic acid.3, Caffeine.2	3.52E-02	-	Up
Phenylalanine, Threonic acid	4.47E-02	-	Up
Threonine	8.06E-03	-	Up
Trimethylamine	9.66E-03	-	Down

REFERENCES

Anderson, P. E., Mahle, D. A., Doom, T. E., Reo, N. V., DelRaso, N. J., & Raymer, M. L. (2011, Jun). Dynamic adaptive binning: an improved quantification technique for NMR spectroscopic data. *Metabolomics, 7*(2), 179-190. https://doi.org/10.1007/s11306-010-0242-7

Bahado-Singh, R. O., Akolekar, R., Mandal, R., Dong, E., Xia, J., Kruger, M., Wishart, D. S., & Nicolaides, K. (2012, Oct). Metabolomics and first-trimester prediction of early-onset preeclampsia. *J Matern Fetal Neonatal Med, 25*(10), 1840-1847. https://doi.org/10.3109/14767058.2012.680254

Beckmann, M., Joosen, A. M., Clarke, M. M., Mugridge, O., Frost, G., Engel, B., Taillart, K., Lloyd, A. J., Draper, J., & Lodge, J. K. (2016, Feb). Changes in the human plasma and urinary metabolome associated with acute dietary exposure to sucrose and the identification of potential biomarkers of sucrose intake. *Molecular Nutrition & Food Research, 60*(2), 444-457. https://doi.org/10.1002/mnfr.201500495

Bouatra, S., Aziat, F., Mandal, R., Guo, A. C., Wilson, M. R., Knox, C., Bjorndahl, T. C., Krishnamurthy, R., Saleem, F., Liu, P., Dame, Z. T., Poelzer, J., Huynh, J., Yallou, F. S., Psychogios, N., Dong, E., Bogumil, R., Roehring, C., & Wishart, D. S. (2013). The human urine metabolome. *PLoS One, 8*(9), e73076. https://doi.org/10.1371/journal.pone.0073076

Butte, N. F., Liu, Y., Zakeri, I. F., Mohney, R. P., Mehta, N., Voruganti, V. S., Goring, H., Cole, S. A., & Comuzzie, A. G. (2015, Aug). Global metabolomic profiling targeting childhood obesity in the Hispanic population. *American Journal of Clinical Nutrition, 102*(2), 256-267. https://doi.org/10.3945/ajcn.115.111872

Cao, D., Lu, H., Lewis, T. L., & Li, L. (2007, Dec 14). Intake of sucrose-sweetened water induces insulin resistance and exacerbates memory deficits and amyloidosis in a transgenic mouse model of Alzheimer disease. *J Biol Chem, 282*(50), 36275-36282. https://doi.org/10.1074/jbc.M703561200

Catalano, P. M., McIntyre, H. D., Cruickshank, J. K., McCance, D. R., Dyer, A. R., Metzger, B. E., Lowe, L. P., Trimble, E. R., Coustan, D. R., Hadden, D. R., Persson, B., Hod, M., Oats, J. J., & Group, H. S. C. R. (2012, Apr). The hyperglycemia and adverse pregnancy outcome study: associations of GDM and obesity with pregnancy outcomes. *Diabetes Care, 35*(4), 780-786. https://doi.org/10.2337/dc11-1790

Chatzi, L., Plana, E., Daraki, V., Karakosta, P., Alegkakis, D., Tsatsanis, C., Kafatos, A., Koutis, A., & Kogevinas, M. (2009, Oct 1). Metabolic syndrome in early pregnancy and risk of preterm birth. *American Journal of Epidemiology, 170*(7), 829-836. https://doi.org/10.1093/aje/kwp211

Chong, J., Soufan, O., Li, C., Caraus, I., Li, S., Bourque, G., Wishart, D. S., & Xia, J. (2018, Jul 2). MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Research, 46*(W1), W486-W494. https://doi.org/10.1093/nar/gky310

Chong, J., Wishart, D. S., & Xia, J. (2019, Dec). Using MetaboAnalyst 4.0 for comprehensive and integrative metabolomics data analysis. *Curr Protoc Bioinformatics, 68*(1), e86. https://doi.org/10.1002/cpbi.86

Cnattingius, S., Villamor, E., Johansson, S., Edstedt Bonamy, A. K., Persson, M., Wikstrom, A. K., & Granath, F. (2013, Jun 12). Maternal obesity and risk of preterm delivery. *JAMA, 309*(22), 2362-2370. https://doi.org/10.1001/jama.2013.6295

Cobb, J., Gall, W., Adam, K. P., Nakhle, P., Button, E., Hathorn, J., Lawton, K., Milburn, M., Perichon, R., Mitchell, M., Natali, A., & Ferrannini, E. (2013, Jan 1). A novel fasting blood test for insulin resistance and prediabetes. *J Diabetes Sci Technol, 7*(1), 100-110. https://doi.org/10.1177/193229681300700112

Craig, A., Cloareo, O., Holmes, E., Nicholson, J. K., & Lindon, J. C. (2006, Apr 1). Scaling and normalization effects in NMR spectroscopic metabonomic data sets. *Analytical Chemistry, 78*(7), 2262-2267. https://doi.org/10.1021/ac0519312

Damm, P., Houshmand-Oeregaard, A., Kelstrup, L., Lauenborg, J., Mathiesen, E. R., & Clausen, T. D. (2016, Jul). Gestational diabetes mellitus and long-term consequences for mother and offspring: a view from Denmark. *Diabetologia, 59*(7), 1396-1399. https://doi.org/10.1007/s00125-016-3985-5

Enquobahrie, D. A., Denis, M., Tadesse, M. G., Gelaye, B., Ressom, H. W., & Williams, M. A. (2015, Nov). Maternal early pregnancy serum metabolites and risk of gestational diabetes mellitus. *Journal of Clinical Endocrinology & Metabolism, 100*(11), 4348-4356. https://doi.org/10.1210/jc.2015-2862

Goodpaster, A. M., Romick-Rosendale, L. E., & Kennedy, M. A. (2010, Jun 1). Statistical significance analysis of nuclear magnetic resonance-based metabonomics data. *Analytical Biochemistry, 401*(1), 134-143. https://doi.org/10.1016/j.ab.2010.02.005

Harazaki, T., Inoue, S., Imai, C., Mochizuki, K., & Goda, T. (2014, May). Resistant starch improves insulin resistance and reduces adipose tissue weight and CD11c expression in rat OLETF adipose tissue. *Nutrition, 30*(5), 590-595. https://doi.org/10.1016/j.nut.2013.10.020

Hedderson, M. M., Malone, K. E., Daling, J. R., & White, E. (2003, Sep). Allergy and risk of breast cancer among young women (United States). *Cancer Causes Control, 14*(7), 619-626. https://doi.org/10.1023/a:1025686913626

Ho, J. E., Larson, M. G., Ghorbani, A., Cheng, S., Chen, M. H., Keyes, M., Rhee, E. P., Clish, C. B., Vasan, R. S., Gerszten, R. E., & Wang, T. J. (2016, Feb 10). Metabolomic profiles of

body mass index in the framingham heart study reveal distinct cardiometabolic phenotypes. *PLoS One, 11*(2). https://doi.org/ARTN e0148361 10.1371/journal.pone.0148361

Holmes, E., Loo, R. L., Stamler, J., Bictash, M., Yap, I. K., Chan, Q., Ebbels, T., De Iorio, M., Brown, I. J., Veselkov, K. A., Daviglus, M. L., Kesteloot, H., Ueshima, H., Zhao, L., Nicholson, J. K., & Elliott, P. (2008, May 15). Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature, 453*(7193), 396-400. https://doi.org/10.1038/nature06882

Horton, T. J., Gayles, E. C., Prach, P. A., Koppenhafer, T. A., & Pagliassotti, M. J. (1997, May). Female rats do not develop sucrose-induced insulin resistance. *Am J Physiol, 272*(5 Pt 2), R1571-1576. https://doi.org/10.1152/ajpregu.1997.272.5.R1571

Horvath, B., Bodecs, T., Boncz, I., & Bodis, J. (2013, Jun). Metabolic Syndrome in normal and complicated pregnancies. *Metabolic Syndrome and Related Disorders, 11*(3), 185-188. https://doi.org/10.1089/met.2012.0086

Kanaya, A. M., Grady, D., & Barrett-Connor, E. (2002, Aug 12-26). Explaining the sex difference in coronary heart disease mortality among patients with type 2 diabetes mellitus: a meta-analysis. *Arch Intern Med, 162*(15), 1737-1745. https://doi.org/10.1001/archinte.162.15.1737

Kanehisa, M. (2019, Nov). Toward understanding the origin and evolution of cellular organisms. *Protein Sci, 28*(11), 1947-1951. https://doi.org/10.1002/pro.3715

Kanehisa, M., & Goto, S. (2000, Jan 1). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research, 28*(1), 27-30. https://doi.org/10.1093/nar/28.1.27

Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K., & Tanabe, M. (2019, Jan 8). New approach for understanding genome variations in KEGG. *Nucleic Acids Research, 47*(D1), D590-D595. https://doi.org/10.1093/nar/gky962

Kaplan, B. J., Giesbrecht, G. F., Leung, B. M., Field, C. J., Dewey, D., Bell, R. C., Manca, D. P., O'Beirne, M., Johnston, D. W., Pop, V. J., Singhal, N., Gagnon, L., Bernier, F. P., Eliasziw, M., McCargar, L. J., Kooistra, L., Farmer, A., Cantell, M., Goonewardene, L., Casey, L. M., Letourneau, N., Martin, J. W., & Team, A. P. S. (2014, Jan). The Alberta Pregnancy Outcomes and Nutrition (APrON) cohort study: rationale and methods. *Matern Child Nutr, 10*(1), 44-60. https://doi.org/10.1111/j.1740-8709.2012.00433.x

Kohler, I., Hankemeier, T., van der Graaf, P. H., Knibbe, C. A. J., & van Hasselt, J. G. C. (2017, Nov 15). Integrating clinical metabolomics-based biomarker discovery and clinical pharmacology to enable precision medicine. *Eur J Pharm Sci, 109S*, S15-S21. https://doi.org/10.1016/j.ejps.2017.05.018

Leddy, M. A., Power, M. L., & Schulkin, J. (2008, Fall). The impact of maternal obesity on maternal and fetal health. *Rev Obstet Gynecol, 1*(4), 170-178. https://www.ncbi.nlm.nih.gov/pubmed/19173021

Leung, B. M., Giesbrecht, G. F., Letourneau, N., Field, C. J., Bell, R. C., Dewey, D., & Team, A. P. S. (2016, Feb). Perinatal nutrition in maternal mental health and child development: Birth of a pregnancy cohort. *Early Hum Dev, 93*, 1-7. https://doi.org/10.1016/j.earlhumdev.2015.11.007

Libert, D. M., Nowacki, A. S., & Natowicz, M. R. (2018). Metabolomic analysis of obesity, metabolic syndrome, and type 2 diabetes: amino acid and acylcarnitine levels change along a spectrum of metabolic wellness. *PeerJ, 6*, e5410. https://doi.org/10.7717/peerj.5410

Litwin, L., Sundholm, J. K. M., Rono, K., Koivusalo, S. B., Eriksson, J. G., & Sarkola, T. (2020, Jan). Transgenerational effects of maternal obesity and gestational diabetes on offspring body composition and left ventricle mass: the Finnish Gestational Diabetes Prevention Study (RADIEL) 6-year follow-up. *Diabet Med, 37*(1), 147-156. https://doi.org/10.1111/dme.14089

Lustgarten, M. S., Price, L. L., Phillips, E. M., & Fielding, R. A. (2013, Dec 31). Serum Glycine Is Associated with regional body fat and insulin resistance in functionally-limited older adults. *PLoS One, 8*(12). https://doi.org/ARTN e84034 10.1371/journal.pone.0084034

Maitre, L., Fthenou, E., Athersuch, T., Coen, M., Toledano, M. B., Holmes, E., Kogevinas, M., Chatzi, L., & Keun, H. C. (2014, Jul 11). Urinary metabolic profiles in early pregnancy are associated with preterm birth and fetal growth restriction in the Rhea mother-child cohort study. *Bmc Medicine, 12*, 110. https://doi.org/10.1186/1741-7015-12-110

McCormack, S. E., Shaham, O., McCarthy, M. A., Deik, A. A., Wang, T. J., Gerszten, R. E., Clish, C. B., Mootha, V. K., Grinspoon, S. K., & Fleischman, A. (2013, Feb). Circulating branched-chain amino acid concentrations are associated with obesity and future insulin resistance in children and adolescents. *Pediatric Obesity, 8*(1), 52-61. https://doi.org/10.1111/j.2047-6310.2012.00087.x

Mokkala, K., Vahlberg, T., Pellonpera, O., Houttu, N., Koivuniemi, E., & Laitinen, K. (2020, Jan 1). Distinct metabolic profile in early pregnancy of overweight and obese women developing gestational diabetes. *J Nutr, 150*(1), 31-37. https://doi.org/10.1093/jn/nxz220

Nerenberg, K. A., Johnson, J. A., & Kaul, P. (2012, Jul). PP157. Risk of preeclampsia in pregnant women with gestational diabetes in Alberta. *Pregnancy Hypertens, 2*(3), 323-324. https://doi.org/10.1016/j.preghy.2012.04.268

Ostlund, I., Haglund, B., & Hanson, U. (2004, Mar 15). Gestational diabetes and preeclampsia. *Eur J Obstet Gynecol Reprod Biol, 113*(1), 12-16. https://doi.org/10.1016/j.ejogrb.2003.07.001

Patel, D. P., Krausz, K. W., Xie, C., Beyoglu, D., Gonzalez, F. J., & Idle, J. R. (2017, May 16). Metabolic profiling by gas chromatography-mass spectrometry of energy metabolism in high-fat diet-fed obese mice. *PLoS One, 12*(5). https://doi.org/ARTN e0177953 10.1371/journal.pone.0177953

Perreault, L., Pan, Q., Mather, K. J., Watson, K. E., Hamman, R. F., Kahn, S. E., & Diabetes Prevention Program Research, G. (2012, Jun 16). Effect of regression from prediabetes to normal glucose regulation on long-term reduction in diabetes risk: results from the Diabetes Prevention Program Outcomes Study. *Lancet, 379*(9833), 2243-2251. https://doi.org/10.1016/S0140-6736(12)60525-X

Psychogios, N., Hau, D. D., Peng, J., Guo, A. C., Mandal, R., Bouatra, S., Sinelnikov, I., Krishnamurthy, R., Eisner, R., Gautam, B., Young, N., Xia, J., Knox, C., Dong, E., Huang, P., Hollander, Z., Pedersen, T. L., Smith, S. R., Bamforth, F., Greiner, R., McManus, B., Newman, J. W., Goodfriend, T., & Wishart, D. S. (2011, Feb 16). The human serum metabolome. *PLoS One, 6*(2), e16957. https://doi.org/10.1371/journal.pone.0016957

Rasmussen, K. M., Catalano, P. M., & Yaktine, A. L. (2009, Dec). New guidelines for weight gain during pregnancy: what obstetrician/gynecologists should know. *Curr Opin Obstet Gynecol, 21*(6), 521-526. https://doi.org/10.1097/GCO.0b013e328332d24e

Razak, M. A., Begum, P. S., Viswanath, B., & Rajagopal, S. (2017). Multifarious Beneficial effect of nonessential amino acid, glycine: a review. *Oxidative Medicine and Cellular Longevity*. https://doi.org/Artn 1716701 10.1155/2017/1716701

Riediger, N. D., & Clara, I. (2011, Oct 18). Prevalence of metabolic syndrome in the Canadian adult population. *CMAJ, 183*(15), E1127-1134. https://doi.org/10.1503/cmaj.110070

Rooney, B. L., & Schauberger, C. W. (2002, Aug). Excess pregnancy weight gain and long-term obesity: One decade later. *Obstetrics and Gynecology, 100*(2), 245-252. https://doi.org/Pii S0029-7844(02)02125-7 Doi 10.1016/S0029-7844(02)02125-7

Scott, H., Buchan, M., Chadwick, C., Leung, B., Montina, T., Metz, G. A. S. (2020). Metabolic dysfunction in pregnancy: fingerprinting the maternal metabolome using proton nuclear magnetic resonance spectroscopy. *Endocrinology, Diabetes & Metabolism*. Manuscript submitted for publication.

Szymanska, E., Saccenti, E., Smilde, A. K., & Westerhuis, J. A. (2012, Jun). Double-check: validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics, 8*(1), S3-S16. https://doi.org/10.1007/s11306-011-0330-3

Tai, E. S., Tan, M. L., Stevens, R. D., Low, Y. L., Muehlbauer, M. J., Goh, D. L., Ilkayeva, O. R., Wenner, B. R., Bain, J. R., Lee, J. J., Lim, S. C., Khoo, C. M., Shah, S. H., & Newgard, C. B. (2010, Apr). Insulin resistance is associated with a metabolic profile of altered protein metabolism in Chinese and Asian-Indian men. *Diabetologia, 53*(4), 757-767. https://doi.org/10.1007/s00125-009-1637-8

Wang, M., Athayde, N., Padmanabhan, S., & Cheung, N. W. (2019, Aug). Causes of stillbirths in diabetic and gestational diabetes pregnancies at a NSW tertiary referral hospital. *Aust N Z J Obstet Gynaecol, 59*(4), 561-566. https://doi.org/10.1111/ajo.12936

Wang, T. J., Larson, M. G., Vasan, R. S., Cheng, S., Rhee, E. P., McCabe, E., Lewis, G. D., Fox, C. S., Jacques, P. F., Fernandez, C., O'Donnell, C. J., Carr, S. A., Mootha, V. K., Florez, J. C., Souza, A., Melander, O., Clish, C. B., & Gerszten, R. E. (2011, Apr). Metabolite profiles and the risk of developing diabetes. *Nat Med, 17*(4), 448-453. https://doi.org/10.1038/nm.2307

Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vazquez-Fresno, R., Sajed, T., Johnson, D., Li, C. R., Karu, N., Sayeeda, Z., Lo, E., Assempour, N., Berjanskii, M., Singhal, S., Arndt, D., Liang, Y. J., Badran, H., Grant, J., Serra-Cayuela, A., Liu, Y. F., Mandal, R., Neveu, V., Pon, A., Knox, C., Wilson, M., Manach, C., & Scalbert, A. (2018, Jan 4). HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Research, 46*(D1), D608-D617. https://doi.org/10.1093/nar/gkx1089

Wishart, D. S., Jewison, T., Guo, A. C., Wilson, M., Knox, C., Liu, Y. F., Djoumbou, Y., Mandal, R., Aziat, F., Dong, E., Bouatra, S., Sinelnikov, I., Arndt, D., Xia, J. G., Liu, P., Yallou, F., Bjorndahl, T., Perez-Pineiro, R., Eisner, R., Allen, F., Neveu, V., Greiner, R., & Scalbert, A. (2013, Jan). HMDB 3.0-The human metabolome database in 2013. *Nucleic Acids Research, 41*(D1), D801-D807. https://doi.org/10.1093/nar/gks1065

Wishart, D. S., Knox, C., Guo, A. C., Eisner, R., Young, N., Gautam, B., Hau, D. D., Psychogios, N., Dong, E., Bouatra, S., Mandal, R., Sinelnikov, I., Xia, J. G., Jia, L., Cruz, J. A., Lim, E., Sobsey, C. A., Shrivastava, S., Huang, P., Liu, P., Fang, L., Peng, J., Fradette, R., Cheng, D., Tzur, D., Clements, M., Lewis, A., De Souza, A., Zuniga, A., Dawe, M., Xiong, Y. P., Clive, D., Greiner, R., Nazyrova, A., Shaykhutdinov, R., Li, L., Vogel, H. J., & Forsythe, I. (2009, Jan). HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Research, 37*, D603-D610. https://doi.org/10.1093/nar/gkn810

Wishart, D. S., Tzur, D., Knox, C., Eisner, R., Guo, A. C., Young, N., Cheng, D., Jewell, K., Arndt, D., Sawhney, S., Fung, C., Nikolai, L., Lewis, M., Coutouly, M. A., Forsythe, I., Tang, P., Shrivastava, S., Jeroncic, K., Stothard, P., Amegbey, G., Block, D., Hau, D. D., Wagner, J., Miniaci, J., Clements, M., Gebremedhin, M., Guo, N., Zhang, Y., Duggan, G. E., MacInnis, G. D., Weljie, A. M., Dowlatabadi, R., Bamforth, F., Clive, D., Greiner, R., Li, L., Marrie, T., Sykes, B. D., Vogel, H. J., & Querengesser, L. (2007, Jan). HMDB: the human metabolome database. *Nucleic Acids Research, 35*, D521-D526. https://doi.org/10.1093/nar/gkl923

Xia, J., & Wishart, D. S. (2010, Sep 15). MetPA: a web-based metabolomics tool for pathway analysis and visualization. *Bioinformatics, 26*(18), 2342-2344. https://doi.org/10.1093/bioinformatics/btq418

Yamada, C., Kondo, M., Kishimoto, N., Shibata, T., Nagai, Y., Imanishi, T., Oroguchi, T., Ishii, N., & Nishizaki, Y. (2015, Jul). Association between insulin resistance and plasma amino acid profile in non-diabetic Japanese subjects. *J Diabetes Investig, 6*(4), 408-415. https://doi.org/10.1111/jdi.12323

Yao, Y. L., Robinson, A. M., Zucchi, F. C. R., Robbins, J. C., Babenko, O., Kovalchuk, O., Kovalchuk, I., Olson, D. M., & Metz, G. A. S. (2014, Aug 7). Ancestral exposure to stress epigenetically programs preterm birth risk and adverse maternal and newborn outcomes. *Bmc Medicine, 12*. https://doi.org/ARTN 121 10.1186/s12916-014-0121-6

Yun, Y. H., Liang, F., Deng, B. C., Lai, G. B., Goncalves, C. M. V., Lu, H. M., Yan, J., Huang, X., Yi, L. Z., & Liang, Y. Z. (2015, Dec). Informative metabolites identification by variable importance analysis based on random variable combination. *Metabolomics, 11*(6), 1539-1551. https://doi.org/10.1007/s11306-015-0803-x

CHAPTER 5: Conclusion

This thesis partially demonstrates the potential of ^{1}H-NMR metabolomics in the impending age of personalized medicine, examining the hypothesis that the health outcomes of stress and pregnancy can be examined using fecal metabolomics.. The development of this technique occurs alongside an increasing understanding of the biochemical metabolic processes underlying health and behaviour in both the research laboratory and the clinic. Techniques such as these can be used for the discovery and exploration of previously unknown systems in the body such as the gut-brain axis. Study of the gut-brain axis requires a technique which allows for the analyses of gastrointestinal microbiota, the host, and their points of interaction, which in turn may have important clinical applications for understanding the long-term health consequences of stress and metabolic disorders. The broad scale of fecal NMR metabolomics allows for a snapshot view into the total metabolic processes of this interaction, mitigating any bias that results from the study of a single factor. This thesis provided an overview of the development and current use of fecal metabolomics, alongside a study displaying how ^{1}H-NMR metabolomics can be used to examine the urinary metabolome.

The optimization of fecal NMR metabolomics techniques

The techniques used to perform NMR fecal metabolomics are still inconsistent across the field, necessitating method optimization studies (Deda et al., 2018; Matysik et al., 2016). Methods of sample preparation can vary from simply analyzing unaltered fecal water to performing a customized version of a Bligh-Dyer extraction. If a ^{1}H-NMR fecal metabolomic technique is going to be widely used for clinical applications, however, a single universal method of preparation is essential to ensure accuracy and replicability. Such a technique must primarily be reliable and informative; once these criteria are met, the technique must then be made as simple and cost-effective as possible to reduce the strain metabolomic analysis would place on

the healthcare system. While the research required to improve the performance of an NMR spectrometer are beyond the scope of this thesis, a comparison of some of the most popular methods of sample preparation was performed. The findings indicate that ultrafiltration and post-analysis baseline correction software present the greatest replicability with the least amount of interference from other compounds present in the sample; both techniques are easy to preform and are not to demanding in regard to preparation time. However, pairing *in silico* baseline correction with unextracted samples or techniques such as Bligh-Dyer extraction reduces the cost per sample compared to ultrafiltration, making the performance of fecal metabolomics a more accessible form of analysis. More studies such as this one must be performed on potential methods of sample preparation until a universally approved, applicable, and effective technique arises.

Fecal metabolomics and the study of prenatal stress

The use of NMR fecal metabolomics to examine the impacts of prenatal stress on the maternal postpartum fecal metabolome provides us with a comparison of how metabolomics can complement other methods of analysis such as 16S rRNA gene sequencing. While no impact of stress was found on the female Long-Evans rats in the study; this finding was consistent across both techniques used. The lack of effect is postulated to be due to a previously observed stability in the maternal gut microbiota during the perinatal period caused by concurrent factors in host physiology.

Research and clinical applications of NMR metabolomics

Urinary metabolomics already has a standardized method of sample preparation and analysis, and as such research into the biochemical metabolic underpinnings of health reflected in the urinary metabolome has already begun. A pre-existing sample preparation method and

metabolite database (Wishart et al., 2018) was used to investigate the urinary metabolomes associated with obesity and gestational diabetes mellitus during pregnancy. Samples were collected from individuals prior to the development of the disease state, and it was possible to distinguish between the metabolic fingerprint of each of the gestational diabetes mellitus (GDM), obesity, and control groups. GDM and obesity are already known to have related underlying mechanisms due to their association with metabolic syndrome, a group of factors associated with increased risk of diabetes, high blood sugar, insulin resistance, and other adverse health outcomes (Chatzi et al., 2009). However, it is still unclear why only some individuals with metabolic syndrome will go on to develop negative health outcomes, and why others will not. This study was not only able to distinguish individuals who would develop GDM or obesity from control subjects, but was able to differentiate between the GDM and obesity groups as well. This demonstrates the potential of NMR metabolomics to investigate the mechanisms behind the development of GDM and obesity, and also its ability to identify biomarkers of disease risk for early risk prediction and diagnosis. In the instance where an early lifestyle or medical intervention can prevent the development of a disease, early screening for these biomarkers in at risk populations is invaluable (Kohler et al., 2017).

Overview and concluding statements

This thesis set out to explore how ^{1}H-NMR metabolomics of excreta, such as stool and urine, could be used to study the metabolic consequences of stress and pregnancy, and be used in other research and clinical applications. Various methods of fecal sample preparation were explored to identify which provided the greatest reliability and efficacy. A study in which fecal metabolomics was used to study stress induced gut-host interactions was undertaken and the results were compared with studies using techniques that examined gastrointestinal microbiota

exclusively. Finally, the use of NMR metabolomics in a human clinical research study was undertaken and the results were analyzed to determine what work must occur before this technique is feasible for further applications at the clinical level. Together, these studies support the use of fecal metabolomics to study gastrointestinal microbiota-host interactions at the core of the gut-brain axis, with the eventual goal of providing a minimally invasive screening tool for biomarkers of risk. Further work, however, is required to determine a standard method for processing fecal samples in a cost effective, time efficient, and highly reproducible manner.

REFERENCES

Chatzi, L., Plana, E., Daraki, V., Karakosta, P., Alegkakis, D., Tsatsanis, C., Kafatos, A., Koutis, A., & Kogevinas, M. (2009, Oct 1). Metabolic syndrome in early pregnancy and risk of preterm birth. *American Journal of Epidemiology, 170*(7), 829-836. https://doi.org/10.1093/aje/kwp211

Deda, O., Gika, H. G., & Theodoridis, G. A. (2018). Rat fecal metabolomics-based analysis. *Methods Mol Biol, 1738*, 149-157. https://doi.org/10.1007/978-1-4939-7643-0_10

Kohler, I., Hankemeier, T., van der Graaf, P. H., Knibbe, C. A. J., & van Hasselt, J. G. C. (2017, Nov 15). Integrating clinical metabolomics-based biomarker discovery and clinical pharmacology to enable precision medicine. *Eur J Pharm Sci, 109S*, S15-S21. https://doi.org/10.1016/j.ejps.2017.05.018

Matysik, S., Le Roy, C. I., Liebisch, G., & Claus, S. P. (2016, Nov). Metabolomics of fecal samples: A practical consideration. *Trends in Food Science & Technology, 57*, 244-255. https://doi.org/10.1016/j.tifs.2016.05.011

Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vazquez-Fresno, R., Sajed, T., Johnson, D., Li, C. R., Karu, N., Sayeeda, Z., Lo, E., Assempour, N., Berjanskii, M., Singhal, S., Arndt, D., Liang, Y. J., Badran, H., Grant, J., Serra-Cayuela, A., Liu, Y. F., Mandal, R., Neveu, V., Pon, A., Knox, C., Wilson, M., Manach, C., & Scalbert, A. (2018, Jan 4). HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Research, 46*(D1), D608-D617. https://doi.org/10.1093/nar/gkx1089